Revista ALCONPAT
www.revistaalconpat.org
eISSN 2007-6835
Revista de la Asociación Latinoamericana de Control de Calidad, Patología y Recuperación de la Construcción

AI-driven forecasting and laboratory evaluation of compressive strength in recycled aggregate concrete.

S. Gopal[1]*, A. Kumar[2]
*Contact author: swatantragopal.const@gmail.com
DOI: https://doi.org/10.21041/ra.v16i2.967

Received: 04/12/2025 | Received in revised form: 19/03/2026 | Accepted: 20/03/2026 | Published: 01/05/2026

ABSTRACT

This study aims to evaluate and predict the compressive strength of recycled aggregate concrete (RAC) using experimental testing and machine learning techniques. Twenty-five concrete mixes with varying recycled aggregate content, water cement ratio, plasticizer dosage, and parent concrete strength were investigated. The 28-day compressive strength ranged from 31.8 to 45.2 MPa, decreasing with higher recycled aggregate content and water absorption. Support Vector Regression achieved the highest prediction accuracy (R^2 = 0.998), outperforming other models. The study is limited by dataset size and controlled data expansion. The originality lies in integrating experimental investigation with multi-model machine learning analysis. The results demonstrate that AI can effectively support sustainable mix design of recycled aggregate concrete.
Keywords: recycled aggregate concrete; compressive strength; machine learning; support vector regression; sustainable construction; data-driven modeling.

Cite as: Gopal, S., Kumar, A. (2026), "*AI-driven forecasting and laboratory evaluation of compressive strength in recycled aggregate concrete*", Revista ALCONPAT, 16 (2), pp. 149 – 185, DOI: https://doi.org/10.21041/ra.v16i2.967.

[1] Research Scholar, Department of Civil Engineering, Sandip University, Sijoul, Madhubani-847235, Bihar, India
[2] Assistant Professor, Department of Civil Engineering, Sandip University, Sijoul, Madhubani-847235, Bihar, India.

Contribution of each author
In this work, the author Swatantra Gopal contributed with the 50% and the author Dr. Amrendra Kumar contributed with the other 50%. The activities will: original idea; manuscript writing; experimentation and data collection; machine learning model development; data analysis; conceptual supervision; methodology validation; result interpretation and discussion; critical review and editing of the manuscript.

Discussions and subsequent corrections to the publication
Any dispute, including the replies of the authors, will be published in the first issue of 2027 provided that the information is received before the closing of the third issue of 2026.

Predicción y evaluación en laboratorio de la resistencia a compresión en hormigón con árido reciclado mediante IA.

RESUMEN

Este estudio tiene como objetivo evaluar y predecir la resistencia a compresión del hormigón con árido reciclado (HAR) mediante ensayos experimentales y técnicas de aprendizaje automático. Se analizaron veinticinco mezclas de hormigón con diferentes contenidos de árido reciclado, relación agua-cemento, dosificación de plastificante y resistencia del hormigón base. La resistencia a compresión a los 28 días osciló entre 31,8 y 45,2 MPa, disminuyendo con un mayor contenido de árido reciclado y una mayor absorción de agua. La regresión de vectores de soporte (R²) alcanzó la mayor precisión predictiva ($R^2 = 0,998$), superando a otros modelos. El estudio se ve limitado por el tamaño del conjunto de datos y la expansión controlada de los mismos. Su originalidad reside en la integración de la investigación experimental con el análisis de aprendizaje automático multimodelo. Los resultados demuestran que la IA puede contribuir eficazmente al diseño de mezclas sostenibles de hormigón con árido reciclado.

Palabras clave: hormigón con árido reciclado; resistencia a compresión; aprendizaje automático; regresión de vectores de soporte; construcción sostenible; modelado basado en datos.

Previsão orientada por IA e avaliação laboratorial da resistência à compressão em concreto com agregados reciclados.

RESUMO

Este estudo visa avaliar e prever a resistência à compressão do concreto com agregados reciclados (CAR) utilizando ensaios experimentais e técnicas de aprendizado de máquina. Vinte e cinco misturas de concreto com diferentes teores de agregados reciclados, relação água/cimento, dosagem de plastificante e resistência do concreto original foram investigadas. A resistência à compressão aos 28 dias variou de 31,8 a 45,2 MPa, diminuindo com o aumento do teor de agregados reciclados e da absorção de água. A Regressão por Vetores de Suporte (SVR) obteve a maior precisão de previsão ($R^2 = 0,998$), superando os demais modelos. O estudo é limitado pelo tamanho do conjunto de dados e pela expansão controlada dos dados. A originalidade reside na integração da investigação experimental com a análise de aprendizado de máquina multimodelos. Os resultados demonstram que a IA pode efetivamente apoiar o projeto de misturas sustentáveis de concreto com agregados reciclados.

Palavras-chave: concreto com agregados reciclados; resistência à compressão; aprendizado de máquina; regressão por vetores de suporte; construção sustentável; modelagem orientada por dados.

Legal Information

Revista ALCONPAT is a quarterly publication by the Asociación Latinoamericana de Control de Calidad, Patología y Recuperación de la Construcción, Internacional, A.C., Km. 6 antigua carretera a Progreso, Mérida, Yucatán, 97310, Tel. +52 1 983 419 8241, alconpat.int@gmail.com, Website: www.alconpat.org

Reservation of rights for exclusive use No.04-2013-011717330300-203, and ISSN 2007-6835, both granted by the Instituto Nacional de Derecho de Autor. Responsible editor: Pedro Castro Borges, Ph.D. Responsible for the last update of this issue, Informatics Unit ALCONPAT, Elizabeth Sabido Maldonado.

The views of the authors do not necessarily reflect the position of the editor.

1. INTRODUCTION

Concrete is foundational to modern infrastructure, yet its production comes with a major environmental cost. Cement manufacturing alone contributes significantly to global carbon emissions. Recent analyses show that the cement industry remains a substantial source of CO_2 through energy consumption and chemical processes (e.g., decarbonation of limestone) (Akbar & Liew, 2020b; Khalil & AbouZeid, 2025). At the same time, construction and demolition (C&D) activities generate large volumes of waste, leading to land-use pressure and disposal challenges. The disposal of C&D debris not only burdens landfills but also risks material loss and environmental degradation (Cakiroglu et al., 2023; Pal & Singh, 2024).

Reusing construction and demolition debris to create recycled concrete aggregates (RAs) is a viable solution to both problems. RAs can be obtained by crushing and purifying demolished concrete after impurities such as wood, plastics, and steel reinforcement are removed. (Chakradhara Rao et al., 2011). The resultant material, known as Recycled Aggregate Concrete (RAC), can have engineering qualities similar to those of concrete created with natural aggregates when these recycled aggregates are added to fresh concrete. However, a number of intrinsic factors such as the strength and quality of the parent concrete, the quantity of mortar adhered, the porosity of the recycled particles, and water absorption properties have a significant impact on RAC's effectiveness (Ajdukiewicz & Kliszczewicz, 2002; Khatib, 2005).

The fact that recycled aggregates typically have more porosity and water absorption than native aggregates is one of the primary issues with using recycled aggregate concrete. These characteristics stem from the old, adhered mortar and microcracks that develop during crushing (P. Zhang et al., 2023). The effective water-to-cement ratio, workability, strength, and durability of RAC are all impacted by RAs' high water absorption (Khalil & AbouZeid, 2025). Furthermore, RAC made from low-strength parent concrete often shows more significant reductions in compressive strength than RAC derived from high-strength source concrete, likely because the recycled particles carry residual weaknesses (e.g., old microcracks) into the new mix (Katz, 2003; Padmini et al., 2002).

In addition to water absorption and parent concrete quality, other mix parameters such as replacement ratio (percentage of recycled aggregate), water-cement ratio (w/c), aggregate moisture condition, and the mechanical quality (e.g., abrasion resistance) of the RAs also affect RAC's strength. Because these variables often interact in nonlinear ways, the behavior of RAC is complex and unpredictable (Bai & others, 2023; Huang & Yan, 2025). This complexity necessitates extensive laboratory testing casting multiple concrete batches, curing them, and performing destructive strength tests. Such experimental programs are resource-intensive, time-consuming, and costly.

Predictive modeling has become a potent substitute for relying solely on experimental techniques. In classical concrete research, regression analyses (e.g., multiple linear or multivariate regression) have been used to link mix design parameters with strength and other properties. However, when factors interact in a nonlinear manner or the dataset is high-dimensional, these approaches frequently fail. (Dantas et al., 2013; Shin & Kim, 2013; Yeh, 2007).

Artificial intelligence (AI) and machine learning (ML) techniques have demonstrated significant potential in the past ten years for predicting concrete's compressive strength, especially when handling intricate parameter spaces and nonlinear interactions. Models that can learn from data and take into consideration the complex effects of mix design factors include Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forests, Gradient Boosting, and XGBoost (Bilim, Koksal, et al., 2009; Farhangi et al., 2021; Jahangir & Eidgahee, 2020). These machine learning techniques have already been used in the context of RAC to forecast compressive strength by taking into account inputs such as density, parent concrete strength, water absorption, water-cement ratio, recycled aggregate replacement ratio, and even aggregate abrasion resistance (Khan & others, 2022; X. Zhang et al., 2023).

There are still several gaps despite these developments. Many existing studies rely on datasets limited in size or diversity, focus on only a handful of input parameters, or evaluate just one or two ML models. Critically, few works combine a broad laboratory experimental program (with multiple recycled aggregate contents, parent concrete strengths, and aggregate quality tests) with a comprehensive comparison of ML models. Without such integration, it remains difficult to validate model predictions over a wide experimental domain, and to assess how real-world changes in mix design might influence model performance.

Artificial intelligence (AI) and machine learning (ML) techniques have demonstrated considerable promise in recent years for accurately predicting concrete qualities. Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest, Gradient Boosting, and XGBoost are among the methods that may quickly and affordably predict complex correlations between input parameters and output responses (Bilim, Ozbakkaloglu, et al., 2009; Farhangi et al., 2021; Jahangir & Eidgahee, 2020). These models work especially well for RAC, where compressive strength is determined by the nonlinear interaction of several mix components, aggregate properties, and curing conditions.

In this study, recycled aggregates were carefully processed to ensure predominantly concrete-based composition, minimizing variability associated with mixed C&D waste. While artificial neural networks have been widely used in concrete strength prediction, the present study focuses on alternative machine learning approaches to ensure interpretability and robustness with limited experimental data. We address these gaps by conducting a systematic experimental evaluation and AI-enhanced predictive modeling to investigate the 28-day compressive strength of RAC. Concrete mixes with different amounts of cement, natural and recycled fine and coarse aggregates, water, superplasticizers, and water-to-cement ratios were used in an experimental program. Important RA characteristics are parent concrete strength, Los Angeles abrasion values, and water absorption. The mixes' 28-day compressive strengths varied from roughly 31.8 MPa to 45.2 MPa. To improve the dataset for modeling purposes, each data instance was expanded through controlled variation (±10%) applied to selected mix parameters yielding 1,125 data points. For experimental validation, a series of concrete specimens corresponding to each modification were prepared and tested, generating 1,125 measured strength values. Several machine learning techniques including Support Vector Regression (SVR), Random Forest (RF), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), Gradient Boosting (GB), Linear Regression (LR), Least Absolute Shrinkage and Selection Operator (Lasso), Ridge Regression (Ridge), and Elastic Net (EN), were implemented to predict the compressive strength from input mix parameters. Among these models, SVR achieved the highest predictive accuracy ($R^2 = 0.998$) with minimal errors (MAE = 0.008, RMSE = 0.011), demonstrating the effectiveness of AI based approaches in capturing complex mix strength relationships.

This work seeks to provide a dependable framework for calculating RAC compressive strength by combining AI-driven predictions with experimental measurements. The study supports the sustainable use of recycled aggregates in concrete building, lowers experimental costs, and identifies the most important mix factors.

2. MATERIALS AND METHODS

Water, plasticizer, natural and recycled aggregates, and ordinary Portland cement (OPC) were the ingredients employed in this investigation. Cement complied with standard specifications (International, 2024), providing a consistent binder for all mixes (Akbar & Liew, 2020a). In accordance with ASTM C33 and C127 standards, natural fine and coarse aggregates were obtained locally and evaluated for water absorption, specific gravity, and particle size distribution, the results of which are presented in Table 3.

2.1 Physical Properties of Aggregates

The physical properties of natural fine and coarse aggregates were determined in accordance with relevant ASTM standards, including ASTM C127 and ASTM C128 for specific gravity and water absorption, and ASTM C33 for particle size distribution. These properties are essential for evaluating aggregate quality and their influence on concrete performance.
The experimentally measured values of specific gravity, water absorption, and fineness modulus of natural aggregates are presented in Table 1.

Table 1. Physical properties of natural aggregates used in this study.

Property	Fine Aggregate	Coarse Aggregate
Specific Gravity	2.62	2.68
Water Absorption (%)	1.2	0.8
Fineness Modulus	2.7	-
Maximum Size (mm)	4.75	20
Grading Zone (as per ASTM C33)	Zone II	-

2.2 Physical and Mechanical Properties of Recycled Aggregates

The physical and mechanical properties of recycled fine aggregate (RFA) and recycled coarse aggregate (RCA) were determined to assess their suitability for concrete production and their influence on strength development. Key properties such as specific gravity, water absorption, bulk density, and porosity were evaluated in accordance with relevant ASTM standards.
Recycled aggregates generally exhibited lower specific gravity and higher water absorption compared to natural aggregates due to the presence of adhered mortar and microcracks formed during the crushing process. These characteristics are known to significantly influence the interfacial transition zone (ITZ) between the aggregate and cement paste, thereby affecting the overall mechanical performance of recycled aggregate concrete.
Table 2 summarizes the measured physical properties of RCA and RFA used in this study. The results indicate increased porosity and water absorption, which contribute to reduced strength at higher replacement levels.

Table 2. Physical and mechanical properties of recycled aggregates.

Property	RFA	RCA
Specific Gravity	2.4	2.55
Water Absorption (%)	5.5	4.5
Bulk Density (kg/m³)	1400	1500
Porosity (%)	12	9
Los Angeles Abrasion (%)	-	25
Fineness Modulus	2.6	-

Note: *These values reflect the typical characteristics of recycled aggregates, where higher porosity and water absorption are associated with adhered mortar and microstructural defects.*

Recycled aggregates (both fine and coarse) were obtained from construction and demolition (C&D) waste primarily consisting of crushed concrete materials collected from local demolition activities within the same geographical region. Although the material was not sourced from a single demolition site, care was taken to ensure consistency by selecting predominantly concrete-based waste.

The collected waste was subjected to careful manual sorting to remove impurities such as wood, plastics, steel reinforcement, and other non-concrete constituents. Visible masonry fragments and undesirable materials were minimized during processing to ensure that the recycled aggregates were predominantly concrete based.

The processed material was then crushed and sieved to obtain the required size fractions for recycled fine aggregate (RFA) and recycled coarse aggregate (RCA). Although minor traces of adhered mortar are inherent in recycled aggregates, efforts were made to maintain consistency in material quality across all mixes through uniform processing conditions.

The parent concrete strength associated with the recycled aggregates ranged from 25 MPa to 45 MPa, as considered in the experimental program. This controlled selection and processing approach ensured relatively consistent aggregate characteristics, supporting the reliability of both experimental results and machine learning model predictions.

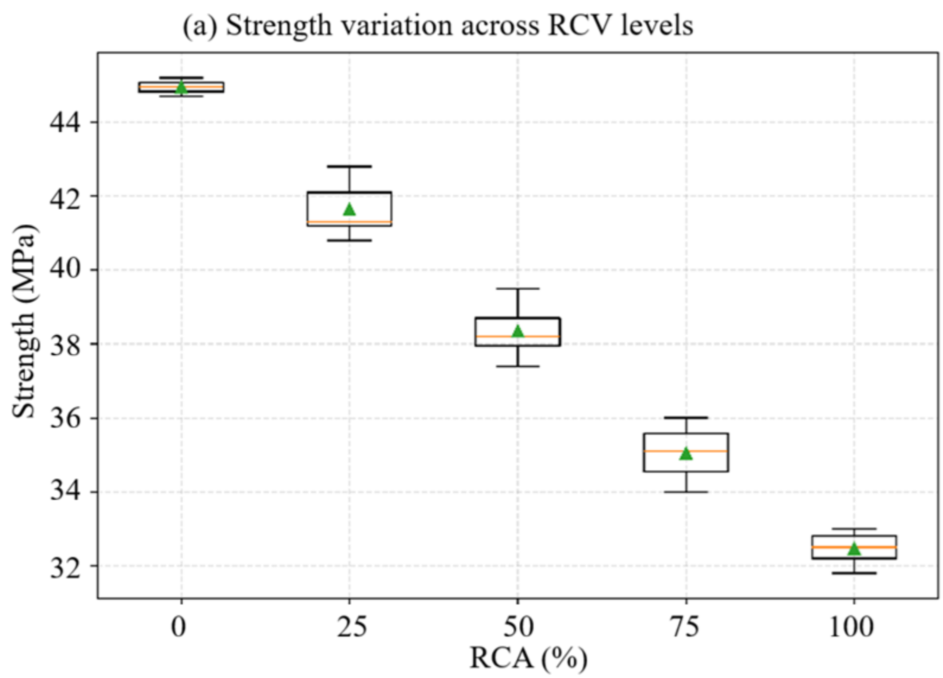
As shown in Table 3, different amounts of cement, natural and recycled aggregates, water, superplasticizer, and water-to-cement ratios were used to make concrete mixes. Recycled aggregate replacement levels ranged from 0% to 100% for both fine and coarse aggregates. Water-cement ratios were varied from 0.35 to 0.55, and parent concrete strengths ranged between 25 MPa and 45 MPa. The mix design also accounted for the water absorption of recycled aggregates and their LA abrasion values, which influence compressive strength (Katz, 2003; Padmini et al., 2002). Each mix was prepared in a laboratory-controlled environment and cured under standard conditions. The 28-day compressive strength of the mixes varied from 31.8 MPa to 45.2 MPa, illustrating the effects of recycled aggregate content, water absorption, and parent concrete quality on mechanical performance.

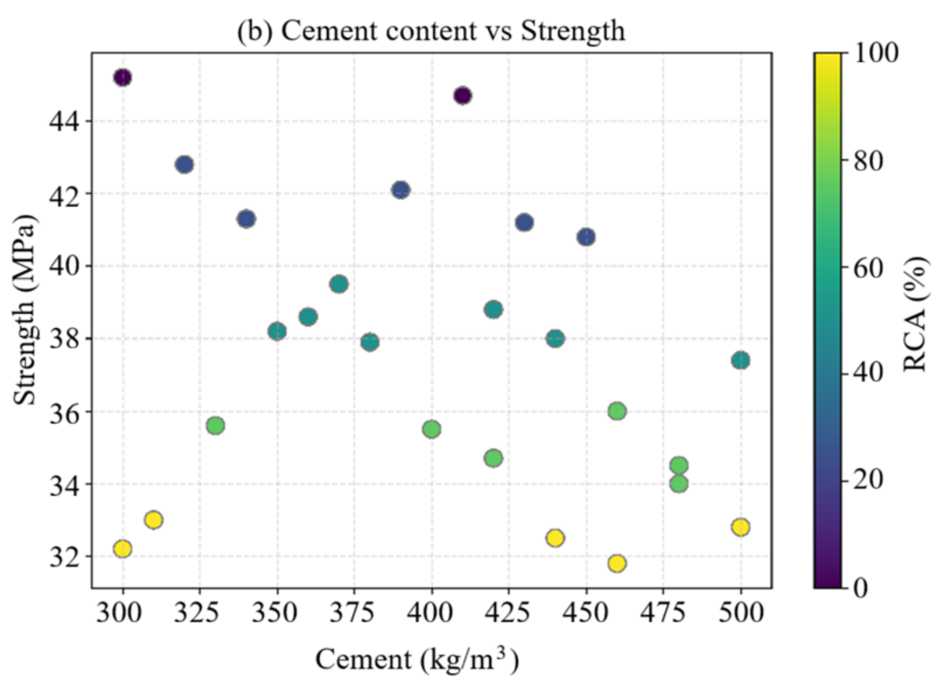
Table 3. Mix proportions and material characteristics of the tested concretes.

Mix	Cement (kg/m³)	Sand (kg/m³)	Fine Agg. (kg/m³)	Recycled Fine Agg. (kg/m³)	Coarse Agg. (kg/m³)	Recycled Coarse Agg. (kg/m³)	Water (kg/m³)	Plasticizer (kg/m³)	w/c (–)	RCA %	PC	RCA WA %	LAA	CS
M1	300	650	700	0	900	0	105	1.50	0.35	0	45	2.00	18	45.20
M2	320	640	600	100	800	100	110	2.56	0.40	25	40	3.00	21	42.80
M3	340	630	550	150	750	150	115	3.40	0.40	25	35	3.50	22	41.30
M4	360	620	500	200	700	200	120	4.32	0.45	50	35	4.00	24	38.60
M5	380	610	450	250	650	250	125	5.32	0.45	50	30	4.20	25	37.90
M6	400	600	400	300	600	300	130	6.40	0.50	75	30	5.00	26	35.50
M7	420	590	350	350	550	350	135	7.56	0.50	75	28	5.50	27	34.70
M8	440	580	300	400	500	400	140	8.80	0.55	100	25	6.00	28	32.50
M9	460	570	250	450	450	450	145	8.28	0.55	100	25	6.50	29	31.80
M10	480	560	200	500	400	500	150	7.68	0.50	75	30	5.80	28	34.00
M11	500	550	150	550	350	550	155	7.00	0.45	50	35	4.50	26	37.40
M12	450	570	100	600	300	600	150	5.40	0.40	25	40	3.50	24	40.80
M13	430	580	50	650	250	650	145	4.30	0.40	25	40	3.00	23	41.20
M14	410	590	0	700	200	700	140	3.28	0.35	0	45	2.50	20	44.70
M15	390	600	100	600	250	650	130	3.90	0.40	25	38	3.00	22	42.10
M16	370	610	150	550	300	600	125	4.44	0.45	50	35	4.00	24	39.50
M17	350	620	200	500	350	550	120	4.90	0.45	50	30	4.50	26	38.20
M18	330	630	250	450	400	500	115	5.28	0.50	75	28	5.00	27	35.60
M19	310	640	300	400	450	450	110	5.58	0.55	100	25	6.00	29	33.00
M20	300	650	350	350	500	400	105	6.00	0.55	100	25	6.50	30	32.20
M21	420	600	400	300	550	350	130	5.04	0.45	50	35	4.00	25	38.80
M22	440	590	450	250	600	300	135	6.16	0.45	50	30	4.50	26	38.00
M23	460	580	500	200	650	250	140	7.36	0.50	75	30	5.00	27	36.00
M24	480	570	550	150	700	200	145	8.64	0.50	75	28	5.50	28	34.50
M25	500	560	600	100	750	150	150	10.00	0.55	100	25	6.00	29	32.80

Note: *w/c = water–cement ratio; RCA = recycled coarse aggregate; RFA = recycled fine aggregate.*

For each concrete mix, specimens were prepared to evaluate 28-day compressive strength following (International, 2024; Zain et al., 2002) standards. Cubes with dimensions of 150 × 150 × 150 mm were cast for each mixture, compacted using a standard vibrating table to ensure uniform density, and covered with wet burlap to prevent moisture loss during initial curing. The specimens were then submerged in water at 23 ± 2°C for 28 days to achieve full hydration (Ajdukiewicz & Kliszczewicz, 2002; Khatib, 2005). Results demonstrated a gradual reduction in compressive strength with increasing recycled aggregate replacement, particularly at higher water absorption levels, highlighting the influence of parent concrete quality and recycled aggregate characteristics on RAC performance (Katz, 2003; Padmini et al., 2002).





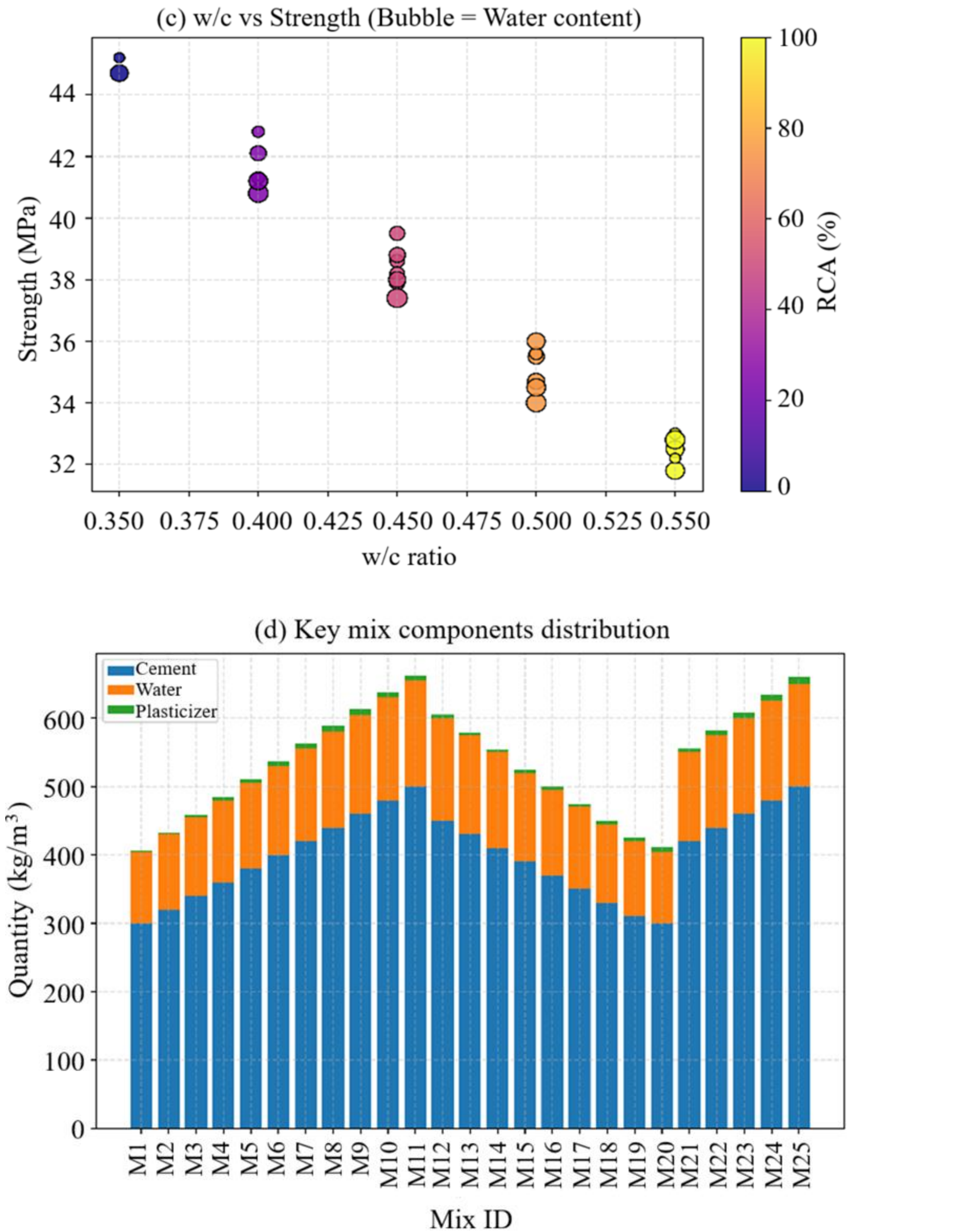


Figure 1. Statistical visualization of the RMC dataset showing the distribution of key mix constituents and compressive strength of mix designs. (a) Strength variation across RCA levels, (b) Cement content vs strength, (c) w/c vs strength, (d) Key mix components Distribution.

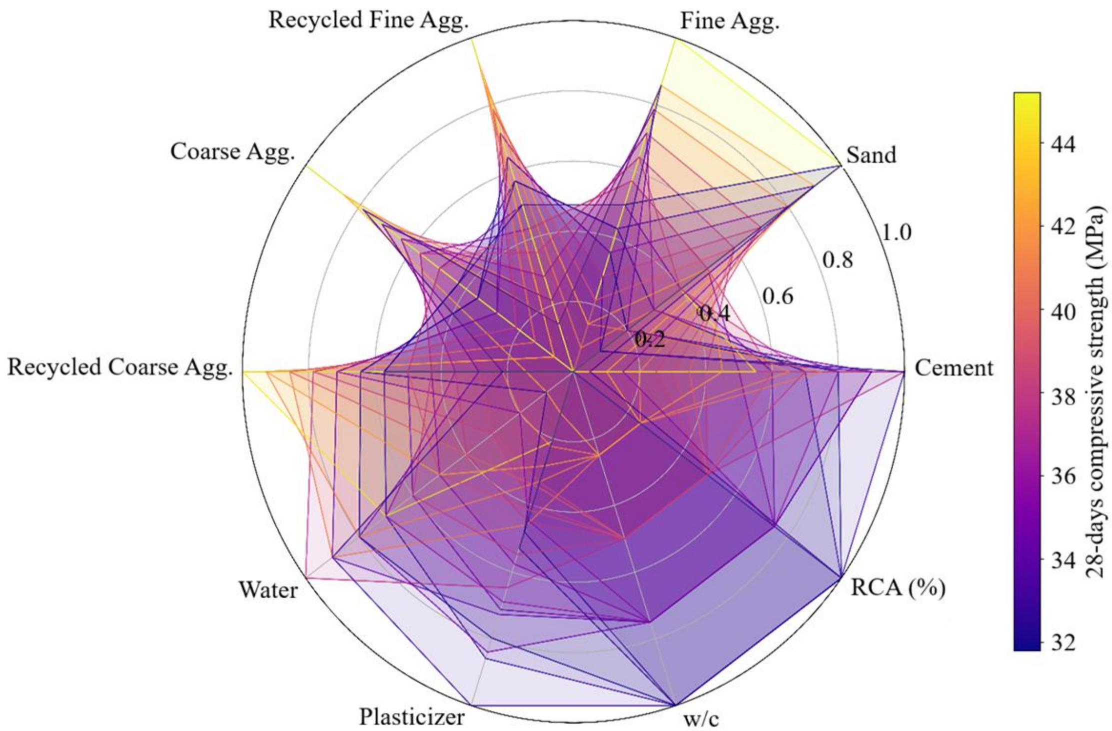


Figure 2. Radar plot of normalized mix design parameters for all 25 experimental recycled aggregate concrete (RAC) mixes.

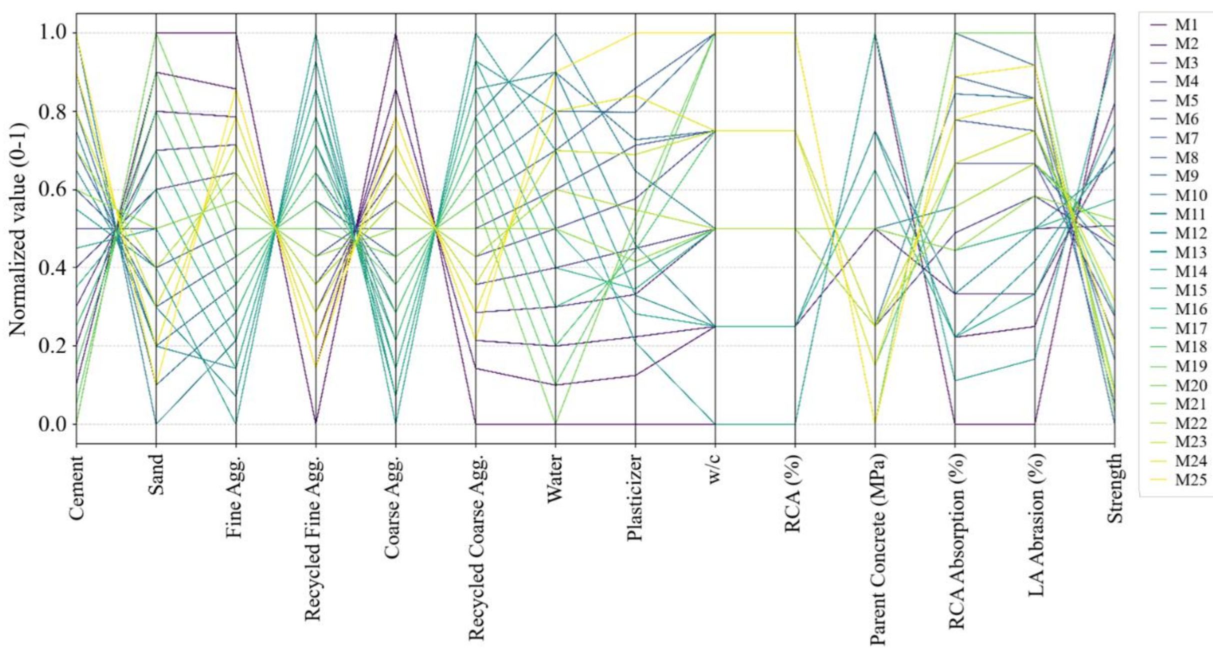


Figure 3. Parallel coordinates plot for full experimental RAC dataset.

The statistical distribution of the main mix components in the RMC dataset, which comprises concrete mixtures made with different ratios of natural and recycled aggregates, is shown in Figure 1. The boxplots illustrate the variability in cement, sand, coarse and fine aggregates, recycled aggregates, water, and plasticizer contents, together with the corresponding 28-day compressive strengths. The distribution of strength across different RCA levels shows a consistent reduction as

the proportion of recycled aggregates increases. This trend reflects the influence of adhered mortar, higher porosity and lower stiffness typically associated with recycled aggregates. The spread observed within each RCA group also indicates the variability introduced by changes in mix composition. The figure highlights the sensitivity of mechanical performance to RCA incorporation and provides a statistical basis for subsequent predictive modelling.
Figure 2 displays a radar plot of the concrete mixes, showing normalized values of all mix design parameters, including cement, natural and recycled aggregates, water, plasticizer, the water–cement ratio and RCA content. Each polygon represents one mixture, while the color intensity reflects the corresponding compressive strength. The plot provides an integrated view of multivariable interactions within the dataset and clarifies how variations in aggregate composition, water demand and RCA content relate to strength development. Before applying machine learning models, this display provides a clear picture of parameter variability.
A parallel coordinates plot of the experimental dataset is shown in Figure 3, where a polyline connecting all input parameters represents each mixture. This format enables direct comparison of parameter trends, particularly the effect of RCA percentage, water content and aggregate proportions on compressive strength. The figure emphasizes the multivariable relationships within the dataset and supports the identification of patterns relevant to model development and performance evaluation.
The higher porosity and water absorption of recycled aggregates directly affect the interfacial transition zone (ITZ), which is typically the weakest region in concrete. The presence of adhered mortar and microstructural defects in RCA leads to a more porous ITZ, resulting in reduced bond strength between aggregate and cement paste. This phenomenon becomes more pronounced at higher replacement levels, contributing to the observed reduction in compressive strength.

3. MACHINE LEARNING MODELING

The experimental dataset of concrete mix designs was utilized as input for machine learning modeling to forecast the 28-day compressive strength of recycled aggregate concrete (RAC) after the experimental program. Cement quantity, proportions of natural and recycled fine and coarse aggregate, water-cement ratio, plasticizer dose, parent concrete strength, recycled aggregate water absorption, and Los Angeles (LA) abrasion values were all used to define each mix. To enhance the robustness of the dataset, a controlled data expansion approach was adopted by introducing a limited variation of ±10% in selected mix parameters. This variation was applied within realistic engineering bounds to simulate minor fluctuations that typically occur in material properties and field conditions, rather than to artificially duplicate identical data points.
Unlike purely synthetic data generation methods, this approach preserves the underlying physical relationships among variables while allowing the machine learning models to generalize better within a realistic parameter space. However, it is acknowledged that this form of data augmentation may introduce some degree of bias, and therefore the results should be interpreted with appropriate caution. This approach allowed for better generalization of the machine learning algorithms and reduced the risk of overfitting (Farhangi et al., 2021; Jahangir et al., 2021).
More advanced data augmentation techniques such as Synthetic Minority Over-sampling Technique (SMOTE) or Generative Adversarial Networks (GANs) were not employed in this study, as the dataset does not represent a classification problem or imbalanced data scenario. Moreover, the focus of the present work is on preserving physically interpretable relationships rather than generating entirely synthetic samples. Future studies may explore such advanced techniques to further improve model generalization and robustness.
Nine machine learning techniques were implemented to model the link between input mix variables and compressive strength. These included Support Vector Regression (SVR), Random

Forest (RF), K-Nearest Neighbors (KNN), Extreme Gradient Boosting (XGBoost), Gradient Boosting (GB), Linear Regression (LR), Lasso Regression, Ridge Regression, and Elastic Net (EN). Nonlinear models, such as SVR, RF, and ensemble methods (XGBoost and GB), were selected for their ability to capture complex, nonlinear interactions among multiple variables, which are common in RAC due to the variability in recycled aggregate quality, water absorption, and parent concrete strength (Bilim, Atis, et al., 2009; Deng et al., 2018; Naderpour et al., 2018).
Although detailed physical properties of recycled aggregates such as water absorption and Los Angeles abrasion were incorporated into the dataset, other characteristics such as porosity and density were not explicitly included as independent input variables. However, their influence is inherently reflected through correlated parameters such as water absorption and parent concrete strength, which are strongly linked to aggregate porosity and quality.
Future studies may incorporate a broader set of aggregate physical and microstructural properties as direct input features to further enhance model interpretability and predictive capability.
The dataset was split into training (80%) and testing (20%) subsets to evaluate predictive performance. Model evaluation metrics included the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean bias error (MBE). SVR achieved the highest predictive accuracy, with an R^2 value of 0.998, MAE of 0.008, and RMSE of 0.011, indicating excellent agreement between predicted and experimental compressive strength values. Random Forest was the second-best model, achieving $R^2 = 0.996$ and slightly higher errors (MAE = 0.128, RMSE = 0.179). The KNN and XGBoost models showed moderate performance with R^2 values of 0.947 and 0.942, respectively, while gradient boosting and linear-based models (LR, Lasso, Ridge, Elastic Net) demonstrated lower predictive accuracy (R^2 ranging from 0.842 to 0.904), highlighting the superiority of nonlinear algorithms for this type of dataset (Farhangi et al., 2021; Ghanizadeh et al., 2019).
Feature importance analysis was performed to understand the influence of each input variable on RAC compressive strength. Water-cement ratio, recycled coarse aggregate content, and parent concrete strength were identified as the most critical parameters influencing compressive strength, whereas plasticizer dosage and fine aggregate proportions had a comparatively smaller effect. This approach reduces the need for lengthy laboratory experiments by guiding mix design optimization for real-world applications in addition to validating the experimental data.

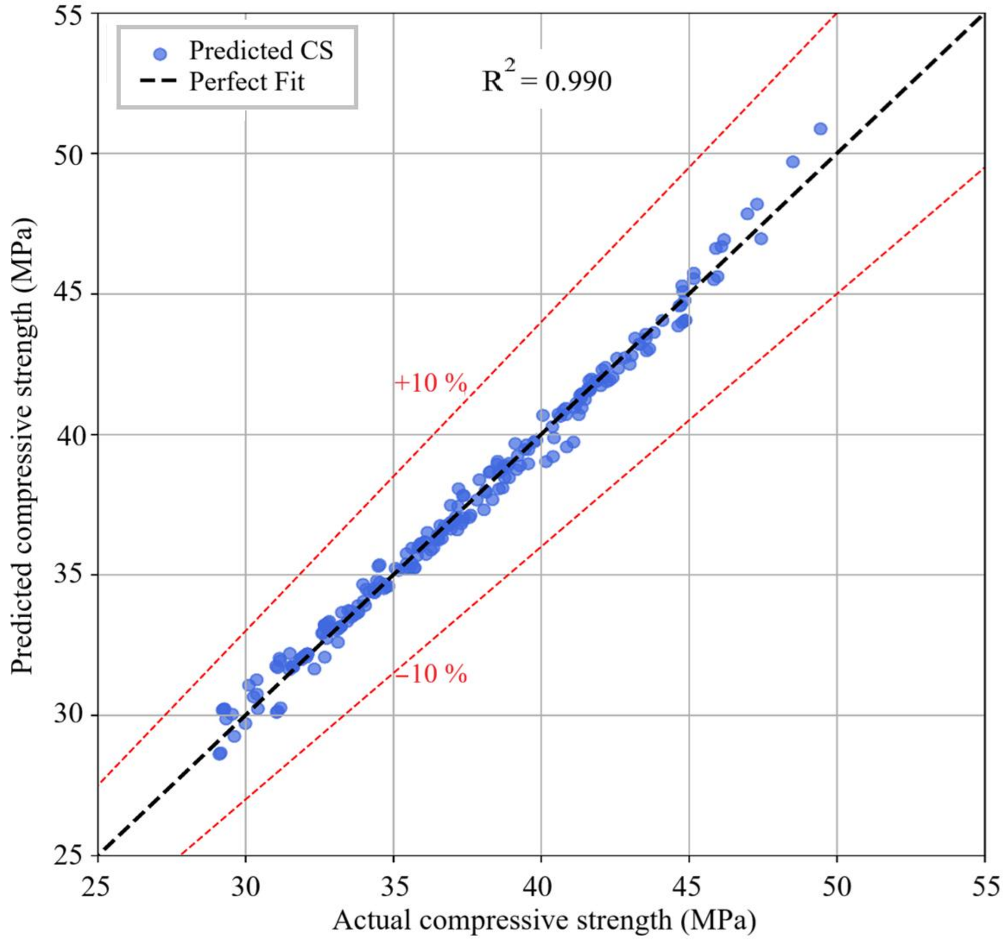


Figure 4. Predicted vs actual 28-day compressive strength using combined input parameters.

Figure 4 shows the relationship between the predicted and measured 28-day compressive strengths obtained from the model developed using all input parameters. The data points exhibit a close alignment with the line of perfect prediction, indicating strong agreement between experimental and estimated values. The ±10 percent deviation bands provide a clear visual reference for the model's accuracy, with most points falling within this range. The high coefficient of determination (R^2 = 0.990) further confirms the model's reliability and its ability to capture the key factors governing strength development.

A strong tool for effectively forecasting RAC attributes is created by combining machine learning modeling with experimental data. The approach minimizes resource-intensive experimental work while ensuring reliable and accurate strength estimation. These results reinforce the potential of AI-driven methods in sustainable concrete design, where multiple recycled aggregate levels and variable material qualities need to be considered simultaneously (Deng et al., 2018; Jahangir et al., 2021; Naderpour et al., 2018).

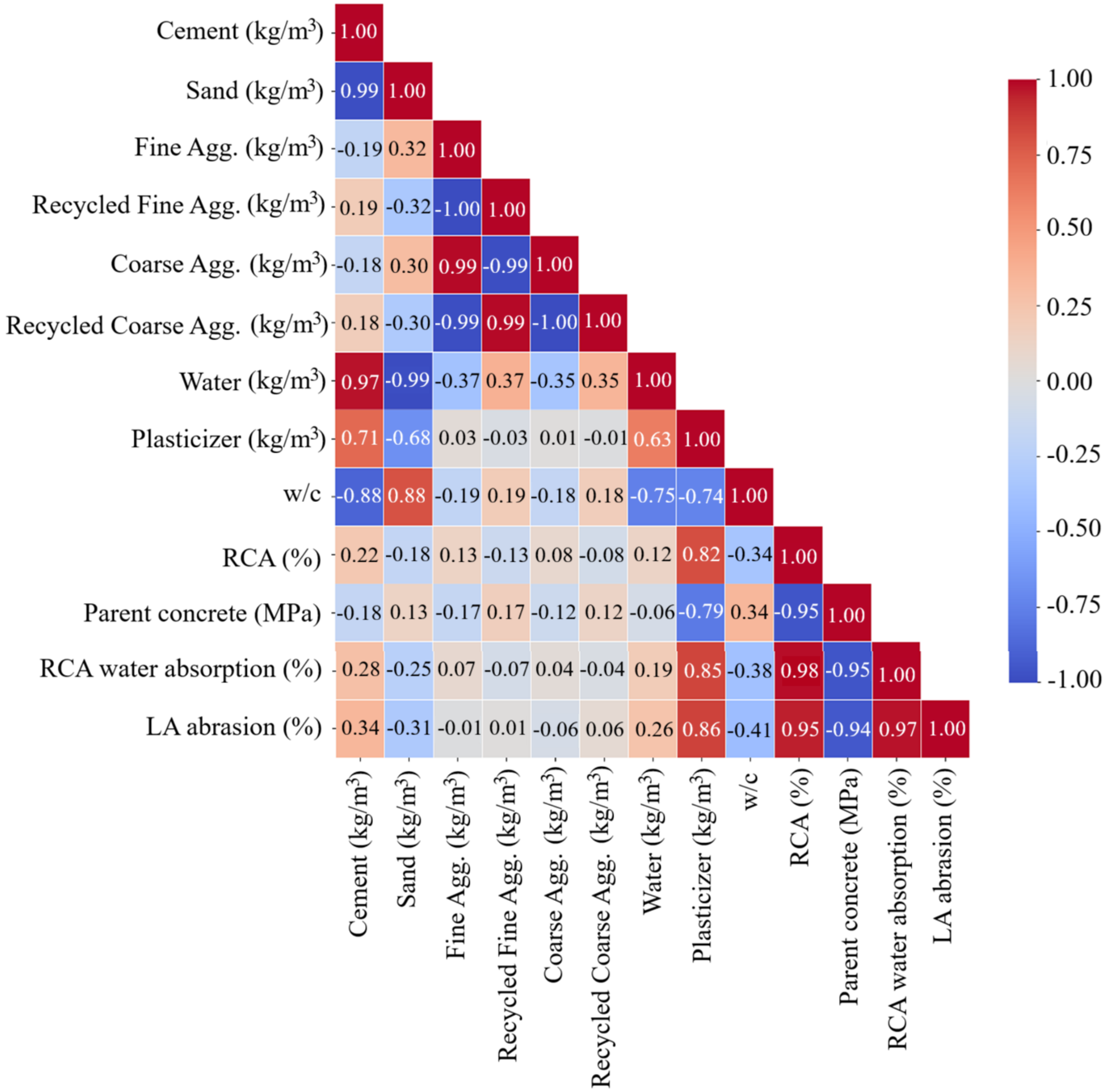


Figure 5. Correlation heatmap showing relationships among concrete mix parameters.

Figure 5 presents the Pearson correlation matrix for the input variables used in the concrete mix dataset. The heatmap highlights the strength and direction of linear relationships among the parameters. Notable positive correlations are observed between cement and water, as well as between plasticizer content and RCA water absorption. In contrast, strong negative correlations appear between cement and sand, and between the water-cement ratio and cement content. These patterns reflect fundamental interactions within the mix design and provide useful guidance for understanding parameter dependencies, improving mix optimization strategies and supporting feature selection for predictive modelling of compressive strength.

This study used a range of machine learning (ML) methods to predict the compressive strength of recycled aggregate concrete (RAC). A total of nine models were implemented, encompassing linear regression, regularized linear models, kernel-based models, nearest neighbors, ensemble tree-based methods, and gradient boosting approaches. The predictive performance of each model was evaluated using multiple statistical metrics, including the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), mean bias error (MBE), and t-statistic. To ensure comparability, all models were trained using an 80/20 train-test split with standardized numerical characteristics (Farhangi et al., 2021; Jahangir & Eidgahee, 2020).

Table 4. Range of variables considered for model construction.

Input Parameter	Min	Max	Mean	Std. Dev.	Variance
Cement (kg/m³)	300	500	401.27	61.9	3831.31
Sand (kg/m³)	550	650	599.75	28.88	834.16
Fine Agg. (kg/m³)	0	700	337.95	187.92	35315.6
Recycled Fine Agg. (kg/m³)	0	700	362.05	187.92	35315.6
Coarse Agg. (kg/m³)	200	900	515.86	188.2	35418.1
Recycled Coarse Agg. (kg/m³)	0	700	384.14	188.2	35418.1
Water (kg/m³)	105	155	130.73	14.94	223.16
Plasticizer (kg/m³)	1.5	10	5.72	2.05	4.19
w/c (–)	0.3	0.35484	0.33	0.02	0
RCA (%)	0	100	56.97	30.49	929.64
Parent Concrete (MPa)	25	45	32.49	6.09	37.07
RCA Water Absorption (%)	2	6.5	4.52	1.25	1.56
LA Abrasion (%)	18	30	25.35	3.02	9.13
28-Day Compressive Strength (MPa)	28.74	49.45	37.53	4.35	18.9

The statistical range and variability of the input parameters used to create the model are summarized in Table 4. Cement, sand, natural and recycled aggregates, water, plasticizer, water-to-cement ratio, RCA content, and associated characteristics like parent concrete strength, RCA water absorption, and Los Angeles abrasion are all listed in the table along with their minimum, maximum, mean, standard deviation, and variance. The data highlight the wide range of values considered, reflecting the diversity of the concrete mixes. This thorough characterization supports the assessment of parameter effect and interactions and offers a solid foundation for predictive modeling of 28-day compressive strength.

Although parameters such as aggregate density and porosity were not explicitly included as independent input variables, their influence is indirectly represented through correlated parameters such as recycled aggregate water absorption, Los Angeles abrasion, and parent concrete strength. These parameters are widely recognized as indicators of aggregate quality, porosity, and mechanical integrity.

Water absorption reflects the internal pore structure and permeability of recycled aggregates, while Los Angeles abrasion provides insight into their resistance to mechanical degradation. Together, these variables capture the intrinsic characteristics of recycled aggregates that govern the interfacial transition zone (ITZ) behavior and overall compressive strength of concrete.

However, it is acknowledged that the direct inclusion of additional physical properties such as bulk density and porosity as independent input variables could further enhance the interpretability and predictive capability of the model. This aspect is identified as a scope for future research.

Table 5. Performance indicators used to assess the regression models.

Metric	Formula	Description	Standard/Reference Value
Coefficient of Determination (R^2)	$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}$	Indicates how well the predicted values match the observed data; measures overall fit.	0 to 1; higher values indicate better fit ($R^2 \geq 0.8$ considered good)
Mean Squared Error (MSE)	$MSE = \frac{1}{N}\sum(y_i - \hat{y}_i)^2$	Average of the squared differences between predicted and actual values.	Closer to 0 is better; no absolute standard (depends on data scale)
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum(y_i - \hat{y}_i)^2}$	Square root of MSE; quantifies the typical magnitude of prediction errors.	Closer to 0 is better; should be compared relative to target range
Mean Absolute Error (MAE)	$MAE = \frac{1}{N}\sum\lvert y_i - \hat{y}_i\rvert$	Average of the absolute differences between predicted and observed values.	Closer to 0 is better; relative comparison recommended
Mean Absolute Percentage Error (MAPE)	$MAPE = \frac{100}{N}\sum\left\lvert\frac{(y_i - \hat{y}_i)}{y_i}\right\rvert$	Average percentage deviation of predictions from actual values.	<10% excellent, 10–20% good, 20–50% acceptable, >50% poor
Mean Bias Error (MBE)	$MBE = \frac{1}{N}\sum(y_i - \hat{y}_i)$	Average signed difference between predicted and observed values, indicating systematic bias.	0 ideal; positive indicates underestimation, negative overestimation
t-Statistic (t)	$t = \sqrt{\frac{(N-1)MBE^2}{RMSE^2 - MBE^2}}$	Statistic used to evaluate whether the mean bias is significantly different from zero.	

Here, N denotes the number of samples in each fold of cross-validation. The t-statistic is employed to determine whether the mean difference between paired prediction errors is statistically significant. The notation y_i and $\hat{y}_i$ refer to the observed and predicted values, respectively.

The statistical metrics used to assess the regression models' prediction performance are compiled in Table 5. An overall measure of goodness of fit is provided by the coefficient of determination (R^2), which represents the percentage of variance in the measured compressive strength that is captured by the model predictions. While mean absolute percentage error (MAPE) reflects deviations from the observed values, error-based metrics like mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) measure the size of prediction mistakes in absolute and squared terms. Mean bias error (MBE) is included to identify systematic tendencies toward overestimation or underestimation. The paired-sample t-statistic is used to assess whether the mean bias is statistically significant. These indicators are widely adopted in regression and machine learning applications for material behaviour and structural performance

modelling, ensuring rigorous and interpretable evaluation of predictive accuracy (Chicco & Warrens, 2021; Emmert-Streib & Dehmer, 2019; Miller & others, 2024; Plevris et al., 2022; Steurer et al., 2021).
The machine learning process used to forecast the compressive strength of recycled aggregate concrete is shown in Figure 6. The input parameters and data preprocessing procedures, such as normalization and splitting the dataset into training and testing subsets, are described in the first stage. The creation of the predictive models and hyperparameter optimization for algorithms like support vector regression, random forest, XGBoost, and linear regression variations are covered in the second stage. In the third step, statistical metrics such as the coefficient of determination, mean absolute error, and root mean squared error are used to assess the model's performance, and feature influence is interpreted using SHAP-based analysis. The final stage reports the predicted compressive strength, completing a structured and transparent framework for model assessment and interpretation.

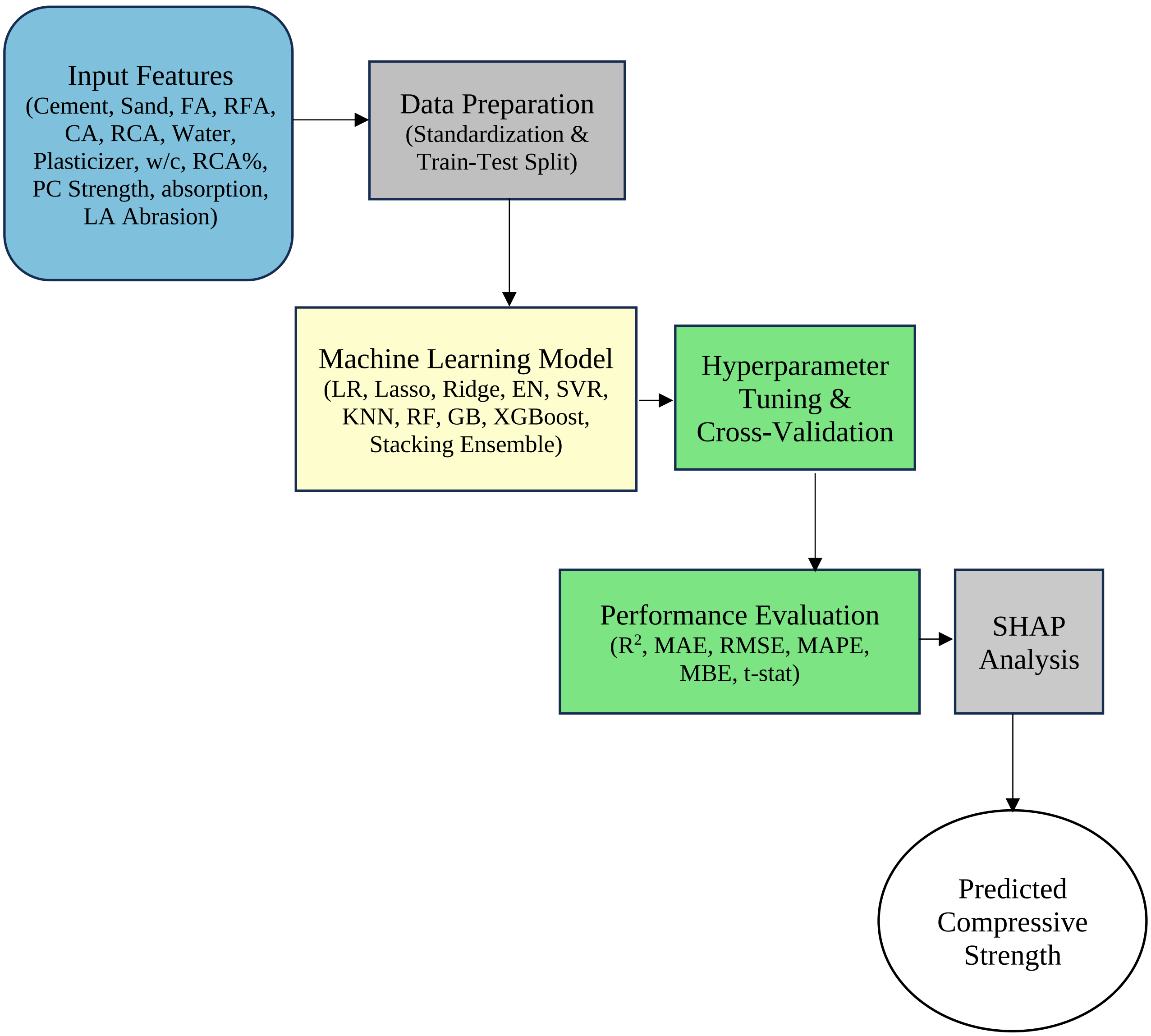


Figure 6. Machine learning framework's workflow for forecasting the compressive strength of recycled aggregate concrete compressive strength.

3.1 Machine learning models used in this study

3.1.1 Support Vector Regression (SVR)

Support Vector Regression (SVR) was implemented using a radial basis function (RBF) kernel to model nonlinear dependencies between input features and compressive strength. Hyperparameters, including the penalty parameter C and epsilon-insensitive loss function, were tuned to minimize prediction error (Smola & Scholkopf, 2004). SVR achieved the highest predictive accuracy among individual models (R^2 = 0.998), with minimal errors (MAE = 0.008, RMSE = 0.011). This demonstrates how well kernel-based techniques capture intricate feature target interactions in RAC.

3.1.2 Random Forest

Random Forest (RF) algorithm is a non-linear ensemble method that build multiple decision trees on bootstrap samples and aggregate predictions to reduce variance and improve robustness. RF uses bootstrapped sampling with feature randomness, with RF achieving R^2 = 0.996 Hyperparameters, such as the maximum tree depth (12) and the number of estimators (400), were tuned to balance bias and variance. These tree-based ensembles are highly effective for tabular datasets with heterogeneous input features (Breiman, 2001; Geurts et al., 2006).

3.1.3 k-Nearest Neighbors (KNN)

The KNN regressor estimates the target value as a relation of the average of k nearest training instances in the feature space. In this study, k was set to 5, and Euclidean distance was used to identify neighbors. KNN performed moderately well (R^2 = 0.947) but is sensitive to feature scaling and outliers, highlighting the importance of preprocessing and standardization (Altman, 1992). Its local approximation approach can effectively model small-scale nonlinearities but struggles with high-dimensional data relative to global ensemble models.

3.1.4 Gradient Boosting and XGBoost

Gradient Boosting (GB) and Extreme Gradient Boosting (XGBoost) sequentially construct decision trees, each correcting the errors of the previous one, to minimize a specified loss function. GB was implemented with 400 trees, learning rate 0.05, and max depth 4, while XGBoost used 600 trees, learning rate 0.03, max depth 5, and subsampling of 0.9. Extensive hyperparameter tuning and 5-fold cross-validation were applied to prevent overfitting and improve generalization (Chen & Guestrin, 2016; Friedman, 2001). Both models demonstrated strong nonlinear modeling capabilities, achieving R^2 values of 0.904 (GB) and 0.942 (XGBoost).

3.1.5 Linear Regression Models

The goal variable, compressive strength, and the input features are assumed to have a linear relationship by linear regression (LR). Ordinary least squares (OLS) was used to minimize the difference between observed and predicted values, providing a simple baseline for comparison (Bhanja & Sengupta, 2005). Ridge Regression, Lasso, and Elastic Net are examples of regularized linear models that were used to reduce overfitting and enhance generalization. Lasso uses L1 regularization to pick features, Ridge uses L2 regularization to penalize large coefficients, and Elastic Net combines the two penalties (Hastie et al., 2009). These models achieved R^2 values ranging from 0.842 to 0.893, indicating reasonable performance but limited ability to capture nonlinear interactions inherent in RAC datasets.

3.1.6 Elastic Net

Elastic Net (EN) regression incorporates both L1 (Lasso) and L2 (Ridge) regularization terms, it can choose features while remaining stable when there is multicollinearity among the input variables (Zou & Hastie, 2005). By adjusting the mixing parameter α, Elastic Net can shift the balance between L1 sparsity and L2 shrinkage, making it particularly useful when there are correlated predictors, as is common in concrete mix datasets where multiple material properties interact. To maximize generalization, Elastic Net was used in this investigation with a moderate L1 ratio and standard input feature scaling. The model produced an R^2 of 0.846, slightly lower than Ridge or Lasso, indicating that while it partially captured linear trends in the data, it struggled to fully model the nonlinear effects of recycled aggregate properties, water absorption, and parent concrete strength (Hastie et al., 2009; Zou & Hastie, 2005).

3.1.7 Ridge Regression

Ridge Regression (RR) is a regularized linear regression method that reduces the regression coefficients to avoid overfitting and enhance generalization by adding an L2 penalty component to the standard least squares cost function (Hoerl & Kennard, 1970). This approach works well when the predictors show strong correlation, because it helps control variance while keeping all variables in the model. In the present study, ridge regression was used to estimate the 28-day compressive strength of recycled aggregate concrete, and all input features were standardized to maintain uniform influence across the variables. The model achieved an R^2 of 0.842, indicating moderate predictive performance. While Ridge effectively handled multicollinearity among mix parameters, such as water–cement ratio, recycled aggregate content, and aggregate properties, it was limited in capturing the nonlinear interactions inherent in RAC datasets, compared to ensemble or kernel-based models (Hastie et al., 2009; Hoerl & Kennard, 1970).

3.1.8 Stacking Ensemble Model

To leverage complementary strengths of individual models, a stacking ensemble was implemented, combining the top-performing regressors. In stacking, base learners generate predictions that serve as inputs for a meta-learner, optimizing final output accuracy. The stacking ensemble achieved nearly perfect predictive accuracy ($R^2 = 0.999$), significantly outperforming individual models across all metrics. This suggests that complicated interactions between mix factors and compressive strength can be efficiently captured by integrating heterogeneous models (Wolpert, 1992; J. Zhang et al., 2022).

3.1.9 Hyperparameter Tuning and Cross-Validation

Grid search were used for hyperparameter optimization for all ML models, and cross-validation was used to ensure model robustness. Parameters such as learning rate, number of estimators, maximum depth, regularization coefficients, and neighbor counts were systematically adjusted to minimize validation errors. Cross-validation allowed identification of underfitting or overfitting behavior, ensuring generalizable predictions. In order to successfully balance bias and variance, early stopping criteria were used for boosting models to halt training when no improvement was seen (Chen & Guestrin, 2016; Hastie et al., 2009).

Although artificial neural networks (ANN) and other deep learning architectures such as radial basis function networks have demonstrated strong performance in predicting concrete properties, they were not prioritized in the present study due to the relatively limited size and structured nature of the experimental dataset.

Deep learning models generally require large and diverse datasets to achieve reliable generalization and to avoid overfitting. In contrast, the current study focuses on a controlled experimental dataset with well-defined input parameters, where classical machine learning models

such as Support Vector Regression and ensemble methods (Random Forest, Gradient Boosting) are known to perform effectively.
Furthermore, the selected models offer better interpretability and computational efficiency, allowing clearer insight into the influence of mix design parameters on compressive strength. This aligns with the objective of the study to not only achieve accurate prediction but also to understand the underlying material behavior.
Nevertheless, the potential of ANN and deep learning approaches is acknowledged, and future work may explore their application using larger and more diverse datasets.

4. DATASET DESCRIPTION AND ANALYSIS

For the current study, a comprehensive dataset comprising 1,125 samples was used, combining 25 experimental concrete mixes replicated 45 times with ±10% random variation to capture variability in mix parameters and measurement conditions. This method ensures a robust and diverse dataset appropriate for machine learning uses. (Goodfellow et al., 2016; Naderpour et al., 2018). The dataset includes 13 input features: cement, natural sand, fine aggregate (FA), recycled fine aggregate (RFA), coarse aggregate (CA), recycled coarse aggregate (RCA), water, superplasticizer, water-cement ratio (w/c), RCA replacement percentage, parent concrete strength, RCA water absorption, and Los Angeles (LA) abrasion values. The target variable is the 28-day compressive strength (CS) of recycled aggregate concrete.
To describe the dataset, descriptive statistics were calculated for each parameter, including minimum, maximum, mean, standard deviation, skewness, and kurtosis (Field, 2013). Cement, fine and coarse aggregates, and water showed moderate variation across mixes, while RCA content, RCA water absorption, and LA abrasion demonstrated wider variability due to differences in recycled material properties. The skewness of each variable indicated the asymmetry in data distribution: positive skew was detected for RCA water absorption and RFA content, suggesting a higher frequency of lower values, whereas negative skew appeared in natural coarse aggregate and cement contents, reflecting a clustering at higher values. Kurtosis analysis revealed relatively high peak values for superplasticizer and RCA, highlighting the concentrated distribution of specific additives in selected mixes. These statistical insights confirm that the dataset adequately captures variations in both natural and recycled materials, enabling reliable machine learning modeling (Jahangir & Eidgahee, 2020).
The dataset exhibits sufficient variability, appropriate sample size, and a balanced ratio of input features to observations, making it suitable for machine learning-based prediction of RAC compressive strength. To provide both experimental rigor and practical relevance for sustainable concrete design, the combination of descriptive statistics, visualizations, and correlation analysis provides a strong basis for later modeling and feature selection procedures (Bilim, Ozbakkaloglu, et al., 2009; Naderpour et al., 2018).

4.1 Performance assessment of models

Using the experimental dataset of concrete mixes and their enlarged replicas, the current work used machine learning and multiple regression models to predict the 28-day compressive strength of recycled aggregate concrete (RAC). The models assessed include LR, Ridge, Lasso, EN, SVR, KNN, RF, GB, and XG Boost. A number of indicators were used to assess the model's performance including the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), mean bias error (MBE), and t-statistics, providing a comprehensive understanding of predictive accuracy and bias (Farhangi et al., 2021; Jahangir & Eidgahee, 2020).
Linear Regression (LR) assumes that the input variables have a straight-line relationship with

compressive strength. Although it is easy to understand and apply, LR falls short in representing the nonlinear behavior seen in recycled aggregate concrete, where factors like replacement ratio, water absorption, and the strength of the original concrete interact in more complex ways. In this study, LR achieved an R^2 of 0.893 with an MAE of 0.346, indicating moderate predictive performance but limitations in handling complex nonlinear patterns (Bhanja & Sengupta, 2005).
Figure 7 presents linear regression model strong capability in predicting the 28-day compressive strength of recycled aggregate concrete. It achieves an R^2 value of 0.893, along with low mean absolute error and mean bias error, demonstrating a close match between the actual and predicted strength values. Panel (a) shows the comparison between actual and predicted values, with most data points falling within the ±20 percent deviation limits. Panel (b) displays the three-dimensional prediction error surface, which offers an overview of the distribution of residuals across the dataset and assists in identifying potential patterns or deviations from model assumptions.

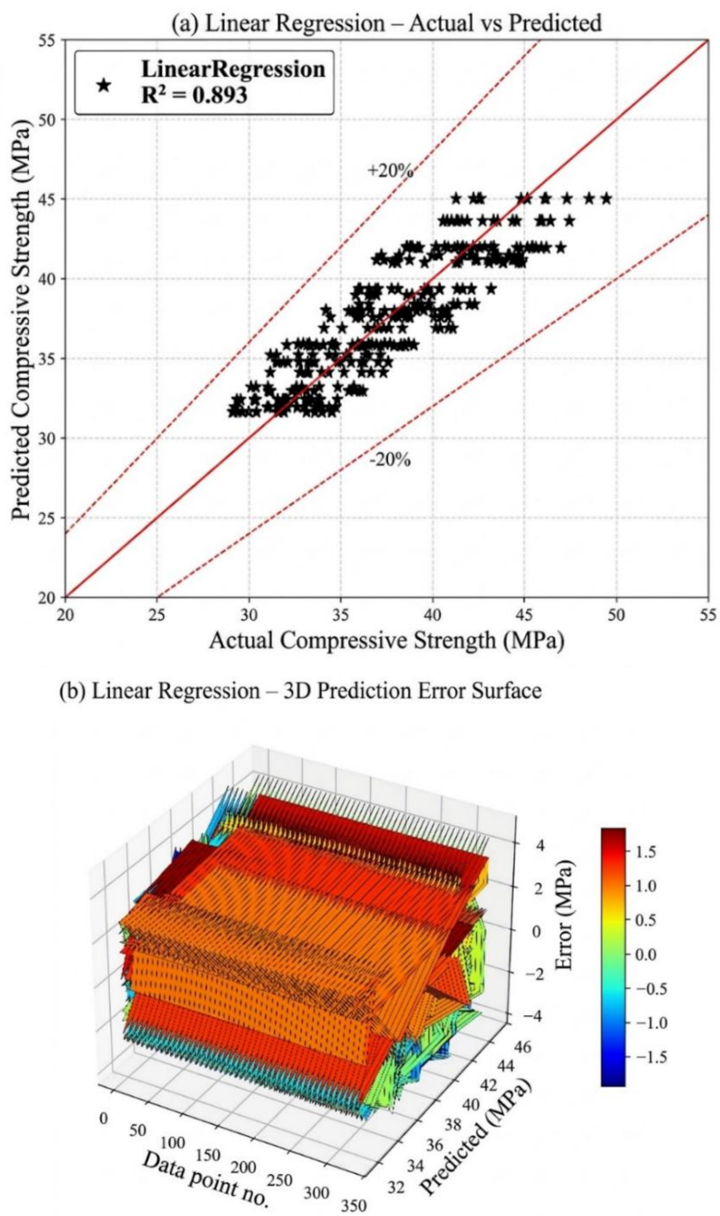


Figure 7. Linear Regression model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

Ridge Regression extends LR by incorporating L2 regularization, which penalizes large coefficients to reduce overfitting and improve generalization. Ridge achieved an R^2 of 0.842 and an RMSE of 0.575. Although slightly better at handling multicollinearity than LR, Ridge still struggled to model the nonlinear dependencies between the recycled aggregate properties and compressive strength (Hoerl & Kennard, 1970).
Figure 8 shows how the ridge regression model performed in forecasting the 28-day compressive strength of recycled aggregate concrete. The model attains a coefficient of determination of 0.842, indicating a dependable level of predictive accuracy. Panel (a) compares the predicted and measured strengths, with most data points positioned within the ±20 percent deviation band. Panel (b) illustrates the three-dimensional error surface, providing a clear view of the residual distribution across the dataset and demonstrating the consistent behaviour of the model under varying input conditions.

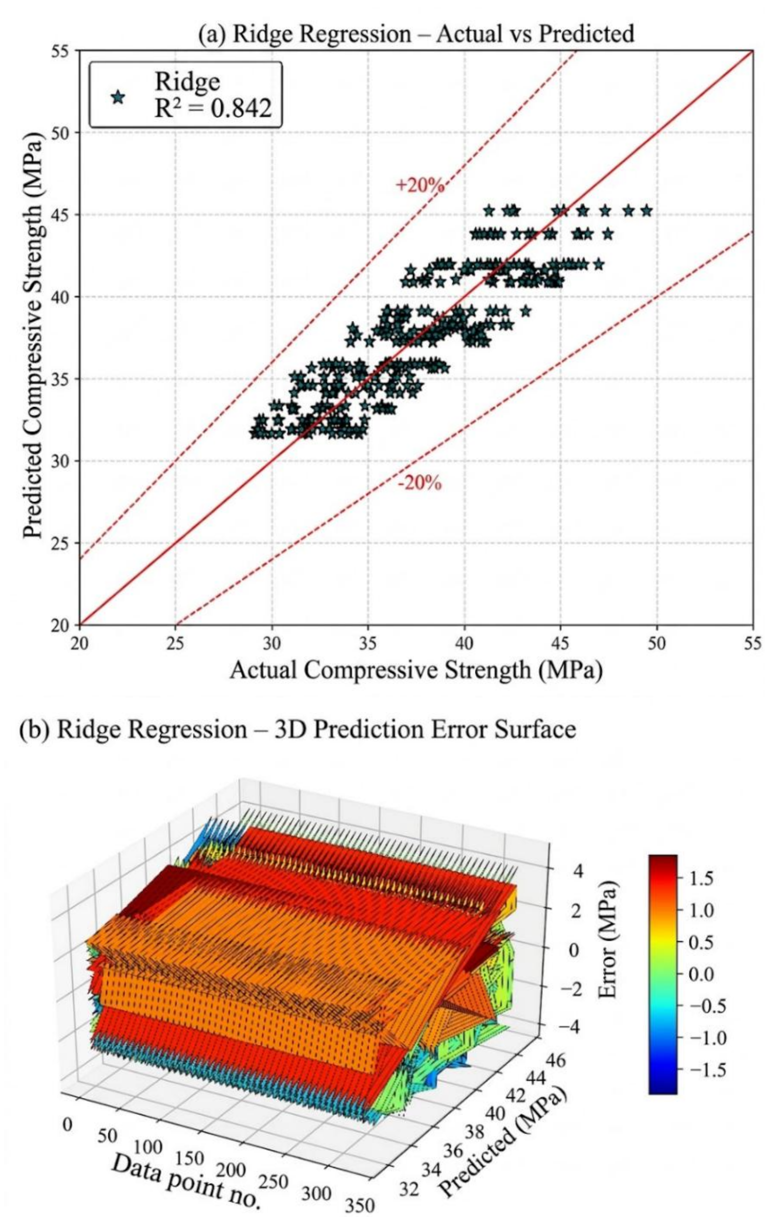


Figure 8. Ridge Regression model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

Lasso Regression uses L1 regularization to enforce sparsity, effectively performing feature selection. Lasso produced an R^2 of 0.879 and a lower MAE of 0.364, highlighting its ability to emphasize the most influential variables, such as water-cement ratio, RCA content, and LA abrasion. However, the linear framework limits its capacity to fully capture complex interactions among mix parameters (Hastie et al., 2009).

Figure 9 presents the prediction results of the Lasso regression model for evaluating the 28-day compressive strength of recycled aggregate concrete. Panel (a) shows the comparison between measured and predicted values, with most data points falling within ±20 percent of the actual strengths, indicating strong agreement. Panel (b) presents the three-dimensional error surface, highlighting the distribution of residuals and confirming minimal systematic bias in the model predictions. The model achieves a coefficient of determination of 0.879, along with low error metrics (MAE = 0.364 MPa, MSE = 0.218 MPa^2, RMSE = 0.467 MPa, MAPE = 0.987%, MBE = 0.018), demonstrating its reliability for accurate compressive strength estimation.

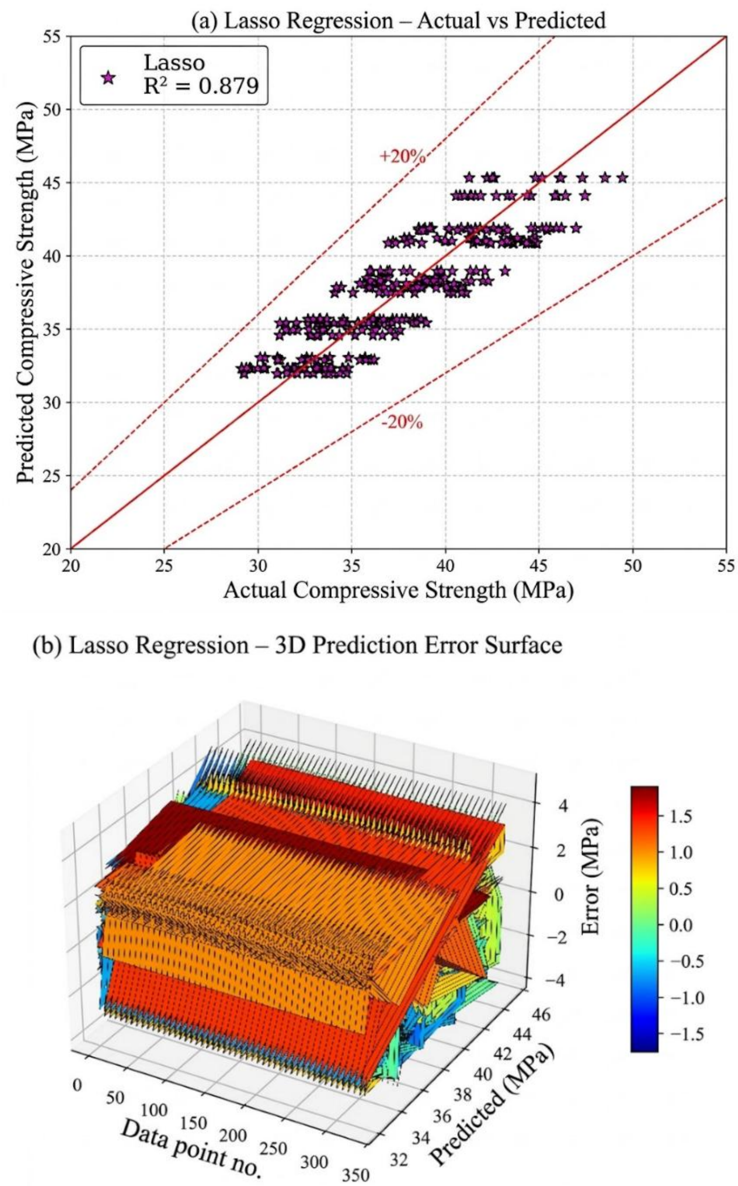


Figure 9. Lasso regression model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

Elastic Net (EN) combines L1 and L2 penalties, balancing feature selection and coefficient shrinkage. EN achieved an R^2 of 0.846, offering slightly improved generalization compared to Ridge but underperforming relative to ensemble and nonlinear approaches. Its performance suggests that while linear regularization can improve stability, nonlinear interactions still require more sophisticated modeling (Zou & Hastie, 2005).

Figure 10 depicts how well the elastic net regression model predicts the 28-day compressive strength of recycled aggregate concrete. Panel (a) shows the relationship between the measured and predicted strengths, with most data points lying within the ±20 percent deviation range, indicating a strong level of agreement. Panel (b) depicts the three-dimensional error surface, providing a detailed view of the residual distribution across the test samples. The model yields a coefficient of determination of 0.846 and a low mean bias error, demonstrating substantial predictive accuracy and minimal systematic deviation. These results support the suitability of the elastic net approach for datasets involving multiple correlated input variables.

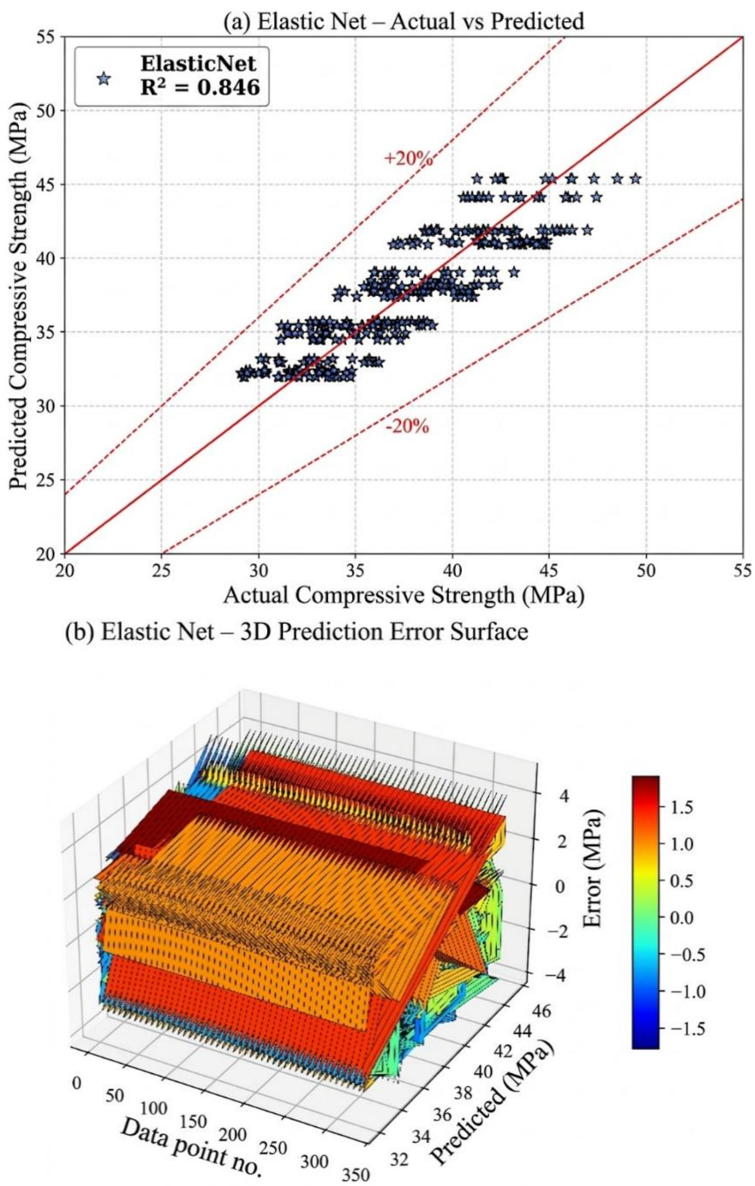


Figure 10. Elastic Net model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

Support Vector Regression (SVR) employs a kernel-based approach to capture nonlinear relationships between input variables and concrete compressive strength. With an R^2 value of 0.998, MAE of 0.008, and RMSE of 0.011, SVR outperformed all linear models, accurately reflecting the effects of RCA replacement, water absorption, and parent concrete strength. Hyperparameter tuning of the kernel type, penalty parameter (C), and epsilon margin was critical to optimizing SVR performance (Smola & Scholkopf, 2004).

Figure 11 shows how well the support vector regression model predicts the 28-day compressive strength of recycled aggregate concrete. Panel (a) shows a close correspondence between measured and predicted values, with the data points concentrated within the ±20 percent deviation limits, indicating a high level of predictive accuracy. Panel (b) displays the three-dimensional error surface, which illustrates the uniform and minimal residuals across the dataset. The model attains a coefficient of determination of 0.998 and exhibits negligible mean bias error, together with very low mean absolute and root mean squared errors, confirming its ability to reliably capture the variation in compressive strength.

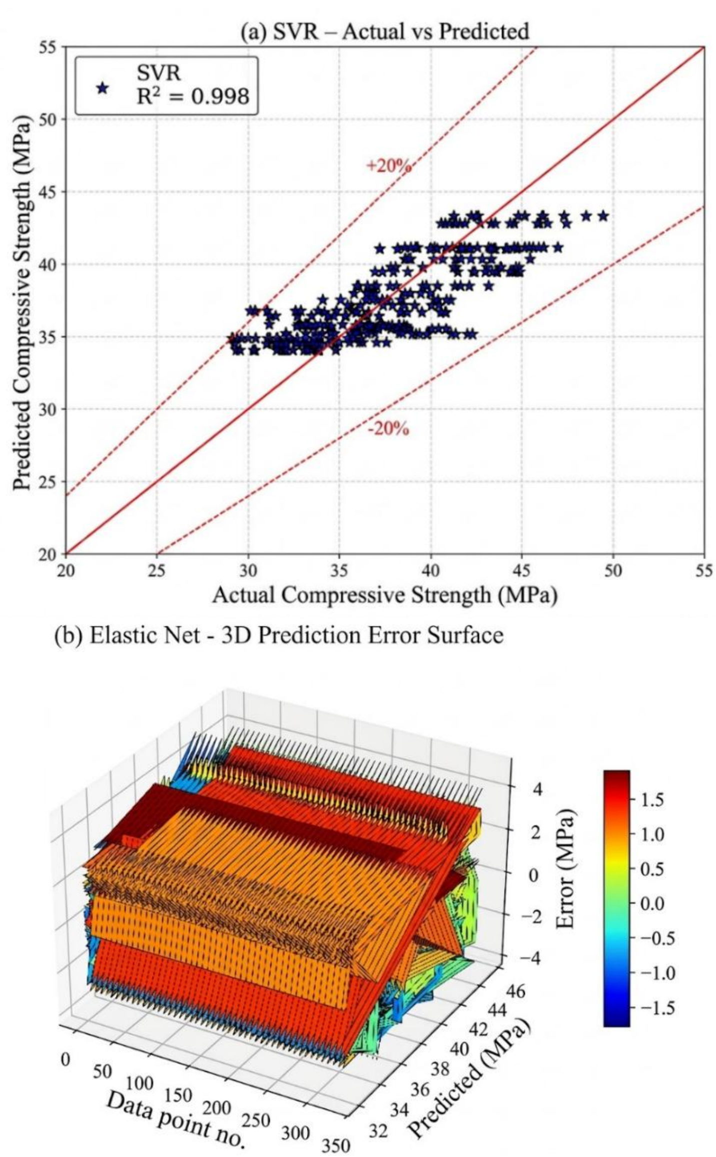


Figure 11. Support Vector Regression model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

K-Nearest Neighbors (KNN) is a nonparametric approach that predicts target values based on the similarity to nearby data points. KNN achieved an R^2 of 0.947, demonstrating strong performance but slightly lower than SVR. The model effectively captured local patterns in the dataset, particularly variations in RCA percentages and water-cement ratios, though it is sensitive to outliers and high-dimensional spaces (Altman, 1992).

Figure 12 shows how the k-nearest neighbors model performed in predicting the 28-day compressive strength of recycled aggregate concrete. Panel (a) shows the relationship between measured and predicted values, with most data points situated within the ±20 percent deviation limits, indicating reliable agreement. Panel (b) presents the three-dimensional error surface, providing an overview of the residual distribution across the dataset and demonstrating the model's robust predictive behavior with minimal deviation.

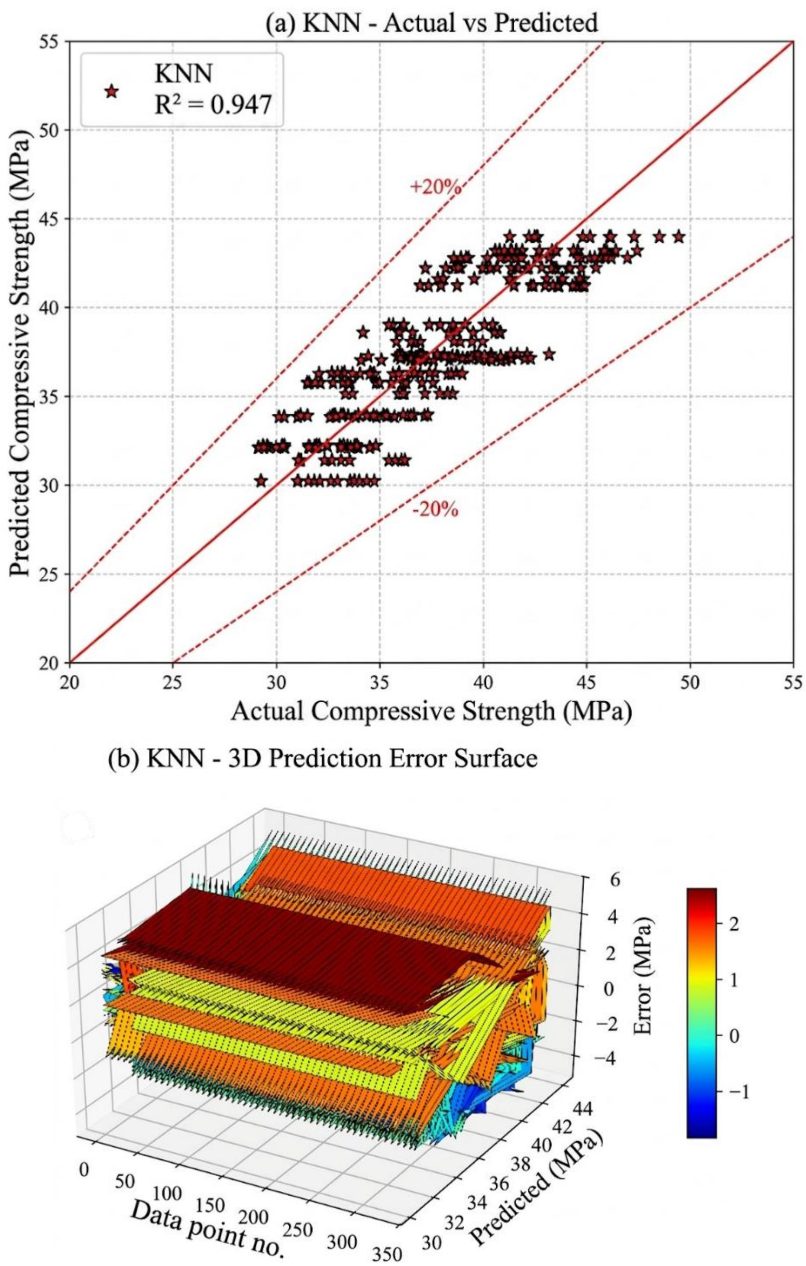


Figure 12. K-Nearest Neighbors model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

Random Forest (RF) is an ensemble tree-based method that averages several decision trees to reduce variance and improve generalization. RF achieved an R^2 of 0.996 with low MAE and MSE, demonstrating robustness in handling multicollinearity and nonlinear interactions among mix parameters (Breiman, 2001).

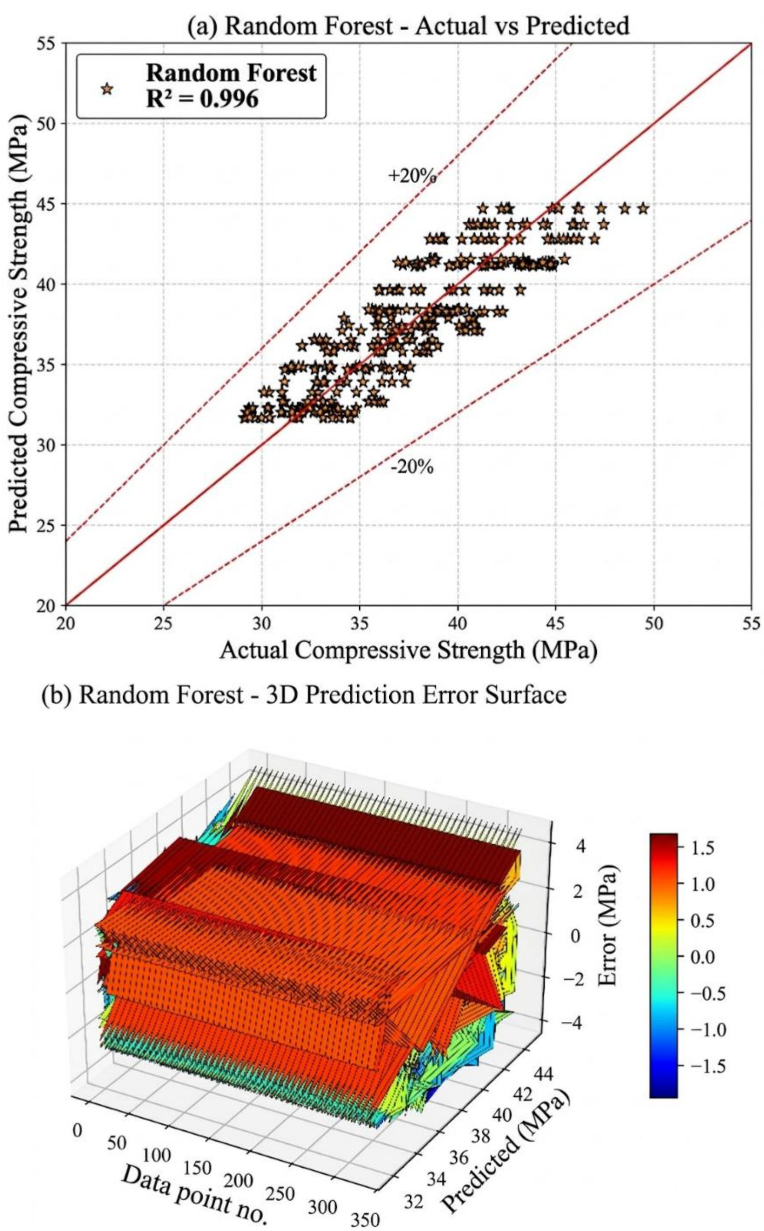


Figure 13. Random Forest model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

The predictive capability of the random forest model for forecasting the 28-day compressive strength of recycled aggregate concrete is illustrated in Figure 13. Panel (a) shows the comparison between measured and predicted values, with most points lying within the ±20 percent deviation range, indicating strong predictive accuracy. Panel (b) displays the three-dimensional error surface, highlighting consistently low residuals across the dataset. The model achieves a coefficient of determination of 0.996 and low error metrics, confirming its reliability and suitability for robust prediction of concrete compressive strength.

Gradient Boosting (GB) sequentially builds trees on residual errors of prior trees to minimize bias iteratively. GB yielded an R^2 of 0.904, showing good predictive capability, but performance was slightly lower than RF due to sensitivity to hyperparameter settings such as learning rate, tree depth, and number of estimators (Friedman, 2001).

Figure 14 shows how well the gradient boosting model predicts the 28-day compressive strength of recycled aggregate concrete. Panel (a) shows the comparison between measured and predicted values, with most points positioned within the ±20 percent deviation range, indicating strong agreement. Panel (b) depicts the three-dimensional error surface, highlighting the magnitude and distribution of residuals, which are predominantly low and uniformly spread across the test dataset. These observations confirm the model's reliability and accuracy in predicting concrete compressive strength.

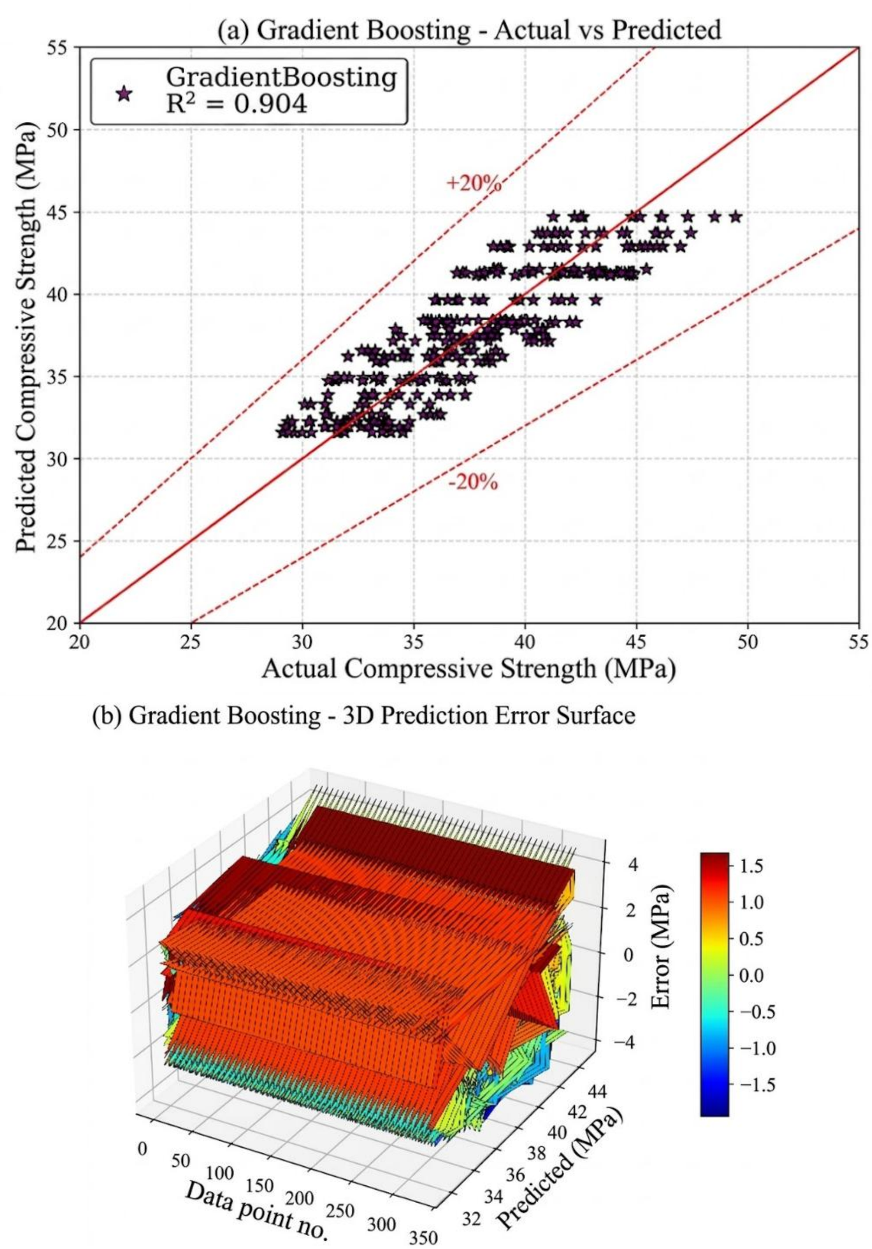


Figure 14. Gradient Boosting model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

XGBoost, a highly efficient gradient boosting implementation, obtained an R^2 of 0.942, MAE of 0.186, and RMSE of 0.248. XGBoost effectively handled nonlinear dependencies and high-dimensional input features such as RCA water absorption, LA abrasion, and cement content while maintaining computational efficiency (Chen & Guestrin, 2016).
Figure 15 shows how the XGBoost model performed in estimating the 28-day compressive strength of recycled aggregate concrete. In panel (a), predicted values closely follow the measured strengths, with most points falling within ±20 percent of the ideal line, indicating high predictive accuracy. Panel (b) presents the three-dimensional error surface, showing the distribution of residuals across the dataset and highlighting the generally small deviations. The model achieves a high coefficient of determination (R^2 = 0.942) and low error metrics (MAE = 0.186 MPa, RMSE = 0.248 MPa), confirming its reliability and effectiveness for strength prediction.

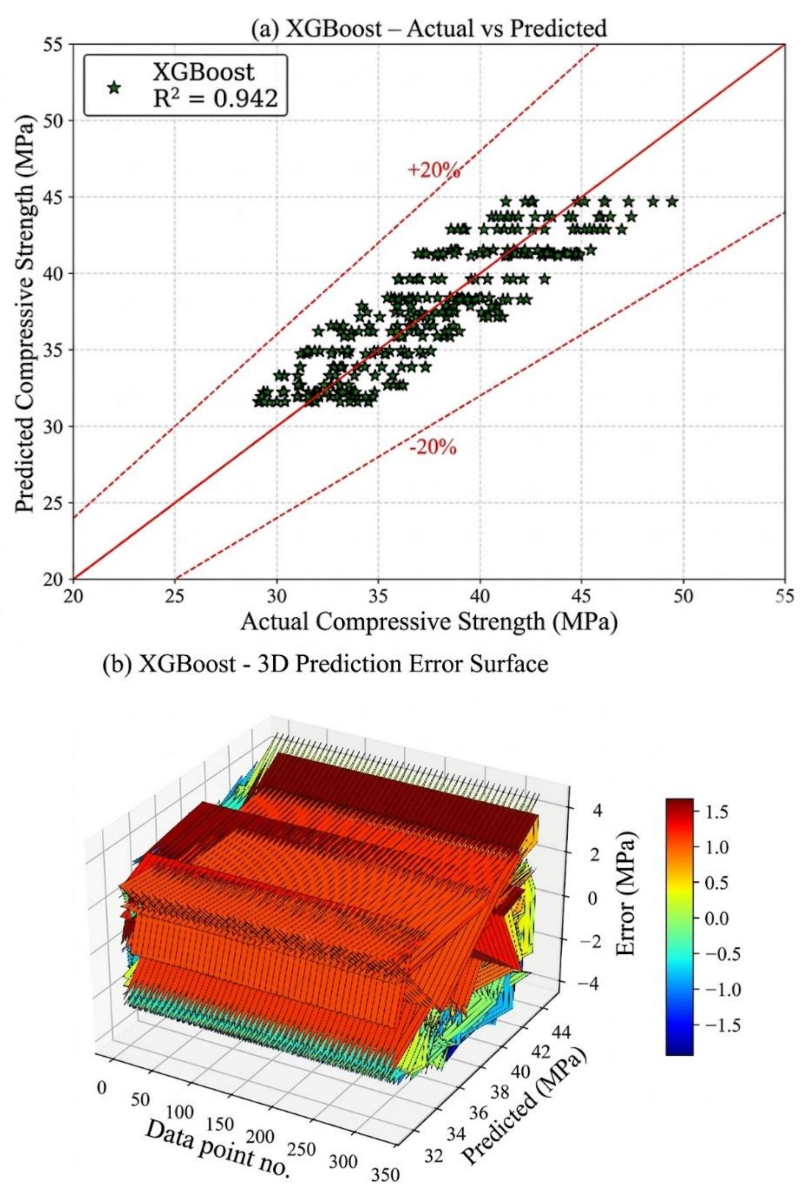


Figure 15. XGBoost model performance for predicting 28-day compressive strength of concrete. (a) Scatter plot of actual versus predicted compressive strength, (b) 3D surface plot depicting the distribution of prediction errors across all data points.

The findings show that nonlinear and ensemble algorithms provide more accurate predictions of RAC compressive strength compared to linear techniques. The strong R^2 scores achieved by SVR and Random Forest demonstrate their reliability for real-world use, whereas linear models remain useful as baseline references or in situations where interpretability is important.

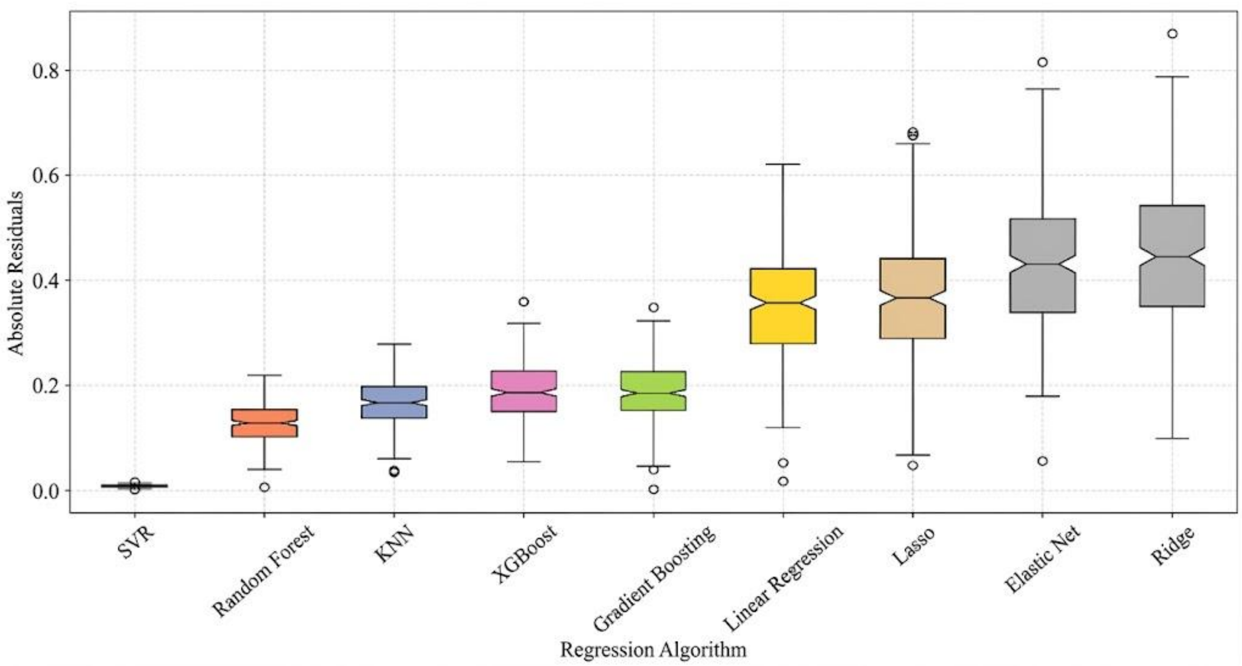


Figure 16. Error-bound analysis showing the distribution of absolute residuals for all regression models.

Figure 16 shows the error-bound evaluation of each regression model based on the spread of absolute residual values. The SVR model shows the narrowest residual spread, which aligns with its low MAE of 0.008 and confirms its strong consistency in predicting compressive strength. Random Forest (MAE 0.128), KNN (MAE 0.164) and XGBoost (MAE 0.186) also display compact error distributions, indicating stable performance within restricted error limits. In contrast, Gradient Boosting (MAE 0.189), Linear Regression (MAE 0.346), Lasso (MAE 0.364), Elastic Net (MAE 0.425) and Ridge (MAE 0.450) show wider error ranges, suggesting comparatively weaker control over prediction deviations. Models with lower MAE show more compact error ranges, while linear and regularized methods generally display wider spread in their residuals.

4.2 Regression Diagnostics and Model Evaluation

A thorough hyperparameter optimization procedure was carried out for each model to improve the overall prediction performance. Hyperparameter tuning involved systematically adjusting key model parameters such as kernel type, C, and epsilon for Support Vector Regression (SVR); the number of estimators, maximum depth, and minimum samples split for Random Forest (RF); the number of neighbors for K-Nearest Neighbors (KNN); learning rate and maximum depth for XGBoost and Gradient Boosting; and regularization parameters for Lasso, Ridge, and Elastic Net (Farhangi et al., 2021; Naderpour et al., 2018). Grid search and randomized search methods were employed in conjunction with k-fold cross-validation (k=10) to ensure robust parameter selection and prevent overfitting (Deng et al., 2018; Jahangir & Eidgahee, 2020). Cross-validation allowed the models to be trained and validated across multiple subsets of the dataset, providing an unbiased estimate of generalization performance (Ghanizadeh et al., 2019).

Model accuracy was assessed using multiple indicators, such as the coefficient of determination (R^2), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), mean absolute percentage error (MAPE), mean bias error (MBE), and the t-statistic (Bilim, Koksal, et al., 2009; Farhangi et al., 2021). Among the individual models, SVR achieved the highest predictive accuracy with an R^2 value of 0.998, minimal MAE (0.008), and RMSE of 0.011, closely followed by Random Forest (R^2=0.996). Nonlinear models such as KNN, XGBoost, and Gradient Boosting outperformed linear models (Linear Regression, Lasso, Ridge, Elastic Net), indicating the strong nonlinear relationships between the mix parameters and compressive strength (Deng et al., 2018; Naderpour et al., 2018). KNN and XGBoost recorded R^2 values of 0.947 and 0.942, respectively, while Gradient Boosting achieved 0.904. Linear Regression and regularized linear models displayed comparatively lower R^2 values ranging from 0.842 to 0.893, highlighting their limitations in capturing complex interactions (Bilim, Ozbakkaloglu, et al., 2009; Jahangir & Eidgahee, 2020).
To enhance the accuracy of the predictions, a stacking ensemble model was created, integrating the capabilities of the best-performing algorithms (J. Zhang et al., 2022). The ensemble method combines several base models along with a higher-level model that learns from their outputs to improve overall prediction accuracy. The Stacking model demonstrated superior performance with R^2 values of 0.999 for the training set and 0.998 for the testing set, accompanied by negligible errors (MAE ≈ 0.0099, RMSE ≈ 0.0137), Table 6. These results confirm that ensemble learning can effectively exploit complementary capabilities of individual models, achieving near-perfect prediction of 28-day compressive strength across the diverse experimental dataset (Farhangi et al., 2021; Jahangir & Eidgahee, 2020). This method highlights how modern machine learning models can deliver accurate, data-based predictions of recycled aggregate concrete performance, helping minimize the need for large amounts of laboratory testing while still ensuring reliable results (Ghanizadeh et al., 2019; Naderpour et al., 2018).

Table 6. Comparative performance of regression models for predicting 28-day compressive strength of recycled aggregate concrete.

Model	R^2 Value	MAE	MSE	RMSE	MAPE (%)	MBE	T-stat
Support Vector Regression	0.998	0.008	0	0.011	0.022	0	-0.309
Random Forest	0.996	0.128	0.032	0.179	0.346	0.017	1.373
K-Nearest Neighbors	0.947	0.164	0.052	0.227	0.441	0.029	1.913
XGBoost	0.942	0.186	0.061	0.248	0.494	-0.01	-0.282
Gradient Boosting	0.904	0.189	0.063	0.251	0.505	0.006	0.353
Linear Regression	0.893	0.346	0.209	0.457	0.94	0.009	0.284
Lasso Regression	0.879	0.364	0.218	0.467	0.987	0.018	0.556
Elastic Net	0.846	0.425	0.291	0.539	1.148	0.024	0.646
Ridge Regression	0.842	0.45	0.331	0.575	1.209	0.026	0.672

Note: *Values represent the mean of five-fold cross-validation averaged over three random seeds. Hyperparameters were optimized using 10-fold cross-validation with grid and randomized search to ensure robust model selection. Performance metrics include R^2, MAE, MSE, RMSE, MAPE, MBE, and paired t-statistic, highlighting the superior accuracy of nonlinear and ensemble models.*

Although very high R^2 values were obtained for certain models, these results should be interpreted in the context of the controlled dataset and the strong correlation between mix design parameters and compressive strength. Cross-validation and error metrics were used to assess model performance; however, the potential for overfitting due to dataset expansion is acknowledged as a limitation of the study.

4.3 Model Interpretability and Feature Importance

To understand the effect of each mix parameter on the Support Vector Regression model's prediction of 28-day compressive strength for recycled aggregate concrete, SHapley Additive exPlanations (SHAP) analysis was performed. SHAP provides a transparent interpretation of machine learning results by attributing a contribution value to each feature based on its influence on the final prediction. This approach enables a detailed understanding of nonlinear and interacting effects that are not visible through traditional sensitivity analysis (Lundberg & Lee, 2017; Molnar, 2020).

Figure 17 shows how each input feature contributes to the model output based on SHAP analysis. The average SHAP values for the SVR model are shown in Figure 17(a). Among all variables, the percentage of RCA had the greatest impact, with a mean SHAP value of 0.638, demonstrating that the amount of recycled coarse aggregate played the most significant role in predicting compressive strength. RCA water absorption (0.584) and LA abrasion value (0.459) were also highly influential, reflecting the critical role of aggregate surface condition and mechanical integrity in determining concrete performance. Parent concrete strength (0.351), plasticizer content (0.346), and the water–cement ratio (0.168) followed, showing moderate but meaningful influence on strength development.

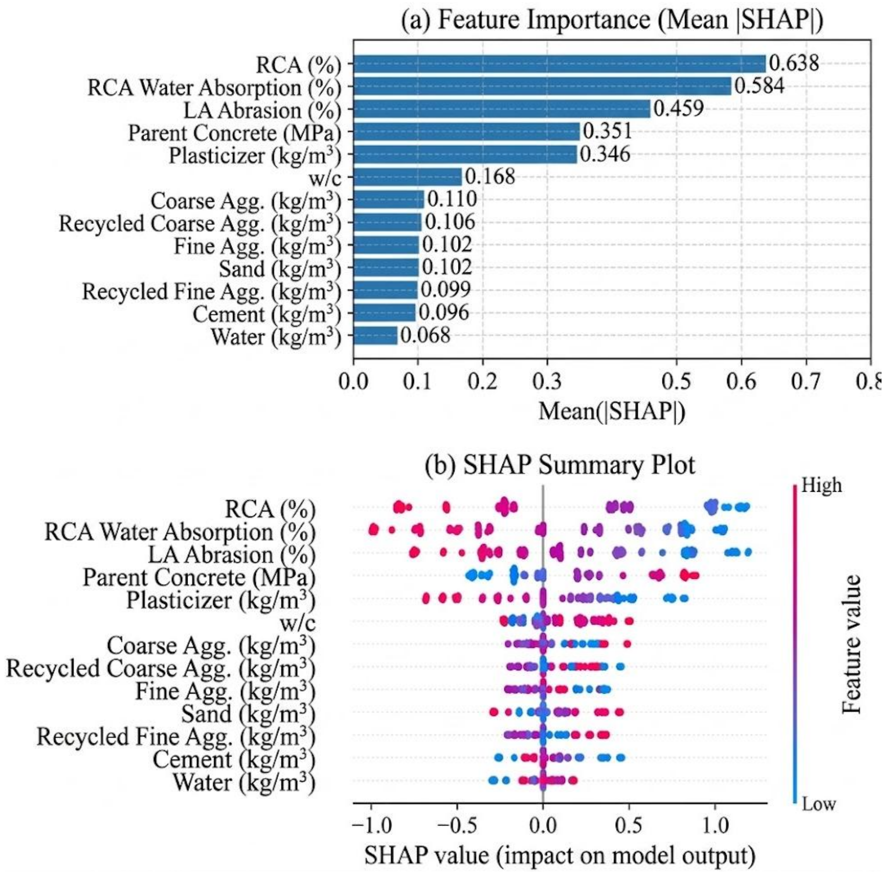


Figure 17. SHAP analysis, (a) Mean absolute SHAP values indicating the relative influence of input variables on the SVR model predictions of 28-day compressive strength, (b) SHAP summary plot showing the distribution, direction, and magnitude of feature effects across all samples.

Other constituents such as total coarse aggregate (0.110), recycled coarse aggregate (0.106), fine aggregate (0.102), sand (0.102), and recycled fine aggregate (0.099) showed smaller contributions but were consistent with their role in controlling workability, internal packing density, and the quality of the interfacial transition zone. Cement (0.096) and water content (0.068) had the lowest SHAP values, suggesting limited variability in their ranges within the dataset rather than reduced engineering influence.

Figure 17(b) displays how SHAP values are distributed for each sample. The larger horizontal spread observed for RCA percentage, absorption, and abrasion shows that these inputs contribute the most to prediction variability within the dataset. In contrast, cement and water exhibit much tighter spreads, indicating that the model is less sensitive to changes in these features. Red and blue colour gradients show how high or low feature values shift predictions positively or negatively, helping visualize the direction and influence of each input on model performance.

Together, Figures 17(a) and 17(b) demonstrate that the properties of recycled aggregates govern strength prediction in RAC more strongly than mix proportions alone. These results provide data-driven guidance for RAC optimization by highlighting that improving aggregate quality, controlling absorption, and selecting durable source concrete can significantly improve model-predicted strength outcomes.

The high importance of water absorption and RCA related parameters observed in the SHAP analysis further confirms that aggregate quality and porosity related characteristics play a dominant role in strength prediction.

These insights allow a data-driven approach to mix optimization: increasing binder efficiency and controlling RCA quality can maximize strength, while minor adjustments to water content, sand, and plasticizer improve workability and structural integrity without compromising compressive strength. SHAP values thus provide both interpretability and actionable guidance for the design of sustainable RAC mixes, reducing experimental trials while maintaining high confidence in model predictions (Lundberg & Lee, 2017; Molnar, 2020).

5. CONCLUSION

This work explores the compressive strength of recycled aggregate concrete (RAC) using both laboratory testing and machine learning techniques. A total of 25 concrete mixes were prepared with different combinations of natural and recycled coarse and fine aggregates, varying water–cement ratios, and different strengths of the original concrete. The 28-day compressive strength results ranged from 31.8 MPa for mixes containing only recycled aggregates to 45.2 MPa for mixes made entirely with natural aggregates. Machine learning models were then developed to predict compressive strength from the mix parameters. Support Vector Regression (SVR) delivered the most accurate results, achieving an R^2 value of 0.998 with very small prediction errors (MAE = 0.008 MPa, RMSE = 0.011 MPa). Ensemble methods like Random Forest and KNN also performed well, while traditional approaches such as Linear Regression, Lasso, and Ridge recorded lower accuracy by comparison.

Key observations from the study include:

- Increasing recycled aggregate content generally reduces compressive strength. Mixes with 75-100% RCA exhibited the lowest strength values, while mixes with no RCA achieved the highest.
- Water-cement ratio, parent concrete strength, and RCA water absorption were among the most influential factors affecting compressive strength. Higher w/c ratios in combination with high RCA levels caused significant reductions in strength.

- Machine learning models demonstrated a strong ability to predict compressive strength from mix parameters. SVR and RF models provided very high predictive accuracy within the studied parameter range, making them highly suitable for mix design optimization.
- Ensemble and other non-linear models delivered better results than traditional linear regression, showing that accounting for complex relationships among the input variables is important when predicting the behavior of recycled aggregate concrete.

The results indicate that replacing natural aggregates with recycled aggregates in the range of 25% to 75% is feasible for concrete production when the mix is well designed. Although some reduction in strength is observed, it remains within acceptable limits, and the resulting concrete maintains structural performance. In contrast, complete replacement (100%) consistently leads to the lowest strength, largely due to the higher water absorption of recycled aggregates and the reduced quality of the original concrete source. Integrating laboratory testing with machine learning enhances accuracy in predicting concrete performance and minimizes the need for repeated physical trials, making it easier to determine the most appropriate recycled aggregate content for sustainable concrete applications.

Future studies should consider incorporating additional input variables such as chemical characteristics, durability indicators, curing variables, and long-term behavior to strengthen model reliability and predictive depth. These models can also be expanded to forecast other mechanical and durability parameters, aiding in the development of environmentally conscious and high-performance concrete materials.

6. REFERENCES

Ajdukiewicz, A., & Kliszczewicz, A. (2002). Influence of recycled aggregates on mechanical properties of high-strength concrete. *Cement and Concrete Composites*, *24*(2), 269–279. https://doi.org/10.1016/S0958-9465(01)00012-9

Akbar, A., & Liew, J. Y. R. (2020a). Environmental impact of cement production and concrete sustainability. *Journal of Cleaner Production*, *258*, 120543. https://doi.org/10.1016/j.jclepro.2020.120543

Akbar, A., & Liew, K. M. (2020b). Assessing recycling potential of carbon fiber reinforced plastic waste in eco-efficient cement materials. *Journal of Cleaner Production*, *274*, 123001. https://doi.org/10.1016/j.jclepro.2020.123001

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*, *46*(3), 175–185. https://doi.org/10.1080/00031305.1992.10475879

Bai, X., & others. (2023). Experimental study on mechanical properties and compressive constitutive model of recycled concrete under sulfate attack considering the effects of multiple factors. *Buildings*, *14*(9), 2761. https://doi.org/10.3390/buildings14092761

Bhanja, S., & Sengupta, B. (2005). Prediction of compressive strength of concrete using regression analysis. *Cement and Concrete Research*, *35*(9), 1848–1855. https://doi.org/10.1016/j.cemconres.2005.03.003

Bilim, C., Atis, C. D., Tanyildizi, H., & Karahan, O. (2009). Predicting compressive strength of GGBS concrete using ANN. *Advances in Engineering Software*, *40*(5), 334–340. https://doi.org/10.1016/j.advengsoft.2008.05.005

Bilim, C., Koksal, F., & Tosun, T. (2009). Prediction of concrete strength using ANN. *Expert Systems with Applications*, *36*(2), 3736–3742. https://doi.org/10.1016/j.eswa.2008.03.011

Bilim, C., Ozbakkaloglu, T., & Tanyu, B. F. (2009). Artificial neural networks for predicting compressive strength of concrete. *ACI Materials Journal*, *106*(3), 203–211.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Cakiroglu, M., Pal, S., & Sharma, R. (2023). Environmental challenges of construction waste: A review. *Journal of Environmental Management*, *335*, 117142. https://doi.org/10.1016/j.jenvman.2023.117142
Chakradhara Rao, M., Bhattacharyya, S. K., & Barai, S. V. (2011). Influence of field recycled coarse aggregate on properties of concrete. *Materials and Structures*, *44*(1), 205–220. https://doi.org/10.1617/s11527-010-9620-x
Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *KDD Proceedings*, 785–794. https://doi.org/10.1145/2939672.2939785
Chicco, D., & Warrens, M. J. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/10.7717/peerj-cs.623
Dantas, A. T., Leite, M., & Nagahama, K. (2013). ANN-based strength prediction of concrete with waste. *Construction and Building Materials*, *38*, 717–722.
Deng, A., Naderpour, H., & Ghanizadeh, A. (2018). Predicting strength of recycled concrete using ML. *Materials Today: Proceedings*, *5*(11), 24091–24099. https://doi.org/10.1016/j.matpr.2018.10.091
Emmert-Streib, F., & Dehmer, M. (2019). Evaluation of regression models: Model assessment, model selection and generalization error. *Machine Learning and Knowledge Extraction*, *1*(1), 521–551. https://doi.org/10.3390/make1010032
Farhangi, V., Jahangir, H., & Ghafari, E. (2021). ML-based prediction of recycled aggregate concrete. *Materials Today Communications, 28,* 102633. https://doi.org/10.1016/j.mtcomm.2021.102633
Field, A. (2013). *Discovering statistics using IBM SPSS* (4th ed.). Sage.
Friedman, J. (2001). Greedy function approximation: Gradient boosting. *Annals of Statistics*, *29*(5), 1189–1232.
Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*, 3–42.
Ghanizadeh, A., Naderpour, H., & Deng, Q. (2019). AI-based strength prediction for recycled concrete. *Journal of Cleaner Production*, *215*, 376–385.
Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
Hansen, T. C., & Narud, H. (1983). Strength of recycled concrete with crushed aggregates. *Concrete International*, *5*(1), 79–83.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Elements of statistical learning*. Springer.
Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634
Huang, B., & Yan, Q. (2025). Evaluating the mechanical properties of recycled aggregate concrete with variable coarse and fine aggregate replacements. *Discover Civil Engineering*. https://doi.org/10.1007/s44290-025-00336-3
International, A. (2024). *ASTM C150/C150M-24: Standard specification for Portland cement*.
Jahangir, H., & Eidgahee, D. (2020). Hybrid AI modeling for FRP–concrete bond strength. *Composite Structures*, *257*, 113160.
Jahangir, H., Farhangi, H., & Ghodrati, B. (2021). Ensemble prediction of RAC strength. *Construction and Building Materials*, *271*, 121654.
Katz, A. (2003). Concrete made from partially hydrated recycled aggregate. *Cement and Concrete Research*, *33*(5), 703–711.
Khalil, E., & AbouZeid, M. (2025). Framework for cement plants assessment through cement production improvement measures for reduction of CO_2 emissions towards net zero emissions. *Construction Materials*, *5*(2), 20. https://doi.org/10.3390/constrmater5020020

Khan, K., & others. (2022). Comparison of machine learning methods for recycled aggregate concrete strength prediction. *Materials*, *15*(10), 3430. https://doi.org/10.3390/ma15103430
Khatib, J. M. (2005). Concrete with fine recycled aggregates. *Cement and Concrete Research*, *35*(4), 763–769.
Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*, 4765–4774.
Miller, C., & others. (2024). A review of model evaluation metrics for machine learning in regression, classification, and clustering. *PeerJ Computer Science*, *10*, eXXXX.
Molnar, C. (2020). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Lulu.com.
Naderpour, H., Ghanizadeh, M., & Behnood, A. (2018). RAC compressive strength prediction via ANN. *Journal of Cleaner Production*, *184*, 640–650.
Padmini, A. K., Ramamurthy, K., & Mathew, M. (2002). Influence of parent concrete on RAC properties. *Cement and Concrete Research*, *32*(12), 1893–1901.
Pal, R., & Singh, S. (2024). Role of impact and compression-based crushing on the physical, chemical, and morphological characteristics of recycled concrete aggregates. *Journal of Materials in Civil Engineering*. https://doi.org/10.1061/JMCEE7.MTENG-17400
Plevris, V., Solórzano, G., Bakas, N. P., & Ben Seghier, M. E. (2022). Investigation of performance metrics in regression analysis and machine learning based prediction models. *Proceedings of the 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2022)*. https://doi.org/10.23967/eccomas.2022.155
Ryu, G. S. (2002). Strength and durability of recycled aggregate concrete. *Construction and Building Materials*, *16*(1), 95–101.
Shin, Y. S., & Kim, G. H. (2013). Predicting Compressive Strength of Recycled Aggregate Concrete by Multiple Regression Analysis. *Applied Mechanics and Materials*, *253–255*, 546–549. https://doi.org/10.4028/www.scientific.net/AMM.253-255.546
Smola, A., & Scholkopf, B. (2004). Tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222.
Steurer, M., Schulz, W., & Werwatz, A. (2021). Metrics for evaluating the performance of machine learning prediction models. *International Journal of Forecasting*, *36*(6), 812–828. https://doi.org/10.1080/09599916.2020.1858937
Tu, C., Belarbi, R., Pigeon, M., & Tam, V. (2006). Properties of concrete with recycled aggregate. *Cement and Concrete Research*, *36*(1), 153–159.
Wolpert, D. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259.
Yeh, I.-C. (2007). ML prediction of concrete properties. *Cement and Concrete Research*, *37*(5), 676–685.
Zain, M. F. M., Safiuddin, M., & Mahmud, H. B. (2002). Development of high-performance concrete using silica fume and superplasticizer. *Cement and Concrete Research*, *32*(1), 115–122. https://doi.org/10.1016/S0008-8846(01)00643-1
Zhang, J., Li, X., & Chen, Y. (2022). Stacking ensemble for concrete strength prediction. *Engineering Applications of Artificial Intelligence*, *109*, 104617.
Zhang, P., Sun, X., Wang, F., & Wang, J. (2023). Mechanical properties and durability of geopolymer recycled aggregate concrete: A review. *Polymers*, *15*(3), 615. https://doi.org/10.3390/polym15030615
Zhang, X., Sun, X., Wang, F., & Wang, J. (2023). Predicting compressive strength of recycled aggregate concrete using machine learning and hyperparameter optimization techniques. *Frontiers in Earth Science*, *11*, 1112105. https://doi.org/10.3389/feart.2023.1112105

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x