Forecasting mechanical properties of concrete containing secondary raw materials

Usama Asif (MS)

Submitted in fulfillment of the requirements

for the degree of Master of Science

in Civil & Environmental Engineering

School of Engineering and Digital Sciences

Department of Civil & Environmental Engineering

Nazarbayev University

53 Kabanbay Batyr Avenue,

Astana, Kazakhstan, 010000

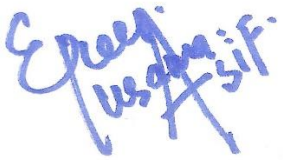
Supervisor: Prof. Shazim Ali Memon
Co-Supervisor: Prof. Jong Kim

March 15, 2024

DECLARATION

"I certify that this manuscript, entitled 'Forecasting mechanical properties of concrete containing secondary raw material,' is my original work, excluding any quoted or cited material that has been appropriately referenced."

"I hereby declare that, to the utmost of my understanding and faith, the work has not been presented, either wholly or partially, for any other academic qualification at Nazarbayev University or any other national or international educational institution, either in the past or at the same time."



Name: Usama Asif
Date: March 15, 2024

Abstract

Climate change due to the significant release of CO_2 from cement industries has become a critical issue worldwide. However, secondary cementitious materials (SCMs), can be used as eco-friendly cement alternatives. Most of the previously conducted studies primarily rely on either experimental investigations or simple regression models to find out the optimal mixture design of concrete made with SCMs. However, in the experimental approach, few tests could be performed for optimization due to time limitations and the availability of resources. On the other hand, simple machine learning (ML) models can't be relied on without extensive validation. Therefore, to overcome these limitations this study aims to use three ML techniques, such as artificial neural network (ANN), adaptive neuro-fuzzy interface system (ANFIS), and gene expression programming (GEP) method with experimental validation for forecasting the compressive strength (CS) and tensile strength (TS) of concrete incorporated with ground granulated blast furnace slag (GGBS) and silica fume (SF) as SCMs. A comprehensive dataset containing the eight most influential inputs of concrete with CS and TS as outputs was collected from the literature and used for model development. The efficiency of the developed model was evaluated using statistical measures and experimental validation. Additionally, sensitivity and parametric analysis were carried out to identify the coherence of developed models with experimental studies. Comparative analysis showed that ANFIS models surpassed other models with higher R^2 and lower errors. Conversely, GEP demonstrated enhanced performance compared to ANFIS and ANN concerning the nearness of statistical measures among the training, validation, and testing sets. Further, GEP also gives predictive formulas that can be utilized for the pre-design of concrete mixtures made with SF and GGBS. Sensitivity and parametric analysis showed strong relevance with experimental studies validating the model's performance. Thus, the recommended models are reliable and can be used to promote the sustainable use of industrial wastes (SF and GGBS) in concrete.

Keywords: Secondary cementitious raw materials; Machine learning; Parametric analysis; Sensitivity analysis; Tensile strength; Compressive strength

Acknowledgments

I wish to express my heartfelt gratitude to my supervisor, **Prof. Shazim Ali Memon**, for his untiring effort, invaluable guidance, and regular meetings that have been instrumental in the successful completion of this thesis. His academic expertise, insightful feedback, and unwavering support have significantly enriched my work quality and helped me develop my research skills.

I want to express my heartfelt appreciation to my co-supervisor, **Prof. Jong Kim**, for his unwavering moral support throughout my thesis. His guidance and encouragement have been invaluable in helping me reach the finish line.

I would also like to thank my friends, **Ayesha Rauf** and **Abdul Moeed Khan**, for their unwavering support, encouragement, and motivation throughout theresearch journey. Their belief in my abilities, and kind gestures have been a great source of inspiration and encouragement.

Lastly, I would like to extend my heartfelt thanks to my **family** for their firm support, steadfast love, and endless prayers throughout my academic journey. Their blessings and encouragementhave been the driving force behind my success.

To summarize, I am profoundly grateful to all individuals who have made significant contributions to the successful completion of this thesis. I deeply appreciate your incredible help, guidance, and motivation, and I am grateful for your unwavering dedication and commitment.

Table of Contents

List of Abbreviations

Compressive strength	CS
Tensile strength	TS
Secondary cementitious raw materials	SCMs
Ground granulated blast furnace slag.	GGBS
Machine learning	ML
Silica fume	SF
Objective function	OF
Random forest	RF
Support vector machine	SVM
Mean absolute error	MAE
Artificial neural network	ANN
Decision tree	DT
Gene expression programming	GEP
Objective function	OF
Multi-expression programming	MEP
Adaptive neuro-fuzzy logic inference system	ANFIS
Deep learning	DL
Root mean square error	RMSE
Performance index	PI
Sensitivity analysis	SA
Coefficient of determination	R^2
Mean square error	MSE

List of Tables

List of Figures

Chapter 1- Introduction

1.1 Background

The construction sector is the primary source of global warming due to the significant CO_2 emission during cement manufacturing [1], [2]. Since the invention of modern-day cement in the early 1800s, its manufacturing has consistently risen as a result of an increase in population, continuously growing urbanization, and extensive infrastructure construction. The global production of cement experienced a significant increase of 54%, rising from 1.6 billion tons in 2000 to 2.55 billion tons in 2006 [3], [4]. In 2021, the worldwide manufacture of cement reached an incredible 4.4 billion tons, resulting in a total of 2.9 billion tons of CO_2 emissions, which is over five times higher than the 0.57 billion tons emitted in 1990, as reported by [5], [6]. The International Energy Agency (IEA) predicts that global cement manufacturing will expand by an additional 12-23% by the year 2050 [7]. According to Hanifa et al. [8], China, India, Europe, and the United States were the leading countries in terms of CO_2 emissions from 2006 to 2021**. Fig 1.1** shows region-wise annual CO_2 emissions from cement industries in the last three decades [9]. It can be seen that Asian countries experienced significant growth in CO_2 emissions, which was 5.3 times more than the initial emissions. In 2019, these countries accounted for 70.3% (1.73 Gt) of the overall CO_2 releases from the cement sector, as opposed to only 32% (0.28 Gt) in 1990. This rise is in line with their growing production of cement. It is known that the production of one ton of cement needs 1.5 tons of raw material, 140kWh electricity, and 4000 MJ energy. Overall, the manufacture of cement accounts for around 8% of the world's manmade CO_2 emissions and consumes around 3% of the world's energy [10].

Cement has a negative long-term impact on the ecosystem due to CO_2 emissions and high energy consumption [11], [12]. Therefore, minimizing CO_2 emissions from the cement manufacturing industry has become a primary concern for researchers [13]. Many approaches could be used to address this problem. One such viable strategy for minimizing the environmental impact associated with the cement industry is the utilization of SCMs as a partial substitute for cement in concrete [14], [15], [16]. The utilization of SCMs as a partial replacement for cement provides numerous benefits, such as higher strength, enhanced durability, financial advantages, and improved sustainability [17].

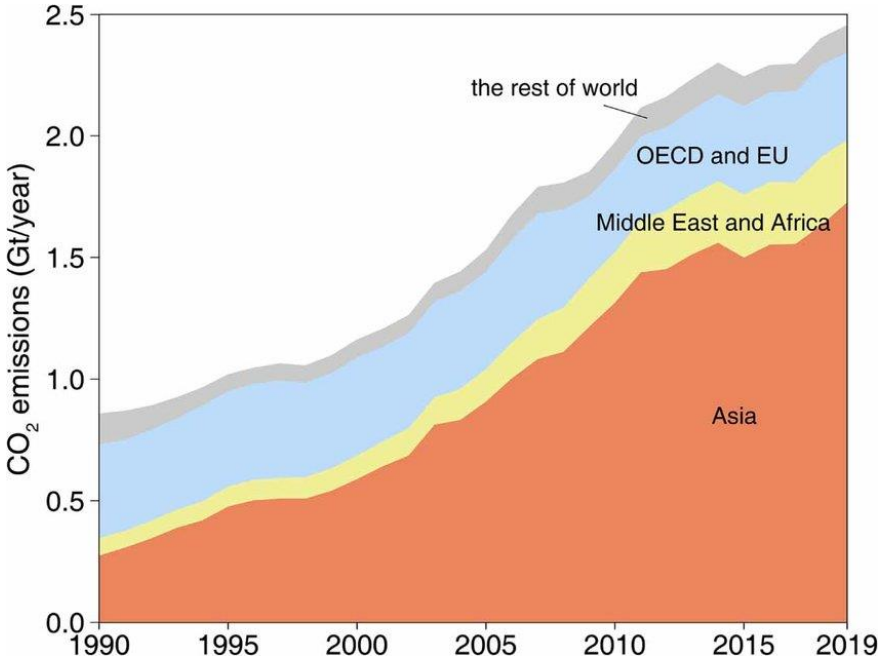


Fig 1.1: Region-wise annual CO_2 from the cement industry.

Industrial by-products such as SF and GGBS are proven to be the most effective SCMs as a substitute for cement in concrete. SF is produced in the smelting process of ferrosilicon or silicon alloy production, whereas GGBS is a by-product of blast furnaces. On average, annually, 1-1.25 million tons of SF and 230-260 million tons of GGBS are produced worldwide [18], [19]. The inadequate management and disposal of these industrial residues can give rise to numerous environmental problems, including the emission of uncontrolled pollutants into the atmosphere during production and transportation, as well as soil and water pollution if appropriate storage and disposal methods are not employed [20]. Furthermore, the excessive accumulation of significant volumes of SF and GGBS waste can exert pressure on waste management systems, resulting in earlier saturation of landfills. This, in turn, can have adverse effects on communities. Therefore, the implementation of efficient methods for handling waste and the exploration of sustainable applications, such as incorporating them as SCM in concrete, can be beneficial by reducing environmental concerns associated with SF, GGBS, and the cement industry. For instance, Akhtar et al. [21] investigated that substituting 10% of SF in 2.2 billion tonnes of concrete decreased CO_2 emissions by 39 million tonnes. In another study [22], it has been stated that replacing 50% of GGBS in concrete with OPC can reduce around 0.5 tonnes of CO_2 emissions per ton of concrete. Considering the environmental drawbacks of the cement industry and the benefits of employing SCMs in concrete, it is imperative to develop robust and reliable methods for sustainable construction practices.

Recent advances in machine learning (ML) techniques led to the development of precise and reliable models for predicting the properties of various cement-based composites containing SCMs [23]. Recently, ANN and ANFIS ML approaches have been used successfully to identify and generalize complex patterns in data. Hence, they can be efficiently utilized to estimate the properties of concrete. Shahmansouri et al. [24] employed ANN models to predict the CS of geopolymer concrete (GPC) using an experimental database of 117 specimens from 39 different mixtures containing varying proportions of GGBS, SF, and natural zeolite. They compared the ANN estimated CS and actual results and observed that the ANN model accurately predicted the CS of GPC. Topçu et al. [25] explore ANN and fuzzy logic algorithms in order to accurately predict the CS and TS of recycled aggregate concrete incorporated with SF. The models were constructed using experimental data from 210 specimens with 35 different mix designs. For the purpose of prediction, the study utilized eight input parameters and examined the strength properties at various time intervals (3, 7, 14, 28, 56, and 90 days). The outcomes of the models indicate that ANN and fuzzy logic algorithms possess considerable potential in accurately forecasting the strength properties of concrete. Dao et al. [26] focused on predicting the CS of GPC through the utilization of two ML techniques: ANN and ANFIS. For the model development, 210 samples were used. The output of ANN and ANFIS were evaluated based on several statistical measures, such as R^2 and MAE. The findings indicate that both models exhibit substantial potential in anticipating the CS of GPC. However, the ANFIS outperforms the ANN model with a marginally lower MAE of 1.655 MPa compared to 1.989 MPa and a higher R^2 of 0.879 compared to 0.851. Overall, the study provides evidence to support the efficacy of ANFIS and ANN as predictive tools for estimating the CS of GPC. Using ANN, Sadowski et al. [27] projected the CS of concrete comprising waste quartz dust. The ANN models revealed strong correlation values (R) of 0.93, 0.91, and 0.94 for the training, testing (evaluation), and validation stages, respectively, indicating that the CS of green concrete utilizing waste mineral dust can be evaluated accurately using these models. Despite the reliable correlation proposed by these models, no empirical formula was suggested for practical use. This is because of the sophisticated architecture of ANN and ANFIS, which is also regarded as a significant barrier to the widespread adoption of ANN and ANFIS models [28]. Multi-collinearity is also another drawback of these modeling techniques. Furthermore, the ANNs and ANFIS are categorized as "black-box models" due to their incapability to demonstrate the fundamental physical mechanisms, limited transparency, and inability to generate closed-form prediction equations [29].

Evolutionary modeling approaches such as GEP, influenced by Darwin's idea of natural selection, have gained the attention of researchers in recent times. Unlike "black-box" models such as ANN and ANFIS, GEP gives predictive equations that provide valuable information about the connections between input parameters and the desired output [30]. The GEP-extracted mathematical formulas can be used in practice with improved prediction efficiency. Nafees et al. [31] used MLPNN, GEP, and ANFIS to forecast the CS and TS of concrete incorporated with silica fume by utilizing a database comprising 283 and 145 data points for CS and TS, respectively. It was reported that all three models showed a notable level of accuracy in their predictive capabilities. Nevertheless, the GEP models demonstrate their superiority by yielding significantly higher R^2 values than other methods. According to the authors, the GEP method developed predictive equations for each outcome, which can be used to pre-design SF concrete mixtures in the future. Awoyera et al. [32] employed two ML models, GEP and ANN, to predict the flexural, CS, and TS of self-compacted geo-polymer concrete. The prediction outcome was compared using statistical values of mean square error (MSE) and R^2. The author concluded that GEP models are more reliable as they gave fewer errors and provided the empirical relationships for forecasting properties of concrete. Shahmansouri et al. [33] developed numerical models using GEP to predict the CS of geopolymer concrete containing GGBS. A total of 351 data records, with the five most important input parameters, were employed to create the models. The efficacy and generalization ability of the models was assessed based on sensitivity and parametric analysis. The study's findings indicate that GEP models can be used as an effective tool for promoting sustainability in the construction sector as they provide robust and strong empirical correlations.

Based on the aforementioned literature analysis, it is evident that numerous researchers have employed ML methodologies to forecast the mechanical characteristics of concrete that incorporate supplementary cementitious materials (SCMs) by employing traditional machine learning techniques. The aforementioned investigations were characterized by certain shortcomings, including a smaller number of data points, a less robust connection among the controlling factors, the non-availability of GEP-based predictive equations, and a lack of comparative analysis and experimental validation. Furthermore, as per the author's knowledge, the literature about predicting the performance of concrete utilizing binary SCMs (silica fume and GGBS) is not available. Therefore, for the first time, this research employed three distinct computational methodologies, including ANN, ANFIS, and GEP, to develop

models that can accurately and precisely forecast the CS and TS of concrete incorporated with SF and GGBS. The models' predictive performance was assessed using various statistical performance measures and experimental testing for cross-validation in the laboratory. Moreover, sensitivity and parametric analysis were conducted to determine the influence of input variables on the CS and TS of concrete incorporated with SF and GGBS. The availability of reliable and precise ML techniques to anticipate the properties of concrete containing SF and GGBS will promote sustainability and save cost and time.

1.2 Objectives

The presented study includes the following objectives.

i) To develop prediction models based on GEP for predicting the mechanical properties of concrete incorporated with SF and GGBS.

ii) To compare the predictive capabilities of the GEP model with those of the ANN and ANFIS models and identify the most influential computational model for forecasting the mechanical properties of concrete containing SF and GGBS.

iii) To cross-validate the developed models by using external experimental validation, statistical metrics, sensitivity, and parametric analysis.

Chapter 2- Literature Review

2.1 Overview

This chapter provides an extensive literature review. The first part will introduce SCMs such as SF and GGBS, followed by a comprehensive overview of the applications of SF and GGBS in concrete. The subsequent part will provide applications of ML techniques in civil engineering, particularly in predicting the various properties of concrete.

2.2 Secondary cementitious materials (SCMs)

Incorporating SCMs in concrete is crucial for promoting sustainable construction methods, as it contributes to environmental preservation and financial viability. The utilization of SCMs as a partial substitute for cement has two major advantages one related to sustainable use of industrial waste and other to reduce carbon footprint associated with cement industries. These materials also contributes to enhanced strength and durability properties of concrete [17]. The advantages of SCMs in concrete are shown in **Fig.2.1**. In this study, industrial by-products, including SF and GGBS, were used as SCMs in concrete and are discussed below.
SF and GGBS

Fig.2.1: Advantages of SCMs in concrete.

2.1.1 Silica fume (SF)

Silica fume (SF), or condensed silica fume, is a by-product derived during silicon metal or ferrosilicon alloy production [34]. The production of SF occurs when the raw materials, such as quartz and coal, undergo a high-temperature treatment in a furnace operating at approximately 2000℃. The furnace is equipped with two outlets, one of which is responsible for the discharge of liquid metal, while the other is responsible for the emission of off-gases. The liquid metal undergoes refinement and processing to obtain various silicon products, while the resulting off-gases are directed through both the boiler and the arc furnace. The effluent from the arc furnace undergoes condensation, forming a finely powdered substance known as silica fume. **Fig.2.2** displays the visual depiction of the process involved in the production of SF [35]. Annually, approximately 100,000 metric tons of SF are generated globally as a by-product of industrial processes [36]. The utilization of this by-product is imperative, as the disposal of SF as waste poses significant health risks to human beings. SF, being a considerably finer substance, can be transported into the surrounding environment through the process of airflow.

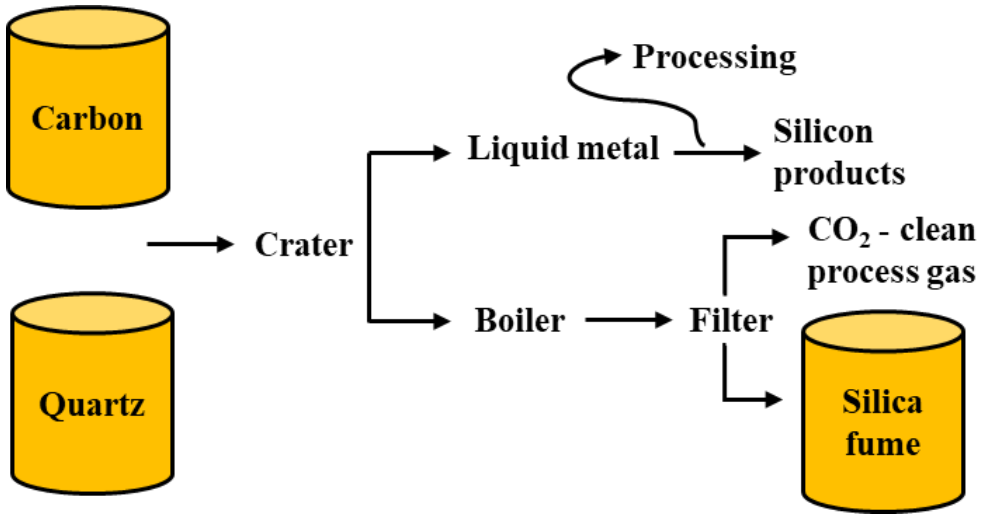


Fig.2.2: Production process of SF

2.1.2 GGBS

Ground Granulated Blast Furnace Slag (GGBS) is a major industrial by-product generated during steel and iron manufacturing processes. The furnace is commonly operated at a heat of 1500 degrees Celsius, which provides a thorough-controlled mixture of limestone, iron ore, and coke. When ore, limestone, iron, and coke are combined and subjected to high temperatures in a blast furnace, a chemical reaction occurs, and iron and slag are formed in a

molten state. The molten slag produced by the blast furnace undergoes rapid cooling through the application of high-pressure water jets, resulting in the formation of GGBS, a finely textured, granular material with a glass-like appearance. **Fig.2.3** illustrates the manufacturing process of GGBS.

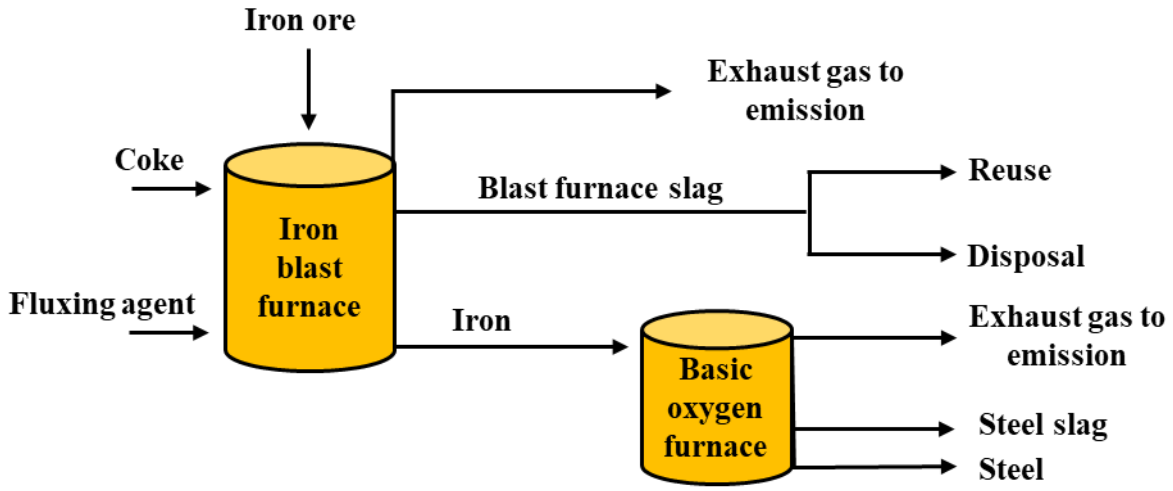


Fig.2.3: Production process of GGBS

2.1.3 Role of Silica fume in concrete

Numerous investigations have demonstrated that including SF in concrete as a partial replacement for cement can be a viable option to reduce the environmental impacts associated with the disposal of SF and the production of cement. The pozzolanic qualities and filling influence of SF particles increase their widespread use in concrete. The reactivity of amorphous silicon dioxide enables its reaction with calcium hydroxide (CH), resulting in the formation of calcium silicate hydrate (CSH). This hydration product is crucial for binding and mechanical properties in cement-based composites. Due to their small dimensions, SF particles effectively occupy the voids within the solidified cement matrix, resulting in enhanced strength and density of the paste [37].

Tak et al. [38] studied the feasibility of incorporating SF as a partial replacement for cement in concrete. Five mix proportions were casted with various SF proportions, 5%, 8%, 11%, and 15%, including the control mix. Concrete specimens were tested at 7, 14, and 28-day curing periods. The CS results of the control mix were compared with modified concrete. Statistical methods were also employed to determine the optimal proportion that yields the highest CS values. The findings indicate that using SF as an SCM in concrete improves its mechanical qualities and reduces the carbon footprint, making the resulting concrete more economical and eco-friendly. The optimal proportion of SF in

concrete was found to be 11%.

Smarzewski [39] evaluated the fracture strength of high-performance concrete (HPC) utilizing SF as an SCM. The six HPC mixtures with a constant water-to-binder ratio of 0.25 and varying replacement levels of SF (0-25%) were prepared to examine the changes in mechanical properties, including characteristic length, CS, tensile strength, elastic modulus, and rupture energy of HPC. The incorporation of SF was shown to improve the mechanical characteristics of the HPC significantly. The inclusion of a 10% SF content led to a significant enhancement of 26% in TS, a 13% rise in CS, and a 5% elevation in elastic modulus. However, it is important to mention that the fragility of high-performance concrete (HPC) was found to be enhanced in the presence of SF. The research outcomes indicate that the utilization of SF as a substitute for cement can yield significant improvements. However, it is suggested that care must be taken to replace the percentage to achieve an appropriate balance between enhanced material characteristics and the risk of increased brittleness. This study highlights the possibility of improving HPC efficiency by including silica fume. It emphasizes the significance of determining the ideal SF content to achieve the required mechanical characteristics.

Ahmed et al. [40] investigated the impacts of incorporating SF in concrete primarily composed of natural pozzolan, a material obtained from volcanic rock. The SF was added with multiple trial combinations to enhance the overall efficacy of the concrete composed of natural pozzolan. The durability of the concrete was evaluated by assessing several concrete qualities, such as CS, drying shrinkage, penetration of water depth, chloride diffusion coefficient, carbonation depth, reinforcement corrosion, and resistance to sulfate and salt exposure. The results of the study demonstrated a notable improvement in the concrete characteristics when SF was incorporated. Despite a marginal decrease in CS at an early age, the incorporation of SF effectively compensated for this effect. Moreover, there was a noticeable enhancement in the concrete's long-term durability, but the shrinkage level remained relatively the same. Notably, the observed elevation in carbonation depth was consistently below the conventional cover thickness employed for preserving reinforcing steel. Consequently, concerns regarding the corrosion of reinforcement can be dismissed. This research highlights the capacity of silica fume to enhance the effectiveness and longevity of cement concrete that is based on natural pozzolan. This study presents an achievable possibility for the advancement of construction materials that are both sustainable and high performing.

Ajileye [41] conducted an experimental study by utilizing SF in various proportions (0%, 5%, 10%, 15%, 20%, and 25%) in C30 grade concrete. The cubes were casted and tested for density and CS at 3, 7, 14, and 28 days. It was observed that the CS of concrete was reduced when the replacement ranged from 15%-25%. Likewise, density was also decreased by using 10% SF as a replacement for cement. However, it was noted that CS enhanced by 15.31% at 10% SF content.

Mazloom et al. [42] assessed the effects of varying levels of SF substitutions for cement in concrete (0%, 6%, 10%, and 15%) with a constant W/C of 0.35 and binder of 500 kg/m^3. A constant slump value of 100 ± 10 mm was used in mixture preparation. The findings indicate that increased amounts of SF resulted in a decrease in workability while simultaneously enhancing mechanical characteristics, specifically the 28-day CS and secant modulus. The utilization of SF resulted in a negligible effect on the overall shrinkage while leading to a rise in internal shrinkage. The presence of increased quantities of SF also resulted in a reduction in basic creep, as seen by the lower expansion observed in swelling tests. It was also observed that the inclusion of increasing levels of SF replacement resulted in an increased demand for superplasticizer doses.

Pradhan & Dutta [43] examine the mechanical characteristics of concrete utilizing SF as a partial replacement of binding material. The study includes the production of five distinct concrete mixtures with differing proportions of SF (0%, 5%, 10%, 15%, and 20%), keeping a consistent W/B ratio of 0.4. The experimental testing was performed on cubes of 100 mm and 150 mm in size. The CS of these cubes was evaluated at three different time intervals: 24 hours, 7, and 28 days. The findings of the study demonstrate that the CS of 100 mm cubes consistently outperforms those of 150 mm cubes throughout all periods and in all mix designs. The maximum CS was observed when there was a 20% substitution of SF, followed by a subsequent decrease in strength. It was noted that the 28-day CS exhibits a substantial rise of 37% and 28% for 150 mm and 100 mm cubes, respectively, in comparison to conventional concrete.

To summarize, the addition of SF to concrete improves its durability, strength, and ability to withstand external influences. However, it is important to have a careful mix design and proper handling management. SF is an essential element in sustainable concrete production, as it aids in the development of durable infrastructure and minimizes negative impacts on the environment [44].

2.1.4 Role of GGBS in Concrete

GGBS has been successfully employed as SCM in concrete. GGBS has natural hydraulic characteristics, hence supporting an extended duration of strength enhancement, making it particularly advantageous for the construction of durable infrastructures. The small particle size and lack of crystalline structure of this material contribute to its improved workability and decreased likelihood of segregation in concrete. Moreover, utilizing GGBS helps reduce the exothermic reaction during the hydration process of concrete, hence diminishing the potential for thermal-induced cracking. The aforementioned characteristics, when paired with its ecologically favorable attributes as a reused industrial byproduct, make GGBS a viable option for improving the performance of concrete and mitigating the carbon emissions associated with the construction sector.

Ganesh et al. [45] examined the potential of GGBS as a substitute for cement in concrete. They prepared five mixtures of Ultra-High-Performance Concrete (UHPC) with varying percentages of GGBS (0%, 20%, 40%, 60%, and 80%) to assess the CS, tensile strength, and durable properties of concrete. A range of qualities, such as flowability, CS, tensile strength, rupture, and durability, were assessed using conventional water and high temperature curing conditions. The findings indicate that the utilization of GGBS-based UHPC leads to substantial enhancements in the strength properties when replacing cement by up to 40% using normal water curing and up to 60% substitution using higher-temperature curing.

Rama & Garikipati [46] investigated the properties of M20 and M40 grade concrete when subjected to different levels of cement substitution by GGBS at percentages of 30%, 40%, and 50%. A series of experiments were performed to evaluate their flexural strength, CS, and split tensile strength. Durability properties were also examined by subjecting the materials to hydrochloric and sulphuric acid. They discovered that a rise in the GGBS content positively impacts the workability of concrete. Further, incorporating GGBS leads to an enhancement in flexural strength, CS, and split tensile strength. The maximum strength values were achieved at a GGBS replacement level of 40%. It was also noted that concrete containing GGBS exhibits improved durability against acid attacks, particularly at an optimum 40% GGBS substitution level. These findings indicate that GGBS incorporation in concrete can enhance mechanical characteristics and improve resistance against acid corrosion, providing considerable implications for the sustainable building material.

Chidiac & Panesar [47] conducted an experimental study to evaluate the impact of incorporating GGBS as a partial substitute for cement on ultrasonic pulse velocity (UPV) and

CS at (1, 3, 7, and 28 days) curing durations. The proportion of GGBS (0%, 10%, 20%, 30%, 40%, 50%, and 60%), the overall amount of binder used (270 kg/m^3 to 450 kg/m^3), W/B, which was either 0.31 or 0.38, and the curing time were considered as key factors. The results of the study indicate that the incorporation of GGBS as an alternative to cement greatly improves the properties of concrete, especially CS, at 28 days of curing when up to 50% of the binder is replaced with GGBS, and W/B of 0.38 was used. The research also highlighted the significance of microstructural uniformity, density, and CS of GGBS particles. These particles were found to have a noticeable effect on pore volume, hydration level, and durability against expansion, particularly after 28 days. However, it is recommended to reassess the 28-day CS in concrete that incorporates GGBS by considering other factors like varying W/B ratios and properties of concrete.

The study by [48] explores the utilization of GGBS as an environmentally friendly partial substitute for cement in concrete. They investigated the effects of different proportions of GGBS (0-50%) as a substitute for cement in concrete. A total of 150 concrete samples were tested in accordance with the standards outlined in the IS Code. The results of the study indicate that the incorporation of 50% GGBS in M25-grade concrete resulted in enhanced workability. It was observed that concrete mixes with 20% GGBS had superior CS and tensile strength properties, whereas with 35% GGBS had higher flexural strength compared to traditional concrete mixes. Mathematical formulations were created to forecast strength values corresponding to varying amounts of GGBS substitution; ultimately, 20% of GGBS was found to be the optimal choice for achieving the desired strength in M25 grade concrete. These findings show that GGBS can be used as a sustainable substitute for cement in concrete without affecting its structural integrity.

Arivalagan [49] studied the efficacy of GGBS at various proportions (20%, 30%, and 40%) as a partial substitute in M35 grade concrete for traditional cement. The cubical and cylindrical shaped specimens were casted to investigate the CS and tensile strength of GGBS-concrete at 7 and 28 days curing duration. It was observed that GGBS concrete exhibited reduced early strength due to its smaller grain size. However, it was noted that GGBS concrete possesses higher later-age strength values than ordinary concrete. The optimum substitution was found to be 20%, which demonstrates a notable increase in strength properties, decreased heat of hydration, enhanced workability, durable properties, and cost-efficiency.

In another study, Arivalagan [50] performed an experimental investigation to evaluate the effectiveness of incorporating GGBS in concrete. M-20 grade of concrete with varied

proportions of GGBS substitution (10%, 20%, 30%, and 40%) was used to assess the mechanical characteristics of GGBS-concrete. The experimental findings indicate that the inclusion of GGBS as a substitute for cement in concrete has the potential to improve its durability and strength properties with observed enhancements ranging from 10% to 30%. The optimum replacement was noted as a 30% GGBS proportion in concrete.

In conclusion, the addition of GGBS to concrete provides significant benefits such as higher durability, decreased permeability, and improved resistance to corrosive substances and environmental factors [51]. However, it is crucial to carefully consider the mixed proportions and handling processes. Utilizing GGBS in the manufacturing of concrete is in line with sustainable construction methods [52].

2.3 Applications of machine learning in concrete

Various machine learning (ML) techniques, including GEP [53], support vector machine (SVM) [54], multilayer perception neural network (MLPNN) [55], deep learning (DL) [56], ANFIS, ANN, multi-expression programming (MEP) [57], random forest (RF), AdaBoost (AB), multi-linear regression (MLR), and gradient boosting machine (GBM), have been successfully adopted to assess the mechanical characteristics of concrete incorporated with SCMs [58], [59], [60], [61], [62], [63], [64]. In this study, neural network-based (ANN and ANFIS) and genetic-based (GEP) modeling techniques were used to forecast the CS and TS of concrete made with SF ad GGBS.

2.3.1 Conventional neural network models

Neural network-based techniques such as ANN and ANFIS have been successfully used to evaluate various concrete properties. They both work on the neural structure of the human brain with the ability to learn complicated patterns and connections present in data. However, ANFIS combines the principles of fuzzy logic with neural networks to improve its ability to change and interpret. These models received instructions on past data sets to forecast concrete characteristics by learning complex relationships among various components that affect concrete performance. ANN and ANFIS have various concrete applications, including the prediction of important parameters such as CS, TS, and slump. A few of the applications of ANN and ANFIS are discussed below.

Dao et al. [65] employed the ANN technique to forecast the CS of foamed concrete. They achieved the R^2 value of 0.972 in the testing stage, which is evidence of the higher accuracy of ANN. Beycioɨlu et al. [66] explored the ANN and MLR to anticipate the CS of clinker mortars. ANN exhibited a strong relationship with actual findings, indicating a potential

alternative method for assessing the CS of clinker mortar by utilizing corresponding input variables. Furthermore, the MLR model has demonstrated inferior predictive performance compared to the ANN.

Amlashi et al. [67] used three soft computing models: M5 model tree (M5MT), ANN, and multivariate regression to estimate elastic modulus, CS, and slump of plastic concrete. The findings indicate that ANN models had superior accuracy in predicting all three parameters compared to the other models. Mahesh et al. [68] used ANN models to estimate the elastic modulus and CS of fiber-reinforced concrete. The models were developed using 158 and 140 experimental results for elastic modulus and CS, respectively. The outcomes of the study show the superior predictive capabilities of the ANN, as it exhibits minimal variation from the actual values with the R^2 values of 0.96 and 0.97 for elastic modulus and CS models, respectively.

Duan et al. [69] focused on the viability of employing ANN to anticipate the modulus of elasticity (Ec) of RAC. The establishment of ANN models involved using a database consisting of 324 data records collected from 21 literature studies. The dataset was partitioned into three distinct subsets, namely training (learning), testing, and validation, using a random allocation method. The anticipated outcomes were compared with empirical observations and models derived from traditional regression analysis. In addition, the authors applied developed models to various experimental databases, including those in the database as well as independent datasets. This was done to assess the adaptability of these ANN models to anticipate the Ec of RAC acquired from different sources. The results indicate that the ANN models developed in this study had a robust capability to effectively forecast the Ec of concrete, including recycled aggregates from diverse origins.

Rehman et al. [70] developed ANN models to accurately forecast geopolymer concrete's mechanical and rheological characteristics (GPC). For the development of the ANN models, 381 records of data, including 13 input and four output parameters, were considered. The model demonstrated notable efficiency in the training set, as evidenced by the MAE values of 0.121 MPa, 2.53 MPa, and 0.72 and 8.9 mm for flexural, compressive, elastic modulus, and slump models. Whereas for these four models, the testing set demonstrated marginally elevated mean absolute error (MAE) values of 0.65 MPa, 4.32 MPa, 1.5 MPa, and 19.7 mm, hence suggesting the model's capacity for generating accurate predictions. This study makes a significant contribution to the optimization of GPC design, thus minimizing the necessity for extensive experimental trials and providing vital insights for actual

implementation.

Van Dao et al. [71] examined the application of geopolymer concrete (GPC) incorporating waste steel slag as aggregates as a partial substitute for OPC in concrete. They used ANN and ANFIS to estimate the CS of GPC. The mix designs were created to achieve the desired CS after 28 days. Both the ANFIS and ANN models exhibited positive results in forecasting CS. The sensitivity analysis results revealed that modifying a single input variable had a negligible effect on the accuracy of the prediction. In summary, the study's findings indicate that artificial intelligence models exhibit high accuracy in predicting the CS of GPC, thereby establishing their potential as a feasible substitute for various construction purposes.

Despite excellent prediction outcomes using ANN and ANFIS models, there are certain shortcomings in these models, such as interpretability, which limits the comprehension of the rationale behind their predictions. In addition, overfitting can be another major concern in neural network-based models. Therefore, it is recommended that the findings of ANN and ANFIS be compared with other ML models, such as GEP and MEP, which are capable of providing insights into the predicting mechanism.

2.3.2 Genetic-based models

Genetic-based models such as GEP is a robust computational methodology that integrates genetic algorithms with expression trees to create mathematical models for many applications, such as forecasting specific characteristics. GEP provides more transparent, comprehensible models by iteratively refining both the structure and parameters of mathematical expressions. In addition, GEP demonstrates a strong ability to generalize, reducing the probability of overfitting and ensuring accurate predictions even with a small amount of training data. In recent years, GEP has been successfully employed to estimate the various properties of concrete [72].

For instance, Alabi et al. [73] Implemented GEP modeling to promote the utilization of SCMs in concrete. The prediction model was constructed using dependable data on CS in ternary mixed cement concrete. The model incorporated four key input factors, namely, rice husk ash (RHA), coal waste powder (CWP), glass waste powder (GWP), and the number of days for curing (CD). The development and reliability assessment of the model was done by using 70% of the data for training, while the remaining 30% was allocated for testing and experimentation purposes. Additionally, statistical and sensitivity analysis were conducted to ensure the consistency and validity of the results. The efficacy of the GEP model is strong,

as evidenced by the R2 of 0.95. The proposed model provides an analytical formula that accurately estimates the CS of sustainable concrete. This approach can be used in the construction industry to improve the usage of pozzolanic materials in concrete.

In one other study Murad [74] utilized the GEP model to anticipate the CS of concrete modified with carbon nanotubes (CNT), nano-clay (NC), nano aluminum (NA), and nano-silica (NS). A comprehensive database of 94 test points was used to create a GEP model considering two (CNT and NS) and four (CNT, NS, NA, and NC) key features of modified concrete. The models demonstrate robust efficiency, as evidenced by their low MAE of 4.6% and 2.9% for the GEP-2 and GEP-4 models, respectively. In addition, a comprehensive examination confirms the validity of the GEP models, highlighting their precision in forecasting the CS of concrete. Parametric investigation revealed that the strength exhibits an upward trend with the inclusion of CNTs, NS, and NC, whereas it decreases with the presence of NA. The GEP modeling performance shows that predicted outcomes are well aligned with experimental findings, hence validating GEP models.

Alabduljabbar et al. [75] collected a comprehensive set of data conducted on UHPC consisting of 810 data records of CS, along with 15 significant input variables to anticipate the CS of the concrete using GEP. The efficacy of the predictive algorithm was thoroughly assessed using statistical metrics, regression, and SHAP analysis. The SHAP study revealed that three parameters, including age, fiber, and SF, exhibited notable effects on CS. The outcomes of this study provided closed-form predictive equations that demonstrate the robustness and efficiency of the GEP model, providing valuable practical knowledge for designing UHPC mixtures. However, to increase validity, the authors recommended comparing the results of GEP with other ML-based models.

Ashrafian et al. [76] utilized a GEP technique to construct novel mathematical models for the CS of Roller-Compacted Concrete Pavement (RCCP). The authors employed an extensive and reliable database consisting of 235 experimental instances collected from the literature. The efficiency of genetic models was evaluated through the use of many metrics, including error and performance measures. Furthermore, the findings were validated through the use of sensitivity and uncertainty analysis. The sensitivity analysis revealed that the sand and W/B ratio were the most important factors. The GEP-based models that have been proposed exhibit characteristics of simplicity, robustness, and practicality. These models provide novel and useful formulations for the prediction of the CS of RCCP.

2.4 Summary

To summarize, the extensive examination of the literature on the use of SCMs in concrete has emphasized many advantages, such as improved mechanical properties of concrete, reduced CO_2 emissions associated with cement, and elimination of industrial wastes. While experimental investigations have experienced difficulties in determining the optimum mix designs due to time and resource constraints. However, machine learning models developed as a possible alternative, demonstrating outstanding predictive capabilities. Notably, a crucial gap in the existing literature is the lack of studies specifically examining concrete, including silica and GGBS. The present study addresses this research gap by combining three different ML algorithms with experimental validation to precisely predict the CS and TC of concrete incorporated with SF and GGBS. Moreover, the research aims to provide GEP-based predictive equations, which can be used as important prediction tools for concrete mix design. The succeeding chapter will provide a detailed discussion on the methods used to construct and validate these models, highlighting the reliability and effectiveness of the recommended method.

Chapter 3- Research Methodology

3.1 Overview

This chapter describes the process used to develop the modeling techniques for forecasting the CS and TS of concrete containing SF and GGBS. Initially, a detailed description of the data collection is provided, followed by a general review of ML techniques (ANN, ANFIS, and GEP) employed in this study. Thereafter, performance measurement and experimental validation criteria for the models are discussed. **Fig. 3.1** shows a detailed overview of the methodological procedure followed in this research.

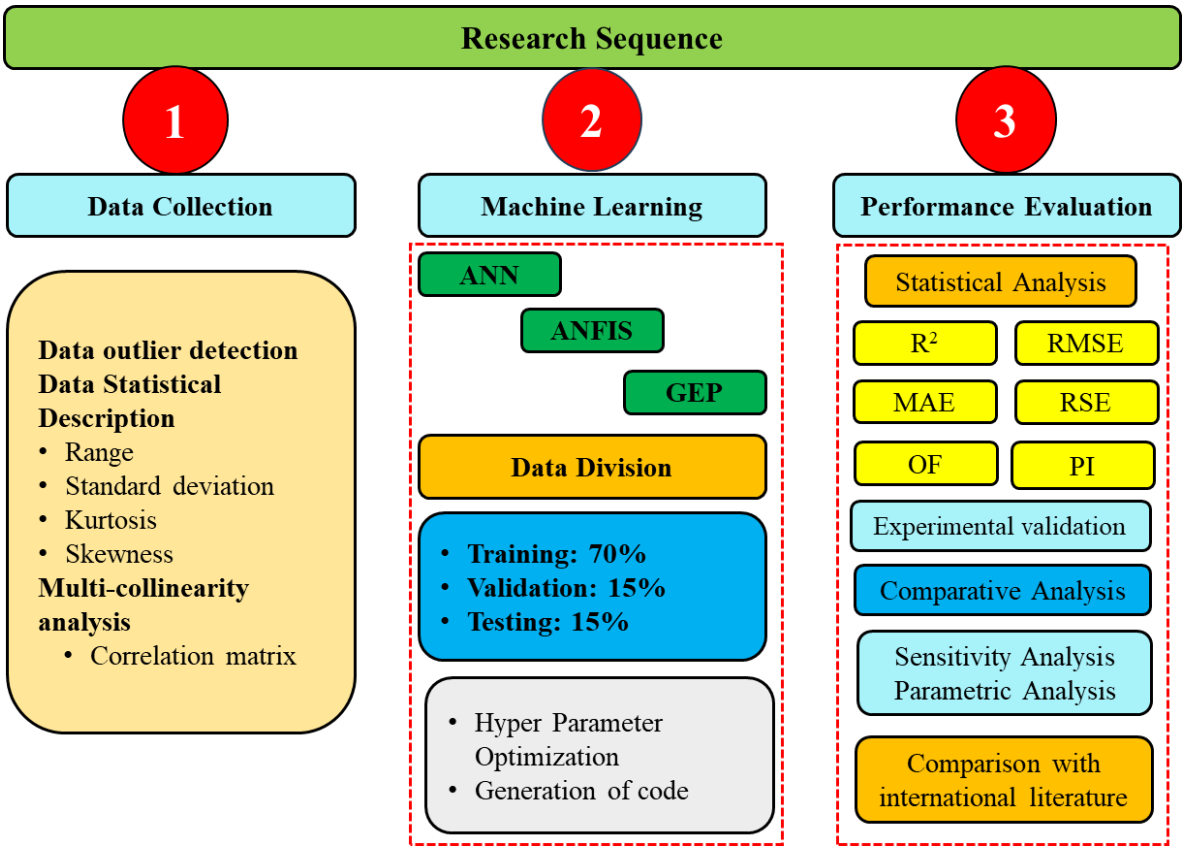


Fig. 3.1: Flow diagram of the methodology

3.2 Data collection

An extensive database of SCMs-based concrete was compiled from articles published in international journals [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [88], [89], [90], [91], [92], [93], [94], [95], [96], [97], [98], [99], [100], [101]. The collected database was mainly obtained from concrete mixes incorporated with a combination of SCMs (SF and GGBS). While developing the database, care was taken to select those data points that provide comprehensive details regarding the mix design and sizes of the specimens used. The total data set consists of 648 data points for CS and 245 data points for TS. The database comprised

two types of concrete samples, i.e., cylinders and cubes. The reported database for CS contains 248 data points obtained from cubic specimens with a size of 150 mm and 400 data points obtained from standard-size cylindrical samples (150 mm Dia and 300 mm height). However, according to previous experimental investigations, the length-to-diameter ratio (L/D) affects the strength properties of concrete [102]. Therefore, for the sake of homogeneity, all the data points were normalized in standard cylindrical form. A normalization factor of 0.8 was used for cubes of 150 mm in size [103]. Similarly, the collected database for TS contains 245 data points obtained from standard-size cylindrical specimens. In this study, eight major input parameters, including water-to-binder ratio (W/B), ground granulated blast furnace slag (GGBS) content, cement content (C), silica fume (SF) content, superplasticizer (SP), specimen age (A), and coarse aggregate (CA) and fine aggregate (FA) proportions were considered to develop the database while CS and TS were used as output parameters. It is known that the distribution of the input data of any model has a significant impact on its performance [104]. Before developing the models, it is important to ensure that data is randomly distributed without biasness, and outliers should be detected. This study used the raincloud visualization technique with a normal distribution curve to determine the potential outliers in the database, as shown in **Fig. 3.2** and **Fig. 3.3** for CS and TS, respectively. It can be noted that only a few data points deviated from the normal trend. Several statistical techniques could be used to eliminate the effect of outliers, such as data trimming, imputation, and data transformation [105]. The selection of the most appropriate outlier detection technique depends on the type of data. The imputation technique can be considered the most effective outlier detection method when the data slightly deviates from the normal trend [106]. Therefore, this study also employs an imputation technique in which outliers are replaced with mean values. After the outlier analysis, the distribution of input parameters and their relationship with the performance indicators (CS and TS) is shown in **Fig. 3.4** and **Fig. 3.5**, respectively. It is evident that the input data is distributed randomly without any biasness.

For analysis, the whole database was split into three categories, i.e., training, validation, and testing subsets. A training database was used for the general evolution of the model, while a validation and testing database was utilized to assess the model's predictive validity. When partitioning a dataset into categories, it is crucial to make sure that the data distribution is uniform in the training and validation phases. This objective was accomplished through the random arrangement of the training (learning), validation, and testing (evaluation) datasets,

ensuring that the variables being used exhibit a satisfactory level of consistency with regard to statistical properties, including mean, median, and standard deviation, as well as the range of data. The present study employed a training dataset comprising 70% of the total data for both CS and TS (454 and 171 data points), while the remaining 30% of the total data points (194 and 74) were allocated to validate the models. Before running models, numerous tests were performed to evaluate whether the consistency of the database was valid.

The statistical analysis of input parameters used for CS and TS models are given in **Table 3.1** and **Table 3.2**, respectively. It provides details about the standard deviation, mode, kurtosis, median, skewness, range, smallest, and highest values of the datasets being used for both CS and TS models. A minimal standard deviation value demonstrates that most values are concentrated in a very narrow range around the mean value in the normal distribution curve. On the other hand, a higher standard deviation shows that the numbers are dispersed more widely. The skewness of a variable is the degree to which the distribution of its probabilities deviates from symmetry with respect to the mean [107]. According to Brown and Greene [108], the acceptable kurtosis value is within the range of -10 to +10 and it indicates the type of probability distribution. The abovementioned statistical measures show that the data is distributed over a wider range, making the models more generalized. Moreover, the efficacy of the model can be enhanced by the distribution of input variables across a wide range. It is essential to carefully examine the mutual dependency of specific parameters utilized in model construction to avoid complexity in the analysis of the model's results. The issue of correlation among particular variables is commonly known as multi-collinearity. To avoid this concern, it is recommended that the correlation coefficient between the two variables should be lower than 0.8 [109]. **Table 3.3** and **Table 3.4** indicate that all variables utilized in the model exhibit a weak correlation, as indicated by both negative and positive correlation values.

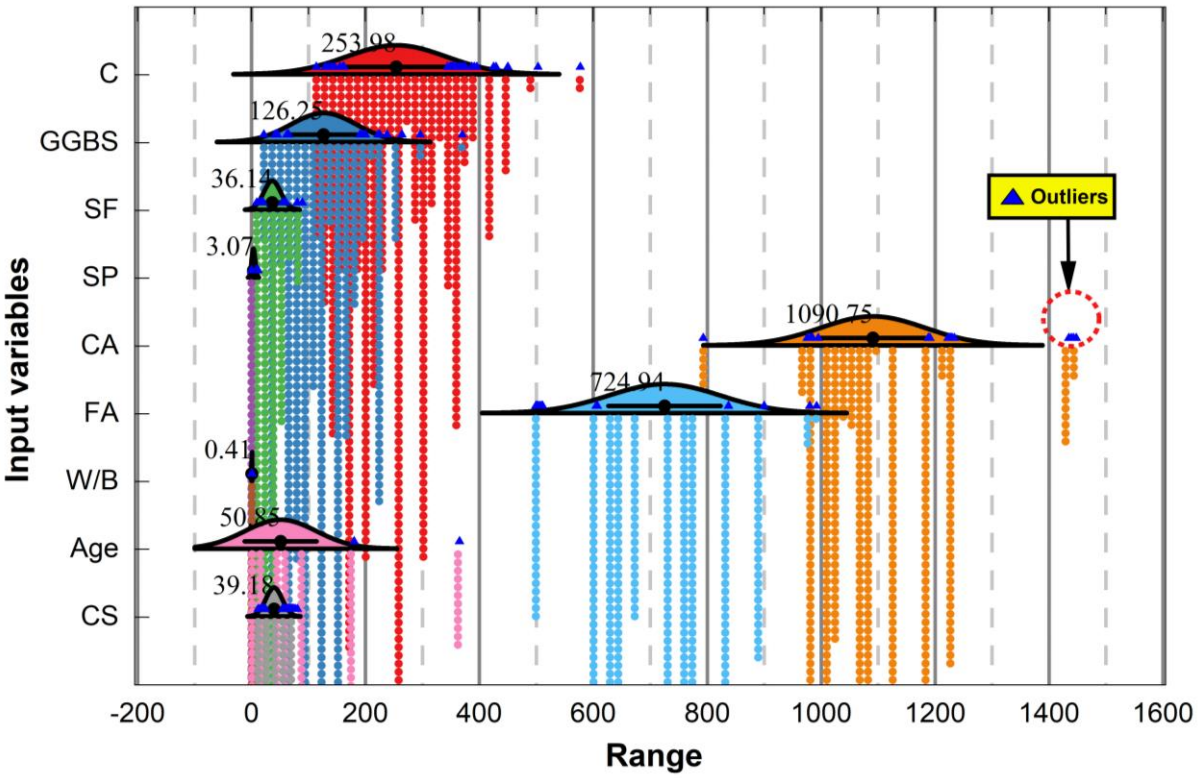


Fig. 3.2: Outliers detection using raincloud plots for CS database.

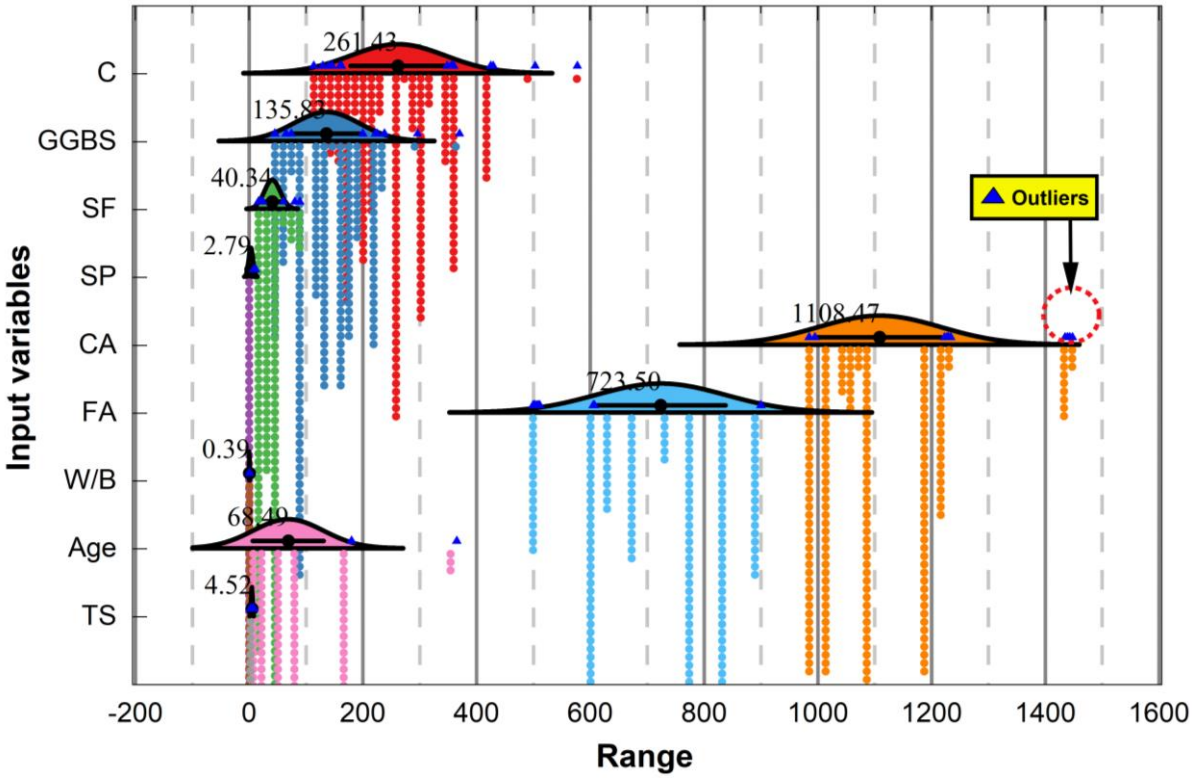


Fig. 3.3: Outliers detection using raincloud plots for TS database.

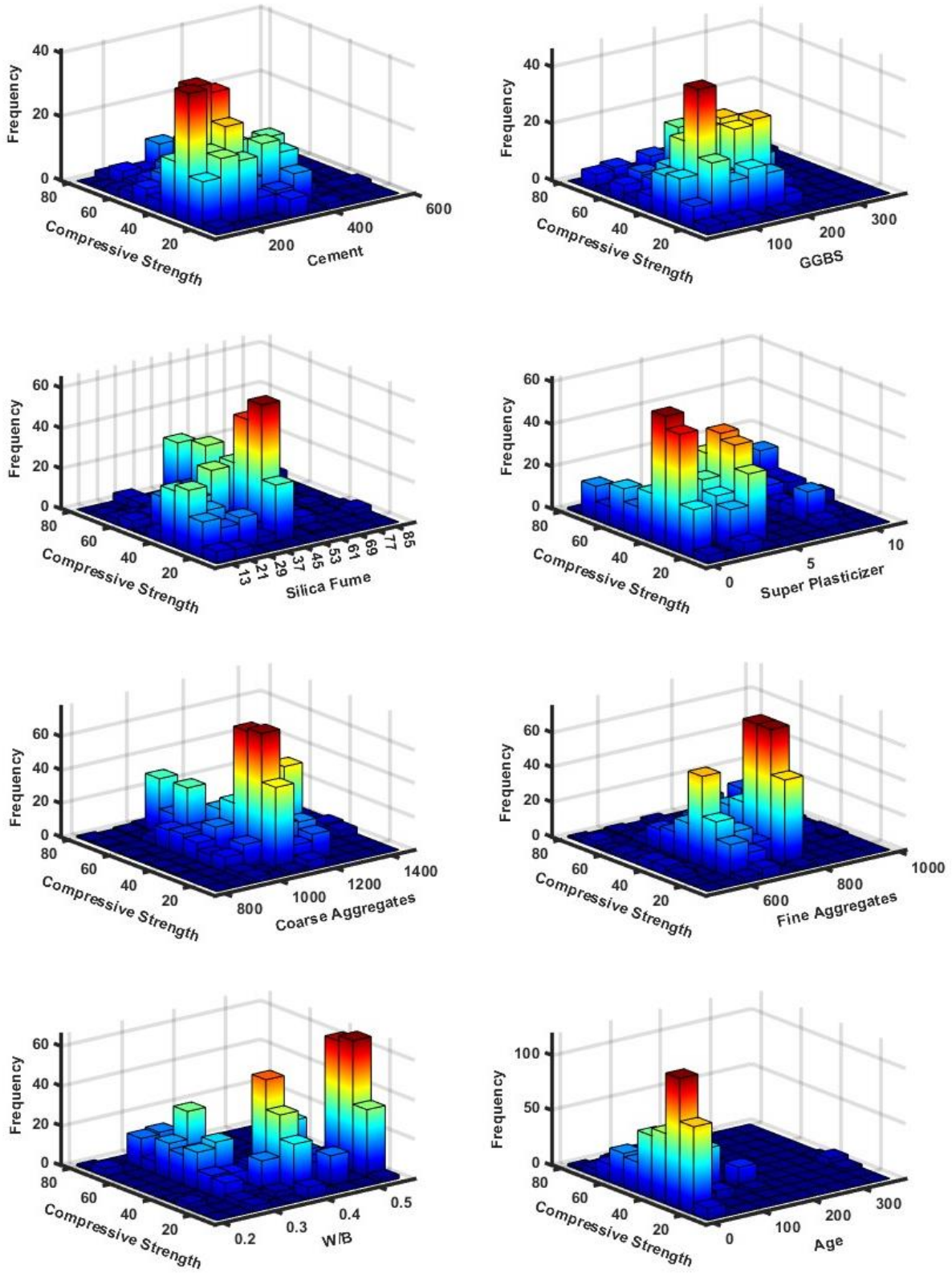


Fig. 3.4: 3D bar charts representing frequency distribution for the CS database.

Note: Measuring unit for all input variables is "kg/m^3" except the age, which is measured in days.

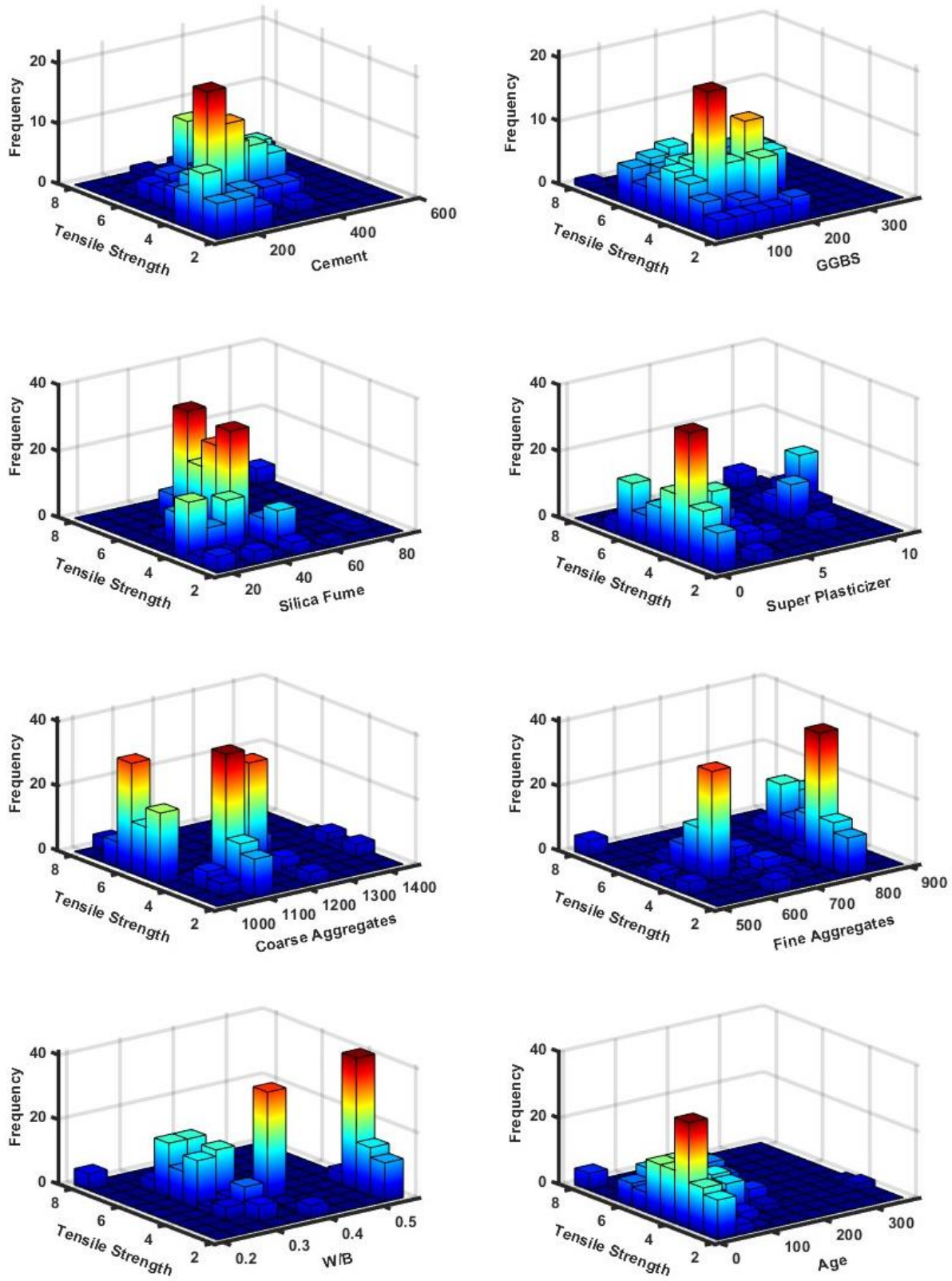


Fig. 3.5: 3D bar charts representing frequency distribution for the TS (MPa) database.

Note: Measuring unit for all input variables is "kg/m^3" except the age which is measured in days.

Table 3.1: Statistical description of CS parameters

Parameters		Minimum	Mean	Range	Median	Kurtosis	Mode	SD	Skewness	Maximum
Inputs	Units									
C	kg/m^3	113.4	254.0	463.6	243.0	-0.3	270.0	87.9	0.5	577.0
GGBS	kg/m^3	21.5	126.2	348.5	129.6	-0.1	64.8	57.6	0.5	370.0
SF	kg/m^3	9.0	36.1	80.0	38.7	0.9	45.0	15.0	0.5	89.0
SP	kg/m^3	0.0	3.1	11.3	3.8	0.6	0.0	3.1	1.0	11.3
CA	kg/m^3	793.0	1090.8	655.0	1093.0	4.0	1093.0	91.7	1.2	1448.0
FA	kg/m^3	499.0	724.9	493.0	768.0	-0.5	785.0	98.5	-0.2	992.0
W/B		0.2	0.4	0.4	0.4	-1.2	0.6	0.1	0.2	0.6
Age	Days	1.0	50.8	364.0	28.0	9.8	28.0	63.0	2.8	365.0
Output										
CS	MPa	11.2	39.2	69.6	35.2	-0.2	34.4	14.3	0.7	80.8

Table 3.2: Statistical description of TS parameters

Parameters		Minimum	Mean	Range	Median	Kurtosis	Mode	SD	Skewness	Maximum
Inputs	**Units**									
C	kg/m^3	113.4	261.4	463.6	270.0	0.0	270.0	83.6	0.4	577.0
GGBS	kg/m^3	45.0	135.8	325.0	135.0	0.0	45.0	58.4	0.4	370.0
SF	kg/m^3	16.2	40.3	72.8	45.0	2.4	45.0	13.9	0.6	89.0
SP	kg/m^3	0.0	2.8	11.3	0.0	-0.1	0.0	3.8	1.1	11.3
CA	kg/m^3	985.0	1108.5	463.0	1093.0	2.0	1093.0	108.1	1.3	1448.0
FA	kg/m^3	499.0	723.5	401.0	785.0	-1.1	785.0	114.3	-0.3	900.0
W/B		0.2	0.4	0.4	0.4	-1.3	0.6	0.1	0.5	0.6
Age	Days	7.0	68.5	358.0	56.0	5.1	28.0	62.2	1.9	365.0
Output										
TS	MPa	2.1	4.5	6.1	4.3	-0.1	3.7	1.2	0.5	8.3

Table 3.3: Correlation coefficient Metrix of CS parameters

	C	GGBS	SF	SP	CA	FA	W/B	Age	CS
C	1								
GGBS	-0.2172	1							
SF	0.0304	0.1853	1						
SP	-0.0681	0.0145	0.2537	1					
CA	-0.0004	-0.0160	-0.1282	-0.0553	1				
FA	-0.4142	-0.1128	0.0059	0.3206	-0.5484	1			
W/B	-0.5972	-0.2312	-0.2535	-0.2416	0.1206	0.3244	1		
Age	-0.1608	-0.0225	0.0425	0.0707	-0.0420	0.1705	0.1828	1	
CS	0.2388	0.2360	0.2262	0.2561	-0.3149	0.2068	-0.5637	0.2442	1

Table 3.4: Correlation coefficient Metrix of TS parameters

	C	GGBS	SF	SP	CA	FA	W/B	Age	TS
C	1								
GGBS	-0.2450	1							
SF	0.3150	0.1287	1						
SP	0.1128	0.1204	0.2327	1					
CA	-0.0548	0.1256	-0.3269	-0.0927	1				
FA	-0.3015	-0.2016	-0.0832	0.2561	-0.6536	1			
W/B	-0.6520	-0.2686	-0.3952	-0.4985	0.0654	0.2200	1		
Age	0.0304	-0.0619	0.0519	0.1025	-0.1778	0.1123	-0.0198	1	
TS	0.5449	0.1457	0.4139	0.5142	-0.3211	0.0702	-0.7511	0.3124	1

3.3 Overview of machine learning techniques

3.3.1 Artificial neural network

Artificial Neural Networks (ANN), initially introduced by McCulloch et al. [110], are computational models utilized for efficient prediction and categorization of non-linear regression problems. Neural networks comprise fundamental computational elements known as neurons, which are grouped into layers. Every single neuron in a particular layer is linked to all the neurons in the adjacent layer. Computational structure is structured hierarchically, with at least three layers comprising its composition: the input layer (input neurons), one or more computational layers (hidden layers), and the output layer, as demonstrated in **Fig. 3.6**. The input layer is accountable for acquiring variables for training and evaluation of the model. Likewise, the computational layer is accountable for linking the input and output layers to help with the processing of data, which is subsequently transmitted to the output layer to generate the model's outcomes [111]. The study employs the process of forward propagation, which is accomplished through a sequential pathway in which the preceding neuron receives and interprets data before transmitting it to the succeeding neurons. Simultaneously, every input is subjected to the influence of weight (Wj w/b, Wj C, Wj GGBS, Wj SF, Wj SP, Wj A, Wj CA, Wj FA) which denotes the varying degrees of the significance of the input data in relation to the output. A threshold value, denoted by j, is added by each node to the sum of weighted signal inputs.

Subsequently, a non-linear conversion function is applied to the integrated input (Ij) that transforms a value from one unit of measurement to another. The activation functions (AF) play a crucial role in ANN and have a notable impact on the efficacy of the models by introducing non-linearity to the networks. Therefore, selecting an appropriate AF is a critical

consideration [112]. Several commonly utilized AFs for improving the efficacy of the ANN models have been identified in the literature [113] [114]. The activation conversion functions that are frequently employed in artificial neural networks (ANNs) include logistic sigmoid, linear, and hyperbolic tangent sigmoid [115]. The present study employs a linear transfer function (PURELIN) and a BPNN transfer function (TRANSIG). These functions result in increased statistical performance and a number of neurons in the training phase while decreasing the performance accuracy in the testing and confirmation phases [116]. The logistic function, which is represented by Equation (3.1), was employed as an AF in this study. The ANN method is mathematically represented by Equations (3.1)– (3.3).

$$z_h(x) = \frac{1}{1+e^{-x}} \tag{3.1}$$

$$I_j = \left\{\left(w_{j\,C} * C + w_{j\,\frac{w}{b}} * \frac{w}{b} + \cdots w_{j\,A} * A\right)\right\} + \theta_j \qquad Summation \tag{3.2}$$

$$CS_j \; or \; TS_j \; = x(I_j) \quad Transfer \tag{3.3}$$

The training process of ANN starts with the propagation of data from the input neuron, while the weights are established according to predetermined rules to produce output with minimal error. Following that, the model that has undergone training is subjected to scrutiny and validation through a distinct subset of data designated for testing purposes. Further details regarding the ANN modeling methodology can be found in [117].

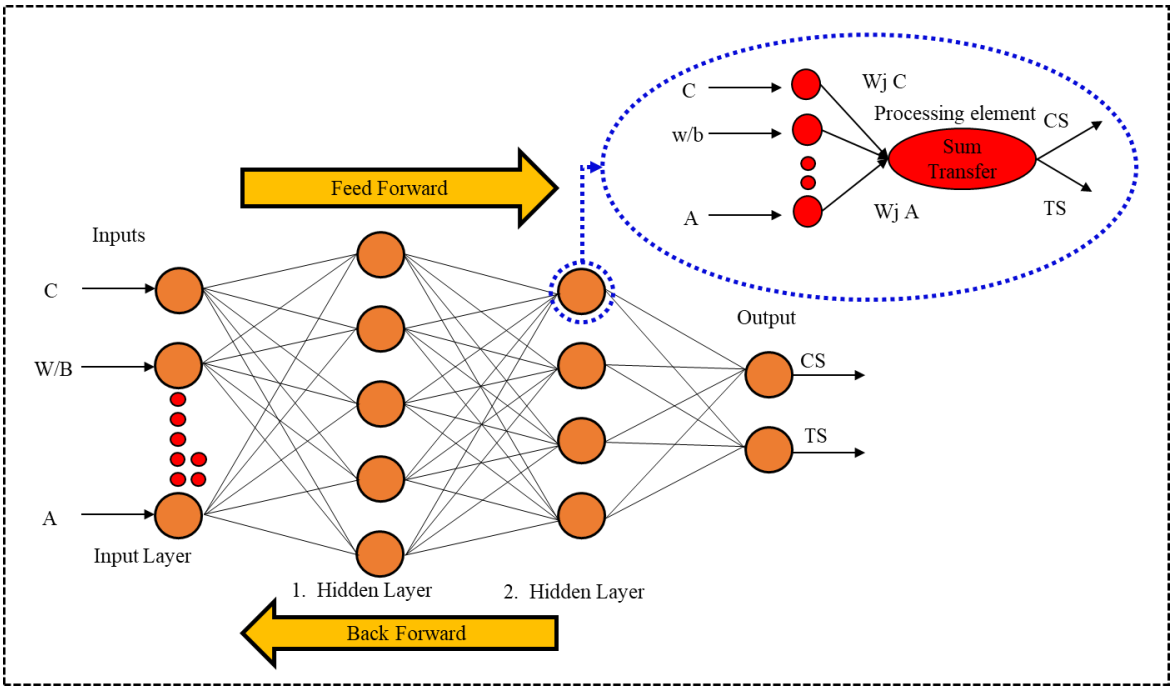


Fig. 3.6: Schematic of ANN models with eight input parameters

3.3.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a computational model that integrates the features of both fuzzy logic and neural networks. It exhibits superior predictive capabilities and represents a more viable option for computing complex non-linear problems with enhanced accuracy [118] [119]. ANFIS learns from a training set to construct a fuzzy inference system that can be fine-tuned and optimized through a learning algorithm, similar to the training process of neural networks.

ANFIS uses a collection of fuzzy if-then regulations to estimate the input-output correlation. The aforementioned regulations are commonly derived from specialized expertise or acquired through a process of learning. In this study, the fuzzy logic toolbox in MATLAB R2022a environment was used for the development of ANFIS models. The fuzzy logic toolbox provides a range of functions and tools for the creation of ANFIS models. This toolbox also facilitates the process of training and optimizing the ANFIS model by utilizing the provided data and parameters. The ANFIS comprises multiple layers, wherein the parameters of each layer undergo modifications through the learning process. The initial stage of a fuzzy system involves the acquisition of input variables by the input layer. These variables are subsequently subjected to a fuzzification process in the fuzzification layer, wherein they are converted into fuzzy sets by utilizing membership functions (MF). Fuzzy if-then rules are utilized in the rule layer to combine the fuzzified inputs. The normalization layer combines the normalized outputs generated by the rule layer, thereby ensuring consistency across the rules. Ultimately, the defuzzification layer transforms the normalized outputs into precise numerical values. During the training process, the fuzzy rules and MF parameters are adapted using a learning algorithm to minimize the differences between predicted and actual outputs. This approach aims to optimize the performance of the model. **Fig. 3.7** illustrates the ANFIS structure for multiple input variables, where the fixed and adaptive nodes are represented by circles and squares, respectively. The ANFIS architecture can be represented by first-order of the Sugeno model, which employs two sets of IF-THEN rules.

Rule 1: applies if C and W/B are A1 and B2, respectively.

THEN, Equation (3.4) indicates that.

$$z_1 = p_1(C) + q_1\left(\frac{w}{b}\right) + r_1 \tag{3.4}$$

Rule 2: applies if C and w/b are A2 and B2, respectively.

THEN, Equation (3.5) says that.

$$z_2 = p_2(C) + q_2\left(\frac{w}{b}\right) + r_2 \quad (3.5)$$

Here,

An and Bn = fuzzy logic sets

Pn, qn, and rn = shape factors (estimated in the training phase)

z1 and z2 = outputs (CS and TS)

The ANFIS model is comprised of five distinct layers, as noted by Golafshani et al. [118]. A comprehensive explanation of the functions of these layers is provided herein.

1st Layer: This layer is also known as the fuzzification layer, in which each input variable is fuzzified or converted into a fuzzy set using membership functions.

Equations (3.6) and (3.7) represent the basic fuzzy rule and the model parameters, respectively.

$$O_i^1 = \mu_{Ai}(C), \qquad i = 1, 2 \quad (3.6)$$

$$O_i^1 = \mu_{Bi-2}(w/b), \quad i = 3, 4 \quad (3.7)$$

Where u is the weight calculated by linking the fuzzy association function, and both Ai (C) and $Bi - 2$ work together to distinguish one method of applying a fuzzy MF from another. The bell-shaped and gaussian MF are represented in Equations (3.8) and (3.9).

$$\mu_{Ai}(C) = e^{-\frac{(C-c_i)^2}{2a_i^2}} \quad (3.8)$$

$$\mu_{Ai}(C) = \frac{1}{1+\left\{\left(\frac{C-c_i}{a_i}\right)\right\}^{b_i}} \quad (3.9)$$

2nd Layer: The resulting output of this layer relates to the firing capacity of the pre-specified regulations applied to a given input series. The points in the second-layer are fixed and execute basic multiplication operations. The resulting output boundaries are presented in Equation (3.10).

$$O_i^2 = w_i = \mu_{Ai}(C).\mu_{Bi}\left(\frac{w}{b}\right), \quad i = 1, 2 \quad (3.10)$$

3rd Layer: In this layer, the outputs obtained from the second layer are normalized to ensure that the overall output is within a specified range. This step is important for consistency and comparability across different rules. Layer outputs are illustrated in Equation (3.11).

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1+w_2} \quad i = 1, 2 \quad (3.11)$$

4th Layer: In the fourth layer, the nodes exhibit adaptability, and their outcomes are represented as the result of standardized firing capacity in conjunction with a first-order polynomial while taking into account the first-order Sugeno model. Consequently, the resulting output is expressed as Equation (3.12):

$$O_i^3 = \overline{w_i} z_i = \overline{w_i}\{(p_i(C) + q_i(w/b) + r_i)\} \quad (3.12)$$

5th Layer: In the fifth layer, only one permanent node is denoted by the symbol Σ. This node is responsible for performing the summation of the weighted consequences of rules that were previously resulting from the preceding layer. Consequently, the outcome of the model can be obtained by utilizing Equation (3.13).

$$O^5 = \sum_{i=1}^{2} \overline{w_i} z_i = \frac{\sum_{i=1}^{2} \overline{w_i} z_i}{w_1 + w_2} \quad (3.13)$$

Notably, within the ANFIS framework, only the first and fourth layers exhibit the ability to adapt. The premise parameters, denoted as {ai, bi, ci}, are linked to the input association functions within the first layer of the system. Similarly, the three consequent parameters {pi, qi, ri} are associated with first-order polynomials and are located in the fourth layer, as reported by Islam et al. [120].

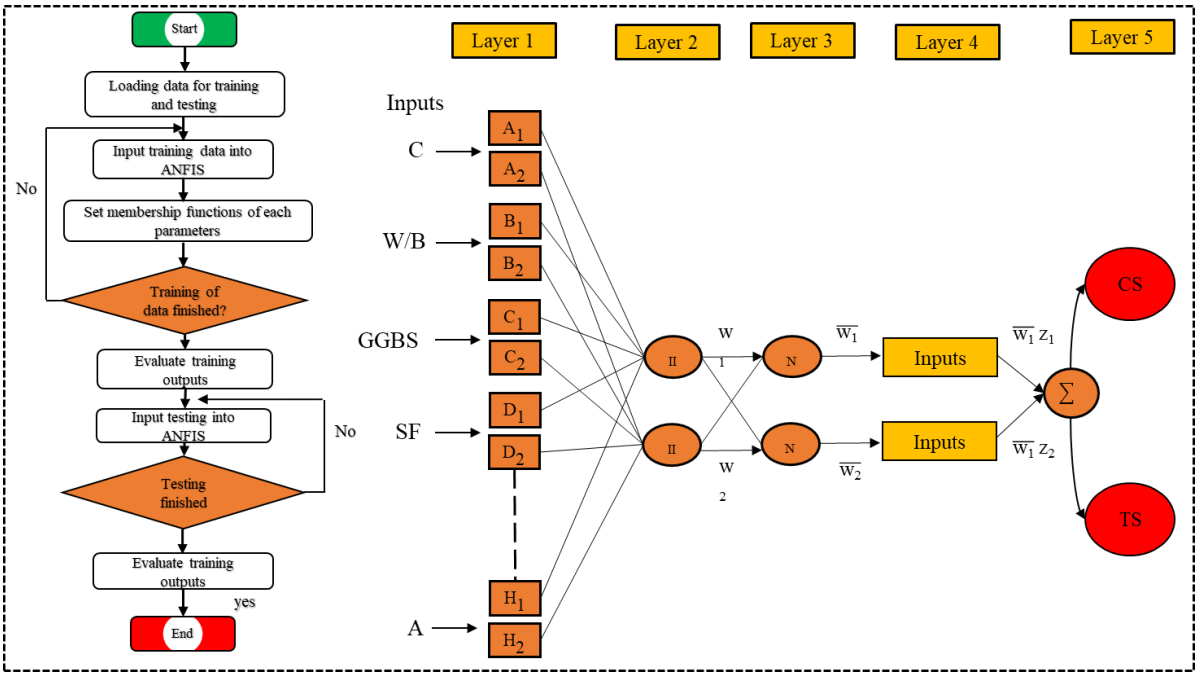


Fig. 3.7: Representation of ANFIS models with parametric settings considered in this study.

3.3.3 Gene expression programming (GEP)

GEP, developed by Ferrira [121], is a modification of genetic programming (GP) and is premised on the evolutionary population hypothesis. It encompasses simple, fixed-length

linear structures like chromosomes and more complex, non-linear structures like parse trees. The essential parameters that must be specified in GEP are identical to those used in GP, such as the function set, set of terminals, fitness measures, controlling parameters, and terminal conditions. Figure 9 shows a graphical description of the GEP working procedure. Chromosomes of fixed length are first generated at random for each individual gene. Static-length linear strings are then represented as non-linear structures of various sizes and forms termed expression trees (ETs), indicating chromosomes of branched structures [122]. It is important to mention that in the GEP, genotype (genetic constitution) and phenotype are separated so that the program can benefit from all of evolution's merits [121]. Then, each individual's fitness is assessed by expressing the chromosomes as ETs. For use in the reproduction phase, the best-fit individuals are chosen. Over many generations, the iteration is repeated with new individuals in search of the optimal solution. Several genetic operations like crossover, mutations, and reproduction are performed for the conversion in populations. **Fig. 3.8** shows the ET representation of the GEP model, crossover, and mutation process. The noteworthy modification in GEP is that the genome (haploid set of chromosomes) is passed down to the succeeding generation, and the whole structure does not need to be replicated and transformed because all modifications occur in a perfectly linear framework. Another important characteristic to note is that each individual is composed of a single chromosome containing several genes that are then separated into head and tail segments [122]. Additionally, genetic algorithms are used to change chromosomes during the reproductive stage [123]. Finally, the GEP models can modify the parameter arrangement based on how well it fits with the results of the experiments.

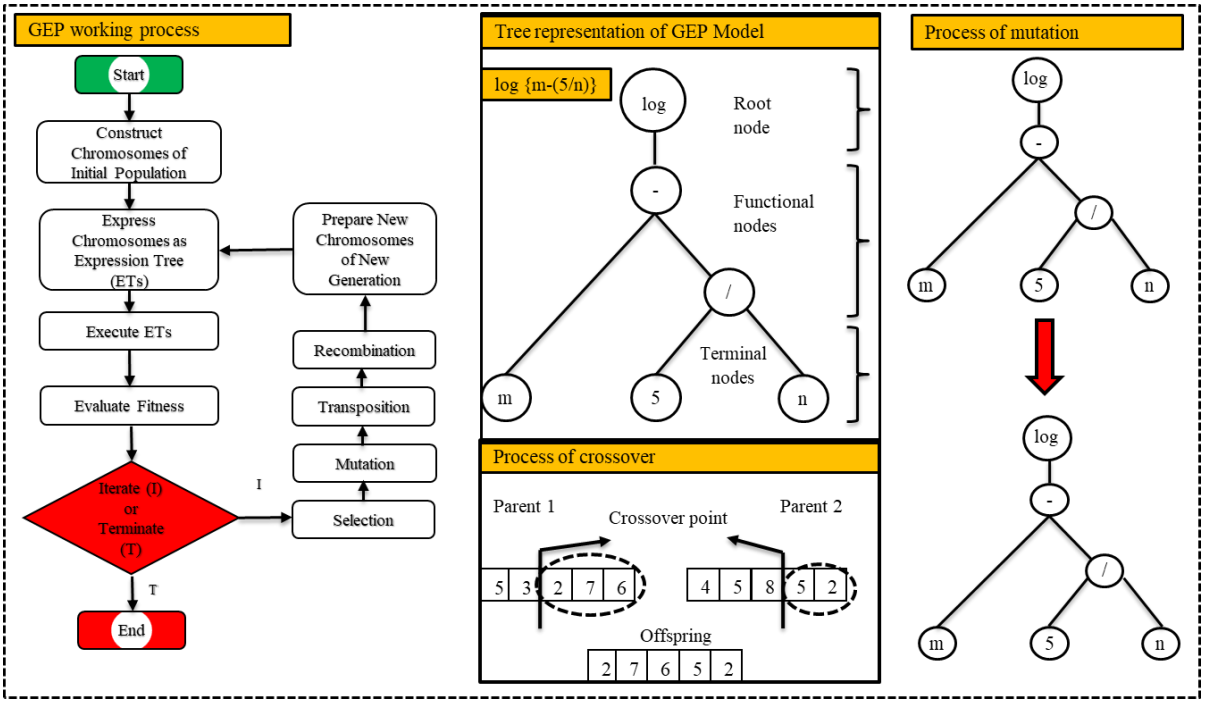


Fig. 3.8: Representation of the GEP working process, Tree structure, crossover, and mutation process.

3.4 Models Structure

Before developing the model, the first step is creating a generalized function. The input parameters that have the most significant effect on the properties of concrete were selected in this function, and parameters having a negligible influence on the properties of concrete were neglected. As a result, the following variables are found to be the function of the mechanical characteristics of concrete containing SF and GGBS.

$$CS\ and\ TS\ (MPa) = f\left(C, GGBS, SF, SP, CA, FA, \frac{W}{B}, A\right) \quad (3.14)$$

The ANN models were developed by using the Levenberg-Marquardt method. The data was portioned randomly, and a total of 10 hidden neurons were used. The Feed-forward back-propagation network type was used for the iteration process. It is crucial to note that, in this study trial and error method was adapted to achieve optimal performance concerning the desired number of hidden layers [124]. **Table 3.5** presents the statistical parameters used for the modeling process using the ANN approach in the present study.

Table 3.5: Parametric settings for training the ANN models.

Parameter type	Value/type	
Data distribution	**CS**	**TS**
Total dataset	682	245
Calibration (Training) (70%)	454	171
Testing (15%)	97	37
Validation (15%)	97	37

General settings	
Hidden neurons	10
Network type	Feed-forward back-propagation
Output layer transfer function	PURELIN
Training method	Levenberg-Marquardt
Computational layer transfer function	TANSIG
Epochs	41
Data division	Random
Rate of learning	0.01
Non-linear parameters	18

ANFIS exhibits a constraint when accommodating a single output compared to ANN. Hence, the outputs were subjected to independent treatment, with identical input variables considered for both the ANN and ANFIS models. Similar data sets of training, testing, and validation were used for ANFIS modeling to observe the optimized outcomes. Initially, the hybrid optimization method, specifically the least square and back-propagation method, was employed to generate the Fuzzy Inference System (FIS) using subtractive clustering, also known as 'sub-cluster.' Subsequently, the FIS was trained utilizing the trim function, as documented by Jalal et al. [125]. It is important to emphasize that this approach was used due to the extensive database. Moreover, Venkatesh and Bind [126] suggested using the grid portioning technique in cases where the total amount of inputs is six or fewer. In this study, trial and error methods were adapted, the ANFIS model was trained, and optimal parameter values were determined using an initial number of epochs as 40. The term "Epoch" refers to the number of iterations or passes over the training set that the program was allowed to run. Each epoch represents a complete iteration through the training data. **Table 3.6** presents a comprehensive list of the different setting parameters used to train ANFIS models.

Table 3.6: Parametric settings for training the ANFIS models.

Parameter type	Value/type	
Data distribution	**CS**	**TS**
Total dataset	648	245
Calibration (Training) (70%)	454	171
Testing (15%)	97	37
Validation (15%)	97	37
General settings		
Number of nodes	10	10
Number of fuzzy rules	7	8
Number of non-linear parameters	50	77
Epochs	40	40
Number of linear parameters	66	120
Number of MFs	18	89
Error goal in training	0	0

Fuzzy structure	Sugeno
Output function	Linear
Optimization technique	Hybrid method
MF types	trimf
FIS type	Sub clustering

In the process of developing the GEP models for the anticipation of CS and TS of concrete containing SF and GGBS, GeneXproTools 5.0, a highly adaptable GEP data modeling program, was used. This is a user-friendly and effective data mining tool created specifically for categorization, time-series analysis, basic functional regression, and logical synthesis. A well-structured Microsoft Excel dataset with eight input parameters and established outputs was uploaded to begin the simulation process.

The generalization capability and robustness of the generated model greatly depend on its fitting parameters. Initially, numeric constants and genetic operators were selected based on the literature, and 20 trial and error-based models were run for each output (CS and TS) to find out the best settings. Different fitting parameter combinations were used by varying the gene number, head size, number of genetic factors (chromosomes), and linking functions [127]. The running time of each model is determined by the number of chromosomes. The basic genetic operators and numeric constants utilized for both CS and TS predictive models are shown in **Table 3.7**. The fitness parameters employed in each of the 20 GEP models for CS are listed in **Table 3.8**, and for TS models, these parameters are listed in **Table 3.9**.

Table 3.7: Genetic operators and numeric constants used in GEP models for CS and TS.

Genetic operators	
RIS transposition rate	0.00541
Permutation	0.00546
Two-point recombination rate	0.00273
Leaf mutation	0.00546
Gene recombination rate	0.00274
Conservative mutation	0.00364
Rate of gene transposition	0.00272
Mutation	0.00138
Rate of mutation	0.00134
Fixed-root mutation	0.00182
Rate of inversion	0.00535
IS transposition rate	0.00531
Numerical constants	
Data type	Floating number
Method	Random selection
Lower bound	−10
Fine-tuning	0.0026
Constant per gene	10
Upper bound	10

Table 3.8: Descriptive summary of parametric settings of GEP models used for CS.

Models	No. of chromosomes	Variable used	Head size	Constant per genes	Number of genes	Linking function	Function set	Duration (minutes)
CS1	30	7	8	10	3	Addition	(÷, ×, +, -)	20
CS2	50	7	10	10	4			23
CS3	80	9	12	10	5			25
CS4	100	7	14	10	6			30
CS5	150	7	16	10	7			40
CS6	30	7	8	10	3	Division	(÷, ×, +, -)	23
CS7	50	9	10	10	4			26
CS8	80	9	12	10	5			29
CS9	100	8	14	10	6			35
CS10	150	7	16	10	7			45
CS11	30	9	8	10	3	Multiplication	(÷, ×, +, -)	18
CS12	50	8	10	10	4			22
CS13	80	9	12	10	5			24
CS14	100	8	14	10	6			27
CS15	150	8	16	10	7			35
CS16	30	7	8	10	3	Multiplication	(÷, ×, +, -, Pow, √)	30
CS17	50	9	10	10	4			34
CS18	80	8	12	10	5			38
CS19	100	9	14	10	6			42
CS20	150	8	16	10	7			50

Table 3.9: Descriptive summary of parametric settings of GEP models used for TS.

Models	No. of chromosomes	Variable used	Head size	Constant per genes	Number of genes	Linking function	Function set	Duration (minutes)
TS1	30	7	8	10	3	Addition	(÷, ×, +, -)	22
TS2	50	8	10	10	4			24
TS3	80	9	12	10	5			26
TS4	100	9	14	10	6			32
TS5	150	9	16	10	7			40
TS6	30	8	8	10	3	Division	(÷, ×, +, -)	24
TS7	50	8	10	10	4			26
TS8	80	7	12	10	5			28
TS9	100	8	14	10	6			33
TS10	150	7	16	10	7			45
TS11	30	9	8	10	3	Multiplication	(÷, ×, +, -)	20
TS12	50	9	10	10	4			22
TS13	80	9	12	10	5			26
TS14	100	8	14	10	6			27
TS15	150	8	16	10	7			37

TS16	30	7	8	10	3	Multiplication	(÷, ×, +, -, Pow, √)	31
TS17	50	7	10	10	4			37
TS18	80	8	12	10	5			38
TS19	100	9	14	10	6			44
TS20	150	9	16	10	7			54

3.5 Models' validation

Performance validation is essential to assess the reliability and generalization capability of ML-based models. In this study, the statistical metrics and experimental validation criteria were used to assess the efficacy of the proposed models.

3.5.1 Statistical validation

Initially, the performance of developed models was evaluated by using statistical error correlations such as mean absolute error (MAE), relative root mean squared error (RRMSE), root mean square error (RMSE), correlation coefficient (R), and root squared error (RSE). Typically, a model with a higher R^2 value and lower RMSE, MAE, and RSE values indicates better results. Equations. (3.15-3.20) presents the mathematical formulation of these statistical measures.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(a_i - m_i)^2}{n}} \tag{3.15}$$

$$MAE = \frac{\sum_{i=1}^{n}|a_i - m_i|}{n} \tag{3.16}$$

$$R = \frac{\sum_{i=1}^{n}(a_i - \bar{a}_i)^2(m_i - \bar{m}_i)^2}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a}_i)^2 \sum_{i=1}^{n}(m_i - \bar{m}_i)^2}} \tag{3.17}$$

$$RSE = \frac{\sum_{i=1}^{n}(m_i - a_i)^2}{\sum_{i=1}^{n}(\bar{a} - a_i)^2} \tag{3.18}$$

$$RRMSE = \frac{1}{|\bar{a}|}\sqrt{\frac{\sum_{i=1}^{n}(a_i - m_i)^2}{n}} \tag{3.19}$$

$$\rho = \frac{RRMSE}{R} \tag{3.20}$$

Here, mi and ai represent model and experimental (actual) output values, while $\bar{m}_i$ and $\bar{a}_i$ represents the mean values of model and experimental outputs, respectively, and n is the total size of the population (database). For an effective and stable correlation between actual and anticipated values, it has been suggested that it is better to have an R-value greater than 0.8 [128]. However, it was excluded due to its negligence to the multiplication and division of outcomes to a constant [128]. Therefore, R^2 was also employed due to its unbiased

evaluation and comparatively higher efficiency. The maximum variation among the input parameters is captured when R^2 values are equal to 1 [129]. There is another statistical correlation commonly used for performance measures, i.e., RMSE. It is most prominent because large errors are addressed more efficiently than small errors, and an error value near zero indicates the best model [130]. However, it does not always guarantee the accuracy of identifying the lowest error in prediction. Due to this, MAE was also estimated, which is very beneficial when dealing with continuous and smooth data [131]. In essence, better and improved models are indicated by lower error statistical measures (MAE, RSE, and RMSE) and higher R and R^2 values.

However, a major issue related to AI approaches is overfitting, which results in higher errors in the testing dataset. Hence, the minimization of the objective function (OBF) is done to select the most suitable prediction model, as demonstrated in Eq. (8) [104]. The OBF is calculated to assess the trained model's effectiveness, including error function and changes in the correlation coefficient. The overfitting problem can be fixed by lowering the OBF value.

$$OBF = \left(\frac{n_T - n_V}{n}\right)\rho_{i\,T} + 2\left(\frac{n_V}{n}\right)\rho_{i\,V} \tag{3.21}$$

V and T subscripts denote validation and training data points, respectively, and n refers to the total size of the population. A model having an OBF value closer to zero refers to the most appropriate predictive model since it comprises the effect of both RRMSE and R.

3.5.2 Experimental validation

3.5.2.1 Materials

In the present study, OPC Grade53, SF, and GGBS were utilized as cementitious materials in the mixture. The chemical composition and physical characteristics of these materials are shown in **Table 3.10.** River sand and granite with a maximum size of 4.75 mm and 19 mm were used as fine aggregates (FA) and coarse aggregates (CA) in the mixtures. The properties of CA and FA were thoroughly examined. The fineness modulus of FA and CA was measured as 2.2 and 5.3, respectively. Likewise, water absorption and specific gravity was determined as 2.1% and 2.57, and 4% and 3.41 for FA and CA, respectively. In addition, the superplasticizer type S-3 was used to maintain the workability of the concrete samples. The schematic of experimental validation is shown in **Fig. 3.9.**

Table 3.10: Properties of cementitious materials [132], [133]**.**

Component	OPC	GGBS	SF
Chemical composition (% mass)			
K_2O	–	0.69	0.56
CaO	62.73	33.06	0.22
Fe_2O_3	3.95	0.34	0.5
SO_3	3.14	2.67	0.12
SiO_2	20.78	35.84	97.5
Na_2O	0.78	1.08	0.25
Al_2O_3	4.82	12.43	0.2
MgO	1.57	12.08	0.56
Physical properties			
Specific gravity	3.15	3.05	2.22
Specific surface area (m^2/kg)	421	550	2300
Loss on ignition	2.08	2.08	2.08

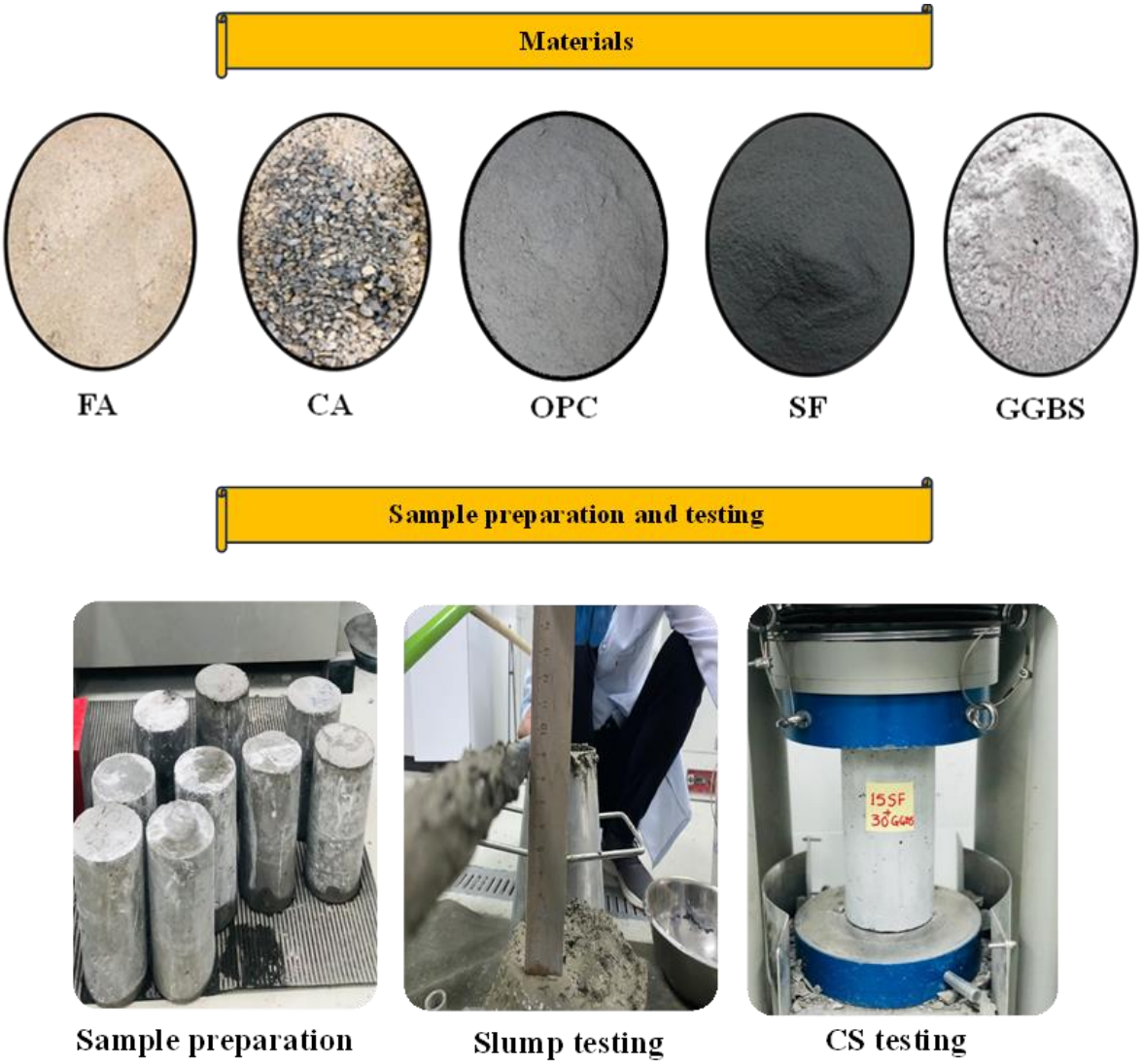


Fig. 3.9: Experimental program.

3.5.2.2 Mix design and specimen preparation

A total of 54 concrete samples, with varying proportions of SF and GGBS, were prepared. The details of the mix design for external validation are presented in **Table 3.11**. In total, 6 different mixes were designed. CM represents the control mix, while mixes M1 to M5 represent mixes with different percentages of SF and GGBS as a replacement for cement. The percentage of SF in these mixes varied from 5 to 20, while the percentage of GGBS varied from 10 to 40. For example, M1 represents a mix containing 5 percent SF and 10 percent GGBS as a replacement. The concrete specimens were tested for workability using ASTM C143. Likewise, the CS of samples was evaluated using the ASTM C39 at the age of 28, 56, and 90 days.

Table 3.11: Mix design used for external validation.

Code	**SCMs**	**Cement (kg/m^3)**	**SP (kg/m^3)**	**SF (kg/m^3)**	**Water (kg/m^3)**	**GGBS (kg/m^3)**	**CA (kg/m^3)**	**FA (kg/m^3)**
				Control mix				
CM	CM	415	10	0	156	0	1050	730
M1	SF5+GGBS10	353	10	20	156	41	1050	730
M2	SF10+GGBS20	291	10	41	156	83	1050	730
M3	SF10+GGBS30	249	10	41	156	125	1050	730
M4	SF15+GGBS30	228	10	62	156	125	1050	730
M5	SF20+GGBS40	166	10	83	156	166	1050	730

Chapter 4- Results and Discussions

4.1 Overview

This section presents and discusses the results of the developed ML models for estimating the CS and TS of concrete incorporated with SF and GGBS. In sections 4.2, 4.3, and 4.4 the efficacy of the ML models (ANN, ANFIS, and GEP) will be evaluated. Section 4.4 will explain the experimental validation of the models. The efficacy of the models will be compared in section 4.6. The best-obtained model in section 4.6 will be validated by using the external validation criteria in section 4.7. Additionally, section 4.8 will discuss the sensitivity and parametric analysis conducted on the best-proposed model in previous sections. In the last section (4.9), the findings of this study will be compared with the literature.

4.2 Performance assessment of ANN models

The statistical characteristics and performance index of the training, validation, and testing sets for optimum models of CS and TS based on ANN models are shown in **Fig. 4.1**. It is obvious that the R^2 values for training, validation, and testing for CS-ANN models are greater than 90% ($R^2_{training}$ = 0.941, $R^2_{validation}$ = 0.937, $R^2_{testing}$ = 0.938). Similarly, for TS-ANN models, R^2 values are 0.958, 0.938, and 0.931 for the training, validation, and evaluation stages, respectively. It is important to mention that the greater difference between the training and validation values is due to the fact that ANN models work on the black box principle. It can be seen that error values (RMSE, MAE, RSE, RRMSE) for each set of both CS and TS models are minimal, approaching zero, indicating the higher accuracy of the optimum models. The RMSE values are equal to 3.328 MPa, 4.049 MPa, and 3.826 MPa for training, validation, and testing phases, respectively, for CS-ANN and 0.269 MPa, 0.321 MPa and 0.343 MPa for TS-ANN models. Consequently, the MAE values are equal to 1.788 MPa, 2.791 MPa, and 2.832 MPa for the training, validation, and testing phases of CS-ANN. For TS-ANN models, these values are 0.202 MPa, 0.255 MPa, and 0.034 MPa, respectively. From these findings, it can be noted that in all three stages for both CS-ANN and TS-ANN models, MAE values are lower than RMSE values, indicating a better efficacy of the models. In addition, the PI and OF values for all the established models in all three stages are close to zero. To further assess the model's accuracy, the greatest error percentage in the model between actual and estimated outcomes is depicted in **Fig. 4.2 (a-d)**. It is evident that experimental and estimated outcomes are comparable with each other with maximum errors

less than 13 MPa and 0.93 MPa and with mean of 2.56 MPa and 0.23 MPa in CS-ANN and TS-ANN models, respectively. **Fig. 4.2 (b and d)** illustrates the error histograms for the CS-ANN and TS-ANN models. These figures demonstrate that greater than 81.17% of the projected results lie within the range of 0 to 4 MPa for the employed data in the case of the CS-ANN model. Meanwhile, greater than 90% of the anticipated TS values fall within the error range of 0 MPa to 6 MPa. It is clear from **Fig. 4.2** that error values are closer to zero, which is evidence for better performance of ANN-based models.

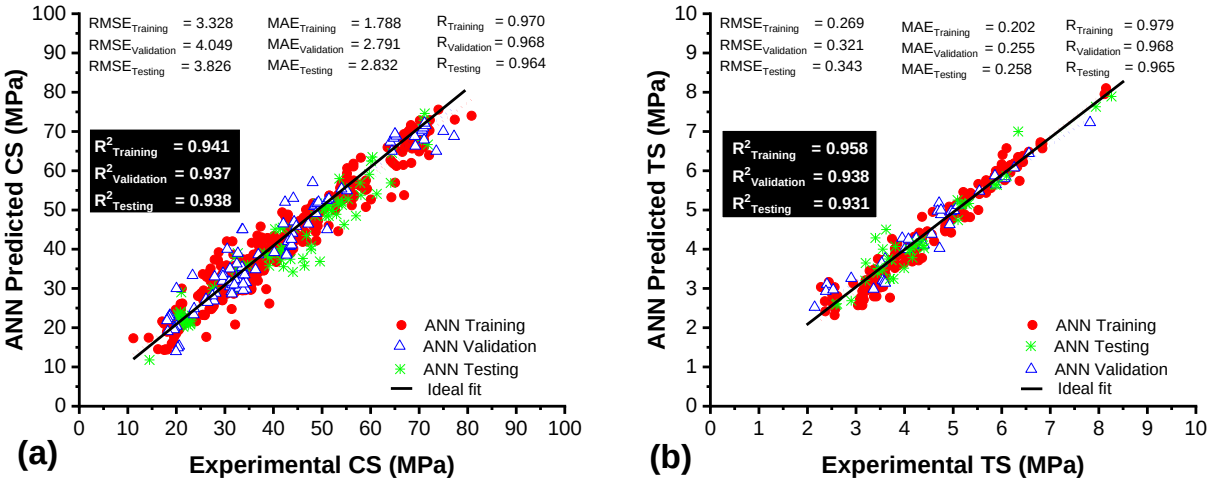


Fig. 4.1: Comparison of experimental and anticipated results using ANN for (a) CS, (b) TS.

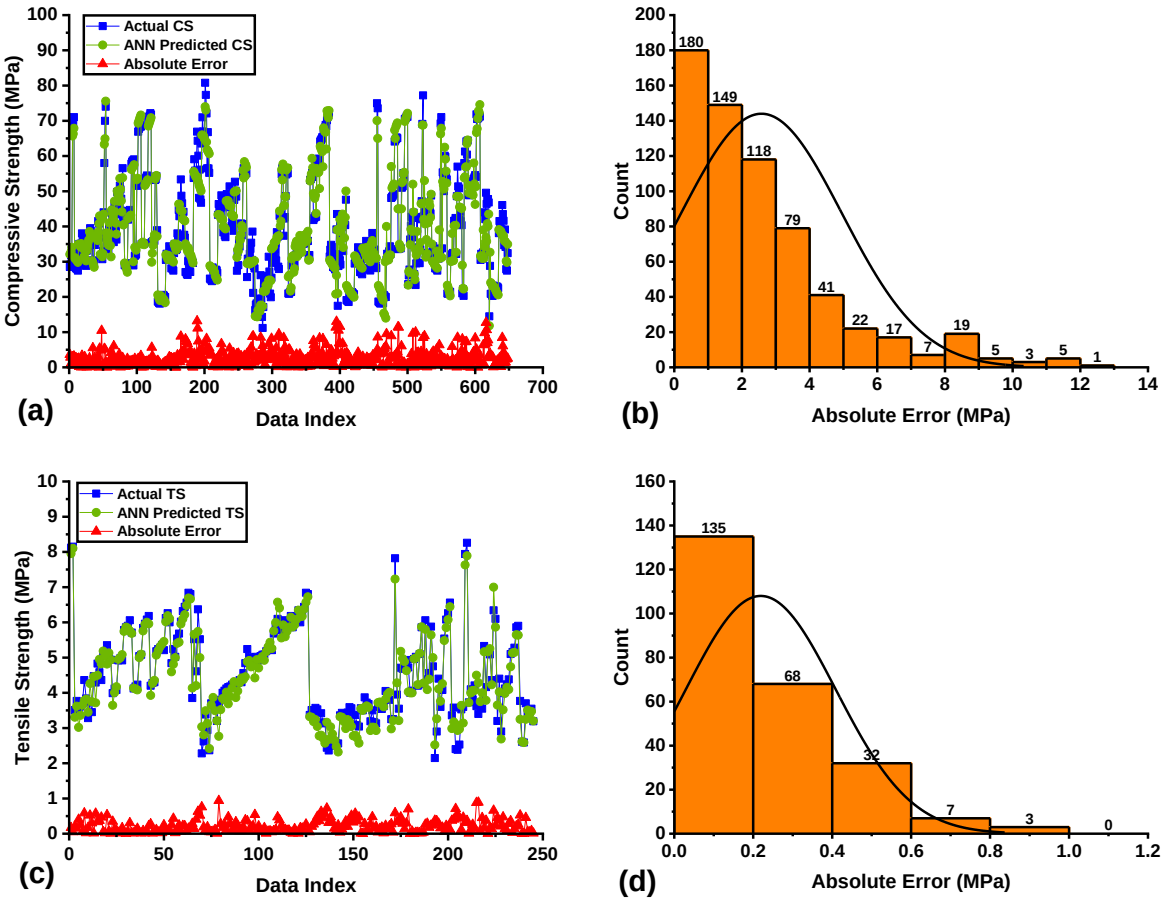


Fig. 4.2: Comparison of experimental and anticipated results using ANN for (a) CS, (b) TS.

4.3 Performance assessment of ANFIS models

In order to enhance efficiency, similar data sets generated by ANN models were used as input for the ANFIS model. However, a difference was observed between the predictive outcomes of both techniques. This difference can be attributed to the use of additional fuzzy logic in the ANFIS models. It is obvious from **Fig. 4.3** that the R^2 values are 0.97, 0.969, and 0.967 for all three stages, i.e., training, validation, and evaluation (testing), respectively, for CS-ANFIS models and 0.995, 0.998, and 0.998 for the TS-ANFIS models. These higher R^2 reveal the strong predictive performance of ANFIS-based models. Like ANN-based models, the RMSE and MAE values for CS-ANFIS and TS-ANFIS models are close to zero, and MAE values are lower than RMSE. Moreover, the ANFIS models show 33% and 30% lower RMSE and MAE values than ANN models for CS and approximately 85% and 93% lower for TS models. Additionally, the RRMSE, OF, and RSE values are close to zero for both CS-ANFIS and TS-ANFIS models. To determine the predictive performance of ANFIS-based models, the absolute error plots between ANFIS anticipated and experimental values are presented in **Fig. 4.4 (a-d)** for CS-ANFIS and for the TS-ANFIS models, respectively. The error values

show that ANFIS-predicted values accurately reflect the experimental values for CS and TS models. It can be seen that ANFIS models exhibit 40% and 86% lower absolute error values for CS and TS models, respectively. The highest and lowest absolute error values for the CS-ANFIS model are equal to 10.73 MPa and 0.009 MPa, while for the TS-ANFIS model, these values are equal to 0.48 Mpa and 0.0003 Mpa, respectively. Similarly, the average absolute error values for these models are 1.68 Mpa and 0.032 Mpa, respectively. **Fig. 4.4 (b and d)** presents the error histograms, which show that 87% of error values lie between 0 Mpa and 4 Mpa for the CS-ANFIS model. Likewise, 98% of absolute error values lie in the 0 Mpa to 0.4 Mpa range for TS-ANFIS. From these findings, it can be inferred that ANFIS-based models exhibit comparatively superior predictions compared to ANN-based models.

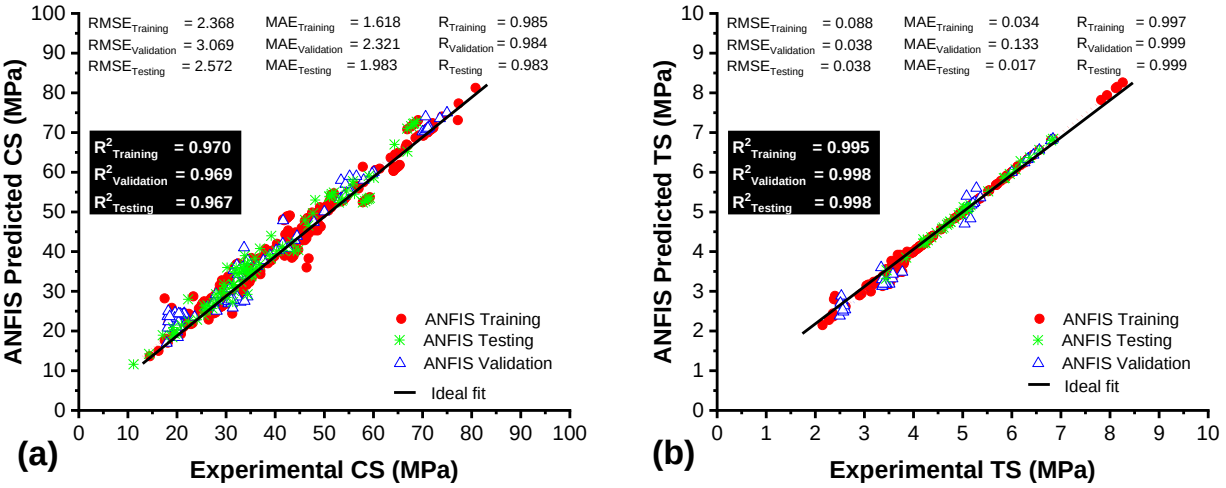


Fig. 4.3: Comparison of experimental and anticipated results using ANN for (a) CS, (b) TS.

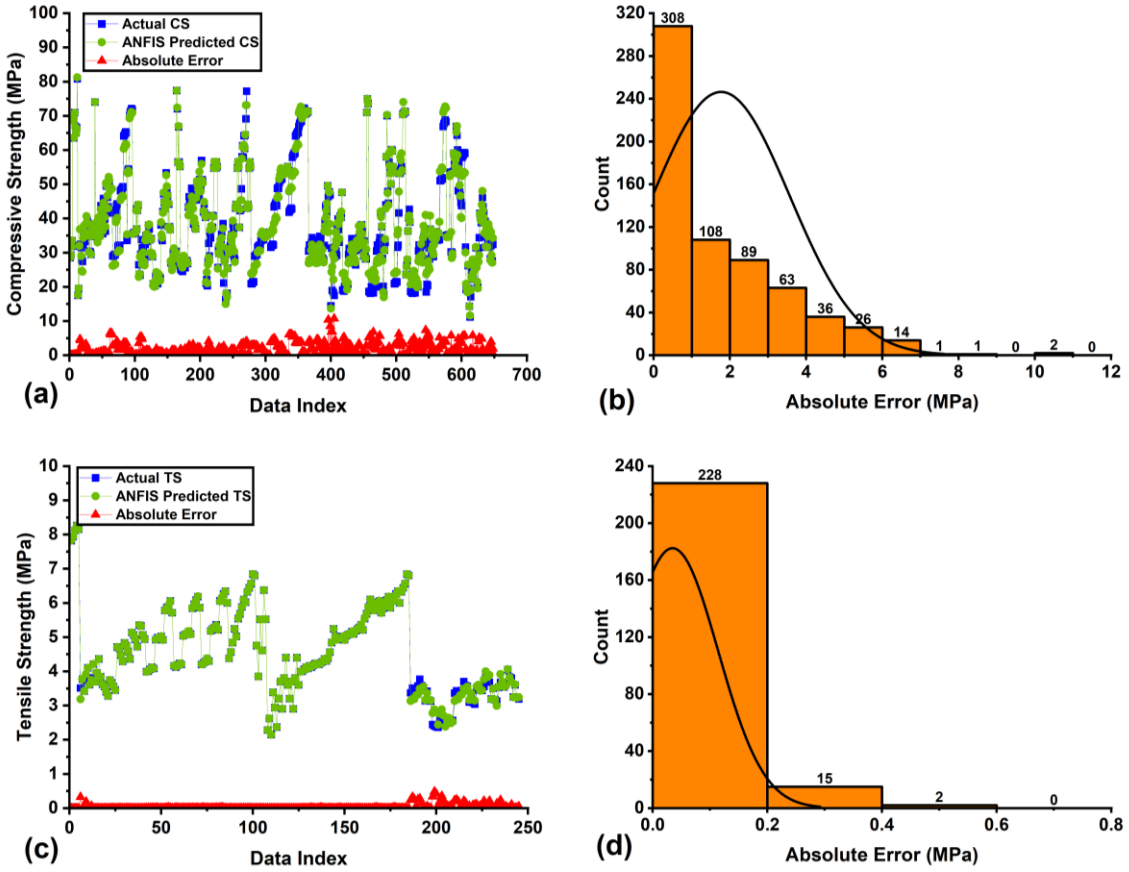


Fig. 4.4: Absolute error plots using ANFIS (a) scatter plot for CS, (b) histogram for CS, (c) scatter plot for TS, (d) histogram for TS.

4.4 GEP models development and performance assessment

Unlike the other models, GEP also provides predictive equations for outputs. So, before discussing the performance of GEP models, the development of predictive equations for CS and TS concrete is described in detail. The efficacy of the GEP-based developed models is significantly influenced by their fitting parameters. Therefore, it's crucial to find the best parametric settings to achieve an appropriate balance between complexity and simplicity, avoid overfitting, and enhance the model's generalization capability. The process of identifying the most suitable parameter values for a model that exhibits strong performance on new data generally involves an iterative procedure. In the present study, the trial-and-error approach was employed. During the experimentation phase of GEP, various combinations of parameters were tested by a varying genetic factor (chromosomes), linking function, head size, and the number of genes to identify the best parametric settings for both CS and TS models, as presented in **Table 3.8** and **Table 3.9**, respectively. The GEP program was designed to run indefinitely due to the significance of maintaining stable values for correlations and fitness functions. The models were evaluated based on several statistical error measures such as R2, RMSE, and MAE values. The statistical measures for the training

and evaluation phase of GEP models for CS are displayed in **Table 4.1**, while these measures for TS models are shown in **Table 4.2**. Results demonstrate that increasing parametric values (head size, genetic factors, linking functions, and gene size) while using different linking functions (+, -, /, x) enhanced model performance, as reflected by lower RMSE and MAE values and higher R2 values. Moreover, running time and the complication of the models may be increased by the increase in these parametric values, thereby making it difficult to comprehend the dynamics of the model. Following the statistical assessments conducted on the suggested GEP models, the best GEP models (CS12 and TS12) were chosen for predicting the CS and TS of concrete incorporated with SF and GGBS, respectively.

Table 4.1: Summary of GEP models for CS

Models	**Training dataset**						**Testing (validation) dataset**					
	R^2	RMSE	MAE	RRSE	R	ρ	R^2	RMSE	MAE	RRSE	R	ρ
CS1	0.861	5.047	4.070	0.374	0.928	0.107	0.839	5.277	4.071	0.402	0.916	0.110
CS2	0.860	5.079	4.069	0.375	0.927	0.108	0.852	5.031	4.098	0.387	0.923	0.111
CS3	0.863	4.960	4.069	0.370	0.929	0.107	0.857	4.997	4.065	0.379	0.926	0.110
CS4	0.871	4.934	4.063	0.357	0.933	0.102	0.861	4.957	4.060	0.366	0.928	0.106
CS5	0.870	4.959	4.061	0.362	0.933	0.104	0.881	4.967	4.087	0.345	0.939	0.108
CS6	0.767	7.691	5.085	0.483	0.876	0.166	0.784	6.738	5.030	0.465	0.885	0.169
CS7	0.798	6.871	5.075	0.463	0.893	0.177	0.805	6.401	5.074	0.457	0.897	0.180
CS8	0.798	6.871	5.075	0.463	0.893	0.123	0.805	6.401	5.064	0.468	0.897	0.127
CS9	0.857	5.046	4.066	0.346	0.926	0.095	0.857	5.035	4.071	0.385	0.925	0.099
CS10	0.866	5.002	4.068	0.366	0.930	0.094	0.869	4.988	4.063	0.366	0.932	0.098
CS11	0.876	4.998	3.98	0.376	0.936	0.085	0.875	4.935	4.062	0.366	0.936	0.088
CS12	**0.883**	**4.917**	**3.821**	**0.341**	**0.940**	**0.064**	**0.876**	**4.972**	**4.021**	**0.356**	**0.936**	**0.067**
CS13	0.882	4.922	3.881	0.344	0.939	0.069	0.872	4.985	4.065	0.359	0.934	0.072
CS14	0.881	4.949	3.897	0.345	0.939	0.069	0.881	4.997	4.058	0.346	0.938	0.072
CS15	0.860	5.759	4.071	0.462	0.927	0.099	0.861	4.967	4.071	0.400	0.928	0.102
CS16	0.778	6.691	5.082	0.483	0.882	0.156	0.787	6.634	5.083	0.467	0.887	0.159
CS17	0.808	5.875	4.945	0.463	0.899	0.167	0.805	6.401	4.865	0.447	0.897	0.170
CS18	0.818	5.781	4.936	0.433	0.905	0.125	0.805	6.401	4.826	0.436	0.897	0.128
CS19	0.860	5.079	4.069	0.375	0.927	0.103	0.852	4.931	4.869	0.387	0.923	0.107
CS20	0.861	5.035	4.068	0.373	0.928	0.102	0.846	5.073	4.888	0.393	0.920	0.106

Table 4.2: Summary of GEP models for TS.

Models	**Training dataset**						**Testing (validation) dataset**					
	R^2	RMSE	MAE	RRSE	R	ρ	R^2	RMSE	MAE	RRSE	R	ρ
TS1	0.896	0.365	0.265	0.312	0.947	0.058	0.871	9.277	0.385	0.353	0.933	0.059
TS2	0.895	0.379	0.275	0.313	0.946	0.058	0.884	8.931	0.395	0.340	0.940	0.060
TS3	0.899	0.396	0.245	0.309	0.948	0.058	0.889	8.747	0.395	0.333	0.943	0.059
TS4	0.907	0.353	0.242	0.298	0.952	0.055	0.892	7.597	0.392	0.322	0.945	0.057
TS5	0.906	0.376	0.242	0.302	0.952	0.056	0.913	7.967	0.382	0.303	0.956	0.058
TS6	0.803	0.469	0.367	0.403	0.896	0.089	0.816	10.738	0.387	0.409	0.903	0.091
TS7	0.834	0.427	0.341	0.387	0.913	0.095	0.836	10.401	0.381	0.401	0.914	0.097
TS8	0.834	0.411	0.342	0.387	0.913	0.067	0.836	10.401	0.384	0.411	0.914	0.068

TS9	0.892	0.388	0.276	0.289	0.945	0.052	0.888	8.346	0.387	0.339	0.942	0.053
TS10	0.902	0.367	0.245	0.306	0.949	0.051	0.901	8.430	0.382	0.321	0.949	0.053
TS11	0.912	0.345	0.239	0.315	0.955	0.046	0.907	8.257	0.379	0.284	0.952	0.047
TS12	**0.919**	**0.321**	**0.234**	**0.285**	**0.959**	**0.035**	**0.908**	**7.972**	**0.374**	**0.312**	**0.953**	**0.036**
TS13	0.918	0.322	0.238	0.287	0.958	0.037	0.904	8.283	0.378	0.315	0.951	0.039
TS14	0.917	0.349	0.237	0.288	0.958	0.037	0.912	7.997	0.377	0.304	0.955	0.039
TS15	0.896	0.359	0.255	0.386	0.946	0.054	0.893	8.967	0.385	0.351	0.945	0.055
TS16	0.814	0.491	0.366	0.403	0.902	0.084	0.818	10.634	0.386	0.410	0.905	0.086
TS17	0.844	0.487	0.356	0.387	0.919	0.090	0.836	10.401	0.396	0.392	0.914	0.092
TS18	0.854	0.478	0.345	0.362	0.924	0.067	0.836	10.401	0.391	0.383	0.914	0.069
TS19	0.895	0.379	0.245	0.313	0.946	0.056	0.884	8.931	0.385	0.340	0.940	0.058
TS20	0.897	0.387	0.247	0.312	0.947	0.055	0.878	9.073	0.387	0.345	0.937	0.057

The output of optimum GEP models used for predicting CS and TS are shown in the form of expression trees (ETs), as displayed in **Fig. 4.5** and **Fig. 4.6**. The ETs for CS and TS include four fundamental mathematical operations, namely +, /, -, and x, as illustrated in **Table 3.7**. These ETs were interpreted to establish the empirical correlations. The empirical expressions for the CS models are developed by considering multiplication as a linking function. The head size and the number of genes are considered as 10 and 4, respectively. The simplified equations EQ 4.1-4.5 are recommended for predicting the CS of concrete containing SF and GGBS. The indicators used in ETs are presented in **Table 4.3**.

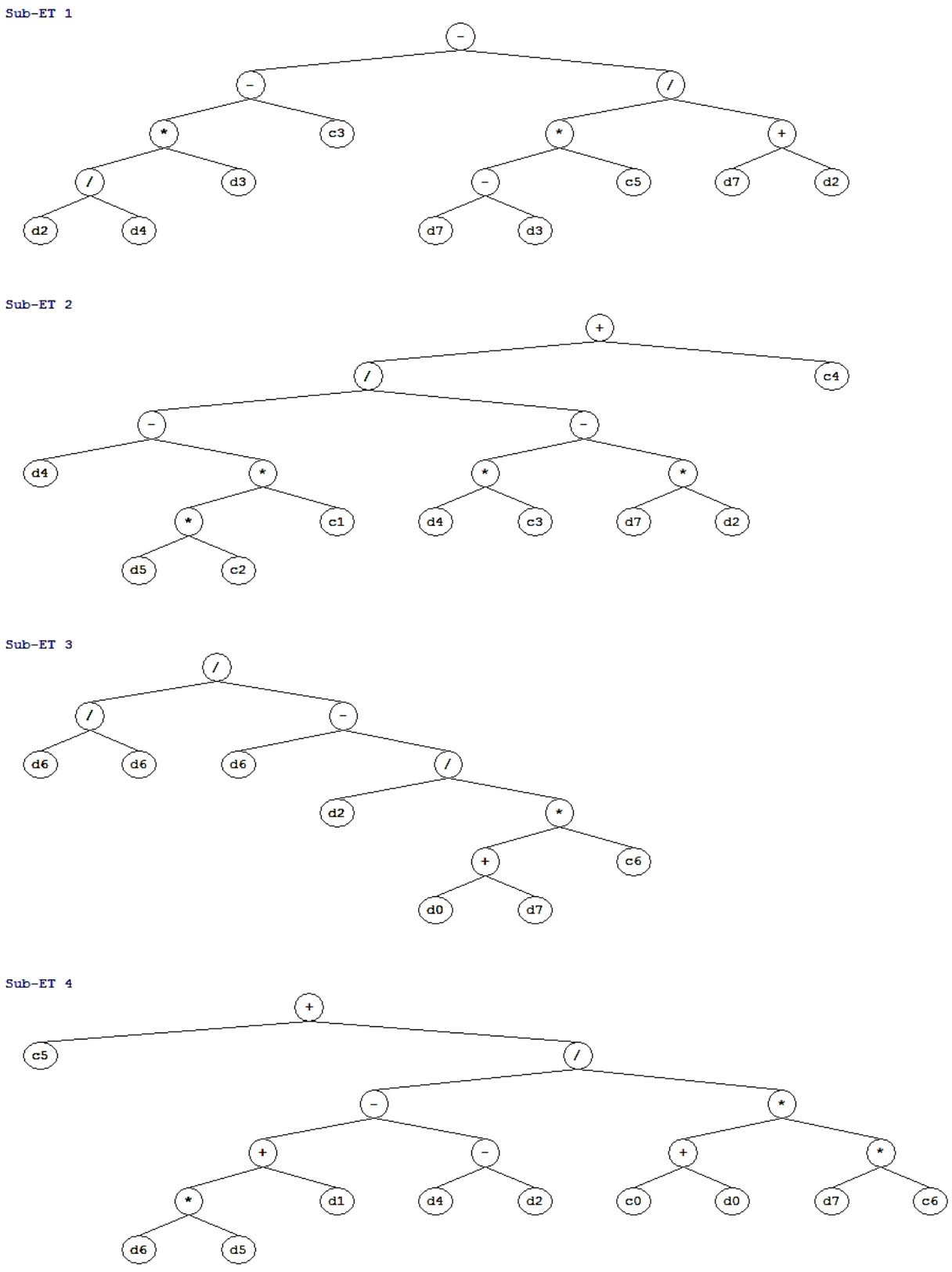


Fig. 4.5: GEP expression tree extracted for CS.

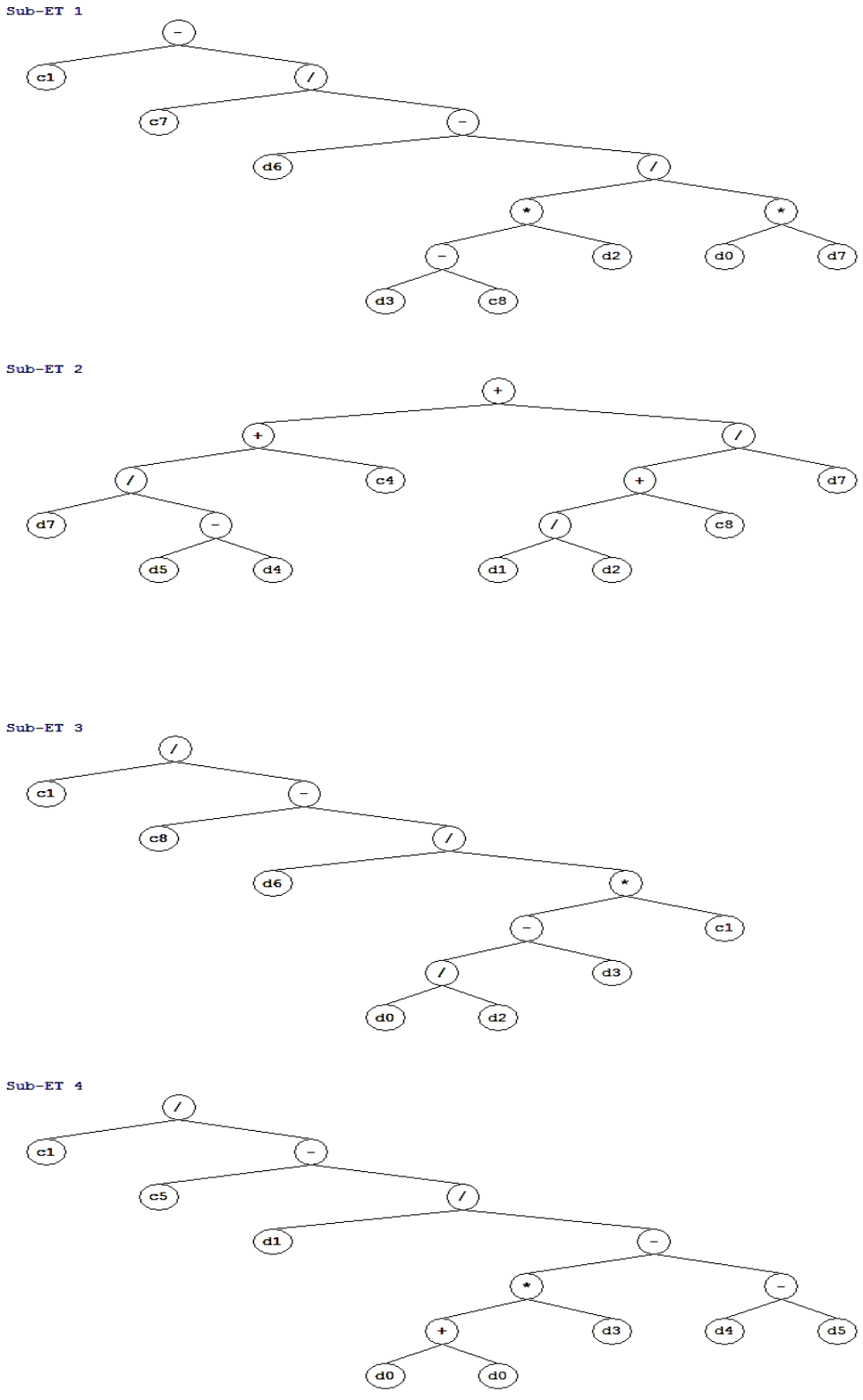


Fig. 4.6: GEP expression tree extracted for TS.

Table 4.3: ET indicators.

Parameter	Unit	Indicator in the expression tree	Description
C	Kg/m^3	d0	Cement content
GGBS	Kg/m^3	d1	Amount of GGBS
SF	Kg/m^3	d2	Amount of silica fume
SP	Kg/m^3	d3	Superplasticizer
CA	Kg/m^3	d4	Coarse aggregates (gravels)
FA	Kg/m^3	d5	Fine aggregates
W/B	-	d6	Water to cement ratio
Age	Days	d7	Age of specimens at the time of testing

$$\mathrm{CS\ (MPa)} = \mathrm{A} \times \mathrm{B} \times \mathrm{C} \times \mathrm{D} \tag{4.1}$$

$$\mathrm{A} = \frac{2.15(\mathrm{d7}-\mathrm{d3})}{\mathrm{d7}+\mathrm{d2}} + \frac{\mathrm{d2}}{\mathrm{d4}}\mathrm{d3} + 2.62 \tag{4.2}$$

$$\mathrm{B} = \left[\mathrm{d4} - \frac{14.\mathrm{d5}}{-4.38.\mathrm{d4}-\mathrm{d7}.\mathrm{d2}}\right] + 9.03 \tag{4.3}$$

$$\mathrm{C} = \frac{4.\mathrm{d6}(\mathrm{d0}+\mathrm{d7})}{4.\mathrm{d6}(\mathrm{d0}+\mathrm{d7})-\mathrm{d2}} \tag{4.4}$$

$$\mathrm{D} = 0.3939 + \frac{(\mathrm{d5}.\mathrm{d6}+\mathrm{d1})-(\mathrm{d4}-\mathrm{d2})}{\mathrm{d7}(5.31.\mathrm{d0}-69.56)} \tag{4.5}$$

The formula for the TS of concrete incorporated with SF and GGBS is created by considering the same parameters as considered for CS. Parametric settings for the optimum proposed model (TS12) are illustrated in **Table 4.2**. The empirical equations (4.6-4.10) are proposed to predict TS.

$$\mathrm{TS} = \mathrm{E} \times \mathrm{F} \times \mathrm{G} \times \mathrm{H} \tag{4.6}$$

$$\mathrm{E} = -1.5 - \frac{1.11.\mathrm{d0}.\mathrm{d7}}{\mathrm{d6}.\mathrm{d0}.\mathrm{d7}-\mathrm{d2}.(\mathrm{d3}-4.67)} \tag{4.7}$$

$$\mathrm{F} = \frac{\mathrm{d7}}{\mathrm{d5}-\mathrm{d4}} + \frac{\mathrm{d1}+9.66.\mathrm{d2}}{\mathrm{d7}.\mathrm{d2}} - 8.32 \tag{4.8}$$

$$\mathrm{G} = \frac{-2.47}{-7.2-\frac{\mathrm{d6}}{(\frac{\mathrm{d0}}{\mathrm{d2}}-\mathrm{d3})(-7.2)}} \tag{4.9}$$

$$\mathrm{H} = \frac{-4.98.\mathrm{d0}.\mathrm{d3}+2.49(\mathrm{d4}-\mathrm{d5})}{-14.32.\mathrm{d0}.\mathrm{d3}+7.16(\mathrm{d4}-\mathrm{d5})-\mathrm{d1}} \tag{4.10}$$

In addition, statistical analysis of GEP-based model outcomes is shown in **Fig. 4.7 (a and b)**. It can be noticed that MAE and RMSE values obtained from the CS-GEP and TS-GEP models are comparatively greater than those obtained from the ANN and ANFIS models. The CS-

GEP models show 22% and 47% higher values of RMSE in the testing phase as compared to ANN and ANFIS, respectively. For TS-GEP, these values are 15% and 20% higher, respectively. Similarly, R^2 values for the evaluation phase are 0.88 and 0.90, respectively, for CS-GEP and TS-GEP models, which are 7% and 3% lower than ANN models, whereas these values are 9% and 10% lower than ANFIS models. Other statistical error values, including RSE and RRMSE calculated for GEP models, are comparable to ANN and ANFIS-based models for CS and TS. The results of the error analysis of GEP-based predictive models for CS and TS are shown in **Fig. 4.8 (a-d)**. It can be noted that the absolute error values between GEP anticipated and actual values are slightly higher for CS and lower for TS models as contrast to those of ANN and ANFIS models. Moreover, as shown in **Fig. 4.8 (b and d)**, 60% of the error values of CS-GEP fall in the range of 0-4 MPa. Similarly, 77% of the error values of TS-GEP fall in the range of 0-0.4 MPa, which is relatively less than ANN and ANFIS models. The possible reason for the discrepancy in the relative efficacy of ANN, ANFIS, and GEP models for CS and TS could be attributed to the variability in the corresponding experimental datasets. The GEP-based empirical algorithm has demonstrated superior predictive performance compared to the ANN and ANFIS when applied to larger datasets.

GEP-based developed models for CS and TS were further evaluated by using the correlation p-values, which represent the likelihood of significance. The statistical analysis was conducted using the SPSS software, which is commonly employed in practical applications. A significance threshold of 5% was used in the computation, as reported by Güllü Güllü & Fedakar [134]. The p-value indicates the degree of confirmation against the null hypothesis, in which a lower p-value corresponds to a greater degree of confirmation. The p-values obtained for GEP models for CS and TS are close to zero, indicating a significant degree of correlation between the observed and projected values.

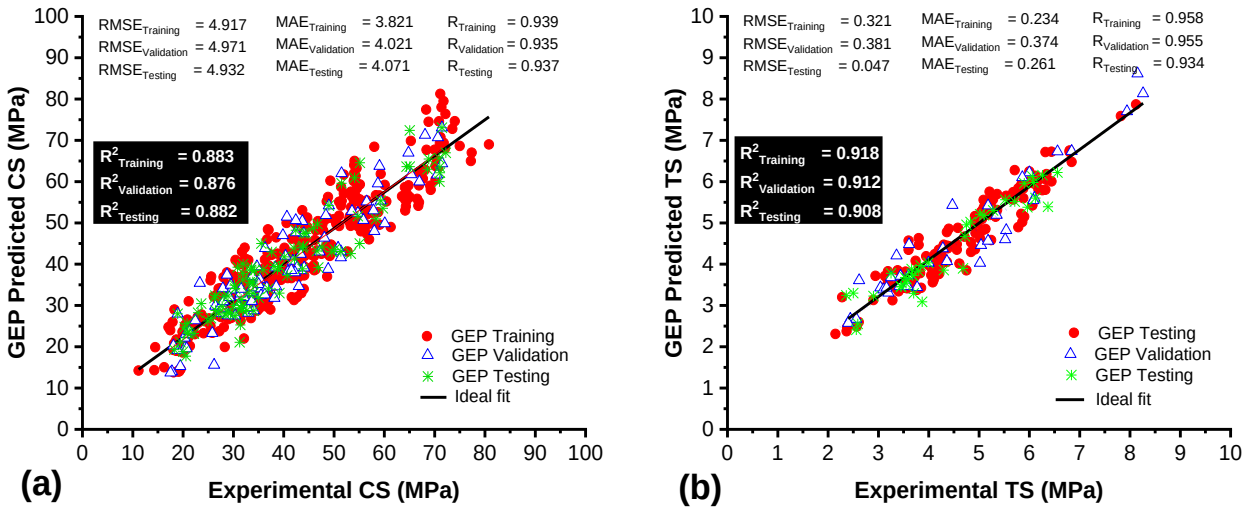


Fig. 4.7: Comparison of experimental and anticipated results using GEP for (a) CS, (b) TS.

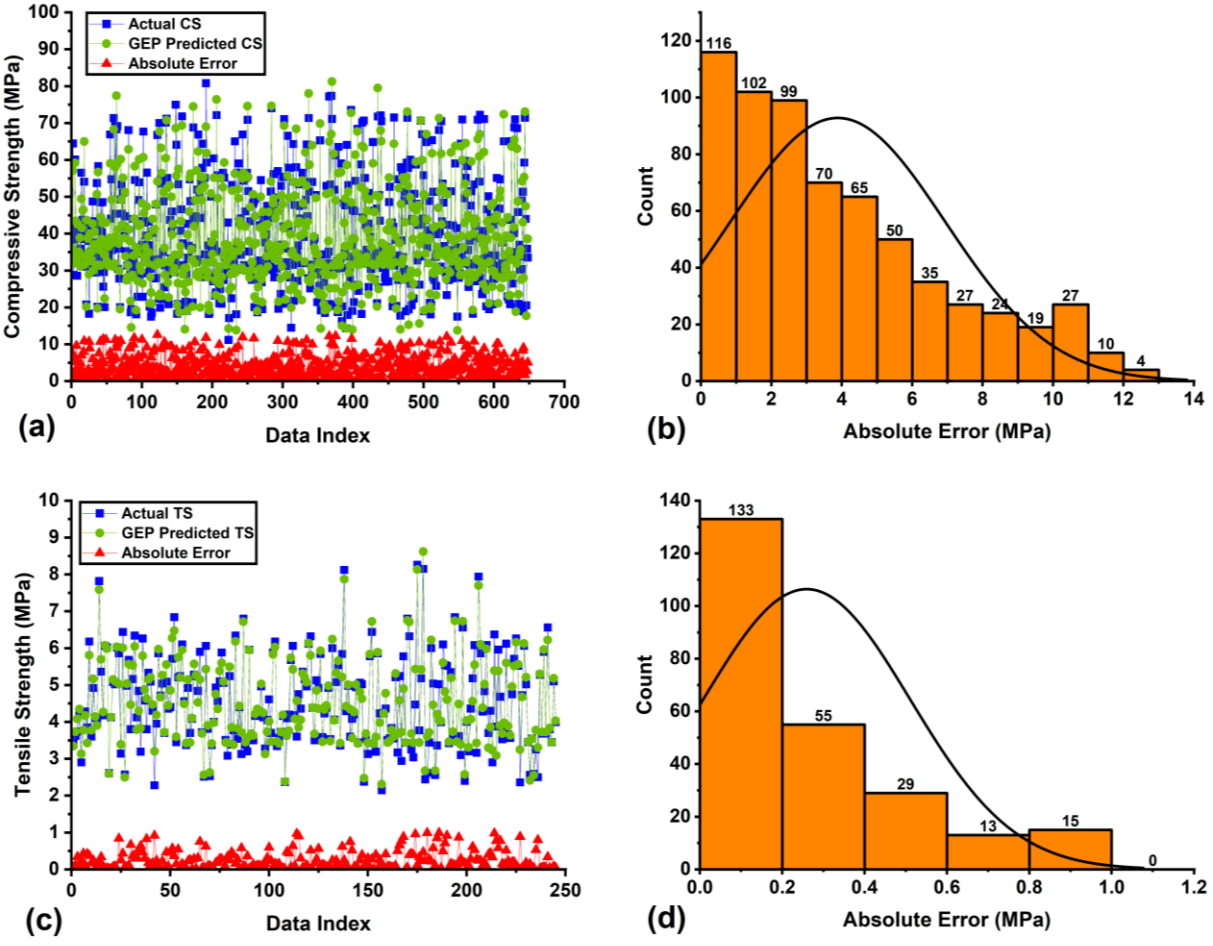


Fig. 4.8: Absolute error plots using GEP (a) scatter plot for CS, (b) histogram for CS, (c) scatter plot for TS, (d) histogram for TS.

4.5 Experimental validation

4.4.1 Workability

Fig. 4.9 displays the slump readings for all six tested combinations. The minimum slump

reading of 42mm was noted in a mixture containing 20% SF and 40% GGBS, which is 55% less than the control mix. It is clear that slump readings decrease with an increase in cement replacement level with SCMs (SF and GGBS) in concrete. This is due to the fact that SF and GGBS have greater specific surface area and more cohesion, which requires extra water to maintain workability as compared to normal concrete [135].

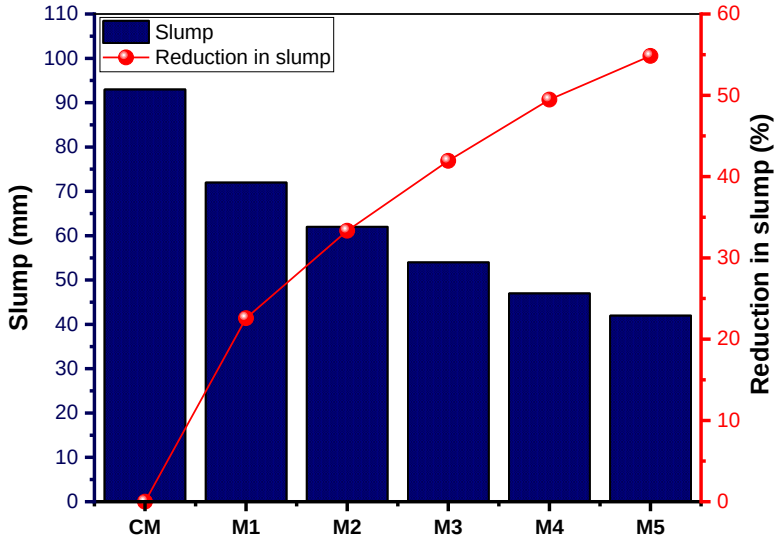


Fig. 4.9: Slump testing results

4.4.2 Compressive strength

The CS was determined using the compression testing machine, following the ASTM C39 guidelines. The results of laboratory-derived CS are shown in **Fig. 4.10**. It can be noted that CS increased with the increase in the age of testing, irrespective of the specific binder mixture employed for the development of the samples. The increment in CS at later ages with the addition of SF and GGBS is more pronounced than in CM. This is because of the pozzolanic reactions and nucleation effects associated with SF that reduce calcium hydrate (CH) content and enhance the densification of samples. Further, the enhanced CS at later age can also be attributed to the fine grain sizes and filler effects due to the inclusion of SF and GGBS [136]. The concrete mix containing SCMs (10%SF and 30% GGBS) has demonstrated the highest CS of 50.4 MPa at the age of 90 days. The optimum replacement of SCMs resulted in 21% enhanced CS as compared to CM. Meanwhile, the addition of SCM exceeding 40% (M5 mix) led to a decrease in CS by 11%. This is due to the fact that excessive substitution of cement with supplementary cementitious materials (SCMs) could reduce the presence of reactive substances needed for hydration, resulting in weaker connections, and weakened durability in the hardened concrete [135]. It is important to mention that ACI-recommended conversion factors were used to convert the CS values into TS values [137].

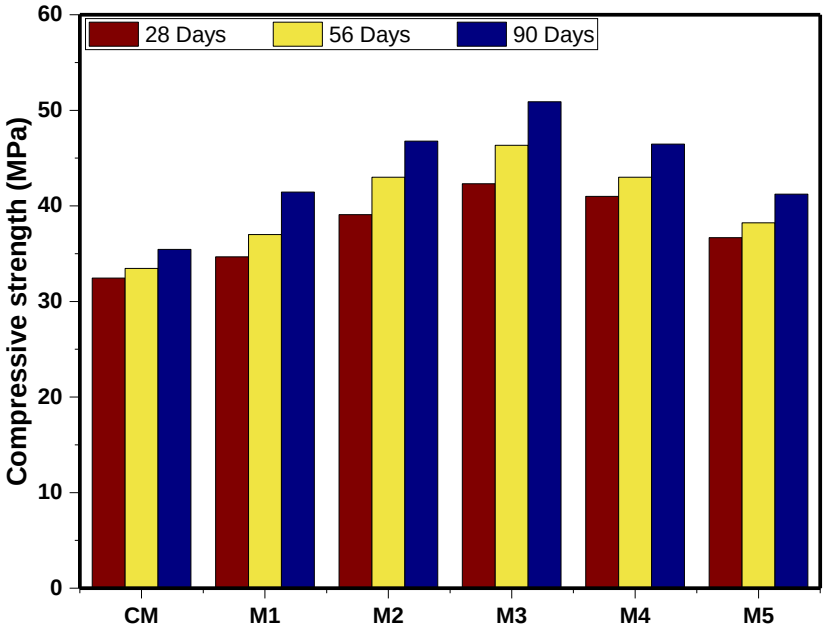


Fig. 4.10: Laboratory-derived CS outcomes.

4.4.3 Comparison of experimental results with proposed models

After thorough evaluations of models using various mathematical approaches, a separate dataset was constructed using experimental testing to evaluate the effectiveness and applicability of the produced ML models. The separate dataset allows a thorough evaluation of the model's ability to make predictions without relying on the initial data used for training. The statistical summary of the constructed dataset using experimental findings is presented in **Table 4.4**. This cross-validation approach is crucial in confirming the models' ability to generalize to unfamiliar data, which is a vital part of their practical applications. It is important to mention that the same parametric settings were used in the development of the models so that hyperparameter optimization can also be verified.

Table 4.4: Statistical summary of the developed database for external validation.

Parameters	Cement (kg/m^3)	Water (kg/m^3)	SF (kg/m^3)	GGBS (kg/m^3)	CA (kg/m^3)	FA (kg/m^3)	Age (Days)	CS (MPa)	TS (MPa)
Mean	283.67	156	41.17	90	1050	730	59.56	40.53	3.16
Median	270	156	104	1050	730	56	41.12	3.17	
SD	84.28	0	57.56	0	0	25	5.1	0.33	
Range	249	0	166	0	0	62	18.45	1.17	
Minimum	166	156	0	1050	730	28	32.44	2.6	
Maximum	415	156	166	1050	730	90	50.89	3.77	

The comparison between the findings of external experimental validation and ML models is depicted in **Fig. 4.11 (a and b)** for CS and TS models, respectively. It can be seen that findings are highly comparable with an R^2 of more than 0.88 in all developed models, as shown in **Fig. 4.12 (a and b)** for CS and TS, respectively. Furthermore, the error percentage

between experimental and estimated values falls in the acceptable range of less than 10%, as shown in **Table 4.5**. Minor differences can be noted between the outcomes of actual and validation models, which can be due to the lower number of databases used for validation. The optimum replacement in experimental validation is found to be 10% SF and 30% GGBS, which is well aligned with the literature. The consistency of the results demonstrated the efficacy of the models in accurately identifying and anticipating the fundamental patterns of mechanical characteristics of concrete incorporated with SF and GGBS. The strong alignment between the model predictions and the findings obtained in the laboratory highlights the reliability of the ML models that have been constructed, thus gaining confidence in their dependability for real-world use.

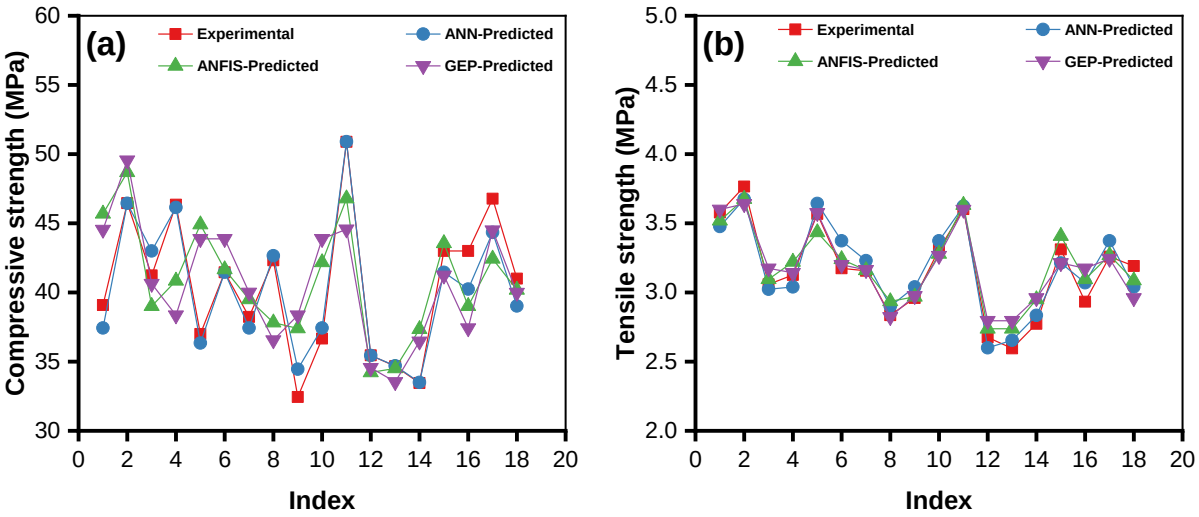


Fig. 4.11: Comparison between experimental and models anticipated values for external validation (a) CS and (b) TS.

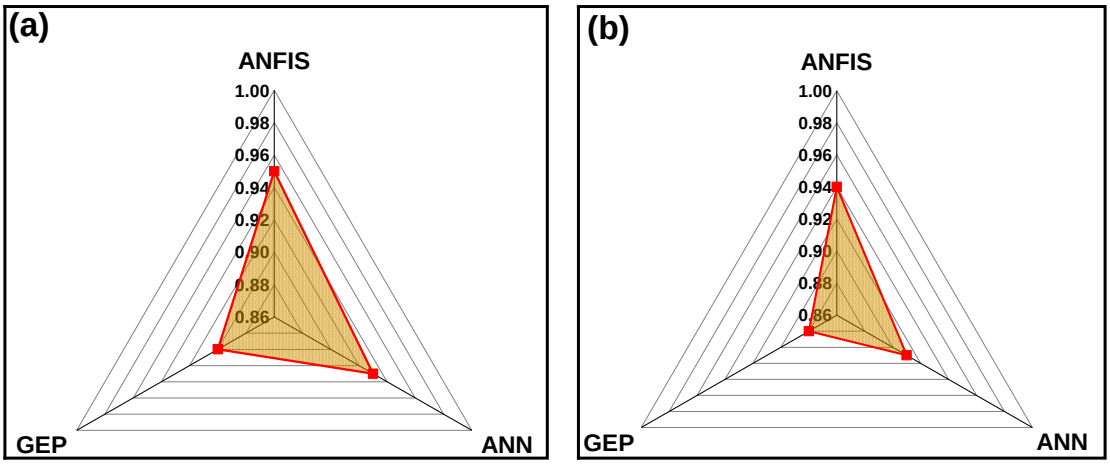


Fig. 4.12: Models R2 values based on the experimental dataset (a) CS and (b) TS.

Table 4.5: Percentage error comparison between models predicted and experimental values.

CS-Experimental Values	Error % in ML-models			TS-Experimental Values	Error % in ML-models		
	ANN	ANFIS	GEP		ANN	ANFIS	GEP
39.09	4.24	1.67	4.12	3.58	2.82	1.64	0.63
46.46	0.03	4.33	2.61	3.77	2.41	2.40	3.34
41.23	4.29	2.03	6.30	3.06	1.16	1.21	3.80
46.34	0.42	4.50	1.88	3.13	2.78	3.02	0.48
37.00	1.78	3.08	5.31	3.57	2.10	3.75	0.22
41.45	0.04	1.06	6.42	3.17	6.25	2.06	0.79
38.23	2.09	3.66	2.44	3.16	2.31	0.10	0.20
42.32	0.78	6.70	3.13	2.84	2.26	3.50	0.44
32.44	6.20	1.08	3.66	2.96	2.71	0.29	0.54
36.67	2.07	1.91	2.04	3.31	1.88	0.96	1.39
50.89	0.03	0.03	4.73	3.60	0.60	0.83	0.09
35.45	0.02	2.32	3.32	2.68	2.84	2.32	4.48
34.67	0.11	6.38	3.38	2.60	2.27	5.51	7.74
33.45	0.15	7.71	5.19	2.77	2.19	6.37	6.73
43.00	3.60	1.28	4.11	3.31	2.96	2.93	2.95
43.00	6.42	6.37	2.05	2.93	4.65	5.51	8.22
46.78	5.20	0.26	5.15	3.26	3.52	0.30	0.42
41.00	4.82	3.04	3.53	3.19	4.74	3.30	7.26

4.6 Statistical evaluation and comparative analysis of models

The comparative analysis based on statistical parameters of the proposed models employing ANN, ANFIS, and GEP is presented in **Table 4.6**. The proposed models efficiently predicted the CS and TS of concrete containing SCMs by considering the impact of all specified input parameters, as obvious from the significantly reduced errors (RMSE, MAE, and RSE). The R^2 values for CS neural models using ANN and ANFIS are around 95% and 99%, respectively, for all three data sets (training, validation, and testing), while they are just below 90% for GEP models. Likewise, the other statistical errors, such as RMSE and MAE for CS models, are lowest for ANFIS, followed by ANN. The mean R-value for all three proposed models is higher than 90%. The higher R values indicate a strong connection between the input parameters, which results in higher prediction accuracy [138]. Similarly, the ANFIS models exhibit higher R^2 values in all three data sets for TS models. All the proposed models for TS showed R^2 values greater than 90% in the following order: R^2_{ANFIS} 0.99 > R^2_{ANN} 0.95 > R^2_{GEP} 0.91. The superior results indicated by ANFIS can be attributed to the integration of the neural networks' training capability and the fuzzy logic's reasoning ability. The predictive capabilities of the established models are also compared based on statistical performance

measures such as PI and OF, as depicted in **Fig. 4.13 (a-d)**. The values of these measures nearer to zero indicate the better precision of the models. In the established models for CS, ANFIS models show the lowest PI equal to 0.031, 0.043, and 0.033 in the training, validation, and testing phases, respectively; for CS-ANN, these values are 0.043, 0.051, and 0.047, respectively. Similarly, PI values in the case of CS using GEP are 0.064, 0.084, and 0.073 for the training, testing, and validation phases, respectively. Additionally, for TS models, PI is lower than 1% for all three testing conditions. According to the criteria discussed by [139], the PI value should be less than 20% for accurate prediction. Hence, all the developed models satisfy the given criteria and can be considered good for the anticipation of CS and TS of concrete incorporated with SF and GGBS. However, in ANN models, significant variation was observed between PI values of training and testing conditions. As discussed in a recent study by Jalal et al. [140], this variation can be attributed to the local minima problem that is commonly associated with ANN models. According to Zhang et al. [141], the optimizing process may cease at a locally optimized state rather than a global termination state, leading to inaccurate predictions. The concept of local optimization involves the closure of the search process for a particular problem when the ideal solution is discovered even before the optimum one. However, in ANFIS models, the aforementioned issue can be addressed by the training capacity of ANN and the fuzzification of logical reasoning in the fuzzy toolbox. Furthermore, it can be noted that OF values obtained from ANN, ANFIS, and GEP models for CS are 0.042, 0.031, and 0.062, respectively, and for TS models, these values are 0.029, 0.009, and 0.035, respectively. The OF values for all cases are close to zero, revealing the validity of the proposed models while controlling the overfitting.

The summary of statistical measures indicates that in all predictive models, namely ANN, ANFIS, and GEP, the projected values are very close to actual values for both CS and TS models. Considering the R and R^2 values, the upward sequence is ANFIS > ANN > GEP. Meanwhile, the ANFIS models for both CS and TS exhibit the lowest values for error measures (RMSE, RSE, MAE, RRMSE) followed by ANN and GEP. However, the GEP models show superior performance compared to ANFIS and ANN with respect to the closeness of RMSE, MAE, RSE, and R^2 values between the training, validation, and testing sets for both CS and TS models. In addition, the GEP represents an evolutionary technique that gives a simple empirical formula for forecasting the CS and TS (refer to Equations (25) and (26)). This approach significantly diminishes the overall duration needed to estimate the CS and TS compared with conventional testing procedures (i.e. experimental investigations),

and the evaluation process can be completed at a significantly faster rate. Therefore, the applications of the presented equations provide a viable and quick method for estimating the CS and TS of concrete incorporated with SF and GGBS.

Table 4.6: The summary of statistical calculations.

Proposed models		Subset type	R^2	RMSE	MAE	RRMSE	RSE	R	ρ	OF
CS	ANN	Trn-set	0.9419	3.3284	1.788	0.0851	0.0583	0.9705	0.0436	0.0427
		Vald-set	0.9377	4.0493	2.791	0.1008	0.0647	0.9683	0.0512	
		Test-set	0.938	3.8269	2.832	0.0925	0.0657	0.9685	0.0472	
	ANFIS	Trn-set	0.9702	2.3688	1.681	0.0731	0.0298	0.985	0.031	0.0311
		Vald-set	0.9698	3.0691	2.321	0.0859	0.0583	0.9848	0.0436	
		Test-set	0.9678	2.5721	1.983	0.0659	0.0339	0.9838	0.03325	
	GEP	Trn-set	0.8831	4.9174	3.821	0.1251	0.1161	0.9392	0.0642	0.0621
		Vald-set	0.8763	4.9716	4.021	0.1263	0.1243	0.9353	0.0841	
		Test-set	0.8821	4.9321	4.071	0.1255	0.1191	0.9371	0.0737	
TS	ANN	Trn-set	0.9585	0.2699	0.202	0.0592	0.2675	0.979	0.0299	0.0299
		Vald-set	0.9384	0.3211	0.255	0.0723	0.0671	0.9681	0.0366	
		Test-set	0.9311	0.3433	0.258	0.0761	0.0692	0.9654	0.0392	
	ANFIS	Trn-set	0.9951	0.0886	0.034	0.0198	0.0167	0.9975	0.0099	0.0099
		Vald-set	0.9981	0.0381	0.133	0.0077	0.0021	0.9991	0.0039	
		Test-set	0.998	0.0383	0.017	0.0076	0.002	0.9992	0.0038	
	GEP	Trn-set	0.9188	0.3213	0.234	0.0712	0.0812	0.9585	0.0362	0.035
		Vald-set	0.9123	0.3812	0.374	0.0782	0.0891	0.9551	0.0426	
		Test-set	0.9081	0.4043	0.261	0.0861	0.0972	0.9342	0.0533	

Note: Trn: Training, Vald: Validation, Test: Testing

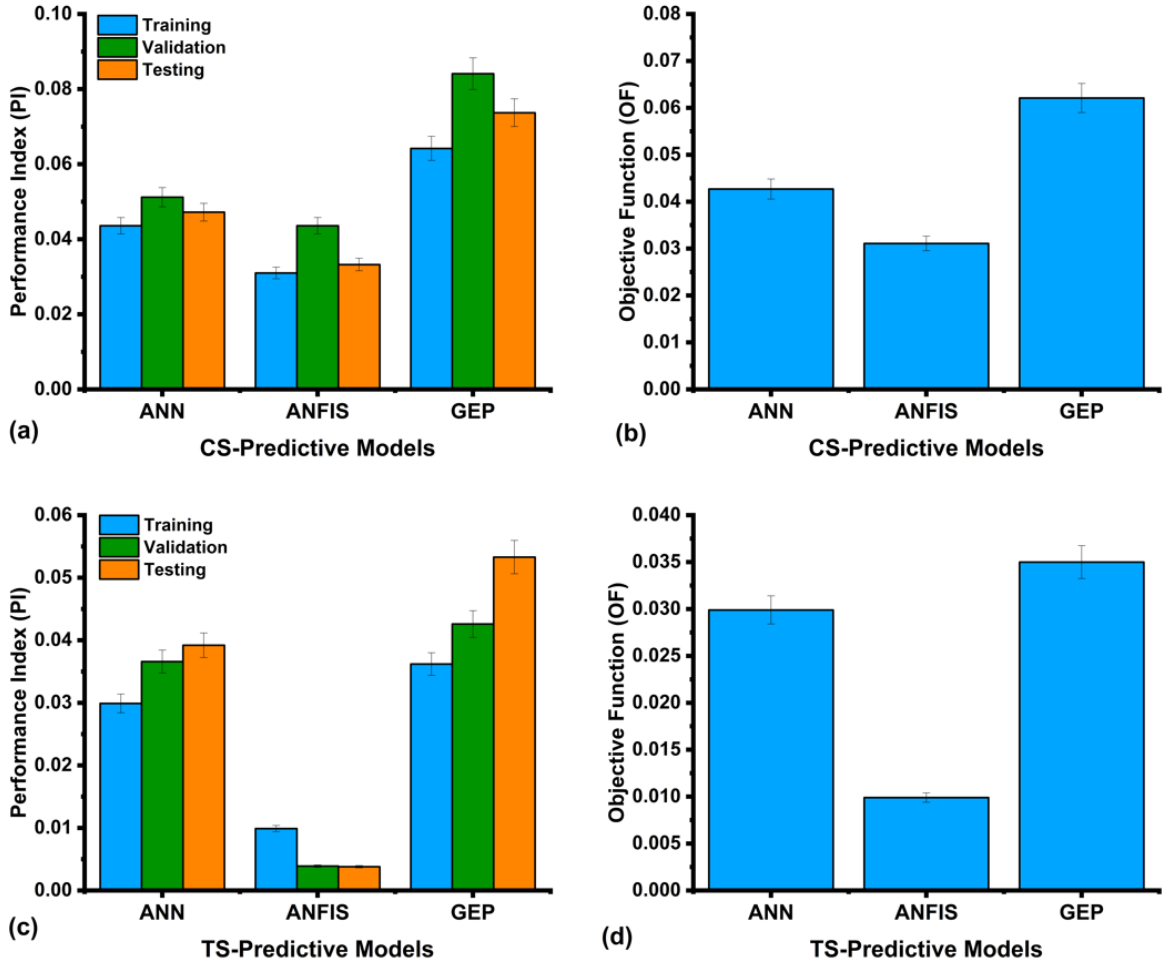


Fig. 4.13: The objective function (OF) and performance index (PI) of; (a)–(b) CS models, (c)–(d) TS models.

4.7 Validation of GEP-based equations

External validation is essential to assess the reliability and generalization capability of ML-based models. Therefore, in this study, several statistical tests are employed for external validation of the best-proposed model (GEP), as shown in **Table 4.7**. Golbraikh et al. [142] recommended that for higher accuracy of the model, the gradient of one of the regression lines (k or k') crossing through the origin must pass close to unity. The gradient of the regression lines for the CS model is 0.971, while for TS, it is 0.973. In addition, the performance indices values (i.e., m and n) of less than 0.1 are considered reliable. This signifies a strong correlation and more accuracy. It can be seen that both CS and TS models exhibit m and n values less than the recommended range. Moreover, several researchers have recommended that the squared coefficient (R02) between actual and estimated values should also be close to 1 [143]. It is evident from the findings that the suggested models satisfy the external verification requirements, demonstrating the models' excellent validity, prediction power, and independence from simple correlations between inputs and output.

Table 4.7: GEP model external validation.

S. No.	Equation	Condition	Model	
			CS	**TS**
1	R	$R > 0.8$	0.991	0.993
2	$k = \sum_{i=1}^{n} \frac{(a_i \times m_i)}{{a_i}^2}$	$0.85 < k < 1.15$	0.971	0.973
3	$k' = \sum_{i=1}^{n} \frac{(a_i \times m_i)}{{m_i}^2}$	$0.85 < k' < 1.15$	1.002	1.001
4	${R_o}^2 = 1 - \frac{\sum_{i=1}^{n}(m_i - {a_i}^o)^2}{\sum_{i=1}^{n}(m_i - {m_i}^o)^2}$, ${a_i}^o = k \times m_i$	${R_o}^2 \cong 1$	0.969	0.970
5	${R'_o}^2 = 1 - \frac{\sum_{i=1}^{n}(a_i - {m_i}^o)^2}{\sum_{i=1}^{n}(a_i - {a_i}^o)^2}$, ${m_i}^o = k' \times a_i$	${R'_o}^2 \cong 1$	0.990	0.991
6	$m = \frac{(R^2 - {R_o}^2)}{R^2}$	$m < 1$	0.0541	0.055
7	$n = \frac{(R^2 - {R'_o}^2)}{R^2}$	$n < 1$	0.005	0.005

4.8 Sensitivity and Parametric Analysis

In the case of modeling based on machine learning, multiple evaluations must be performed to ensure that models are reliable and work efficiently across various data combinations. Firstly, sensitivity analysis (SA) was performed to determine the relative influence of various input parameters to the properties of concrete incorporated with SF and GGBS using Equations. 4.11 and 4.12.

$$M_i = f_{max}(m_i) - f_{min}(m_i) \tag{4.11}$$

$$SA = \frac{M_i}{\sum_{n}^{j=1} M_j} \tag{4.12}$$

Here, fmax(mi) and represents fmin(mi) maximum and minimum anticipated output based on the input variable; other input variables are fixed at average values. **Fig. 4.14 (a and b)** display the findings of the SA for CS and TS, respectively. It can be seen that W/B, age and cement has the highest, whereas SP has the lowest impact on CS and TS of concrete containing binary SCMs (SF and GGBS). The effect of other input parameters on CS and TS is relatively small and quite similar from a structural materials perspective [144], [145].

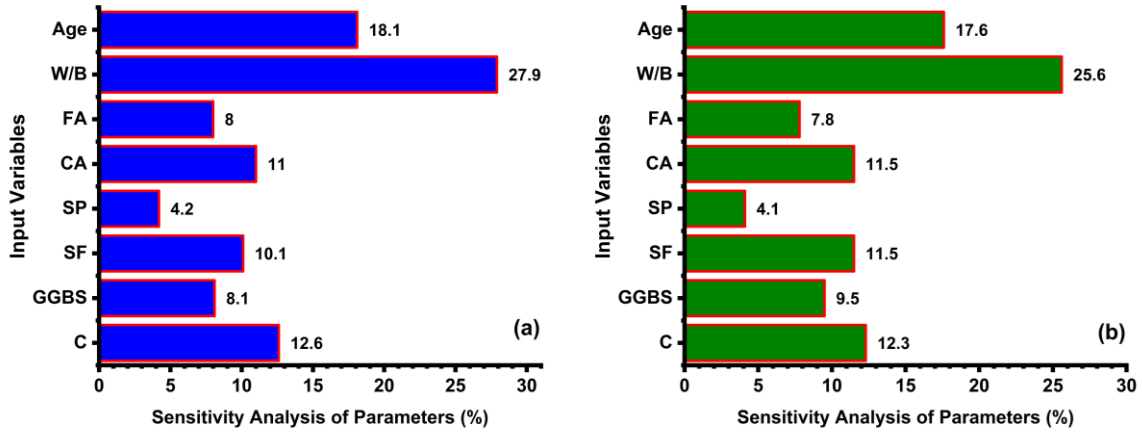


Fig. 4.14: Sensitivity analysis outcomes of (a) CS models and (b) TS models

Furthermore, numerous research studies have suggested parametric analysis to evaluate the efficiency of the most impactful input parameters in predicting outcomes. In parametric analysis, each individual input variable was changed by using definite increments, while the remaining variables were fixed at their mean values. The resulting changes in the output were then observed and recorded. **Fig. 4.15** and **Fig. 4.16** show the results obtained from parametric analysis for the proposed CS and TS models. It is clear that the CS and TS increased with a reduction in the W/B ratio and vice versa. Similar results have been found in the experimental investigation carried out by Bhaskar et al. [95]. They compared the mechanical characteristics of concrete incorporated with SF and GGBS by using varying W/B ratios and discovered that concrete exhibited higher strength properties with W/B of 0.35 as opposed to W/B of 0.45 and 0.55. GGBS and SF trends show that CS and TS of the concrete increase up to the optimum percentage of around 30% and 12% of GGBS and SF replacement, respectively, and then decrease, which is in line with the available literature [94]. In several experimental studies, the authors investigated 30% and 10% GGBS and SF as optimum replacement proportions in concrete. For instance, Bhaskar et al. [93] conducted an experimental study to determine the optimum replacement of SF and GGBS in concrete. They used various combinations of SF+GGBS (by using constant W/B of 0.45) and found cement replacement with 30% GGBS and 10% SF to be the optimum dosage for CS and TS of concrete. SP has very little influence on concrete strength, which is also observed in experimental studies [146]. The effect of coarse and fine aggregates greatly depends on their shape, size, and type; however, the increase in coarse-to-fine aggregates ratio increases the CS and TS of concrete, which is evident from experimental studies [147]. Finally, the CS and TS of concrete increased with its age, which is in line with the available literature [148]. For example, Suda

and Rao reported that CS and TS of ternary blended concrete incorporated with SF and GGBS increased with age [93]. Based on the parametric analysis, it can be concluded that the recommended models are appropriate and consistent in predicting the CS and TS.

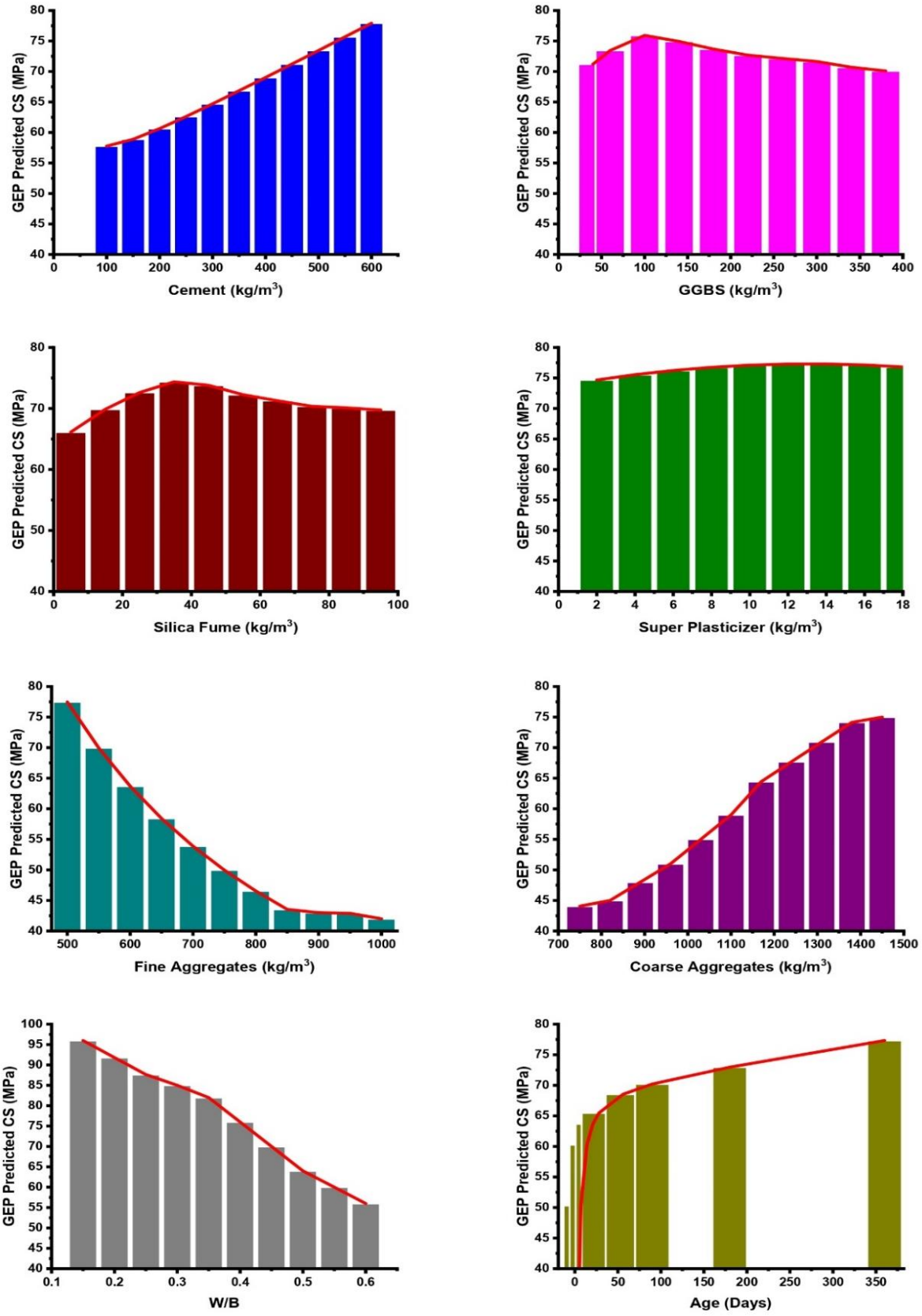


Fig. 4.15: Parametric analysis of CS input parameters

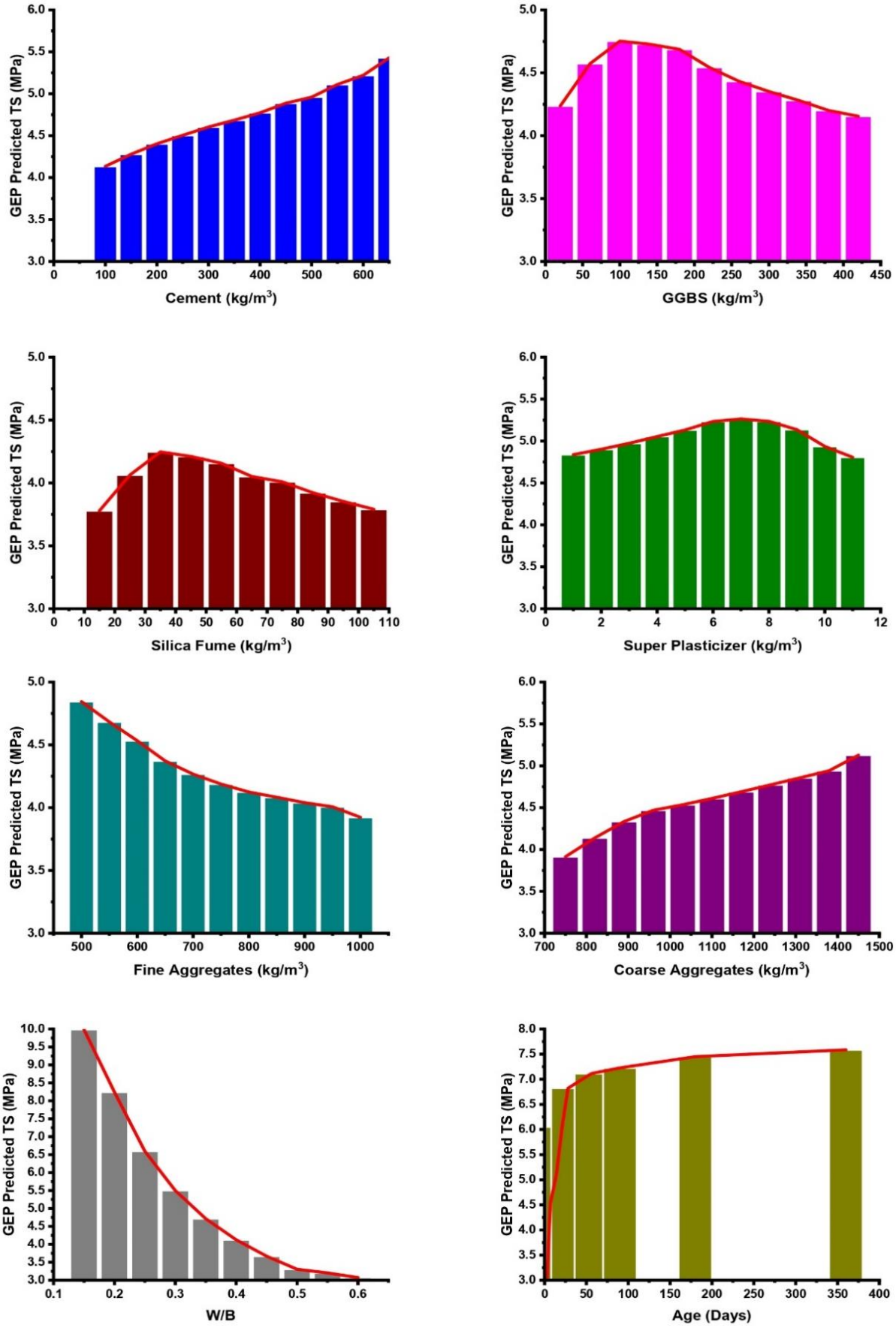


Fig. 4.16: Parametric analysis of TS input parameters.

4.9 Comparison of proposed models with literature

According to previous studies, no empirical models have been developed to estimate the mechanical characteristics of concrete containing binary SCMs (SF and GGBS), including a wide range of input parameters that were taken into account in this study. However, in this section, the outcomes of this study are compared with a few of the existing studies that used similar ML techniques for other SCMs. The descriptive summary of statistical metrics and performance indicators of proposed models and models recommended by the available literature is presented in **Table 4.8**. It can be seen that the findings of the key statistical measures such as R^2, MAE, and RRME for the best-proposed models from this study and existing models are highly comparable. Therefore, it can be concluded that the suggested models can be effectively utilized in practical applications to estimate the CS and TS of concrete incorporated with SF and GGBS. Additionally, GEP-based predictive equations can be used to pre-design the concrete mixture containing SF and GGBS, which will promote sustainable materials in concrete.

Table 4.8: Comparison of proposed models with literature.

Description	ML techniques	Data sub-type	MAE	RMSE	R^2	PI	References
Predictive modeling for CS of fly ash-based concrete	**ANN**	Trn-set	2.7	6.12	0.9571	0.033	[149]
		Vald-set	0.3	2.19	0.956	0.03	
		Test-set	0.5	2.35	0.9487	0.035	
	ANFIS	Trn-set	2	6.12	0.9747	0.036	
		Vald-set	0.3	2.19	0.9644	0.036	
		Test-set	0.4	2.38	0.9539	0.04	
	GEP	Trn-set	1.8	6.12	0.9849	0.035	
		Vald-set	0.3	2.18	0.9592	0.04	
		Test-set	0.5	2.39	0.9539	0.04	
Predictive modeling for slump of fly ash-based concrete	**ANN**	Trn-set	0.3	1.5	0.9539	0.029	[149]
		Vald-set	0.1	0.76	0.9487	0.033	
		Test-set	0.1	0.65	0.9434	0.03	
	ANFIS	Trn-set	0.4	1.53	0.9644	0.03	
		Vald-set	0.1	0.73	0.9539	0.031	
		Test-set	0.1	0.66	0.9487	0.03	
	GEP	Trn-set	0.3	1.49	0.9798	0.034	
		Vald-set	0.1	0.71	0.9747	0.032	
		Test-set	0.1	0.62	0.9747	0.051	
Predictive modeling for CS of hemp-based bio composite concrete	**ANN**	Trn-set	0.5332	0.9968	0.9505	0.2523	[150]
		Test-set	0.4985	0.984	0.9214	0.2496	
	ANFIS	Trn-set	0.2579	0.4568	0.9893	0.1145	
		Test-set	0.2466	0.539	0.9694	0.135	

Predictive modeling for thermal conductivity of hemp-based bio composites concrete	**ANN**	Trn-set	0.0125	0.0513	0.9113	0.1572	[150]
		Test-set	0.0077	0.0126	0.9911	0.0354	
	ANFIS	Trn-set	0.0014	0.004	0.9992	0.0121	
		Test-set	0.0012	0.0018	0.9998	0.0051	
Predictive modeling for CS of fly ash-based concrete	**ANN**	Trn-set	5.925	6.029	0.9214	0.095	[151]
		Vald-set	4.127	4.986	0.9649	0.043	
	ANFIS	Trn-set	3.286	4.086	0.9618	0.046	
		Vald-set	2.084	2.593	0.9889	0.019	
	GEP	Trn-set	5.823	5.971	0.9263	0.091	
		Vald-set	2.057	2.643	0.9818	0.025	
Predictive modeling for CS of SF and GGBS based concrete	**ANN**	Trn-set	1.788	3.3284	0.9419	0.0436	This study
		Vald-set	2.791	4.0493	0.9377	0.0512	
		Test-set	2.832	3.8269	0.938	0.0472	
	ANFIS	Trn-set	1.681	2.3688	0.9702	0.031	
		Vald-set	2.321	3.0691	0.9698	0.0436	
		Test-set	1.983	2.5721	0.9678	0.0333	
	GEP	Trn-set	3.821	4.9174	0.8831	0.0642	
		Vald-set	4.021	4.9716	0.8763	0.0841	
		Test-set	4.071	4.9321	0.8821	0.0737	
Predictive modeling for TS of SF and GGBS based concrete	**ANN**	Trn-set	0.202	0.2699	0.9585	0.0299	This study
		Vald-set	0.255	0.3211	0.9384	0.0366	
		Test-set	0.258	0.3433	0.9311	0.0392	
	ANFIS	Trn-set	0.034	0.0886	0.9951	0.0099	
		Vald-set	0.133	0.0381	0.9981	0.0039	
		Test-set	0.017	0.0383	0.998	0.0038	
	GEP	Trn-set	0.234	0.3213	0.9188	0.0362	
		Vald-set	0.374	0.3812	0.9123	0.0426	
		Test-set	0.261	0.0474	0.9081	0.0533	

Note: Trn: Training, Vald: Validation, Test: Testing

Chapter 5- Conclusion and Recommendations

In this study, three machine-learning algorithms, namely ANN, ANFIS, and GEP, were used to develop a mathematical formulation for accurate predictions of the CS and TS of concrete containing binary SCMs (SF and GGBS). An extensive database considering the eight most influencing input parameters was used for developing the models. The predictive accuracy of the recommended models was evaluated by using several statistical measures, performance indices, and external experimental validation criteria. In addition to that, sensitivity and parametric analysis were performed to determine the coherence of the best-proposed model with published literature. Based on these analyses, the following key findings and recommendations can be drawn.

a) Statistical analysis indicated that in all developed models (ANN, ANFIS, and GEP), the projected values are very near to actual values for both CS and TS models.

b) The performance index of ANN, ANFIS, and GEP created for CS and TS is less than 0.15, indicating that these models are classified as excellent. Furthermore, it can be noted that OF values obtained from ANN, ANFIS, and GEP models for CS are 0.042, 0.031, and 0.062, correspondingly, and for TS models, these values are 0.029, 0.009, and 0.035, respectively. The OF values for all cases are close to zero, demonstrating the validity of the proposed models while controlling the overfitting.

c) Comparative analysis shows that ANFIS models exhibit higher predictive performance compared to ANN and GEP. The mean R^2 value of all three testing conditions (training, testing, and validation) for CS models are 0.988 (ANFIS), 0.944 (ANN), and 0.887 (GEP), whereas 0.998 (ANFIS), 0.954 (ANN), and 0.903 (GEP) for TS models.

d) Based on MAE values, the ANFIS models showed enhanced performance by 29% and 48%, as compared to CS models of ANN and GEP, respectively, whereas ANFIS models for TS showed better predictive performance by 35% and 49% compared to ANN and GEP models. However, the GEP models showed superior performance compared to ANFIS and ANN with respect to the closeness of statistical measure values between the training, validation, and testing sets for both CS and TS models.

e) GEP is an evolutionary technique that also gives a simple empirical formula for forecasting the CS and TS. This approach significantly diminishes the overall duration needed to estimate the CS and TS compared with conventional testing procedures, the

evaluation process can be completed at a significantly faster rate. Therefore, the applications of the presented equations give a viable and quick method for the estimation of CS and TS of concrete containing SF and GGBS.

f) External validation based on experimental investigations showed strong evidence for the applicability of the proposed models with R^2 of 0.88 and error percentages of less than 10%.

g) The sensitivity analysis reveals the significance of input variables in the following increasing trend: W/B (27.9%) > Age (18.1%) > C (12.6%) > CA (11%) > SF (10.1%) > GGBS (8.1%) > FA (8%) > SP (4.2%) for CS models whereas W/B (25.6%) > Age (17.6%) > C (12.3%) > CA (11.5%) > SF (11.5%) > GGBS (9.5%) > FA (7.8%) > SP (4.1%) in case of TS models. These findings are highly comparable with the actual database. The parametric study shows that all the input variables consistently follow the trend mentioned in the experimental database.

h) The developed models successfully fulfilled various criteria that were considered for their external validation.

Finally, it can be deduced that soft computing techniques can provide an efficient basis to promote the utilization of industrial waste in civil engineering applications. The GEP models that have been suggested have the potential to serve as a standard for forecasting the CS and TS of concrete containing GGBS and SF. Additionally, they may be employed during the initial stages of designing concrete mixes. This can play a key role in sustainable development as green concrete reduces energy usage, greenhouse gas emissions, disposal, and building costs by using leftover industrial wastes. The findings of this study emphasize the significance of AI techniques as robust and efficient tools for addressing complex challenges in the field of cement-based composites. However, the applicability of the suggested GEP formulations for CS and TS is constrained solely to the range of the input parameters in their corresponding databases. Under this limitation, it is possible to make further adjustments to the existing forecasting models by incorporating an increased number of data points. In addition, the outcomes of this study should be validated by using the increased database for external validation and also by comparing with other machine learning algorithms such as MLPNN, MEP, MLR, DT, SVM, and ensemble methods.

References:

[1] A. Arrigoni *et al.*, "Life cycle greenhouse gas emissions of concrete containing supplementary cementitious materials: cut-off vs. substitution," *J Clean Prod*, vol. 263, p. 121465, Aug. 2020, doi: 10.1016/j.jclepro.2020.121465.

[2] A. Rauf, A. Tussupbekova, Jong-Kim, and S.-W. Moon, "Effect of Drying-Wetting Cycles on the Mechanical Behavior of Cement-Treated Soil," 대한토목학회 학술대회, pp. 259–260, 2023, Accessed: Mar. 04, 2024. [Online]. Available: https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11627589

[3] "IEA: Cement Technology Roadmap 2009–Carbon emissions... - Google Scholar." Accessed: Dec. 17, 2023. [Online]. Available: https://scholar.google.com/scholar_lookup?title=Cement%20Technology%20Roadmap%202009%3A%20Carbon%20Emissions%20Reductions%20up%20to%202050&publication_year=2009&author=IEA

[4] U. National Minerals Information Center, "Mineral commodity summaries 2020," *Mineral Commodity Summaries*, 2020, doi: 10.3133/MCS2020.

[5] C. O. Rusănescu *et al.*, "Recovery of Sewage Sludge in the Cement Industry," *Energies 2022, Vol. 15, Page 2664*, vol. 15, no. 7, p. 2664, Apr. 2022, doi: 10.3390/EN15072664.

[6] N. Chatziaras, C. S. Psomopoulos, and N. J. Themelis, "Use of waste derived fuels in cement industry: a review," *Management of Environmental Quality: An International Journal*, vol. 27, no. 2, pp. 178–193, Mar. 2016, doi: 10.1108/MEQ-01-2015-0012/FULL/XML.

[7] "Technology Roadmap - Low-Carbon Transition in the Cement Industry – Analysis - IEA." Accessed: Dec. 17, 2023. [Online]. Available: https://www.iea.org/reports/technology-roadmap-low-carbon-transition-in-the-cement-industry

[8] M. Hanifa, R. Agarwal, U. Sharma, P. C. Thapliyal, and L. P. Singh, "A review on CO2 capture and sequestration in the construction industry: Emerging approaches and commercialised technologies," *Journal of CO2 Utilization*, vol. 67, p. 102292, Jan. 2023, doi: 10.1016/J.JCOU.2022.102292.

[9] C. Chen *et al.*, "A striking growth of CO 2 emissions from the global cement industry driven by new facilities in emerging countries A striking growth of CO 2 emissions from the global cement industry driven by new facilities in emerging countries".

[10] A. Sivakrishna, A. Adesina, P. O. Awoyera, and K. R. Kumar, "Green concrete: A review of recent developments," *Mater Today Proc*, vol. 27, pp. 54–58, 2020, doi: 10.1016/j.matpr.2019.08.202.

[11] A. B. M. A. Kaish, T. C. Odimegwu, I. Zakaria, and M. M. Abood, "Effects of different industrial waste materials as partial replacement of fine aggregate on strength and microstructure properties of concrete," *Journal of Building Engineering*, vol. 35, no. November 2020, p. 102092, 2021, doi: 10.1016/j.jobe.2020.102092.

[12] B. Iftikhar *et al.*, "Experimental study on the eco-friendly plastic-sand paver blocks by utilising plastic waste and basalt fibers," 2023, doi: 10.1016/j.heliyon.2023.e17107.

[13] N. Bheel, M. H. W. Ibrahim, A. Adesina, C. Kennedy, and I. A. Shar, "Mechanical performance of concrete incorporating wheat straw ash as partial replacement of cement," *Journal of Building Pathology and Rehabilitation 2020 6:1*, vol. 6, no. 1, pp. 1–7, Nov. 2020, doi: 10.1007/S41024-020-00099-7.

[14] A. Paul and M. Hussain, "Sustainable Use of GGBS and RHA as a Partial Replacement of Cement in the Stabilization of Indian Peat," *International Journal of Geosynthetics and Ground Engineering 2020 6:1*, vol. 6, no. 1, pp. 1–15, Mar. 2020, doi: 10.1007/S40891-020-0185-7.

[15] N. Bheel and A. Adesina, "Influence of Binary Blend of Corn Cob Ash and Glass Powder as Partial Replacement of Cement in Concrete," *Silicon 2020 13:5*, vol. 13, no. 5, pp. 1647–1654, Jun. 2020, doi: 10.1007/S12633-020-00557-4.

[16] R. Garg and R. Garg, "Effect of zinc oxide nanoparticles on mechanical properties of silica fume-based cement composites," *Mater Today Proc*, vol. 43, pp. 778–783, Jan. 2021, doi:

10.1016/J.MATPR.2020.06.168.
[17] W. Ahmad, A. Ahmad, K. A. Ostrowski, F. Aslam, P. Joyklad, and P. Zajdel, “Sustainable approach of using sugarcane bagasse ash in cement-based composites: A systematic review,” *Case Studies in Construction Materials*, vol. 15, p. e00698, Dec. 2021, doi: 10.1016/J.CSCM.2021.E00698.
[18] J. M. Paris, J. G. Roessler, C. C. Ferraro, H. D. Deford, and T. G. Townsend, “A review of waste products utilized as supplements to Portland cement in concrete,” *J Clean Prod*, vol. 121, pp. 1–18, May 2016, doi: 10.1016/J.JCLEPRO.2016.02.013.
[19] N. M. Piatak, M. B. Parsons, and R. R. Seal, “Characteristics and environmental aspects of slag: A review,” *Applied Geochemistry*, vol. 57, pp. 236–266, Jun. 2015, doi: 10.1016/J.APGEOCHEM.2014.04.009.
[20] S. Gupta and S. Chaudhary, “State of the art review on supplementary cementitious materials in India – II: Characteristics of SCMs, effect on concrete and environmental impact,” *J Clean Prod*, vol. 357, p. 131945, Jul. 2022, doi: 10.1016/J.JCLEPRO.2022.131945.
[21] M. N. Akhtar, M. Jameel, Z. Ibrahim, and N. M. Bunnori, “Incorporation of recycled aggregates and silica fume in concrete: an environmental savior-a systematic review,” *Journal of Materials Research and Technology*, vol. 20, pp. 4525–4544, Sep. 2022, doi: 10.1016/J.JMRT.2022.09.021.
[22] E. Özbay, M. Erdemir, and H. I. Durmuş, “Utilization and efficiency of ground granulated blast furnace slag on concrete properties – A review,” *Constr Build Mater*, vol. 105, pp. 423–434, Feb. 2016, doi: 10.1016/J.CONBUILDMAT.2015.12.153.
[23] 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, Jan. 2017, doi: 10.1016/J.CONBUILDMAT.2016.10.114.
[24] 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, Jan. 2021, doi: 10.1016/J.JCLEPRO.2020.123697.
[25] I. B. Topçu and M. Saridemir, “Prediction of mechanical properties of recycled aggregate concretes containing silica fume using artificial neural networks and fuzzy logic,” *Comput Mater Sci*, vol. 42, no. 1, pp. 74–82, Mar. 2008, doi: 10.1016/J.COMMATSCI.2007.06.011.
[26] D. Van Dao, H. B. Ly, S. H. Trinh, T. T. Le, and B. T. Pham, “Artificial Intelligence Approaches for Prediction of Compressive Strength of Geopolymer Concrete,” *Materials 2019, Vol. 12, Page 983*, vol. 12, no. 6, p. 983, Mar. 2019, doi: 10.3390/MA12060983.
[27] Ł. Sadowski, M. Piechówka-Mielnik, T. Widziszowski, A. Gardynik, and S. Mackiewicz, “Hybrid ultrasonic-neural prediction of the compressive strength of environmentally friendly concrete screeds with high volume of waste quartz mineral dust,” *J Clean Prod*, vol. 212, pp. 727–740, 2019, doi: 10.1016/j.jclepro.2018.12.059.
[28] H. Sebaaly, S. Varma, and J. W. Maina, “Optimizing asphalt mix design process using artificial neural network and genetic algorithm,” *Constr Build Mater*, vol. 168, pp. 660–670, Apr. 2018, doi: 10.1016/J.CONBUILDMAT.2018.02.118.
[29] U. Asif, S. A. Memon, M. F. Javed, J. Kim, and A. Luísa Velosa, “Predictive Modeling and Experimental Validation for Assessing the Mechanical Properties of Cementitious Composites Made with Silica Fume and Ground Granulated Blast Furnace Slag,” *Buildings 2024, Vol. 14, Page 1091*, vol. 14, no. 4, p. 1091, Apr. 2024, doi: 10.3390/BUILDINGS14041091.
[30] A. H. Gandomi and A. H. Alavi, “Multi-stage genetic programming: A new strategy to nonlinear system modeling,” *Inf Sci (N Y)*, vol. 181, no. 23, pp. 5227–5239, Dec. 2011, doi: 10.1016/J.INS.2011.07.026.
[31] A. Nafees *et al.*, “Predictive modeling of mechanical properties of silica fume-based green concrete using artificial intelligence approaches: MLPNN, ANFIS, and GEP,” *Materials*, vol. 14, no. 24, pp. 1–28, 2021, doi: 10.3390/ma14247531.
[32] 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, Jul. 2020, doi: 10.1016/J.JMRT.2020.06.008.

[33] A. A. Shahmansouri, H. Akbarzadeh Bengar, and S. Ghanbari, “Compressive strength prediction of eco-efficient GGBS-based geopolymer concrete using GEP method,” *Journal of Building Engineering*, vol. 31, p. 101326, Sep. 2020, doi: 10.1016/J.JOBE.2020.101326.
[34] R. Siddique and M. I. Khan, “Supplementary Cementing Materials,” vol. 37, 2011, doi: 10.1007/978-3-642-17866-5.
[35] R. Siddique and M. Iqbal Khan, “Silica Fume,” *Engineering Materials*, vol. 37, pp. 67–119, 2011, doi: 10.1007/978-3-642-17866-5_2/COVER.
[36] M. Türköz, S. U. Umu, and O. Öztürk, “Effect of Silica Fume as a Waste Material for Sustainable Environment on the Stabilization and Dynamic Behavior of Dispersive Soil,” *Sustainability 2021, Vol. 13, Page 4321*, vol. 13, no. 8, p. 4321, Apr. 2021, doi: 10.3390/SU13084321.
[37] A. Mehta and D. Kumar Ashish, “Silica fume and waste glass in cement concrete production: A review,” 2019, doi: 10.1016/j.jobe.2019.100888.
[38] S. Tak, P. Gupta, A. Kumar, A. Sofi, and C. Mei Yun, “Effect of using silica fume as a partial replacement of cement in concrete,” *Mater Today Proc*, Apr. 2023, doi: 10.1016/J.MATPR.2023.04.205.
[39] P. Smarzewski, “ScienceDirect Peer-review under responsibility of the Scientific Committee of Influence of silica fume on mechanical and fracture properties of high performance concrete,” 2019, doi: 10.1016/j.prostr.2019.08.002.
[40] S. Ahmad, O. S. B. Al-Amoudi, M. S. Khan, and M. Maslehuddin, “Effect of silica fume inclusion on the strength, shrinkage and durability characteristics of natural pozzolan-based cement concrete,” 2022, doi: 10.1016/j.cscm.2022.e01255.
[41] F. V. Ajileye, “Global Journal of researches in engineering Civil And Structural engineering Investigations on Microsilica (Silica Fume) As Partial Cement Replacement in Concrete Investigations on Microsilica Silica FumeAs Partial Cement Replacement in Concrete GJRE-E Classification : FOR Code: 090503 Investigations on Microsilica (Silica Fume) As Partial Cement Replacement in Concrete,” 2012.
[42] M. Mazloom, A. A. Ramezanianpour, and J. J. Brooks, “Effect of silica fume on mechanical properties of high-strength concrete,” *Cem Concr Compos*, vol. 26, no. 4, pp. 347–357, May 2004, doi: 10.1016/S0958-9465(03)00017-9.
[43] D. Pradhan and D. Dutta, “Influence of Silica Fume on Normal Concrete,” *Journal of Engineering Research and Applications www.ijera.com*, vol. 3, no. 5, pp. 79–82, 2013.
[44] H. M. Hamada *et al.*, “Effect of silica fume on the properties of sustainable cement concrete,” *Journal of Materials Research and Technology*, vol. 24, pp. 8887–8908, May 2023, doi: 10.1016/J.JMRT.2023.05.147.
[45] P. Ganesh and A. Ramachandra Murthy, “Tensile behaviour and durability aspects of sustainable ultra-high performance concrete incorporated with GGBS as cementitious material,” 2018, doi: 10.1016/j.conbuildmat.2018.11.240.
[46] V. Rama and R. Garikipati, “Strength and Durability Studies on GGBS Concrete,” *Article in International Journal of Civil Engineering*, 2015, doi: 10.14445/23488352/IJCE-V2I10P106.
[47] S. E. Chidiac and D. K. Panesar, “Evolution of mechanical properties of concrete containing ground granulated blast furnace slag and effects on the scaling resistance test at 28 days,” *Cem Concr Compos*, vol. 30, no. 2, pp. 63–71, Feb. 2008, doi: 10.1016/J.CEMCONCOMP.2007.09.003.
[48] A. Professor, “Studies on Optimum Usage of GGBS in Concrete,” *Int J Innov Sci Res Technol*, vol. 2, no. 5, 2017, Accessed: Nov. 08, 2023. [Online]. Available: www.ijisrt.com
[49] S. Arivalagan, “Sustainable studies on concrete with GGBS as a replacement material in cement,” *Jordan Journal of Civil Engineering*, vol. 8, no. 3, pp. 263–270, 2014.
[50] S. Arivalagan, “Sustainable studies on concrete with GGBS as a replacement material in cement,” *Jordan Journal of Civil Engineering*, vol. 8, no. 3, pp. 263–270, 2014.
[51] J. Ahmad *et al.*, “A Comprehensive Review on the Ground Granulated Blast Furnace Slag (GGBS) in Concrete Production,” *Sustainability 2022, Vol. 14, Page 8783*, vol. 14, no. 14, p. 8783, Jul. 2022, doi: 10.3390/SU14148783.
[52] Y. Li, C. Qiao, and W. Ni, “Green concrete with ground granulated blast-furnace slag activated by desulfurization gypsum and electric arc furnace reducing slag,” *J Clean Prod*, vol.

269, p. 122212, Oct. 2020, doi: 10.1016/J.JCLEPRO.2020.122212.
[53] M. I. Shah, M. F. Javed, and T. Abunama, “Proposed formulation of surface water quality and modelling using gene expression, machine learning, and regression techniques,” *Environmental Science and Pollution Research*, vol. 28, no. 11, pp. 13202–13220, Mar. 2021, doi: 10.1007/S11356-020-11490-9/FIGURES/19.
[54] J. Sun, J. Zhang, Y. Gu, Y. Huang, Y. Sun, and G. Ma, “Prediction of permeability and unconfined compressive strength of pervious concrete using evolved support vector regression,” *Constr Build Mater*, vol. 207, pp. 440–449, May 2019, doi: 10.1016/J.CONBUILDMAT.2019.02.117.
[55] F. Farooq *et al.*, “A Comparative Study for the Prediction of the Compressive Strength of Self-Compacting Concrete Modified with Fly Ash,” *Materials 2021, Vol. 14, Page 4934*, vol. 14, no. 17, p. 4934, Aug. 2021, doi: 10.3390/MA14174934.
[56] 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, Jun. 2018, doi: 10.1016/J.CONBUILDMAT.2018.04.169.
[57] U. Asif, M. F. Javed, M. Alyami, and A. W. Hammad, “Performance Evaluation of Concrete Made with Plastic Waste Using Multi-Expression Programming,” *Mater Today Commun*, p. 108789, Apr. 2024, doi: 10.1016/J.MTCOMM.2024.108789.
[58] K. Kim, W. Kim, J. Seo, Y. Jeong, M. Lee, and J. Lee, “Prediction of Concrete Fragments Amount and Travel Distance under Impact Loading Using Deep Neural Network and Gradient Boosting Method,” *Materials*, vol. 15, no. 3, Feb. 2022, doi: 10.3390/MA15031045.
[59] K. Khan, W. Ahmad, M. N. Amin, and A. Ahmad, “A Systematic Review of the Research Development on the Application of Machine Learning for Concrete,” *Materials 2022, Vol. 15, Page 4512*, vol. 15, no. 13, p. 4512, Jun. 2022, doi: 10.3390/MA15134512.
[60] G. Khambra and P. Shukla, “Novel machine learning applications on fly ash based concrete: An overview,” *Mater Today Proc*, Jul. 2021, doi: 10.1016/J.MATPR.2021.07.262.
[61] 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.
[62] H. Song *et al.*, “Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms,” *Constr Build Mater*, vol. 308, p. 125021, Nov. 2021, doi: 10.1016/J.CONBUILDMAT.2021.125021.
[63] M. I. Shah, M. F. Javed, F. Aslam, and H. Alabduljabbar, “Machine learning modeling integrating experimental analysis for predicting the properties of sugarcane bagasse ash concrete,” *Constr Build Mater*, vol. 314, p. 125634, Jan. 2022, doi: 10.1016/J.CONBUILDMAT.2021.125634.
[64] Y. Liu, Y. Li, J. Mu, H. Li, and J. Shen, “Modeling and analysis of creep in concrete containing supplementary cementitious materials based on machine learning,” *Constr Build Mater*, vol. 392, p. 131911, Aug. 2023, doi: 10.1016/J.CONBUILDMAT.2023.131911.
[65] D. Van Dao, H. B. Ly, H. L. T. Vu, T. T. Le, and B. T. Pham, “Investigation and Optimization of the C-ANN Structure in Predicting the Compressive Strength of Foamed Concrete,” *Materials 2020, Vol. 13, Page 1072*, vol. 13, no. 5, p. 1072, Feb. 2020, doi: 10.3390/MA13051072.
[66] A. Beyciołlu, M. Emirołlu, Y. Kocak, and S. Subaşi, “Analyzing the compressive strength of clinker mortars using approximate reasoning approaches - ANN vs MLR,” *Computers and Concrete*, vol. 15, no. 1, pp. 89–101, Jan. 2015, doi: 10.12989/CAC.2015.15.1.089.
[67] 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.
[68] R. R. Mahesh and D. Sathyan, “Modelling the hardened properties of steel fiber reinforced concrete using ANN,” *Mater Today Proc*, vol. 49, pp. 2081–2089, Jan. 2022, doi: 10.1016/J.MATPR.2021.08.311.
[69] Z. H. Duan, S. C. Kou, and C. S. Poon, “Using artificial neural networks for predicting the elastic modulus of recycled aggregate concrete,” *Constr Build Mater*, vol. 44, pp. 524–532, Jul. 2013, doi: 10.1016/J.CONBUILDMAT.2013.02.064.

[70] F. Rehman, S. A. Khokhar, and R. A. Khushnood, “ANN based predictive mimicker for mechanical and rheological properties of eco-friendly geopolymer concrete,” *Case Studies in Construction Materials*, vol. 17, p. e01536, Dec. 2022, doi: 10.1016/J.CSCM.2022.E01536.
[71] D. Van Dao, H. B. Ly, S. H. Trinh, T. T. Le, and B. T. Pham, “Artificial Intelligence Approaches for Prediction of Compressive Strength of Geopolymer Concrete,” *Materials 2019, Vol. 12, Page 983*, vol. 12, no. 6, p. 983, Mar. 2019, doi: 10.3390/MA12060983.
[72] U. Asif, M. F. Javed, M. Abuhussain, M. Ali, W. A. Khan, and A. Mohamed, “Predicting The Mechanical Properties of Plastic Concrete: An Optimization Method by Using Genetic Programming and Ensemble Learners,” *Case Studies in Construction Materials*, p. e03135, Apr. 2024, doi: 10.1016/J.CSCM.2024.E03135.
[73] S. A. Alabi, C. Arum, A. P. Adewuyi, R. C. Arum, J. O. Afolayan, and J. Mahachi, “Mathematical model for prediction of compressive strength of ternary blended cement concrete utilizing gene expression programming,” *Sci Afr*, vol. 22, p. 1954, 2023, doi: 10.1016/j.sciaf.2023.e01954.
[74] Y. Murad, “Compressive strength prediction for concrete modified with nanomaterials,” 2021, doi: 10.1016/j.cscm.2021.e00660.
[75] H. Alabduljabbar, M. Khan, H. Awan, S. M. Eldin, R. Alyousef, and M. Mohamed, “Predicting ultra-high-performance concrete compressive strength using gene expression programming method,” 2023, doi: 10.1016/j.cscm.2023.e02074.
[76] A. Ashrafian, A. H. Gandomi, M. Rezaie-Balf, and M. Emadi, “An evolutionary approach to formulate the compressive strength of roller compacted concrete pavement,” 2019, doi: 10.1016/j.measurement.2019.107309.
[77] S. Park, S. Wu, Z. Liu, and S. Pyo, “The role of supplementary cementitious materials (Scms) in ultra high performance concrete (uhpc): A review,” *Materials*, vol. 14, no. 6, pp. 1–24, 2021, doi: 10.3390/ma14061472.
[78] Y. Balakrishna, V. Lavanya, A. Naresh, and S. B. Reddy, “Triple Blending of Cement Concrete with Micro Silica and Ground Granulated Blast Furnace Slag,” pp. 1500–1508, 2016.
[79] M. R. Akram and S. Raza, “Effect of Micro Silica and GGBS on Compressive Strength and Permeability of Impervious Concrete as a Cement Replacement,” *Eur Acad Res*, vol. III, no. 7, pp. 7456–7468, 2015.
[80] G. Rajagopalan and C. Komarasamy, “Influence of silica fume on strength characteristics of high strength concrete Influence of Silica Fume on Strength Characteristics of High Strength Concrete,” no. December, 2018.
[81] K. A. Reddy, V. R. Kode, P. Malasani, S. Satish, and K. Darapu, “Influence of Addition of Micro Silica on Strength Properties of Basalt Fiber Reinforced Multi Blended Concrete,” no. October, 2019, doi: 10.35940/ijeat.A1627.109119.
[82] S. S. Vivek and G. Dhinakaran, “Engineering Science and Technology , an International Journal Fresh and hardened properties of binary blend high strength self compacting concrete,” *Engineering Science and Technology, an International Journal*, vol. 20, no. 3, pp. 1173–1179, 2017, doi: 10.1016/j.jestch.2017.05.003.
[83] R. Rajeshwari and S. Mandal, “Prediction of compressive strength of high-volume fly ash concrete using artificial neural network,” *Lecture Notes in Civil Engineering*, vol. 25, pp. 471–483, 2019, doi: 10.1007/978-981-13-3317-0_42.
[84] V. B. R. Suda and S. Priyatham Paul, “Relationship between compressive, split tensile and flexural strengths of ternary blended concrete,” *Mater Today Proc*, vol. 65, pp. 1112–1119, 2022, doi: 10.1016/j.matpr.2022.04.162.
[85] S. V. B. Reddy and P. R. Mounika, “Strength and Durability Studies of Ternary Concrete,” *... Journal of Applied Engineering …*, vol. 13, no. June, pp. 12161–12177, 2018, [Online]. Available: https://www.ripublication.com/ijaer18/ijaerv13n15_65.pdf
[86] “2 . Ground Granulated Blast Furnace Slag (Ggbs): 5 . Mix Design Based on Is :”.
[87] A. Babu, “Study on the Strength and Durability Properties of Ternary Blended Concrete,” no. September, 2016.
[88] A. Marani, A. Jamali, and M. L. Nehdi, “Predicting Ultra-High-Performance Concrete Compressive Strength Using Tabular Generative Adversarial Networks,” 2020.
[89] W. Haroon, N. Ahmad, and R. R. Akram, “DEVELOPING OF IMPERVIOUS CONCRETE

USING SILICA FUME AND GGBS AS CEMENT REPLACEMENT MATERIALS Department of Civil Engineering," 2012.
[90] W. Zhang and H. Ba, "Effect of ground granulated blast-furnace slag (GGBFS) and silica fume (SF) on chloride migration through concrete subjected to repeated loading," *Sci China Technol Sci*, vol. 55, no. 11, pp. 3102–3108, 2012, doi: 10.1007/s11431-012-5027-y.
[91] "21-Geena George.pdf."
[92] S. Prakash, S. Kumar, R. Biswas, and B. Rai, "Influence of silica fume and ground granulated blast furnace slag on the engineering properties of ultra-high-performance concrete," *Innovative Infrastructure Solutions*, vol. 7, no. 1, pp. 1–18, 2022, doi: 10.1007/s41062-021-00714-7.
[93] V. B. R. Suda and P. S. Rao, "Experimental investigation on optimum usage of Micro silica and GGBS for the strength characteristics of concrete," in *Materials Today: Proceedings*, Elsevier Ltd., 2020, pp. 805–811. doi: 10.1016/j.matpr.2019.12.354.
[94] S. V. B. Reddy and P. S. Rao, "ScienceDirect Experimental studies on compressive strength of ternary blended concretes at different levels of micro silica and ggbs," *Mater Today Proc*, vol. 3, no. 10, pp. 3752–3760, 2016, doi: 10.1016/j.matpr.2016.11.024.
[95] V. Bhaskar, R. Suda, and P. S. Rao, "Influence of Mineral admixtures on Compressive strength of Ternary concrete with different Water binder ratios Influence of Mineral admixtures on Compressive strength of Ternary concrete with different Water binder ratios," no. July, 2020, doi: 10.9790/1684-16053014856.
[96] P. Vasanthi, A. Sujaatha, A. Bhaskar, and L. Gaur, "No Title," vol. 15, pp. 85–88, 2018.
[97] A. Prof, "INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH AND ENGINEERING TRENDS STRENGTH AND DURABILITY OF CONCRETE BY USING INTERNATIONAL JOURNAL OF ADVANCE SCIENTIFIC RESEARCH," vol. 5, no. 8, pp. 182–191, 2020.
[98] S. Chandini, P. Mohammed, and S. Nusari, "Green Concrete - A Low Cost and Sustainable Solution for a Better Environment," vol. 19, no. 2, pp. 261–269, 2020.
[99] I. Publication, "A Ternary Blended Concrete Mix with Partial Replacement of OPC with GGBS & Micro Silica and Their Effects on Str ... A Ternary Blended Concrete Mix with Partial Replacement of OPC with GGBS & Micro Silica and Their Effects on Strength - An Experimental St".
[100] F. Ghasemzadeh, M. K. Moradllo, and S. Sadati, "Effect of Silica Fume and Ground Granulated Blast-Furnace Slag on Shrinkage in High Performance Concrete Effect of silica fume and ground granulated blast-furnace slag on shrinkage in the high performance concrete," no. June, 2010.
[101] S. S. Deshmukh, G. R. Gandhe, and M. Studies, "ISSN NO : 2347-6648 Volume IX , Issue II , February / 2020 Page No : 247 ISSN NO : 2347-6648 Page No : 248," vol. IX, no. 2347, pp. 247–250, 2020.
[102] A. Mardani-Aghabaglou, S. H. Bayqra, and A. Nobakhtjoo, "Specimen size and shape effects on strength of concrete in the absence and presence of steel fibers," *Revista de la Construccion*, vol. 20, no. 1, pp. 128–144, 2021, doi: 10.7764/RDLC.20.1.128.
[103] D. J. Elwell and G. Fu, "Compression testing of concrete: cylinders vs. cube," *Special Report 119*. p. 21, 1995.
[104] A. H. Gandomi and D. A. Roke, "Assessment of artificial neural network and genetic programming as predictive tools," *Advances in Engineering Software*, vol. 88, pp. 63–72, Oct. 2015, doi: 10.1016/J.ADVENGSOFT.2015.05.007.
[105] A. C. Atkinson, M. Riani, and A. Cerioli, "The forward search: Theory and data analysis," *J Korean Stat Soc*, vol. 39, no. 2, pp. 117–134, Jun. 2010, doi: 10.1016/J.JKSS.2010.02.007.
[106] S. I. Khan and A. S. M. L. Hoque, "SICE: an improved missing data imputation technique," *J Big Data*, vol. 7, no. 1, pp. 1–21, Dec. 2020, doi: 10.1186/S40537-020-00313-W/FIGURES/9.
[107] C. Sharma and C. S. P. Ojha, "Statistical Parameters of Hydrometeorological Variables: Standard Deviation, SNR, Skewness and Kurtosis," *Lecture Notes in Civil Engineering*, vol. 39, pp. 59–70, 2020, doi: 10.1007/978-981-13-8181-2_5.
[108] S. C. Brown and J. A. Greene, "The wisdom development scale: Translating the conceptual to the concrete," *J Coll Stud Dev*, vol. 47, no. 1, pp. 1–19, 2006, doi: 10.1353/CSD.2006.0002.

[109] “GN Smith - Probability & Statistics in Civil Engineering PDF | PDF | Matrix (Mathematics) | Normal Distribution.” Accessed: May 31, 2023. [Online]. Available: https://www.scribd.com/document/338933868/GN-Smith-Probability-Statistics-in-Civil-Engineering-pdf#
[110] W. Pitts and W. S. McCulloch, “How we know universals the perception of auditory and visual forms,” *Bull Math Biophys*, vol. 9, no. 3, pp. 127–147, Sep. 1947, doi: 10.1007/BF02478291/METRICS.
[111] N. N. Kourgialas, Z. Dokou, and G. P. Karatzas, “Statistical analysis and ANN modeling for predicting hydrological extremes under climate change scenarios: The example of a small Mediterranean agro-watershed,” *J Environ Manage*, vol. 154, pp. 86–101, May 2015, doi: 10.1016/J.JENVMAN.2015.02.034.
[112] Y. Koçak and G. Üstündağ Şiray, “New activation functions for single layer feedforward neural network,” *Expert Syst Appl*, vol. 164, p. 113977, Feb. 2021, doi: 10.1016/J.ESWA.2020.113977.
[113] “Searching for activation functions,” *6th International Conference on Learning Representations, ICLR 2018 - Workshop Track Proceedings*, 2018.
[114] K. V. Naresh Babu and D. R. Edla, “New Algebraic Activation Function for Multi-Layered Feed Forward Neural Networks,” *http://dx.doi.org/10.1080/03772063.2016.1240633*, vol. 63, no. 1, pp. 71–79, Jan. 2016, doi: 10.1080/03772063.2016.1240633.
[115] S. K. Das, “Artificial Neural Networks in Geotechnical Engineering: Modeling and Application Issues,” *Metaheuristics in Water, Geotechnical and Transport Engineering*, pp. 231–270, 2013, doi: 10.1016/B978-0-12-398296-4.00010-6.
[116] M. Tahani, M. Vakili, and S. Khosrojerdi, “Experimental evaluation and ANN modeling of thermal conductivity of graphene oxide nanoplatelets/deionized water nanofluid,” *International Communications in Heat and Mass Transfer*, vol. 76, pp. 358–365, Aug. 2016, doi: 10.1016/J.ICHEATMASSTRANSFER.2016.06.003.
[117] S. Khan *et al.*, “Predicting the Ultimate Axial Capacity of Uniaxially Loaded CFST Columns Using Multiphysics Artificial Intelligence,” *Materials 2022, Vol. 15, Page 39*, vol. 15, no. 1, p. 39, Dec. 2021, doi: 10.3390/MA15010039.
[118] E. M. Golafshani, A. Behnood, and M. Arashpour, “Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer,” *Constr Build Mater*, vol. 232, p. 117266, Jan. 2020, doi: 10.1016/J.CONBUILDMAT.2019.117266.
[119] H. Moayedi and A. Mosavi, “A water cycle-based error minimization technique in predicting the bearing capacity of shallow foundation,” *Eng Comput*, vol. 38, pp. 3993–4006, Dec. 2022, doi: 10.1007/s00366-021-01289-8.
[120] M. R. Islam, W. Z. W. Jaafar, L. S. Hin, N. Osman, A. Hossain, and N. S. Mohd, “Development of an intelligent system based on ANFIS model for predicting soil erosion,” *Environ Earth Sci*, vol. 77, no. 5, Mar. 2018, doi: 10.1007/S12665-018-7348-Z.
[121] “Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence - Candida Ferreira - Google Books.” Accessed: Oct. 09, 2022. [Online]. Available: https://books.google.kz/books?hl=en&lr=&id=NkG7BQAAQBAJ&oi=fnd&pg=PR7&ots=Y_orvEYgF2&sig=nbBd7c2pns60WPMczu9EAjAjPQA&redir_esc=y#v=onepage&q&f=false
[122] M. Saridemir, “Genetic programming approach for prediction of compressive strength of concretes containing rice husk ash,” *Constr Build Mater*, vol. 24, no. 10, pp. 1911–1919, Oct. 2010, doi: 10.1016/J.CONBUILDMAT.2010.04.011.
[123] “(2) (PDF) A GENE EXPRESSION PROGRAMMING SYSTEM FOR TIME SERIES MODELING.” Accessed: Oct. 09, 2022. [Online]. Available: https://www.researchgate.net/publication/253404813_A_GENE_EXPRESSION_PROGRAMMING_SYSTEM_FOR_TIME_SERIES_MODELING
[124] 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, Jan. 2021, doi: 10.1016/J.JCLEPRO.2020.123697.
[125] F. E. Jalal, Y. Xu, M. Iqbal, M. F. Javed, and B. Jamhiri, “Predictive modeling of swell-

strength of expansive soils using artificial intelligence approaches: ANN, ANFIS and GEP," *J Environ Manage*, vol. 289, Jul. 2021, doi: 10.1016/J.JENVMAN.2021.112420.
[126] K. Venkatesh and Y. K. Bind, "ANN and Neuro-Fuzzy Modeling for Shear Strength Characterization of Soils," *Proceedings of the National Academy of Sciences India Section A - Physical Sciences*, vol. 92, no. 2, pp. 243–249, Jun. 2022, doi: 10.1007/S40010-020-00709-6.
[127] 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, Jan. 2017, doi: 10.1016/J.CONBUILDMAT.2016.10.114.
[128] M. F. Iqbal *et al.*, "Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming," *J Hazard Mater*, vol. 384, p. 121322, Feb. 2020, doi: 10.1016/J.JHAZMAT.2019.121322.
[129] W. Zhang *et al.*, "State-of-the-art review of soft computing applications in underground excavations," *Geoscience Frontiers*, vol. 11, no. 4, pp. 1095–1106, Jul. 2020, doi: 10.1016/J.GSF.2019.12.003.
[130] I. O. Alade, M. A. Abd Rahman, and T. A. Saleh, "Modeling and prediction of the specific heat capacity of Al2 O3/water nanofluids using hybrid genetic algorithm/support vector regression model," *Nano-Structures & Nano-Objects*, vol. 17, pp. 103–111, Feb. 2019, doi: 10.1016/J.NANOSO.2018.12.001.
[131] M. A. Shahin, "Use of evolutionary computing for modelling some complex problems in geotechnical engineering," *http://dx.doi.org/10.1080/17486025.2014.921333*, vol. 10, no. 2, pp. 109–125, Apr. 2014, doi: 10.1080/17486025.2014.921333.
[132] G. Kareken *et al.*, "Geopolymer as a key material to utilize basic oxygen furnace slag (BOFS) as an aggregate," *Mater Today Proc*, Oct. 2023, doi: 10.1016/J.MATPR.2023.10.093.
[133] U. Bakhbergen, C. S. Shon, D. Zhang, J. Ryeol Kim, and J. Liu, "Optimization of mixture parameter for physical and mechanical properties of reactive powder concrete under external sulfate attack using Taguchi method," *Constr Build Mater*, vol. 352, p. 129023, Oct. 2022, doi: 10.1016/J.CONBUILDMAT.2022.129023.
[134] H. Güllü and H. İ. Fedakar, "On the prediction of unconfined compressive strength of silty soil stabilized with bottom ash, jute and steel fibers via artificial intelligence," *Geomechanics and Engineering*, vol. 12, no. 3, pp. 441–464, Mar. 2017, doi: 10.12989/GAE.2017.12.3.441.
[135] S. Prakash, S. Kumar, R. Biswas, and B. Rai, "Influence of silica fume and ground granulated blast furnace slag on the engineering properties of ultra-high-performance concrete," *Innovative Infrastructure Solutions*, vol. 7, no. 1, pp. 1–18, Feb. 2022, doi: 10.1007/S41062-021-00714-7/METRICS.
[136] I. Y. Hakeem, F. Althoey, and A. Hosen, "Mechanical and durability performance of ultra-high-performance concrete incorporating SCMs," *Constr Build Mater*, vol. 359, p. 129430, Dec. 2022, doi: 10.1016/J.CONBUILDMAT.2022.129430.
[137] ACI Committee 363, "ACI PRC-363-10 Report on High-Strength Concrete," *American Concrete Institute*, 2010.
[138] E. M. Golafshani, A. Behnood, and M. Arashpour, "Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer," 2019, doi: 10.1016/j.conbuildmat.2019.117266.
[139] F. E. Jalal, Y. Xu, M. Iqbal, M. F. Javed, and B. Jamhiri, "Predictive modeling of swell-strength of expansive soils using artificial intelligence approaches: ANN, ANFIS and GEP," *J Environ Manage*, vol. 289, p. 112420, Jul. 2021, doi: 10.1016/J.JENVMAN.2021.112420.
[140] F. E. Jalal, Y. Xu, M. Iqbal, M. F. Javed, and B. Jamhiri, "Predictive modeling of swell-strength of expansive soils using artificial intelligence approaches: ANN, ANFIS and GEP," *J Environ Manage*, vol. 289, Jul. 2021, doi: 10.1016/J.JENVMAN.2021.112420.
[141] W. Zhang *et al.*, "State-of-the-art review of soft computing applications in underground excavations," *Geoscience Frontiers*, vol. 11, no. 4, pp. 1095–1106, Jul. 2020, doi: 10.1016/J.GSF.2019.12.003.
[142] A. Golbraikh and A. Tropsha, "Beware of q2!," *J Mol Graph Model*, vol. 20, no. 4, pp. 269–276, Jan. 2002, doi: 10.1016/S1093-3263(01)00123-1.
[143] A. H. Alavi, M. Ameri, A. H. Gandomi, and M. R. Mirzahosseini, "Formulation of flow number of asphalt mixes using a hybrid computational method," *Constr Build Mater*, vol. 25,

no. 3, pp. 1338–1355, Mar. 2011, doi: 10.1016/J.CONBUILDMAT.2010.09.010.
[144] M. Cyr, “Influence of supplementary cementitious materials (SCMs) on concrete durability,” *Eco-Efficient Concrete*, pp. 153–197, Jan. 2013, doi: 10.1533/9780857098993.2.153.
[145] S. Gupta and S. Chaudhary, “State of the art review on supplementary cementitious materials in India – II: Characteristics of SCMs, effect on concrete and environmental impact,” *J Clean Prod*, vol. 357, p. 131945, Jul. 2022, doi: 10.1016/J.JCLEPRO.2022.131945.
[146] S. H. Chu, “Effect of paste volume on fresh and hardened properties of concrete,” *Constr Build Mater*, vol. 218, no. 2, pp. 284–294, 2019, doi: 10.1016/j.conbuildmat.2019.05.131.
[147] M. S. Meddah, S. Zitouni, and S. Belâabes, “Effect of content and particle size distribution of coarse aggregate on the compressive strength of concrete,” *Constr Build Mater*, vol. 24, no. 4, pp. 505–512, 2010, doi: 10.1016/j.conbuildmat.2009.10.009.
[148] S. Samad and A. Shah, “Role of binary cement including Supplementary Cementitious Material (SCM), in production of environmentally sustainable concrete: A critical review,” *International Journal of Sustainable Built Environment*, vol. 6, no. 2, pp. 663–674, Dec. 2017, doi: 10.1016/J.IJSBE.2017.07.003.
[149] S. Nazar *et al.*, “Machine learning interpretable-prediction models to evaluate the slump and strength of fly ash-based geopolymer,” *Journal of Materials Research and Technology*, vol. 24, pp. 100–124, May 2023, doi: 10.1016/J.JMRT.2023.02.180.
[150] M. A. Khan, F. Aslam, M. F. Javed, H. Alabduljabbar, and A. F. Deifalla, “New prediction models for the compressive strength and dry-thermal conductivity of bio-composites using novel machine learning algorithms,” *J Clean Prod*, vol. 350, no. August 2021, p. 131364, 2022, doi: 10.1016/j.jclepro.2022.131364.
[151] M. A. Khan *et al.*, “Geopolymer Concrete Compressive Strength via Artificial Neural Network, Adaptive Neuro Fuzzy Interface System, and Gene Expression Programming With K-Fold Cross Validation,” *Front Mater*, vol. 8, no. May, pp. 1–19, 2021, doi: 10.3389/fmats.2021.621163.