Computers and Concrete, *Vol. 38, No. 4 (2026) 779-818*
https://doi.org/10.12989/cac.2026.38.4.08

Techno-Press

Machine learning-based prediction of shear capacity in corroded reinforced concrete beams

Hasan Cem Akkaya*[1], Sema Alacalı[2]

[1]*Department of Civil Engineering, Kirklareli University, 39000, Kırklareli, Türkiye*
[2]*Department of Civil Engineering, Yildiz Technical University, 34220, Istanbul, Türkiye*

(Received April 29, 2026, Revised June 16, 2026, Accepted September 15, 2026)

Abstract. In this study, the reliability of machine learning (ML) models for predicting the shear capacity of reinforced concrete (RC) beams exposed to corrosion was investigated. A dataset involving 204 experimental results collected from the literature was constructed for this investigation. The prediction performance of ML models was evaluated for the training and testing datasets. The mean value (MV), standard deviation (STD), root mean square error (RMSE), coefficient of determination (R^2), and coefficient of variation (COV) were utilized to compare the prediction performance of ML models. It was concluded this comparison that overfitting and underfitting problems were overcome in the predictive performance of ML models. Furthermore, another conclusion from the ML model's prediction results was that the predictions obtained using ML models were accurate and robust according to the statistical indicators mentioned above. The new equation, which was obtained with the help of the Gene Expression Programming (GEP) model, was proposed. Moreover, Extreme Gradient Boosting (XGBoost) was selected as the best model out of eleven ML models. The MV, STD, RMSE, R^2 and COV values of the XGBoost model for testing dataset are 1.028, 0.116, 13.725, 0.965 and 0.113, respectively. The prediction results of the XGBoost was superior to those of the five analytical models available in the literature. Finally, SHapley Additive Explanations (SHAP) analysis was performed on the XGBoost model.

Keywords: corrosion; extreme gradient boosting (XGBoost); gene expression programming (GEP); machine learning (ML); shapley additive explanations (SHAP)

1. Introduction

Reinforced concrete (RC) beams are required to provide adequate shear capacity in structural engineering. In other words, engineers design RC beams to undergo flexural failure rather than shear failure in structural projects. Because the shear failure occurs so quickly, it hardly gives a chance to take precautions [1, 2]. In addition, the reliability of RC beam design depends on the compatibility between concrete and reinforcement. Corrosion is one of the most important factors that negatively affects the compatibility between concrete and reinforcement [3, 4]. The degree of corrosion increases over time due to environmental factors [5]. As the corrosion degree increases, the cross-sectional area of the reinforcement decreases, and a rust layer forms inside the concrete. Eventually, the concrete cover cracks or even spalls. The reduced compatibility between concrete and reinforcement, the loss in the cross-sectional area of reinforcement, and the formation of

*Corresponding author, Assistant Professor, E-mail: hasancemakkaya@klu.edu.tr

Copyright © 2026 Techno-Press

ISSN: 1598-8198 (Print), 1598-818X (Online)

cracks in the concrete cover diminish the service life and load-bearing capacity of reinforced concrete beams and increase maintenance costs [2-7]. It was seen in the existing experimental studies that the corrosion occurred more rapidly in stirrup bars with a smaller diameter and a thinner concrete cover compared to longitudinal bars [8-11]. Furthermore, the experimental study [12] showed that although the failure mode of the uncorroded control beams was flexure, the failure mode of the corroded RC (CRC) beams with the same properties shifted to shear failure. Even if RC beams are well-designed for flexural failure, the failure mode could potentially change to shear failure due to corrosion in the reinforcement [2, 4, 13, 14]. Therefore, it is necessary to predict the shear capacity of CRC beams with a high degree of accuracy to guarantee the structural integrity during its service life [2, 3, 6, 7].

Various models are available in the literature for predicting the shear capacity of CRC beams. These models include analytical [15-22], finite element [23-25], and empirical models [26, 27]. Analytical models in the literature are derived from traditional RC beam theories (such as limit equilibrium theory [28], the strut and tie model (STM) [17], modified compression field theory [21], equivalent truss theory [29]) and design codes (such as ACI 318-08 [30], GB 50010-2010 [31], and Eurocode 2 [32]). The effect of corrosion is considered by including corrosion parameters or reduction factors in updated analytical models. Higgins et al. [21] proposed a method to predict the shear capacity of CRC beams by modifying several existing analytical models. It was shown that the predicted shear capacities were in good agreement with the experimental results. In addition, Lu et al. [27] investigated the experimental parameters affecting the shear capacity of CRC beams. Their study includes a dataset collected from experimental studies, comprising various design parameters. They also proposed a new analytical model for predicting the shear capacity of CRC beams. When compared with the experimental results, the predictions obtained using their proposed analytical model outperformed nine different analytical models in the literature. In another analytical study [33], it was observed that the proposed shear model provided accurate predictions based on the compiled dataset. It is known that the shear behaviour of beams is in a highly complex manner. Selecting different parameters and accepting simplifying assumptions can limit the accuracy of predictions beyond the specific prepared datasets. However, concerns have been raised regarding their applicability and the potential for unreliable predictions when applied to comprehensive new experimental datasets. [2, 3, 6, 17, 28, 34]. In addition, although experimental methods yield important insights into the behavior of structural elements, conducting experimental studies on corroded elements,which are highly difficult to fabricate, requires specialized facilities, expensive equipment, and a significant amount of time. Moreover, finite element modeling (FEM) is highly challenging in terms of computational effort and model formulation, particularly for simulating the shear behavior of CRC beams. Therefore, ML models have recently been focused on by researchers. These are part of the field of artificial intelligence that has become widespread, and complex relationships between parameters can be easily solved by them.

Recently, researchers have increasingly adopted ML models to address structural engineering problems. In the literature, the Extreme Gradient Boosting (XGBoost) model has been reported to provide highly accurate predictions for the flexural capacity of RC beams [35], RC beams strengthened with fabric-reinforced cementitious matrix (FRCM) composites [36], and steel fiber-reinforced concrete [37]. Similarly, another ML model, CatBoost, was reported to achieve robust performance in predicting the flexural capacity of reinforced ultra-high-performance concrete (UHPC) beams [38]. Wakjira et al. [39] also highlighted the potential of ML models for predicting the flexural capacity of FRP-reinforced concrete (RC) beams. Furthermore, the predictive

performance of ML models has proved to be highly accurate in estimating axial compression [40] and torsional [41] capacities, which exemplify other key research areas in structural engineering. ML models are also often used in fields related to structural health monitoring, earthquake engineering, and structural reliability [42-45]. A review of recent studies [46-59] in the literature reveals an increase in the number of studies utilizing ML models to predict the shear capacity of RC beams. Among these studies, Tang et al. [46] utilized ML models to predict the shear capacity of RC beams strengthened with FRP sheets. In this scope, ML models such as Artificial Neural Networks (ANN), Decision Trees (DT), Extremely Randomized Trees (ERT), K-Nearest Neighbors (KNN), Random Forests (RF), and XGBoost were employed. It was concluded that the use of ML models provided highly accurate predictions for the shear capacity of RC beams. Furthermore, it was reported that the XGBoost model provided the highest prediction accuracy. In another study [53], seven ML models, including ANN, Support Vector Machine (SVM), DT, RF, Adaptive Boosting (AdaBoost), Gradient Boosting Regression Trees (GBRT), and XGBoost, were employed to predict the shear capacity of slender RC beams without stirrups. The results of that study indicated that the GBRT and XGBoost models achieved superior predictive performance compared to the other ML models evaluated. In a similar study, Yehia et al. [56] focused on predicting the shear capacity of RC T-beams. It was understood that the XGBoost model was selected as the best model for shear capacity prediction in this study. Furthermore, the predictive performance of XGBoost was reported to outperform the results obtained using traditional methods and code equations. Another similar study was carried out by Megahed [58] on reinforced concrete deep beams, and the results demonstrated that ML models offer superior predictive accuracy compared to traditional empirical methods and design code equations. Although the use of ML models in the prediction of shear capacity of RC beams is increasing, such studies for the prediction of shear capacity of corroded RC beams remain limited. A ML model using GBRT technique was developed by Fu and Feng [2] to predict the shear capacity of CRC beams. Consequently, the developed ML model demonstrated significantly higher accuracy than the existing equations in the literature. Another study [3] presented a hybrid ML model by combining the gradient boosting (GB) and Hunger Games Search (HGS) algorithms. This study also used the dataset containing 158 experimental samples. The hybrid ML model was observed to have high prediction accuracy. Sensitivity analysis and optimal design solutions were also conducted in this study. Dogan [5] also demonstrated the applicability of two different ML models, XGBoost and AdaBoost, for the prediction of shear capacity of CRC beams on 157 experimental samples. Kumar et al. [7] predicted the shear capacity of corroded RC beams using four different ML models, namely ANN, adaptive neuro-fuzzy inference systems (ANFIS), DT and XGBoost, based on a dataset of 140 experimental samples compiled from the literature. Based on the outcome of the prediction results, it was concluded that XGBoost was the most superior predictive performance ML model. Furthermore, the sensitivity analysis indicated that beam depth, stirrup spacing, and the shear span-to-depth ratio were the most significant parameters governing the shear capacity of the corroded beams.

As evidenced by the literature, only a limited number of machine learning models have been applied to predict the shear capacity of corroded RC beams, and the size of existing experimental datasets remains highly constrained. Therefore, this study aims to conduct a comprehensive investigate on the application various ML models in the shear capacity prediction of CRC beams using an updated dataset. To achieve this, an updated dataset comprising 204 experimental samples is compiled, and the predictive performances of eleven distinct ML models are evaluated. These models are Decision Tree (DT), Random Forest (RF), Extremely Randomized Trees (ET),

Histogram-Based Gradient Boosting (HGB), Adaptive Boosting (AdaBoost), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), Categorical Boosting (CatBoost), Support Vector Regression (SVR), Gaussian Process Regression (GPR) and Gene Expression Programming (GEP). The difference between the GEP model and other ML models is that it gives an equation as output. Numerous studies [60-64] have demonstrated that GEP-derived equations provide better accuracy than the existing empirical formulations and standard design codes in the literature. This study also proposes a novel equation by using GEP model to predict the shear capacity of CRC beams. Moreover, the prediction results of ML models are compared with the prediction results obtained from the existing equations in the literature with the help of statistical indicators. Based on this comprehensive comparison, the ML model exhibiting the highest predictive accuracy was successfully identified. Finally, SHapley Additive Explanations (SHAP) analysis was performed to investigate the significance of the experimental input parameters on the best-performing ML model.

2. Experimental database and methods

2.1 Data collection from literature

To determine the best, perform ML model, an existing database containing 158 experimental samples collected by Lu et al. [27] from studies [11, 18, 26, 28, 29, 65-70] was updated by removing the control samples in the existing database and adding CRC beams obtained from recent experimental studies [4, 16, 25, 71-77] through a literature review. The updated dataset contains a total of 204 experimental samples. Corrosion in the samples had occurred either in only the longitudinal reinforcement, in only the stirrups, or in both. All beams in the updated dataset were rectangular in cross-section, and the experimental setup was designed for three-point or four-point bending tests. The experimental parameters included in the updated dataset in this study are parallel with those included in similar studies [2, 3, 5, 6] in the literature. This study uses a total of 13 parameters. These consist of 12 input parameters and 1 output parameter. The input parameters are the beam width (b (mm)), the beam height (h (mm)), the beam effective depth (d (mm)), the shear span-to-effective depth ratio (a/b), the concrete cylinder compressive strength (f_c (MPa)), the yield strength of longitudinal reinforcement (f_y (MPa)), the yield strength of stirrups (f_{yw} (MPa)), the longitudinal reinforcement ratio (ρ_l (%)), the stirrups reinforcement ratio (ρ_w (%)), the spacing of stirrups (s (mm)), the loss of cross-sectional area of longitudinal reinforcement ($n_{l,sn}$ (%)), the loss of cross-sectional area of stirrups ($n_{w,sn}$ (%)), while the output parameter is the experimental shear capacity ($V_{u,exp.}$ (kN)). The experimental samples were considered simply supported beams. The shear span was measured as the distance between the center of the concentrated load and the center of the support plate. The b, h and d represent the geometric properties of the CRC beam. The geometric properties of CRC beams are important parameters influencing the size effect on shear behavior. [78, 79]. As is well known, the h is obtained by adding the concrete cover to the d. The concrete cover is a parameter that directly influences the corrosion mechanism [127-129]. The inclusion of both h and d among the input parameters of the model enables non-linear machine learning algorithms to successfully analyze the mathematical relationship between these two parameters, thereby allowing the effect of concrete cover on the shear behavior of corroded beams to be implicitly reflected in the models. If the a/b is equal or

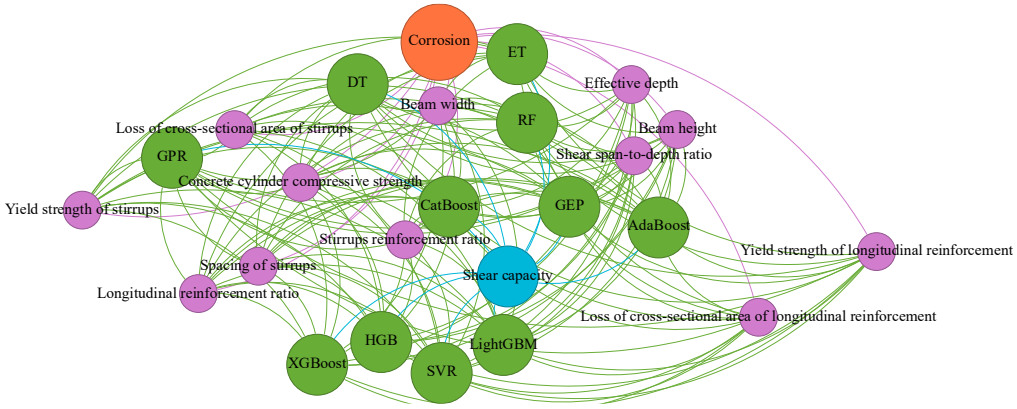


Figure 1. Structure of the database and mapping of parameters

less than 2, the RC beams are named as a deep beam, whereas if a/b exceeds 2, the RC beams are named as a slender beam [80]. The main shear transfer mechanisms differ between RC deep and slender beams [78]. Therefore, the a/b is considered an important parameter for understanding shear behaviour. The f_c, f_y and f_{yw} are parameters related to the materials properties of CRC beam. In addition, the studies [81, 82] contain proof of the effect of longitudinal reinforcement on the shear capacity of the beam, the delay of widening of the shear crack and the increase of the ductility of the beam. It is well known that an increase in the stirrup ratio leads to an improvement of the shear capacity [83, 84]. Furthermore, reducing stirrup spacing promotes the development of finer shear cracks along the shear span of the beam and contributes to a reduction in crack widths [85]. The $n_{l,sn}$ (%) and $n_{w,sn}$ (%) represent the corrosion degree of the CRC beam. The degree of corrosion in stirrups is a parameter that affects the shear behavior of CRC beams, as seen in experimental studies in literature. When the degree of corrosion in the stirrups is at a level that does not cause cracks in the concrete, the decrease in the shear capacity of the CRC beam is negligible. However, it has been observed that an increase in the degree of corrosion in the stirrups leads to significant decreases in shear capacity [11, 19, 21, 75, 86]. In addition, experimental studies in the literature indicate that corrosion in longitudinal reinforcement reduces the bond between the concrete and the reinforcement, thereby decreasing the shear capacity [87] and the effect of corrosion in longitudinal reinforcement on the shear capacity of RC beams may change depending on the a/d [34, 88, 89]. Finally, it is concluded that the selected parameters are appropriate for the ML models developed in this study according to ML, experimental, analytical, and theoretical researches in the literature mentioned above. Fig. 1 visually illustrates the structure of the database and the mappings between parameters.

Concrete compressive strength can be obtained from different sample types. These samples are typically cube (150×150×150 mm) and cylinder (150×300 mm) shaped. In the dataset created in this study, cylinder concrete compressive strength was used as the basis. In the experimental studies where the cube compressive strength of concrete was given, these values were converted to cylinder compressive strength using the conversion relationships recommended by Eurocode 2 [32]. The conversion relationship between the types of concrete compressive strength is presented in Table 1. In addition, the tensile strength values of concrete are also given in Table 1.

The degree of corrosion is expressed in two different terms in literature. These are the weight loss ratio (n_{wt}) and the cross-sectional area loss ratio (n_{sn}). In the dataset used in this study, the

Table 1. Conversion relationship between f_c, f_{cu} and f_t

Type														
f_c (MPa)	12	16	20	25	30	35	40	45	50	55	60	70	80	90
f_{cu} (MPa)	15	20	25	30	37	45	50	55	60	67	75	85	95	105
f_t (MPa)	1.6	1.9	2.2	2.6	2.9	3.2	3.5	3.8	4.1	4.2	4.4	4.6	4.8	5

Table 2. Descriptive statistics of input and output parameters

Parameters	Mean	STD	Min	Max	Q1	Q2	Q3	Skewness	Kurtosis
b (mm)	155.44	39.54	100	254	120	150	200	0.71	-0.24
h (mm)	256.18	93.12	150	610	200	230	300	1.82	3.94
d (mm)	216.55	87.04	118	521	167	184	265	1.60	2.88
a/d	2.35	0.88	1.00	4.70	1.75	2.20	3.00	0.89	0.67
f_c (MPa)	29.59	9.48	14.30	56.30	21	27.60	35.00	0.67	0.11
f_y (MPa)	454.03	111.99	210	706	390	435	493.25	0.48	0.34
f_{yw} (MPa)	395.49	100.62	275	626	324.25	369.60	464	0.72	-0.56
ρ_l (%)	2.31	0.69	1.22	4.56	1.92	2.17	2.62	1.38	2.29
ρ_w (%)	0.39	0.20	0.11	0.90	0.25	0.32	0.48	1.15	0.67
s (mm)	147.77	46.44	70	305	120	150	180	0.36	-0.07
$n_{l,sn}$ (%)	4.86	9.06	0	72.13	0	0	7.15	3.59	18.17
$n_{w,sn}$ (%)	21.84	21.23	0	97.20	4.87	18.74	30.23	1.50	2.46
$V_{u,exp.}$ (kN)	115.72	80.19	26.60	507	72.90	95.50	124.80	2.68	8.13

degree of corrosion in both longitudinal reinforcement and stirrups is represented by the cross-sectional area loss ratio. Some statistical studies have provided conversion relationships between these two loss ratio types. The conversion relationship between the cross-section area loss ratio and weight loss ratio is calculated according to Eqs. (1)-(3) that are given in study [27], [20] and [75], respectively. In this study, the conversion relationship between the two loss ratios is calculated according to Eq. (3).

$$n_{sn} = \begin{cases} 0.013 + 0.987n_{wt} & 0\% \le n_{wt} < 10\% \\ 0.061 + 0.939n_{wt} & 10\% \le n_{wt} < 20\% \\ 0.129 + 0.871n_{wt} & 20\% \le n_{wt} < 30\% \\ 0.099 + 0.801n_{wt} & 30\% \le n_{wt} \le 40\% \end{cases} \quad (1)$$

$$n_{sn} = 1.36n_{wt} \quad (2)$$

$$n_{sn} = \begin{cases} 0.015 + 0.97n_{wt} & n_{wt} \le 10\% \\ 0.062 + 0.95n_{wt} & n_{wt} > 10\% \end{cases} \quad (3)$$

Table 2 provides a comprehensive summary of the statistical properties of all the input and output parameters. It also provides quartile values for the distributions, as well as skewness and kurtosis values. As can be seen in Table 2, the minimum values of the geometric parameters of the CRC beams, such as b (mm), h (mm), d (mm) and a/b are 100, 150, 118 and 1.00, respectively, while the maximum values of these parameters are 254, 610, 521 and 4.70, respectively. The

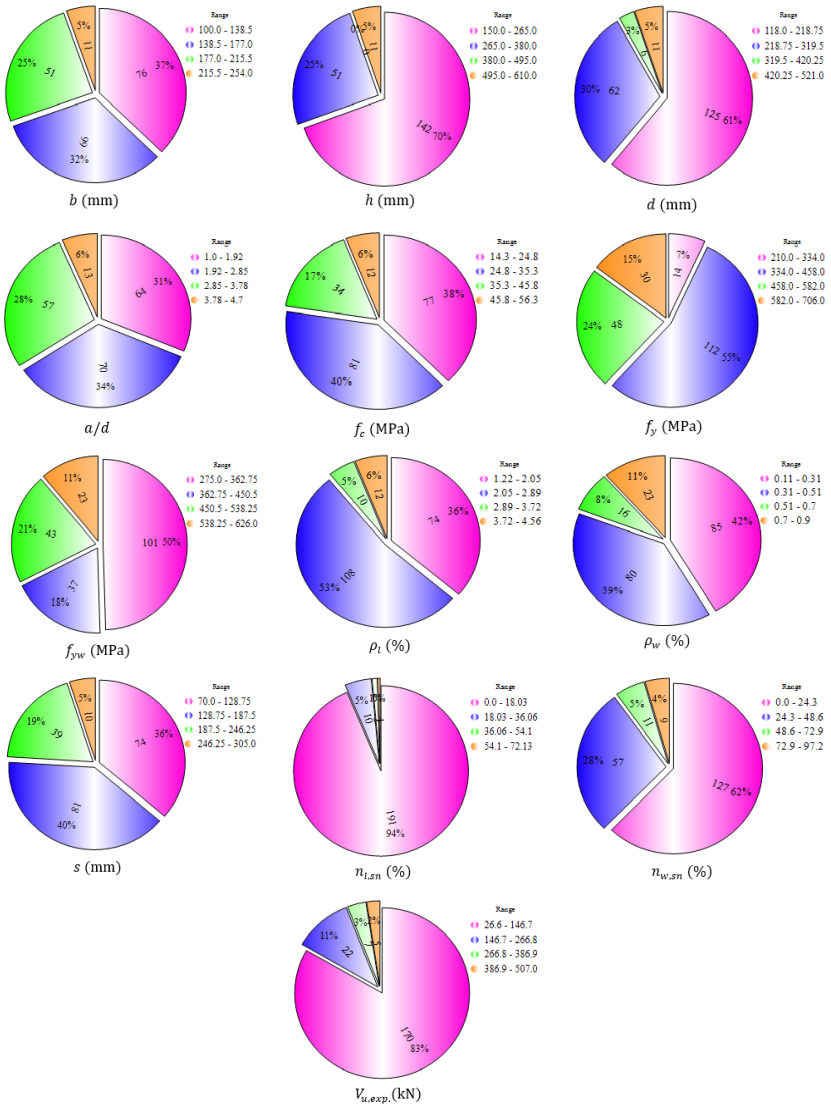


Figure 2. Frequency distribution of parameters

minimum-maximum values for the parameter-related material values f_c (MPa), f_y (MPa) and f_{yw} (MPa) are 14.30-56.30, 210-706 and 275-626, respectively. The ρ_l (%), ρ_w (%) as well as the s (mm), ranged from 1.22 to 4.56, 0.11 to 0.90, and 70 to 305, respectively. Moreover, the minimum and maximum values of $n_{l,sn}$ (%) and $n_{w,sn}$ (%) that are the corrosion-related parameters in Table 2 range from 0 to 72.13, and from 0 to 97.20, respectively. The ML models analyzed in this study and proposed GEP equation can be applied within the range of minimum and maximum values of the mentioned parameters as seen Table 2. However, in the collected dataset, corrosion exists either in the longitudinal reinforcement, the stirrups, or both. Therefore, the developed machine learning models and the proposed GEP equation are applicable exclusively

to corroded reinforced concrete beams. They are not valid for uncorroded control beams. There are significant differences between the updated dataset used in the present study and the dataset reported in Lu et al. [27], which has been widely used as a basis for studies on CRC beams in the literature. The first notable difference is related to the loss of cross-sectional area of the longitudinal reinforcement ($n_{l,sn}$ (%)). While the maximum loss in the updated dataset reaches 72.13%, the corresponding value in the dataset of Lu et al. [27] is limited to 26%. This indicates that severely corroded RC beams, which have not been sufficiently represented in previous studies, are included in the updated database. A second important difference can be observed in the range of concrete compressive strength (f_c (MPa)). The minimum and maximum concrete compressive strengths in the updated dataset are 14.30 MPa and 56.60 MPa, respectively, whereas the corresponding values in the dataset of Lu et al. [27] range from 20.0 MPa to 44.40 MPa. This suggests that the updated dataset includes CRC beams with both lower and higher concrete strength, thereby providing a broader representation of concrete material properties. A third significant difference between the two datasets is the range of longitudinal reinforcement ratios (ρ_l (%)). In the updated dataset, the longitudinal reinforcement ratio reaches a maximum value of 4.56%, whereas it is limited to 3.27% in the dataset reported by of Lu et al. [27]. Consequently, a wider range of reinforcement ratio is represented in the updated database. Overall, the broader coverage of these key parameters, which play an important role in the shear behavior of CRC beams, has enabled the machine learning models to be trained and tested over a wider parameter space. As a result, the generalization capability and potential applicability of the developed machine learning models to engineering practice have been improved. Furthermore, Fig. 2 presents the frequency distribution of the parameters. As shown in Fig. 2, the frequency of each parameter is indicated within its specified range. According to Fig. 2, approximately 32 % of the beams in the dataset are deep beams, and approximately 68 % are slender beams. The number of CRC beams having high-strength concrete (f_c≥50 MPa) is quite limited compared to the total number in dataset. Furthermore, the majority of the $n_{l,sn}$ is below 20%. Approximately 62% of the $n_{w,sn}$ is to be below 25%. These results suggest that future experimental studies examining the impact of corrosion on deep beams or those made of high-strength concrete could significantly contribute to existing literature on the subject. Furthermore, it is thought that experimental studies in which corrosion is severe in both the longitudinal reinforcement and the stirrups could improve our understanding of the effect of corrosion on the behavior of RC beams.

2.2 Pearson correlation analysis

Pearson Correlation Analysis is used to determine the linear relationship between the parameters in the dataset. Fig. 3 shows the results of the Pearson correlation analysis for input and output parameters. If the Pearson Correlation value is close to 1 or -1, this suggests a strong positive or negative correlation between the two parameters, whereas if this value is close to 0, it indicates a weak correlation between the two parameters. According to the heat map in Fig. 3, The parameter pairs showing the strongest correlation among the input parameters are h and d (0.99), h and b (0.86) and d and b (0.85). It is observed that a very strong correlation exists between the input parameters h and d (0.99). This situation can lead to multicollinearity. However, the complex relationship arising from such a very strong correlation between these parameters has been successfully analyzed using artificial intelligence-based machine learning algorithms that employ non-linear regression.

The correlation values of the input parameter pairs vary over a wide range. In addition, the

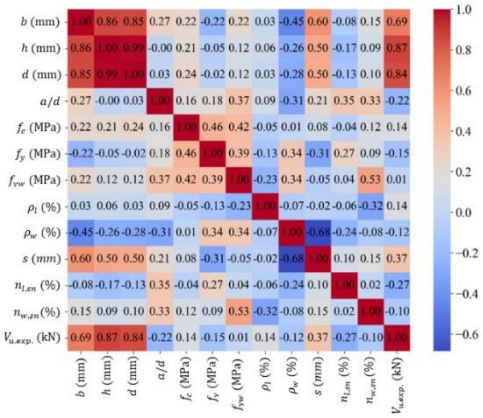


Figure 3. Pearson correlation heatmap of parameters

parameters that are most strongly correlated having the shear capacity of CRC beams are h, d and b respectively. In contrary, except for s, the correlation coefficients of other variables are low. The parameter having the weakest correlation is f_{yw}. This situation proves that there is no direct multivariate linear relationship between the input and output parameters. Therefore, it is appropriate to determine the nonlinear correlation relationship between each input parameter using the ML method for prediction of the shear capacity of CRC beams.

2.3 Existing prediction models in the literature

Five analytical models were examined, all of which were proposed by researchers in the literature. The models to predict the shear capacity of CRC beams were taken from Xu and Niu [28], Li et al. [90], Xue et al. [18], Tang et al. [75] and Lu et al. [27] studies. The equations of these five analytical models are observed in Table 3.

The shear capacity of the CRC beam is generally determined [28] by calculating the contributions of both the concrete and the stirrups, separately. The models listed in Table 3 are based on different theoretical frameworks. The limit equilibrium theory forms the basis of the equation proposed by Xu and Niu [28]. This model takes the effect of corrosion into account using reduction coefficients. The first reduction coefficient for the capacity provided by the concrete is expressed as an exponential. This exponential expression considers a/d, $n_{w,sn}$, the concrete cover of the stirrups and the stirrup diameter. This exponential expression considers a/d, $n_{w,sn}$, the concrete cover of the stirrups and the stirrup diameter. The second reduction coefficient for the capacity provided by the stirrups in these models is determined by taking into account the stirrup's yield strength and the reduction in its cross-sectional area. The model proposed by Li et al. [90] is obtained by updating the equation that predicts the shear capacity RC beams in the GB 50010-2010 code [31]. The design equation in the GB 50010-2010 code [31] code was proposed based on the traditional truss model. The updated model proposed by Li et al. [90] takes into account the reduction in beam width, effective depth, stirrup cross-sectional area and stirrup yield strength caused by corrosion in an RC beam. Furthermore, Xue et al. [18] developed a new model based on the outcomes of their experimental study. Xue et al. [18] concluded that as the corrosion ratio in longitudinal reinforcement bars increases, the critical shear crack behavior was similar to those of

Table 3. Models used for predicting the shear capacity of CRC beam

Ref.	Equations
[28]	$V_u = V_c + V_s;\ \alpha(n_{w,sn}) = 1 - 1.077 n_{w,sn};\ A_{swc} = (1 - n_{w,sn}) A_{sw}$ $V_u = \xi\left(\frac{a}{d}, n_{w,sn}\right) * \left(\frac{0.08+4\rho_l}{\frac{a}{d}-0.3}\right) f_{cp} b d + \alpha(n_{w,sn}) * (0.25 + 0.4 * \frac{a}{d}) A_{swc} f_{yw} d/s$ $\xi\left(\frac{a}{d}, n_{w,sn}\right) = \begin{cases} 1 & n_{w,sn} \le n_{cr,sn} \\ \left(\frac{n_{w,sn}}{n_{cr,sn}}\right)^{\left(\frac{0.069a}{d} - 0.43\right)} & n_{w,sn} > n_{cr,sn} \end{cases};\ n_{cr,sn} = 10.4\left(\frac{c_v}{\emptyset_v{}^2}\right) + \frac{f_{cu}}{\emptyset_v}$
[90]	$V_{cs} = \frac{1.75}{\frac{a}{d}+1} f_t b_c d_c + \frac{d_c}{s} A_{swc} f_{ywc};\ f_{ywc} = \left(1 - 1.1219 n_{w,wt}\right) f_{yw} / (1 - n_{w,wt})\ ;\ b_c = b - c_{v1} - c_{v2};$ $d_c = d - c_{sc}$
[18]	$V_u = V_c + V_s;\ V_c = \max(V_{c1}, V_{c2}) + f_{yw} A_{sw} j \frac{d}{s};\ j = \frac{1}{1.15}$ $V_{c1} = 0.2\sqrt[3]{100 f_c \rho_l}\sqrt[4]{\frac{1000}{d}}\left(0.75 + 1.4\frac{d}{a'}\right) b d;\ a' = (1 - 0.016 n_{l,wt}) a$ $V_{c2} = 0.24\sqrt[3]{f_c{}^2\left(1 + \sqrt{100\rho_l}\right)}\left(1 + 3.33\frac{r}{d}\right) b d / \left(1 + \left(\frac{a'}{d}\right)^2\right)$
[27]	$V_u = \emptyset V_c + V_s;\ V_c = \max(V_{c1}, V_{c2})$ $V_{c1} = 0.2\sqrt[3]{100 f_c \rho_{lc}}\sqrt[4]{\frac{1000}{d}}\left(0.75 + 1.4\frac{d}{a}\right) b d;\ \rho_{lc} = (1 - n_{l,sn})\rho_l$ $V_{c2} = \frac{0.24\sqrt[3]{f_c{}^2}\left(1+\sqrt{100\rho_{lc}}\right)\left(1+3.33\frac{r}{d}\right) b d}{1+\left(\frac{a}{d}\right)^2};\ V_s = f_{yw} A_{swc} j \frac{d}{s};\ j = \frac{1}{1.15}$ $\emptyset = \begin{cases} 0.008 e^{(-0.122\frac{a}{d})} - 0.003 n_{w,sn} + 1.01 & \frac{a}{d} < 2.5 \\ 0.1 e^{(-0.122\frac{a}{d})} - 0.003 n_{w,sn} + 1.38 & \frac{a}{d} \ge 2.5 \end{cases}$ $r = the\ width\ of\ the\ loading\ plate\ in\ mm$
[75]	$V_{Rd} = \emptyset V_{Rd,c} + \gamma V_{Rd,s};\ V_{Rd} = \frac{1}{\beta}\left(\emptyset C_{Rd,c} k (100\rho_l f_{ck})^{\frac{1}{3}} b_c d + \gamma \frac{0.9 d f_{ywc} A_{sw}}{s} \cot\theta\right);$ $\beta = \begin{cases} 0.25 & a < 0.5d \\ \frac{a}{2d} & 0.5d \le a \le 2d \\ 1 & a > 2d \end{cases};\ \cot\theta = \sqrt{\frac{a_{cv} v_1 f_{cd}}{\rho_{wc} f_{ywc}} - 1};\ a_{cv} = 1\ (nonprestressed\ structures)$ $k = 1 + \frac{\sqrt{200}}{d} \le 2;\ f_{cd} = \frac{a_{cc} f_{ck}}{\gamma_c};\ a_{cc} = 0.85;\ v_1 = \begin{cases} 0.6 & f_{ck} \le 60\ MPa \\ 0.9 - \frac{f_{ck}}{200} > 0.5 & f_{ck} > 60\ MPa \end{cases};$ $b_c = \begin{cases} b & n_{w,wt} < 30\% \\ b - 2(c + d_{ws}) + \frac{s}{5.5} & n_{w,wt} > 30\%; s \le 5.5c \\ b - \frac{5.5}{s}(c + d_{ws})^2 & n_{w,wt} > 30\%; s > 5.5c \end{cases};\ n_{sn} = \begin{cases} 0.015 + 0.97 n_{wt} & n_{wt} \le 10\% \\ 0.062 + 0.95 n_{wt} & n_{wt} > 10\% \end{cases};$ $\emptyset = \begin{cases} 1 & n_{l,wt} \le 5\% \\ 1.098 - 1.96 n_{l,wt} & n_{l,wt} > 5\% \end{cases}$ $\gamma = \begin{cases} 1 & n_{w,wt} \le 10\% \\ 1.076 - 0.76 n_{w,wt} & n_{w,wt} > 10\% \end{cases}$

shear crack behavior observed in deep beams. In light of this outcome, Xue et al. [18] proposed a model to predict the shear capacity of CRC beams based on modified truss analogy theory. In this model, the shear span is reduced by a coefficient dependent on $n_{l,sn}$. Lu et al. [27] carried out an

analytical study by using data from 158 experimental samples and performed comparative study. The results of comparation study showed that the model proposed by Xue et al. [18] was the best model in terms of minimum variability and mean shear capacity. However, the over residual shear capacity of beams with corroded at longitudinal reinforcement and stirrups is considered unreliable. For this reason, Lu et al. [27] proposed a model by modifying the model proposed by Xue et al. [18]. The shear capacity obtained to using model by proposed Lu et al. [27] given better and more reliable results in their study. Lu et al. [27] included a coefficient in the proposed model that reduces the contribution of the shear capacity from the concrete. This reduction coefficient is related to the a/d and $n_{w,sn}$ parameters. Furthermore, unlike the model proposed by Xue et al. [18], Lu et al. [27] did not reduce the shear span and took into account the reduction in the cross-sectional area of the longitudinal reinforcement due to corrosion. Tang et al. [75] proposed a new model based on the equations given in Eurocode 2 [32] for the design of the shear capacity of RC beams. The equations in Eurocode 2 [32] were established by adopting the variable-angle truss model method. In the variable angle truss model, the concrete was considered a compression bar, and the angle at which the compression bar and the longitudinal reinforcement bar can vary. Stirrup was also considered as a tension bar. Tang et al. [75] used effective beam width in Eurocode equations due to cracks in the concrete cover caused by corrosion. Moreover, Tang et al. [75] included corrosion coefficients in the equations to determine the effect of corrosion.

2.4 Machine learning models

In this study, eleven ML models were employed to predict the shear capacity of corroded reinforced concrete beams: DT [91], RF [92], ET [93], HGB [94], AdaBoost [95], XGBoost [96], LightGBM [94], CatBoost [97], SVR [98], GPR [99], and GEP [100]. The analyses were performed using the Python programming language within the JupyterLab environment [125] through the Anaconda Navigator platform [126]. These models are briefly introduced below.

2.4.1 Decision tree (DT)

Decision trees are used in machine learning for classification and regression problems where the target variable takes continuous values. The basic principle of the algorithm is to divide the dataset into more homogeneous subsets through iterative binary splitting. An important advantage of the model is that its predictions can be easily explained through the rules followed along the branches. On the other hand, since decision trees are non-parametric methods, they can capture non-linear relationships and work directly with both categorical and numerical variables. They are practical in application because they do not require scaling, normalization, or complex preprocessing steps [101, 102]. However, one disadvantage is that they tend to overfit the training data, especially deep trees, which leads to a decrease in the generalization performance of the model. Moreover, even small changes in the dataset can significantly alter the structure of the trees and cause the model to have high variance [91, 102]. Thus, pruning methods are used to prevent the tree from growing beyond a certain depth or to remove branches that contribute little.

2.4.2 Random forest (RF)

Random Forest (RF) is a powerful ensemble learning method proposed by Breiman [92] to overcome the limitations of decision trees. It is based on bagging, where multiple decision trees are trained on random bootstrap samples taken from the training data and in regression problems, the final prediction is obtained by averaging the outputs of all trees [92, 103]. In each split, only a

randomly selected subset of features is considered, which increases the diversity among the trees. The RF model can capture nonlinear and complex relationships, offers high accuracy and stability, and is relatively robust to noisy and partially missing data. However, using multiple trees reduces interpretability compared to a single decision tree and while its parallel processing-friendly structure allows efficient use on large datasets, it increases computational and memory costs [103, 104]

2.4.3 Extremely randomized trees (ET)

The Extremely Random Trees (Extra Trees, ET) algorithm, developed by Geurts et al. [93] as an alternative to the Random Forest model, is a decision tree-based ensemble method that introduces more randomness into the tree structure to reduce variance and overfitting. Unlike Random Forest, each tree is trained on the entire dataset without bootstrap sampling, and at each split a random subset of features is considered, while split thresholds are chosen completely at random. The best split from these candidates is then selected [93]. This leads to greater tree diversity and lower correlation within the ensemble. ET offers strong generalization ability and is computationally efficient, making it suitable for fast analysis of large-scale datasets. However, its high level of randomness can limit accuracy on small datasets, and like other tree-based ensembles, it has lower interpretability than a single decision tree and may require more memory and storage due to the large number of trees [104].

2.4.4 Histogram-based gradient boosting (HGB)

Histogram-based gradient boosting is a faster and more memory-efficient version of the classic boosting approach. Ke et al. [94] proposed discretizing continuous features into a fixed number of bins, creating histograms. This reduces the number of candidate split points, speeding up train considerably. At the same time, memory consumption is much lower, making it well-suited for large datasets. Despite the efficiency gains, accuracy is typically preserved. The method is flexible and adapts well to different data types. For these reasons, histogram-based boosting has become a standard tool in modern machine learning applications.

2.4.5 Adaptive boosting (AdaBoost)

The fundamental idea behind the model proposed by Freund and Schapire in 1997 [95] is to create a strong predictive model by sequentially combining weak learners, such as shallow decision trees. AdaBoost ensures that subsequent learners focus on difficult instances by assigning higher weights to misclassified or incorrectly predicted samples. In this way, each new model compensates for the shortcomings of the previous one, systematically reducing the overall error [105]. AdaBoost is a flexible algorithm that can be applied with various weak learners. Even when the train error theoretically approaches zero, the literature reports that the method exhibits greater resilience against overfitting than expected, owing to the controlled progression of the generalization error [95, 106]. However, the algorithm is highly sensitive to noisy data and outliers because of the increased weight assigned to misclassified examples [107]. In addition, excessive iterations may lead to the risk of overfitting. Therefore, careful tuning of hyperparameters such as the number of learners and the learning rate is critical.

2.4.6 Extreme gradient boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a highly efficient and scalable implementation of gradient boosting that has gained considerable popularity in both academic research and industry.

Introduced by Chen and Guestrin [96], XGBoost enhances classical boosting algorithms through advanced optimization strategies and computational improvements, making it particularly suitable for large-scale and high-dimensional datasets. The fundamental principle of XGBoost, as with other boosting methods, is the sequential train of an ensemble of weak learners, where each new model aims to correct the errors of its predecessors. Unlike traditional boosting, XGBoost optimizes a regularized objective function that balances model fit and complexity. This objective function improves generalization by combining a differentiable convex loss with an explicit regularization term that penalizes overly complex models [96].

2.4.7 Light gradient boosting machine (LightGBM)

Light Gradient Boosting Machine (LightGBM), introduced by Ke et al. [94] and developed by Microsoft Research, is an efficient implementation of gradient boosting decision trees designed for large-scale datasets. Its main innovation is the leaf-wise tree growth strategy, which expands the leaf that provides the largest error reduction, enabling faster convergence but with a potential risk of overfitting if depth is not properly constrained. To improve efficiency, LightGBM employs Gradient-based One-Side Sampling (GOSS), which prioritizes high-gradient samples, and Exclusive Feature Bundling (EFB), which reduces dimensionality by combining sparse features. These innovations make LightGBM significantly faster and more memory-efficient than conventional boosting algorithms [94, 108].

2.4.8 Categorical boosting (CatBoost)

CatBoost is a gradient boosting algorithm developed by Prokhorenkova et al. [109] for efficient processing of categorical data. One of its most important innovations is the ordered boosting approach, which divides the training data into random permutations and uses only the information of previous examples for each example. This feature reduces target leakage and improves generalization performance [109]. It also stands out with its ability to directly process categorical variables. Unlike most boosting algorithms, it applies ordered target encoding without the need for preprocessing steps such as one-hot encoding, thus both improving accuracy and reducing preprocessing overhead [97]. CatBoost is widely preferred for its robustness to noisy data, high accuracy even with default hyperparameters, and enabling efficient train on large datasets with GPU support [108].

2.4.9 Support vector regression (SVR)

Proposed by Cortes and Vapnik, Support Vector Regression (SVR) became widely adopted for its effectiveness in handling high-dimensional data, aiming to identify a regression function that best fits the target values within a predefined tolerance margin [110, 111]. By introducing the ε-insensitive zone, the generalization ability of the model is improved while the risk of overfitting is reduced. SVR is effective not only for linear regression but also for capturing non-linear relationships through the use of kernel functions. Commonly employed kernels include the polynomial kernel, sigmoid kernel, and, most notably, the Radial Basis Function (RBF) kernel [112]. The kernel approach enables SVR to achieve strong performance even in complex and high-dimensional problems. However, a notable drawback of SVR is that train time and memory requirements increase very rapidly with the number of train instances, which restricts its application to very large datasets. Moreover, selecting the appropriate kernel function and tuning hyperparameters often requires expertise and extensive search. In terms of interpretability, SVR is more limited compared to tree-based methods, as the model can primarily be analyzed through

support vectors.

2.4.10 Gaussian process regression (GPR)

Gaussian processes, comprehensively described by Rasmussen and Williams [99], define a prior distribution over functions that are updated with observations to obtain a posterior distribution. Each observation is assumed to be a Gaussian random variable, and together they form a multivariate Gaussian distribution. This structure is governed by the chosen covariance (kernel) function. Commonly used kernels include the Radial Basis Function (RBF), Matérn, and periodic kernels, and the appropriate selection of these functions increases the flexibility of the model. Gaussian Process Regression (GPR) provides not only predictions but also confidence intervals, enabling the evaluation of model reliability in engineering applications. Its non-parametric nature allows it to capture complex relationships and achieve high accuracy on small-to-medium-sized datasets. Hyperparameters are typically optimized via the marginal likelihood function [113, 121]. However, due to its high computational complexity, GPR becomes impractical for very large datasets. To address this limitation, approximate methods such as inducing points and sparse Gaussian processes have been developed [114].

2.4.11 Gene expression programming (GEP)

Gene Expression Programming (GEP) is an evolutionary algorithm that combines ideas from genetic algorithms and genetic programming. Proposed by Cândida Ferreira in 2001 [100], this algorithm is widely used for symbolic regression and modeling complex relationships with mathematical expressions [100]. In GEP, candidate solutions are encoded as fixed-length strings called genotypes, which are then translated into phenotypes, which are interpreted as branched mathematical expressions. Using Karva notation, individuals are evolved through operators such as crossover, mutation, and transposition to efficiently explore the search space [115]. A key advantage of GEP is that it produces clear and interpretable mathematical formulas, making it valuable in contexts where black-box models are criticized. Thanks to its fixed-length gene structure, it allows for fast evolutionary processes and outperforms classical genetic algorithms in terms of speed [100, 115]. During derivation of the proposed GEP equation, the physical parameters and their units (%, mm, MPa and kN), which were previously mentioned and given in Table 2, were directly used. Standardization and normalization processes were not applied. Therefore, an engineer applying the equation does not need to perform any preprocessing or non-dimensionalization; they should directly substitute the parameters value into the proposed GEP equation.

2.5 Machine learning workflow

The machine learning workflow followed in this study is shown in Fig. 4. First, the dataset was created, relevant parameters were selected and the data were split into training (80%) and testing (20%) subsets. Gene Expression Programming (GEP) and different machine learning methods were used in the study. In GEP, chromosomes were evolved by selection, mutation, and crossover until a fixed number of generations was reached or convergence was achieved [116]. For the machine learning models, hyperparameters were tuned with GridSearchCV using five-fold cross-validation [117]. The performance of the models was evaluated using statistical metrics. These metrics comprehensively reveal model accuracy, error magnitude, and explanatory power [118, 119]. In the final stage, the model with the best performance was selected, and SHAP analysis was

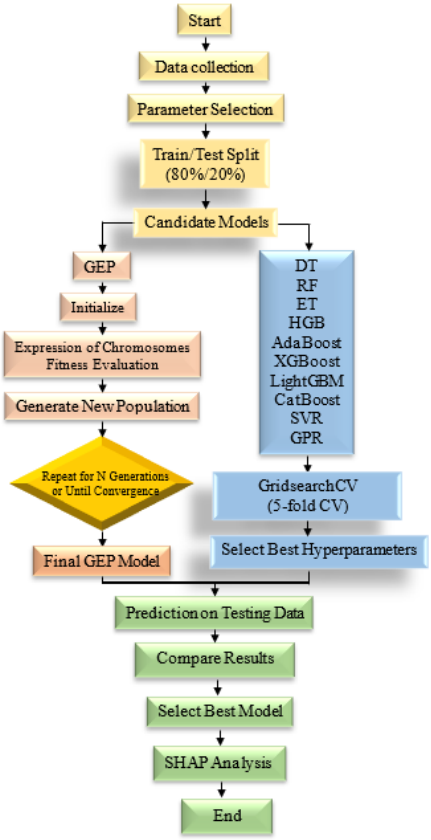


Figure 4. Machine learning workflow

performed to examine the relative importance of the parameters.

2.6 Performance indicators of models

To quantitatively evaluate and compare the predictive performance of the developed models, five widely used statistical indicators were employed: the mean value (MV), standard deviation (STD), root mean square error (RMSE), coefficient of determination (R^2), and coefficient of variation (COV) [118, 119]. Their mathematical formulations are given below:

$$\mathrm{MV} = \frac{\sum_{i=1}^{n} \frac{experimental\ value}{predicted\ value}}{n} \tag{4}$$

$$\mathrm{STD} = \sqrt{\frac{\sum_{i=1}^{n} (\frac{experimental\ value}{predicted\ value} - \mathrm{MV})^2}{n-1}} \tag{5}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

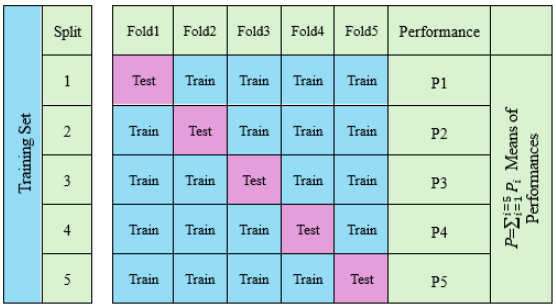


Figure 5. Schematic representation of 5-fold cross-validation [124]

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7}$$

$$\mathrm{COV} = \frac{\mathrm{STD}}{\mathrm{MV}} \tag{8}$$

where y_i represents the experimental values, $\hat{y}_i$ denotes the predicted values, $\bar{y}$ is the mean of the experimental values, and n is the number of samples. RMSE presents errors in the same unit as the actual values and is particularly sensitive to large deviations. R^2, on the other hand, indicates the extent to which the model explains the variability in the data.

2.7 Hyperparameter tuning and cross-validation

In this study, hyperparameter optimization and cross-validation were used together to improve the performance of machine learning models and allow a fair comparison. Choosing them poorly can cause overfitting or underfitting. Therefore, all inputs were first standardized using the StandardScaler method to eliminate the effects of variables with different scales on model train. Subsequently, Hyperparameter optimization was then conducted for each model using the GridSearchCV method, combined with 5-fold cross-validation. As shown in Fig. 5, the train data was divided into five equal parts; in each iteration, one part was used as the validation set, while the other four parts were reserved for train. This process was repeated five times, with each fold evaluated once for validation, and the average of the performance indicators obtained from each iteration was accepted as the overall performance.

Table 4 summarizes the optimal hyperparameter settings for each algorithm. Hyperparameters for machine learning models were determined through a systematic screening process. For each model, different hyperparameter combinations were tested using the GridSearchCV method, and after five-fold cross-validation, the settings that provided the lowest error and highest accuracy were selected. The models were run with the most appropriate hyperparameter configurations. In tree-based methods (DT, RF, ET, HGB), the optimal results were achieved by adjusting the basic settings that affect the model's complexity and diversity, such as tree depth, number of leaves, learning rate, and number of predictors. In boosting-based models (AdaBoost, XGBoost, LightGBM, CatBoost), performance depended largely on the optimal adjustment of the learning rate, iteration count, depth limit, and regularization terms. In kernel-based approaches (SVR, GPR), the regularization coefficients and kernel structure played a decisive role in both sensitivity

Table 4. Optimal hyperparameter values for each ML model

Model	Hyperparameters
DT	ccp_alpha=0.0, max_depth=None, max_features=None, min_samples_leaf=2, min_samples_split=5
RF	max_depth: None, max_features: 0.6, max_samples: 0.9, min_samples_leaf: 2, min_samples_split: 5, n_estimators: 1200
ET	max_depth: None, max_features: sqrt, min_samples_leaf: 2, min_samples_split: 2, n_estimators: 200
HGB	l2_regularization: 1.0, learning_rate: 0.05, max_depth: 4, max_iter: 1000, max_leaf_nodes: 31, min_samples_leaf: 30
AdaBoost	learning_rate: 0.1, loss: exponential, n_estimators: 200
XGBoost	colsample_bytree: 0.75, gamma: 1.0, learning_rate: 0.05, max_depth: 3, min_child_weight: 10, n_estimators: 600, reg_alpha: 2.0, reg_lambda: 15.0, subsample: 0.75
LightGBM	colsample_bytree: 0.75, learning_rate: 0.04, max_depth: 5, min_child_samples: 30, n_estimators: 1500, num_leaves: 31, reg_alpha: 0.5, reg_lambda: 10.0, subsample: 0.75
CatBoost	bootstrap_type=Bernoulli, depth=3, iterations=1000, l2_leaf_reg=15.0, learning_rate=0.03, min_data_in_leaf=75, random_strength=1.0, rsm=0.7, subsample=0.7
SVR	C: 100, epsilon: 0.01, gamma: auto, kernel: rbf
GPR	alpha: 0.05, kernel: 1**2 * Matern(length_scale=1, nu=2.5) + WhiteKernel(noise_level=0.5), n_restarts_optimizer: 2
GEP	chromosomes=150, head_size=10, genes=4, linking_function=Addition, mutation=0.044, one_point_recombination=0.3, two_point_recombination=0.3, gene_recombination=0.25, IS_transposition=0.1, RIS_transposition=0.1, constants_per_gene=10, constant_range=[-1000,1000], functions=[Addition, Subtraction, Multiplication, Square root, Power2]

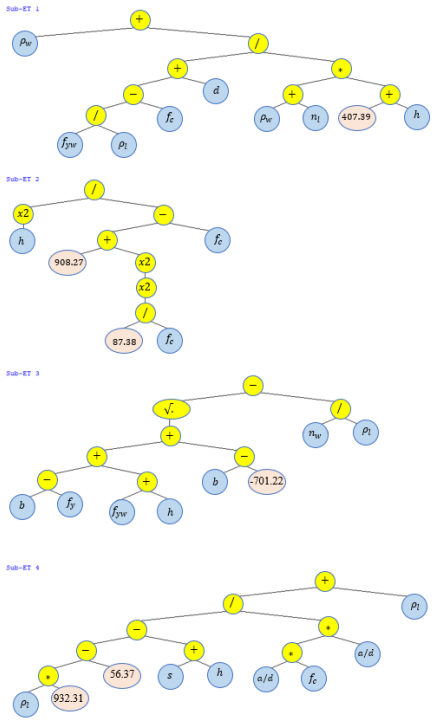


Figure 6. Sub-expression trees (Sub-ETs) generated by the GEP model

and generalization. All basic settings for the GEP model, such as the chromosome number, gene structure, function set, and evolutionary operator ratios, were determined by optimizing them.

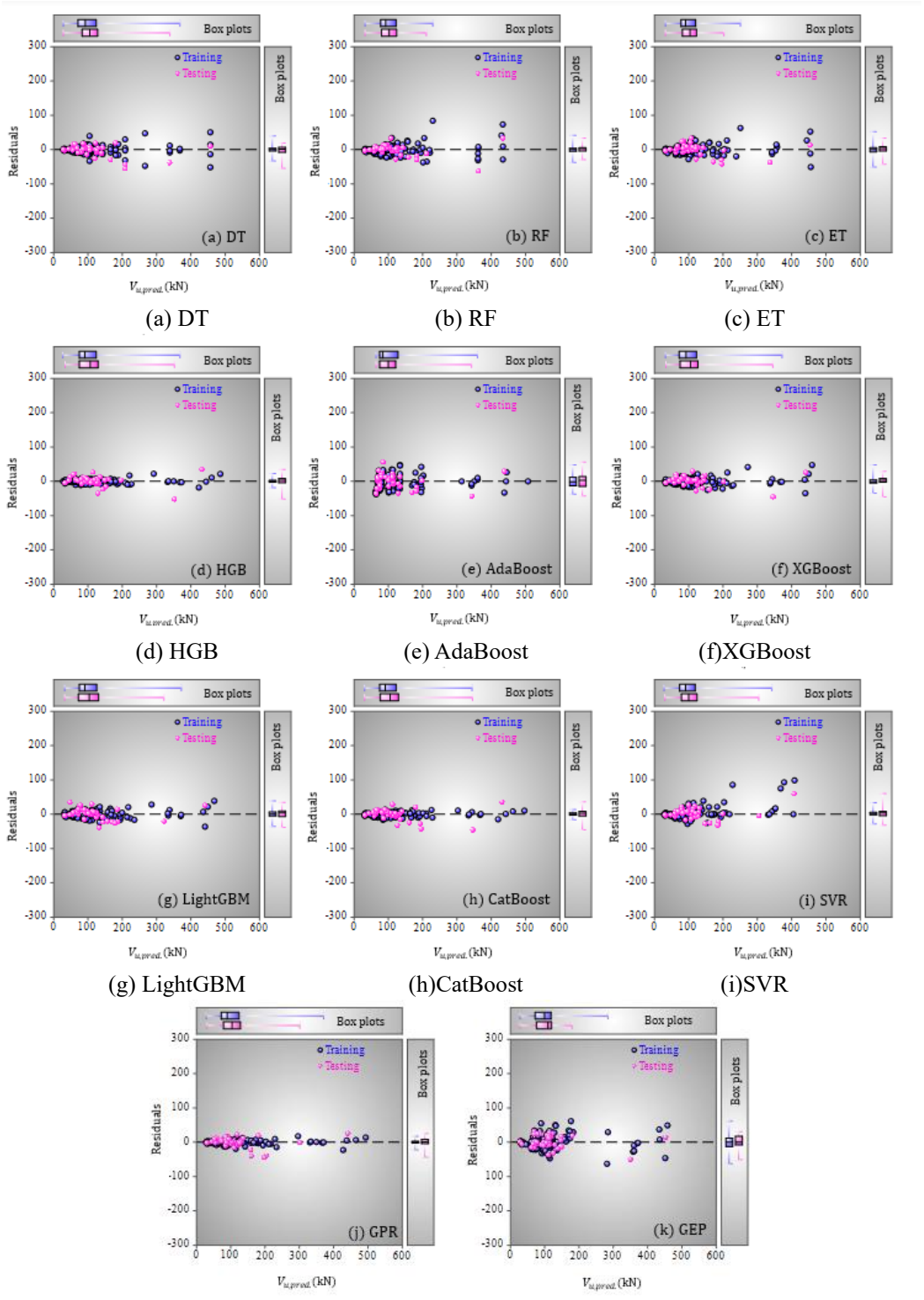
To ensure reproducibility and fair model comparison, the training and testing datasets used in the GEP analysis were also employed for all machine learning models. In addition, a fixed random seed value of 42 was used during the five-fold cross-validation procedure to ensure consistent fold generation.

2.8 GEP expression tree and derivation of the mathematical model

Fig. 6 represents four sub-expression trees (Sub-ETs) generated by the GEP model. Each Sub-ET illustrates a distinct combination of input parameters and mathematical operations, where nodes represent functions and terminals correspond to variables or constants. They were obtained under the optimal parameter settings summarized in Table 4, which ensured sufficient diversity while avoiding excessive complexity. By combining these trees with the addition linking function, the mathematical expression given in Eq. (9) was derived, which explicitly defines the shear capacity of the corroded beams.

Table 5. Statistical performance results of ML models

Models	Dataset	MV	STD	RMSE	R^2	COV
DT	Training	1.000	0.082	11.411	0.981	0.082
	Testing	1.002	0.119	16.299	0.956	0.119
RF	Training	0.991	0.096	14.041	0.971	0.097
	Testing	1.009	0.105	15.844	0.955	0.104
ET	Training	0.992	0.097	12.365	0.977	0.098
	Testing	1.018	0.129	15.899	0.956	0.126
HGB	Training	0.998	0.093	10.076	0.985	0.094
	Testing	1.044	0.178	16.209	0.951	0.171
AdaBoost	Training	0.978	0.200	17.659	0.954	0.204
	Testing	1.015	0.231	21.022	0.916	0.228
XGBoost	Training	0.997	0.080	9.643	0.986	0.080
	Testing	1.028	0.116	13.725	0.965	0.113
LightGBM	Training	0.997	0.100	10.842	0.983	0.100
	Testing	1.033	0.167	16.578	0.949	0.161
CatBoost	Training	0.994	0.096	11.009	0.982	0.097
	Testing	1.020	0.117	14.949	0.959	0.115
SVM	Training	1.015	0.100	16.893	0.970	0.099
	Testing	1.029	0.117	15.917	0.954	0.114
GP	Training	0.995	0.095	11.642	0.980	0.095
	Testing	1.011	0.127	14.443	0.963	0.125
GEP	Training	0.996	0.186	19.831	0.941	0.187
	Testing	1.025	0.176	17.851	0.941	0.171



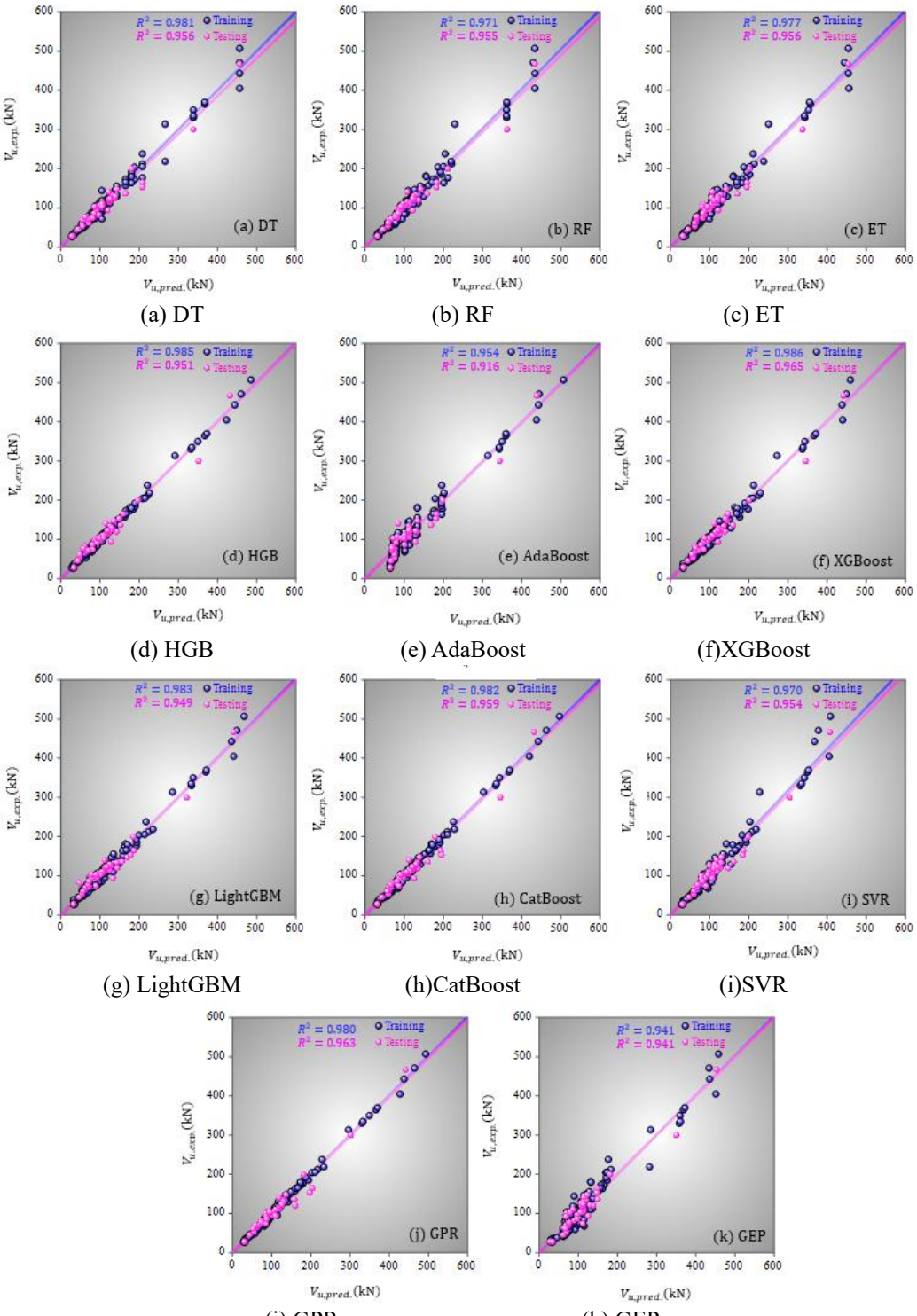
(a) DT (b) RF (c) ET

(d) HGB (e) AdaBoost (f)XGBoost

(g) LightGBM (h)CatBoost (i)SVR

(j) GPR (k) GEP

Figure 7. Residual analysis (V_{exp}-$V_{pred.}$) with respect to $V_{pred.}$ among ML models



(a) DT (b) RF (c) ET

(d) HGB (e) AdaBoost (f)XGBoost

(g) LightGBM (h)CatBoost (i)SVR

(j) GPR (k) GEP

Figure 8. Evaluation of $V_{u,pred.}$ against $V_{u,exp.}$ for ML models

$$V_{u,pred.} = \frac{[(f_{yw}/\rho_l)-f_c]+d}{(\rho_w+n_{l,sn})(h-407.39)} + \frac{h^2}{[(87.38/f_c)^4+908.27]-f_c} + \frac{(932.31\rho_l-s-h-56.37)}{f_c(a/d)^2} + (2b - f_y + f_{yw} + h + 701.22)^{1/2} + \rho_w+\rho_l - \left(n_{w,sn}/\rho_l\right) \quad (9)$$

3. Result and discussion

3.1 Predictive performances of ML models

The data collected for this study were randomly divided into two subparts: training set (80%) and testing set (20%). The prediction results obtained from the ML models used were compared with the experimental results. As a result of this comparison, the values of the above-mentioned statistical indicators calculated separately for training and testing set, are presented in Table 5. Furthermore, the residuals between the experimental and predicted results for each ML model are presented in Fig. 7. Finally, the scatter plots for all models are presented in Fig. 8 to better evaluate their prediction accuracies.

In this study, the shear capacity prediction capabilities of eleven ML algorithms, including DT, RF, ET, HGB, AdaBoost, XGBoost, LightGBM, CatBoost, SVR, GPR and GEP, were comprehensively investigated. It was observed in Fig. 7 that the residuals of the predicted shear capacity of CRC beams for both the training and testing datasets are concentrated around the zero line for XGBoost, HGB, GPR, CatBoost and LightGBM. Although the DT, RF, ET, AdaBoost, SVR and GEP models exhibit slight deviations from the zero line, the majority of their residuals remain tightly gathered near the center. In addition, as illustrated by the scatter plots in Fig. 8, the predicted shear capacities closely align with the experimental values, demonstrating a strong correlation between them. Therefore, it can be concluded that the developed ML models are robust and exhibit a high degree of accuracy in predicting the shear capacity of corroded RC beams.

Eleven ML models were developed using the training dataset. The robustness of these ML models was then statistically verified using the testing dataset, with the R^2 values exceeding 0.90 across all evaluated models (Table 5). Based on the statistical indicators in Table 5, the XGBoost model achieved the highest R^2 values, yielding 0.986 for the training dataset and 0.965 for the testing dataset. As detailed in Table 5, the R^2 values in the training dataset for the DT, RF, ET, HGB, AdaBoost, LightGBM, CatBoost, SVR, GPR, and GEP models are 0.981, 0.971, 0.977, 0.985, 0.954, 0.983, 0.982, 0.970, 0.980 and 0.941, respectively. For the testing dataset, the corresponding R^2 values are 0.956, 0.955, 0.956, 0.951, 0.916, 0.949, 0.959, 0.954, 0.963 and 0.941, respectively. It can be concluded that the R^2 values in the training and testing datasets for each ML model are very close to each other. Furthermore, the values of other statistical indicators, namely MV, STD, RMSE and COV, in the training and testing datasets for each ML model are also close to each other when evaluated between them in Table 5, separately. These similar statistical values obtained in the train and testing datasets prove that the ML models used show similar and consistent performance in the training and testing datasets. This result demonstrates that the ML models overcome the problems of overfitting or underfitting thereby possessing superior generalization capabilities, confirming their robustness and applicability to unseen datasets. Moreover, the high R^2 values and low STD, RMSE and COV values exhibit that ML models have a high degree of accuracy and robustness in predicting the shear capacity of CRC beams. Additionally, the MV of the ML models used remains remarkably close to 1. These

outcomes indicate that the ML prediction models successfully capture the actual shear capacity behavior without underestimating or overestimating the results.

The statistical values presented in Table 5 show that the prediction results of the ML models investigated are quite close to each other for the training and testing sets, and that the performance of ML models is superior prediction capability. Despite the similar statistical performance among the employed machine learning models, the XGBoost model was determined to achieve the highest predictive accuracy based on the testing dataset evaluation. Table 5 shows that the STD, RMSE, R^2 and COV values of the XGBoost model are 0.116, 13.725, 0.965 and 0.113, respectively, for testing dataset. Taking all these statistical values into account, XGBoost was selected as the best ML model in terms of robustness and prediction accuracy among the ML models used in this study.

To further evaluate the robustness of the best-performing model, a repeated random split validation was performed for the XGBoost model. The dataset was randomly divided into 80% training and 20% testing sets over 30 independent runs using different random seed values. In each run, the model was retrained and evaluated on the corresponding testing set. The repeated random split validation yielded mean±standard deviation values of R^2=0.924±0.051 and RMSE=18.127±3.153. These results indicate that the predictive performance of the XGBoost model remained stable across different data partitions, further supporting the robustness of the proposed modeling approach. Moreover, to further assess the uncertainty associated with the predictive performance of the proposed model, a test-set bootstrap analysis was conducted for the XGBoost model, which exhibited the highest prediction accuracy among the evaluated machine learning algorithms. In this procedure, the trained model was kept fixed, and the observed and predicted values in the test dataset were resampled 5000 times. The R^2 and RMSE metrics were recalculated for each bootstrap sample, and the 2.5th and 97.5th percentiles were used to determine the 95% confidence intervals.

The results of the bootstrap analysis are presented in Table 6. The mean R^2 value was 0.9476 with a 95% confidence interval ranging from 0.8510 to 0.9847, while the mean RMSE value was 13.5575 with a 95% confidence interval of 9.4465-17.6677. The original test-set performance values (R^2=0.9636 and RMSE=13.7248) were found to lie within the corresponding confidence intervals, indicating that the predictive performance of the XGBoost model remains stable under repeated resampling conditions. These findings provide an explicit quantification of prediction uncertainty and further demonstrate the robustness of the proposed model. In addition to predictive performance, the computational efficiency of the ML models was also evaluated. For this purpose, the training and prediction times of each model were measured under the same computational environment. Training time was calculated by including the GridSearchCV-based hyperparameter optimization process with five-fold cross-validation, whereas prediction time was measured on the testing dataset. All analyses were performed on a computer equipped with an Intel® Core™ i7-1255U processor, 16 GB RAM and a 64-bit Windows operating system. The obtained training and prediction times are presented in Table 7. The results indicate notable differences in computational cost among the ML models. LightGBM exhibited the highest training time, whereas SVM and Decision Tree required substantially lower computational effort. The comparatively longer training time observed for LightGBM is primarily associated with the extensive hyperparameter search space evaluated during the GridSearchCV optimization process. Despite these differences in training cost, prediction times were generally low for all models. In addition to the machine learning models, a GEP model was also utilized. Owing to its evolutionary search process for deriving explicit formulation, GEP required a higher computational cost than the other models.

Table 6. Bootstrap-based uncertainty quantification results for the XGBoost model

Metric	Test Value	Mean	95% Confidence Interval
R^2	0.9636	0.9476	0.8510-0.9847
RMSE	13.7248	13.5575	9.4465-17.6677

Table 7. Model Training and prediction times of the ML models

Model	Training Time (s)	Prediction Time (s)
DT	0.7770	0.0014
RF	73.9440	0.1022
ET	39.4971	0.0272
HGB	15.3007	0.0214
AdaBoost	7.7576	0.0147
XGBoost	6.2384	0.0059
LightGBM	484.6070	0.0073
CatBoost	8.3235	0.0030
SVM	0.2831	0.0012
GP	1.5921	0.0020

The prediction results of the XGBoost model on the testing dataset were analyzed. Among the 41 samples included in the test dataset of the XGBoost model, 22 samples have an a/d between 1.22 and 2. Therefore, these 22 samples can be classified as deep beams. The statistical metrics obtained from the prediction results for these deep beams are R^2=0.984 and COV=0.071. As can be seen, the R^2 value is very high, while the COV value is very low. Therefore, it can be stated that the prediction accuracy of the XGBoost model is very high for deep beams within this dataset range. For beams with an a/d greater than 2, which are classified as slender beams, the statistical metrics obtained from the prediction results are R^2=0.971 and COV=0.150. It can also be stated that the prediction accuracy for slender beams is high. When comparing the prediction performance of deep and slender beams, the XGBoost model yields predictions with a higher accuracy rate for deep beams on the testing dataset. Another important experimental parameter is concrete compressive strength. In the testing dataset, there are 6 samples with concrete compressive strengths lower than 20 MPa and only 2 samples with compressive strengths higher than 50 MPa. Due to the limited number of samples, a statistical evaluation could not be performed for corroded beams with either low-strength or high-strength concrete. As the number of experimental studies on corroded beams increases in future research, datasets containing a greater number of low-strength and high-strength corroded beams may be developed, thereby providing a basis for more comprehensive investigations. The dataset contains 33 samples with concrete compressive strengths between 20 and 50 MPa. The statistical metrics obtained from the prediction results of these samples are R^2=0.970 and COV=0.111. Therefore, it can be stated that the XGBoost model provides highly accurate predictions for beams with concrete compressive strengths ranging from 20 to 50 MPa. Another parameter is the level of corrosion in the longitudinal reinforcement. The corrosion level of the longitudinal reinforcement does not exceed 30% in the test dataset. Therefore, no assessment could be made for beams with severe longitudinal reinforcement corrosion. Regarding stirrup corrosion, there are 21 samples with

Table 8. Statistical performance results of XGBoost ML model and existing models

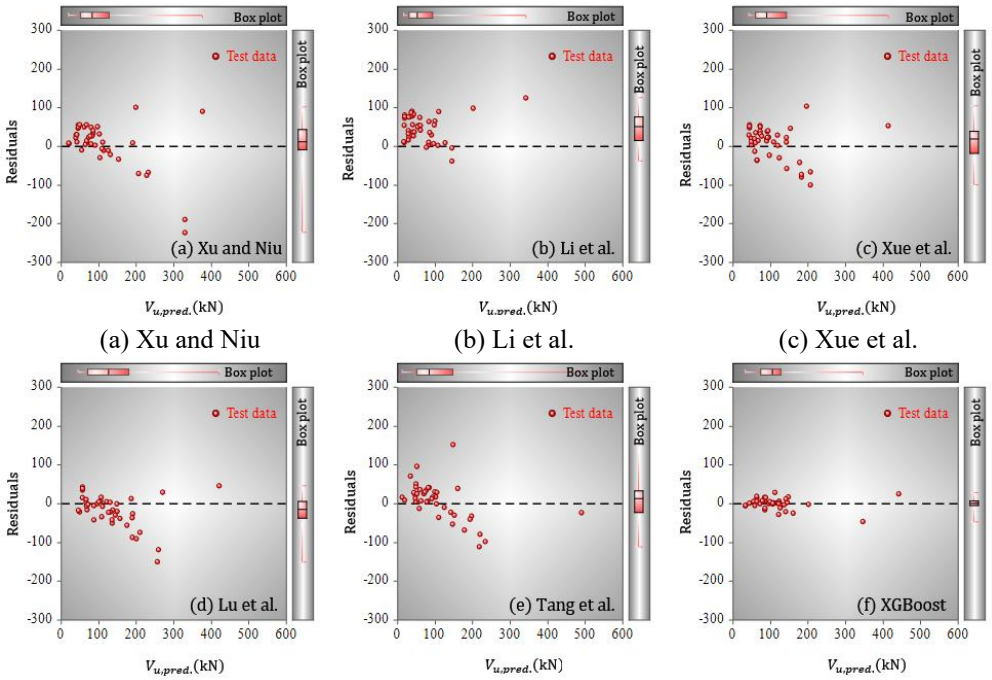
Models	Dataset	MV	STD	RMSE	R^2	COV
XGBoost	Testing	1.028	0.116	13.725	0.965	0.113
Xu and Niu [28]	Testing	1.267	0.466	60.924	0.506	0.368
Li et al. [90]	Testing	2.166	1.169	57.967	0.774	0.540
Xue et al. [18]	Testing	1.207	0.455	43.544	0.670	0.377
Lu et al. [27]	Testing	0.916	0.304	44.961	0.710	0.332
Tang et al. [75]	Testing	1.295	0.584	48.875	0.662	0.451

corrosion levels equal to or less than 20% and 20 samples with corrosion levels greater than 20%. For beams with stirrup corrosion levels of 20% or less, the R^2 and COV values are 0.904 and 0.103, respectively. For beams with stirrup corrosion levels greater than 20%, the corresponding values are 0.975 and 0.120. As can be seen from these statistical metrics, although the COV values are relatively close, the higher R^2 value obtained for beams with severe stirrup corrosion indicates that the XGBoost model can provide highly accurate predictions for beams with higher levels of stirrup corrosion.

3.2 Comparison between existing models and XGBoost

In this section, the experimental shear capacities of the corroded RC beams are compared against the predictions from the XGBoost model, which was selected as the optimal ML algorithm due to its accuracy and robustness, and various analytical models from the literature, including those developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], and Tang et al. [75]. Subsequently, the statistical indicators derived from this comparative study are thoroughly evaluated. The statistical indicators values of the models for testing dataset were presented in Table 8. Furthermore, Fig. 9 shows the residual values of the analytical models developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], Tang et al. [75], as well as those of the XGBoost model. Fig. 10 presents scatter plots to better understand the differences between the XGBoost ML model's prediction results and those obtained from the analytical models developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], Tang et al. [75]. Figs. 9 and 10 were plotted using experimental values in testing dataset.

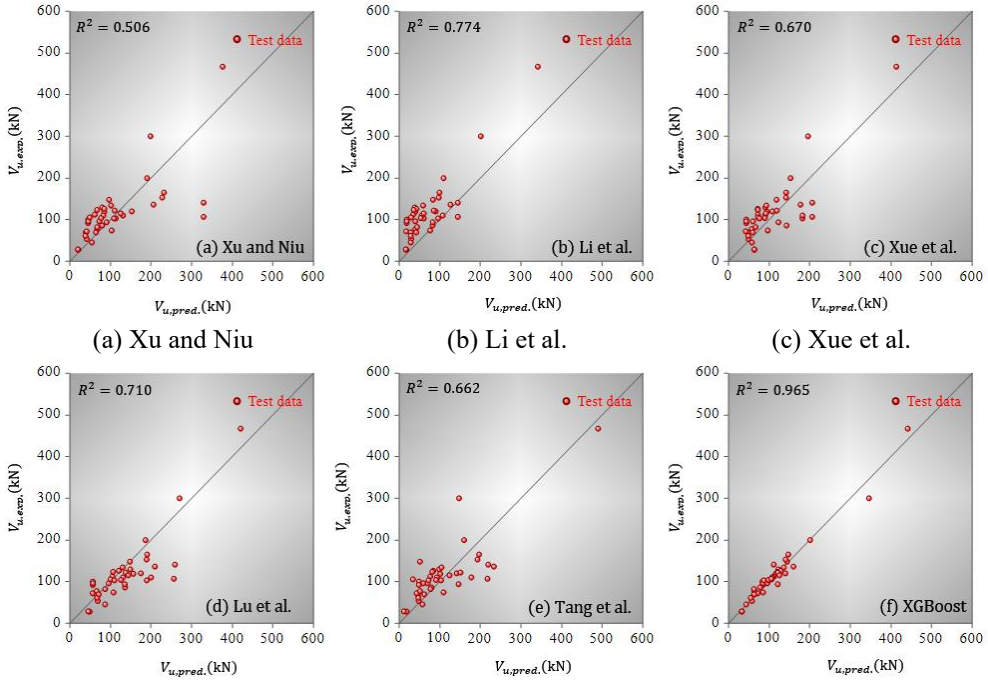
Fig. 9 shows that the residuals for testing datasets are significantly further from the zero line for analytical models developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], Tang et al. [75], while the residuals for the XGBoost model remain tightly clustered around the zero line. Furthermore, Fig. 10 shows that the prediction results of the analytical models developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], Tang et al. [75] are extremely scattered compared to the prediction results of the XGBoost model, meaning that the predicted values calculated using analytical models are far from the experimental results. Based on the statistical indicators in Table 8, the R^2 value of XGBoost is 0.965, while the closest R^2 value to this value belongs to Li et al. [90] with 0.774. Similarly, the STD, RMSE and COV values of XGBoost are 0.116, 13.725 and 0.113, respectively, while the closest values to these statistical indicators belong to Li et al. [90] with 0.169 for STD, Xue et al. [18] with 43.544 for RMSE and Lu et al. [27] with 0.332 for COV. It was understood that there are considerable differences even among analytical models with statistical indicators closest to the statistical indicators of the



(a) Xu and Niu (b) Li et al. (c) Xue et al.

(d) Lu et al. (e) Tang et al. (f) XGBoost

Figure 9. Residual analysis (V_{exp}-V_{pred}) with respect to V_{pred} between exiting models and XGBoost model



(a) Xu and Niu (b) Li et al. (c) Xue et al.

(d) Lu et al. (e) Tang et al. (f) XGBoost

Figure 10. Evaluation of $V_{u,pred.}$ against $V_{u,exp.}$ for existing models and XGBoost Model

XGBoost model, especially for STD and COV values. In addition, Table 8 shows that the model with the MV value closest to 1 belongs to XGBoost. Consequently, as evidenced by Figs. 9 and 10, as well as Table 8, the XGBoost model demonstrates clear superiority over the analytical models investigated in this study in terms of prediction accuracy and robustness. In addition, the statistical indicators obtained using the GEP model, which differs from other ML models in that it proposes the equation (Eq. (9)) as an output, are presented in Table 5 for testing dataset. According to Table 5 for testing dataset, the MV, STD, RMSE, R^2 and COV values of the GEP model are 1.025, 0.176, 17.851, 0.941, and 0.171, respectively. These statistical indicators of the proposed GEP equation exhibit clear superiority over the statistical indicators of the five analytical equations in the literature investigated in this study. As another result, it was demonstrated that the predictions based on the GEP model equation proposed in this study are more accurate than the predictions based on analytical equations developed by Xu and Niu [28], Li et al. [90], Xue et al. [18], Lu et al. [27], Tang et al. [75]. Table 8 shows that The MV value of Li et al. [90] is 2.166 for testing dataset. The value of this MV is the highest recorded in Table 8. This value means that the predictions are overly conservative. The MV value of Lu et al. [27] is 0.897 for all data. The value of this MV is the lowest recorded in Table 6. The MV value of Lu et al. [36] is 0.916 for testing dataset. This MV value also means that the predictions obtained using Lu et al. [36] model may be higher than the actual value. As shown in Table 8, the lowest COV value among the analytical models is 0.332, which is achieved analytical model proposed by Lu et al. [27]. Of the five analytical models analyzed in this study, analytical model proposed by Lu et al. [27] was determined to have the higher predictive accuracy, based on the statistical indicators in Table 8.

Table 9 provides information about the prediction tendency of the models on the testing dataset based on the average value. According to Table 9, it can be concluded that the model providing the closest prediction to the experimental value is the XGBoost model. In the XGBoost model, the MV value was between 0.95 and 1.05 in 16 out of 41 experimental samples. Among the predictions made with the XGBoost model, 1 remains on the highly unsafe side and 8 remains on the moderately unsafe side. It can be said that the prediction values of the remaining samples consist of 13 on the moderately conservative side and 3 on the highly conservative side. The tendency of the predictions made with XGBoost can be described as highly accurate and conservative. Although the predictions made by Xu and Niu [28], Li et al. [90], Xue et al. [18] and Tang et al. [75] tend to remain on the safe side, overly conservative results were obtained in many samples, especially in the model of Li et al. [90]. In the predictions made with the model of Lu et al. [27], highly unsafe predictions are observed in 15 samples and moderately unsafe predictions are seen in 9 samples. In the predictions made with the models of Xu and Niu [28], Xue et al. [18] and Tang et al. [75], the number of samples remaining on the highly unsafe side is remarkable.

Table 9. Conservativeness and reliability evaluation of model predictions in the testing dataset

Models	MV≤0.8	0.8<MV<0.95	0.95≤MV<1.05	1.05≤MV<1.2	MV≥1.2
XGBoost	1	8	16	13	3
Xu and Niu [28]	7	5	2	4	23
Li et al. [90]	1	-	4	3	33
Xue et al. [18]	11	-	1	7	22
Lu et al. [27]	15	9	5	8	4
Tang et al. [75]	8	4	2	7	20

Consequently, when Table 9 is examined, it can be stated that the XGBoost model makes more reliable and balanced predictions compared to the models in the literature.

3.3 SHAP Analysis

SHAP (SHapley Additive Explanations) is a game theory-based method that enhances the transparency and interpretability of complex machine learning models by decomposing predictions into the positive and negative contributions of the input parameters [120]. In this study, SHAP analysis was performed on the XGBoost model, which achieved the best performance in predicting the shear strength of CRC beams. During the analysis, SHAP dependency diagrams were first employed to evaluate the interactions of the key parameters, followed by a beeswarm plot to illustrate the relative importance of all input parameters. Subsequently, SHAP force plots were used to explain individual predictions, while waterfall plots were applied to step-by-step track parameter contributions in specific cases. Finally, a parameter contribution diagram was presented to show the percentage contributions of the parameters. All figures were created using testing data to reflect the model's generalization performance. This approach allows for a more reliable and unbiased assessment of the effects of parameters on cutting capacity [122, 123].

3.3.1 SHAP dependency analysis

In this section, SHAP dependency analysis was performed to examine the relationships between model input parameters and predicted shear capacity [120]. It should be noted that SHAP values reflect statistical relationships learned from the database and should not be interpreted as direct mechanical causality. Fig. 11 shows the effects of parameters on capacity and their interactions through color scales. Fig. 11(a) demonstrates that an increase in beam width (b) leads to higher SHAP values, thereby enhancing shear capacity. This effect becomes particularly pronounced once the width exceeds 200 mm, and the color scale (d) indicates that the positive influence is more substantial at greater effective depths. As illustrated in Fig. 11(b), SHAP contributions for beam height (h) increase sharply beyond 300 mm, with the color scale (d) revealing that the effect intensifies at larger effective depths. It is evident from Fig. 11(c) that SHAP contributions remain negative at small and medium beam effective depth (d) but become strongly positive above 300 mm, making a significant contribution to capacity; the color scale ($n_{w,sn}$) indicates that this effect is more pronounced at low loss rates. According to Fig. 11(d), SHAP values decrease as the shear span-to-effective depth ratio (a/d) increases, and the capacity-reducing effect becomes dominant, especially above 2. The color scale (d) reveals that this negative influence is more severe at greater effective depths. The results presented in Fig. 11(e) highlight the role of concrete cylinder compressive strength (f_c). Many points in the 15-25 MPa range remain at negative SHAP values, whereas contributions shift to distinctly positive above 30 MPa. The color scale (d) further demonstrates that the positive influence of (f_c) is more evident at greater effective depths. Fig. 11(f) demonstrates that SHAP values increase with the yield strength of stirrups (f_{yw}) and this effect becomes particularly pronounced at strengths above 500 MPa. As the longitudinal reinforcement ratio ρ_l increases, the positive contribution of stirrup yield strength (f_{yw}) to shear capacity becomes significantly stronger. In Fig. 11(g), positive contributions are concentrated at optimal levels (2-3%) of the longitudinal reinforcement ratio (ρ_l), while at both low and high ratios the effect is limited or negative. This positive influence is particularly pronounced at low shear span ratios (a/d). As depicted in Fig. 11(h), the effect of yield strength of

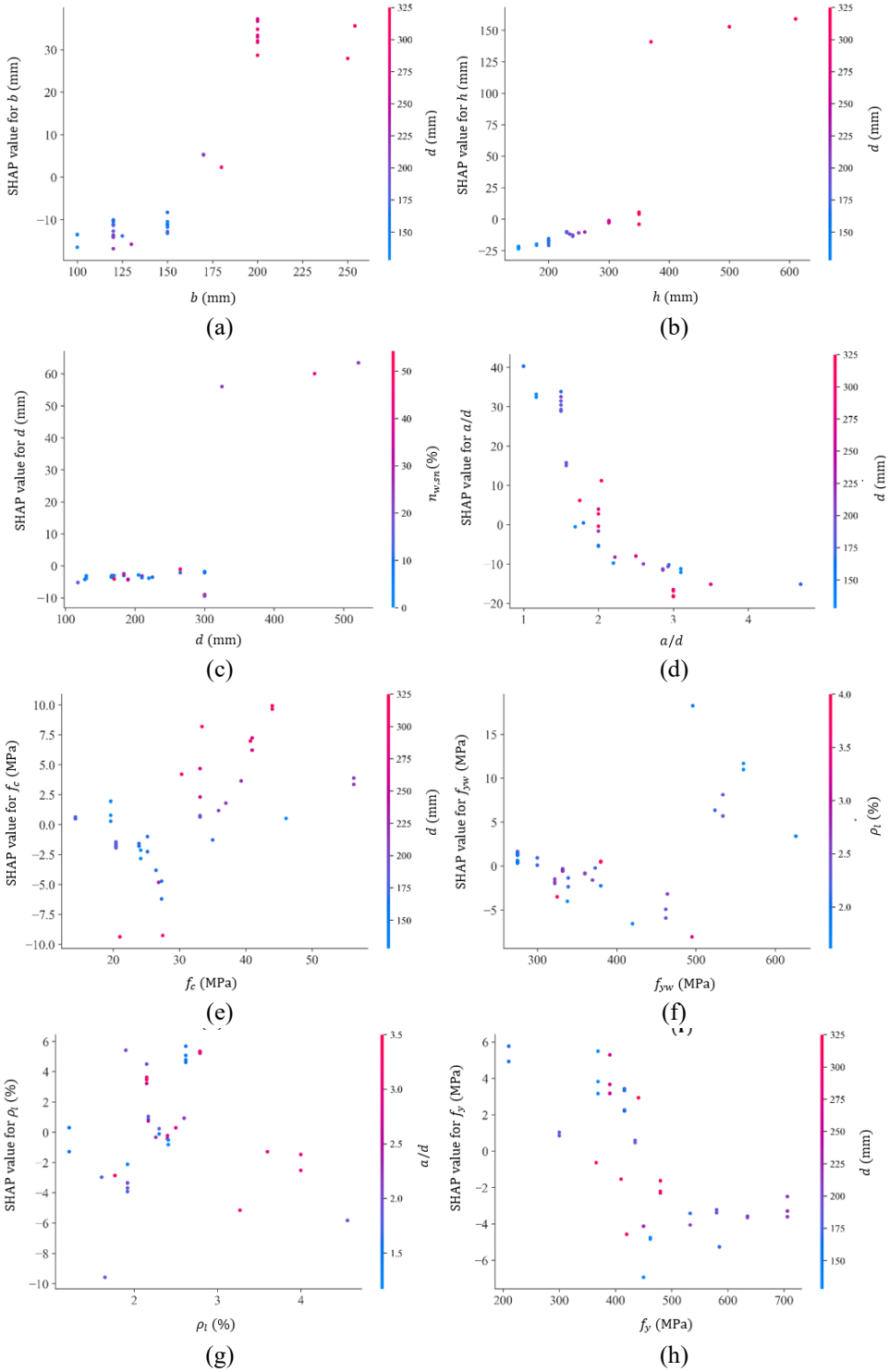


Figure 11. SHAP dependence plots of key input parameters (a) b, (b) h, (c) d, (d) a/d, (e) f_c, (f) f_{yw}, (g) ρ_l, (h) f_y, (i) ρ_w, (j) s, (k) $n_{l,sn}$, (l) $n_{w,sn}$

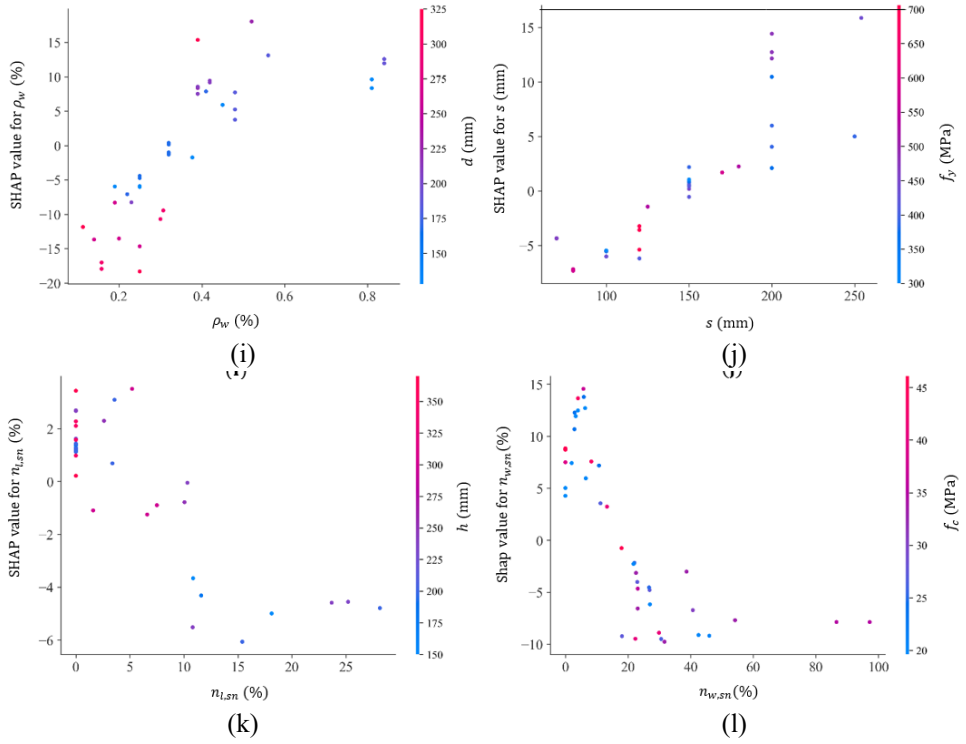


Figure 11. Continued

longitudinal reinforcement (f_y) is nonlinear: positive contributions emerge at medium strength levels, while both low and high values show restricted influence. This effect is less significant at greater effective depths (d). Fig. 11(i) shows that SHAP values become significantly positive as the stirrup reinforcement ratio (ρ_w) increases, and especially as the effective internal depth (d) increases. Fig. 11(j) shows that some larger stirrup spacing (s) values are associated with positive SHAP values. However, this should not be interpreted as a mechanical increase in shear capacity due to wider stirrup spacing, since such an interpretation is inconsistent with conventional shear-resistance mechanisms. This trend is more likely related to the distribution of data and correlations with other parameters, particularly the frequent coexistence of larger s values with higher longitudinal reinforcement yield strength (f_y) or other capacity-related variables. Fig. 11(k) shows that as the loss of longitudinal reinforcement area ($n_{l,sn}$) increases, SHAP values become clearly negative and shear capacity decreases, with the effect being stronger for beams with larger height (h). Finally, Fig. 11(l) shows limited positive contributions at low levels of the loss of cross-sectional area of stirrups ($n_{w,sn}$); however, above a 20% loss ratio, SHAP values quickly become negative and strongly reduce capacity. Here, higher concrete cylinder compressive strength (f_c) partly mitigates the negative effect.

To summarize, the SHAP dependency diagrams in Fig. 11 generally indicate that larger section dimensions (b, h, d), higher concrete compressive strength (f_c) and higher stirrup ratio (ρ_w) are associated with higher predicted shear capacity, whereas a high shear span-to-effective depth ratio (a/d) and corrosion-induced section losses in longitudinal and stirrup reinforcement ($n_{l,sn}$ and

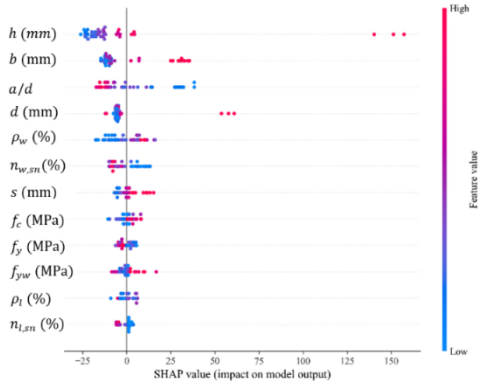


Figure 12. SHAP summary plot for feature contributions

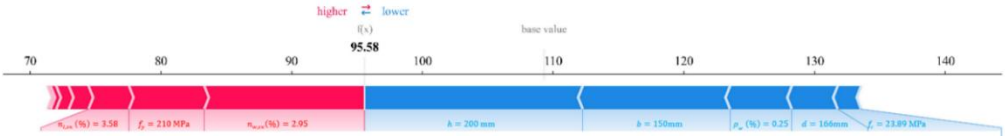


Figure 13. SHAP force plot for a single prediction

$n_{w,sn}$) are associated with lower predicted capacity. The color scales reveal that these effects do not occur independently but are influenced by interactions with other parameters.

3.3.2 SHAP-based interpretation of feature importance

Fig. 12 presents the SHAP plot showing the contribution of the parameters to the model prediction. Since SHAP values reflect the statistical relationships learned from the dataset, they should not be interpreted as direct mechanical causality. According to the results, beam height (h), beam width (b), effective depth (d), concrete compressive strength (f_c) and stirrup reinforcement ratio (ρ_w) are associated with higher predicted shear capacity. This trend is consistent with conventional shear behavior in terms of section size, concrete contribution, load transfer, and crack control. In contrast, the shear span-to-effective depth ratio (a/d) and the cross-sectional area losses of longitudinal and stirrup reinforcements ($n_{l,sn}$ and $n_{w,sn}$) negatively affect the predicted capacity. This can be explained by the reduction in arching action with increasing a/d and by the adverse effects of corrosion on reinforcement area, bond deterioration, and diagonal crack development. Stirrup spacing (s), yield strengths (f_y and f_{yw}), and longitudinal reinforcement ratio (ρ_l,) show relatively limited effects. In particular, the positive SHAP values observed for some large stirrup spacing values should not be interpreted as a mechanical increase in shear capacity; rather, they should be evaluated in relation to data distribution, correlations among variables, and the combined effects of other parameters. A possible mechanical explanation for this behavior is that, when the stirrups are severely corroded, their contribution to the shear capacity may remain limited even in beams with small stirrup spacing. Therefore, it can be considered that the SHAP analysis yielded positive values for beams with larger stirrup spacing.

3.3.3 Local interpretability analysis

Figs. 13 and 14 present the same prediction using different visualization methods. As seen in

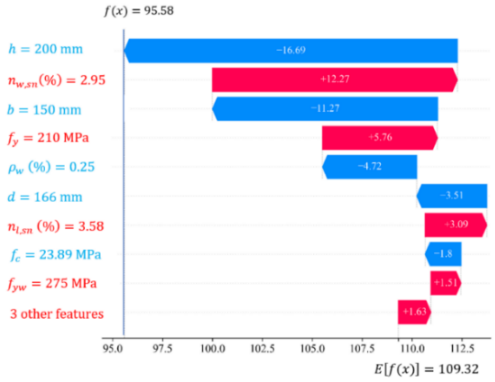


Figure 14. SHAP waterfall plot showing the impact of input parameters on a single shear capacity prediction

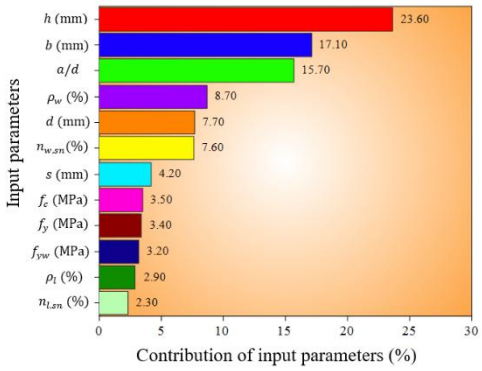


Figure 15. Relative contribution of input parameters

Fig. 13, the force plot illustrates the positive and negative contributions of the parameters with colored arrows, while Fig. 14, the waterfall plot, presents these contributions step by step with their numerical values. Both figures indicate that beam height (h) and beam width (b) have a capacity-reducing effect, whereas reinforcement and material parameters enhance capacity. In this way, the formation of a single prediction is explained transparently, both visually and analytically.

3.3.4 Relative contribution of parameters

The parameter contribution diagram presented in Fig. 15 quantitatively illustrates the relative importance of the variables used in predicting the shear capacity of corroded reinforced concrete beams. The results reveal that the highest contributions are provided by beam height h (23.60%), beam width b (17.10%) and the shear span-to-effective depth ratio a/d (15.70%). Moderately influential parameters include the stirrups reinforcement ratio ρ_w (8.70%), beam effective depth d (7.70%) and the loss of cross-sectional area of stirrups $n_{w,sn}$ (7.60%). More limited contributions are observed for the spacing of stirrups (4.20%), concrete cylinder compressive strength f_c (3.50%), yield strength of longitudinal reinforcement f_y (3.40%) and yield strength of stirrups f_{yw} (3.20%). The lowest contributions are found for the longitudinal reinforcement ratio ρ_l (2.90%) and the loss of cross-sectional area of longitudinal reinforcement $n_{l,sn}$ (2.30%). Overall, these results indicate that beam dimensions are the primary factors in predicting shear capacity,

whereas reinforcement properties and material strengths play a secondary role.

4. Conclusions

This study investigated the effectiveness of using eleven ML models to predict the shear capacity of CRC beams. Among these eleven ML models, the ML model that provided the statistically best prediction was selected. Furthermore, the prediction results of the selected best ML model were compared statistically with the prediction results of models in the literature. Finally, the SHAP analysis was conducted on the ML model that provided the best results. The results obtained are given below:

1. Analyzing the experimental studies in the dataset prepared for this study, it was observed that limited number of studies have conducted on understanding the effect of corrosion about deep beams, RC beams with high-strength concrete, RC beams with severe corrosion in both longitudinal reinforcement and stirrups. It is suggested for future experimental studies on corroded beams may focus on these issues to help us to better understand the effect of corrosion on the behavior of RC beams.
2. The ML models analyzed in this study and the proposed GEP equation are valid and applicable within the specified parameter ranges of the updated dataset. Given that the database includes corrosion in the longitudinal reinforcement, stirrups, or both, the developed predictive frameworks are tailored specifically for corroded reinforced concrete beams, rather than standard uncorroded control members.
3. The residual values of the predictions obtained from the eleven ML models were close to zero, meaning that the predicted values closely matched the experimental data. This demonstrated a strong correlation between the ML model predictions and the experimental results. Consequently, it can be concluded that the deployed ML models are highly reliable and exhibit an exceptional degree of accuracy in predicting the shear strength of corroded RC beams.
4. Evaluating the ML models individually, the statistical indicators (MV, STD, RMSE, and COV) show remarkable consistency between the training and testing subsets. This balanced performance confirms that the models are free from overfitting or underfitting, thereby verifying their robust generalization capacity for unseen data.
5. The high R^2 and low STD, RMSE and COV statistical indicators obtained from the prediction results of the eleven ML models indicate that these models are effective and usable for predicting the shear capacity of CRC beams. In addition, the MV values of the ML models used are very close to 1. In this situation, it demonstrates that ML prediction models are successful in capturing the actual value.
6. Among the ML models evaluated in this study, XGBoost was selected as the optimal model owing to its superior robustness and predictive accuracy
7. The XGBoost model demonstrates clear superiority over the five analytical models investigated in this study in terms of both prediction accuracy and robustness.
8. The XGBoost model makes more reliable and balanced predictions compared to five analytical the models.
9. The proposed GEP equation outperforms those of the five analytical formulations evaluated in this study.
10. Of the five analytical models analyzed in this study, the model proposed by Lu et al. [27]

was determined to have higher predictive accuracy, based on the statistical indicators.

11. In this study, SHAP analysis was performed on the XGBoost model, which achieved the best performance in predicting the shear strength of CRC beams.

12. SHAP dependency analysis revealed that the model produces consistent predictions. The b, h, d, f_c and ρ_w parameters consistently exhibit capacity-enhancing effects, whereas a high a/d, $n_{l,sn}$ and $n_{w,sn}$ parameters systematically reduce capacity. Moreover, the interactions demonstrate that the effects of these parameters are not isolated but are shaped in conjunction with other variables.

13. According to SHAP analysis, the beam height (h) exerts the most dominant effect on the model output, followed by the beam width (b) and shear span-to-effective depth ratio (a/d). These findings clearly demonstrate that structural behavior under corrosion is predominantly governed by section geometry. By contrast, the effects of material properties such as concrete cylinder compressive strength (f_c), yield strength of longitudinal reinforcement (f_y), and yield strength of stirrups (f_{yw}) are comparatively limited.

14. According to SHAP analysis, the lowest contributions are associated with the longitudinal reinforcement ratio (ρ_l) and loss of cross-sectional area of longitudinal reinforcement ($n_{l,sn}$).

Data availability statement

The dataset and source code are available from the corresponding author upon reasonable request.

References

1. Akkaya, H. C., Aydemir, C., Arslan, G. (2022). Investigation on shear behavior of reinforced concrete deep beams without shear reinforcement strengthened with fiber reinforced polymers. Case Studies in Construction Materials, 17, e01392. https://doi.org/10.1016/j.cscm.2022.e01392.
2. Fu, B., Feng, D.C. (2021). A machine learning-based time-dependent shear strength model for corroded reinforced concrete beams. Journal of Building Engineering, 36, 102118. https://doi.org/10.1016/j.jobe.2020.102118.
3. Nguyen, T.A., Ly, H.B. (2024). Hybrid machine learning techniques-aided design of corroded reinforced concrete beams. Computers and Structures, 298, 107388. https://doi.org/10.1016/j.compstruc.2024.107388.
4. Ye, Z., Zhang, W., Gu, X.D. (2018). Deterioration of shear behavior of corroded reinforced concrete beams. Engineering Structures, 168, 708-720. https://doi.org/10.1016/j.engstruct.2018.05.023.
5. Dogan, G. (2023). A simplified approach to determine shear strength for corroded RC beams. Düzce University Journal of Science and Technology, 11(4), 1958-1971. https://doi.org/10.29130/dubited.1293075.
6. Asteris, P.G., Nguyen, T.A. (2022). Prediction of shear strength of corrosion reinforced concrete beams using artificial neural network. Journal of Science and Transportation Technology, 2(2), 1-12. https://doi.org/10.58845/jstt.utt.2022.en.2.2.1-12.
7. Kumar, A., Arora, H.C., Kapoor, N.R., Kumar, K., Hadzima-Nyarko, M., Radu, D. (2023). Machine learning intelligence to assess the shear capacity of corroded reinforced concrete beams. Scientific Reports, 13(1), 2857. https://doi.org/10.1038/s41598-023-30037-9.
8. Val, D.V. (2007). Deterioration of strength of RC beams due to corrosion and its influence on beam

reliability. Journal of Structural Engineering, 133(9), 1297-1306. https://doi.org/10.1061/(ASCE)0733-9445(2007)133:9(1297).
9. Papadakis, V.G. (2000). Effect of supplementary cementing materials on concrete resistance against carbonation and chloride ingress. Cement and Concrete Research, 30(2), 291-299. https://doi.org/10.1016/S0008-8846(99)00249-5.
10. Vu, K.A.T., Stewart, M.G. (2000). Structural reliability of concrete bridges including improved chloride-induced corrosion models. Structural Safety, 22(4), 313-333. https://doi.org/10.1016/S0167-4730(00)00018-7.
11. Xia, J., Jin, W.L., Li, L.Y. (2011). Shear performance of reinforced concrete beams with corroded stirrups in chloride environment. Corrosion Science, 53(5), 1794-1805. https://doi.org/10.1016/j.corsci.2011.01.058.
12. Zhao, Y.X., Chen, J., Jin, W.L. (2009). Design of shear strengths of corroded reinforced concrete beams. International Journal of Modelling, Identification and Control, 7(2), 190-198. https://doi.org/10.1504/IJMIC.2009.027074.
13. Zhang, W., Ye, Z., Gu, X. (2017). Effects of stirrup corrosion on shear behaviour of reinforced concrete beams. Structural Infrastructure Engineering, 13(8), 1081-1092. https://doi.org/10.1080/15732479.2016.1243563.
14. Stewart, M.G. (2009). Mechanical behaviour of pitting corrosion of flexural and shear reinforcement and its effect on structural reliability of corroding RC beams. Structural Safety, 31(1), 19-30. https://doi.org/10.1016/j.strusafe.2007.12.001.
15. El-Sayed, A.K. (2017). Shear capacity assessment of reinforced concrete beams with corroded stirrups. Construction and Building Materials, 134, 176-184. https://doi.org/10.1016/j.conbuildmat.2016.12.118.
16. Li, W.W., Huang, J.Q., Lu, Y., Zhou, Y.W., Mansour, W., Kai, M. F., ..., Wang, P. (2024). Steel corrosion induced shear performance deterioration of RC beams: Experimental investigation and numerical simulation. Case Studies in Construction Materials, 20, e03266. https://doi.org/10.1016/j.cscm.2024.e03226.
17. Khan, I., François, R., Castel, A. (2014). Experimental and analytical study of corroded shear-critical reinforced concrete beams. Materials and Structures, 47(9), 1467-1481. https://doi.org/10.1617/s11527-013-0129-y.
18. Xue, X., Seki, H., Chen, Z.W. (2013). Shear capacity of RC beams containing corroded longitudinal bars. The thirteenth East Asia-Pacific conference on structural engineering and construction (EASEC-13), Sapporo, Japan, September.
19. Xue, X., Seki, H., Song, Y. (2014). Shear behavior of RC beams containing corroded stirrups. Advanced Structural Engineering, 17(2), 165-177. https://doi.org/10.1260/1369-4332.17.2.165.
20. Cladera, A., Marí, A., Ribas, C. (2021). Mechanical model for the shear strength prediction of corrosion-damaged reinforced concrete slender and non-slender beams. Engineering Structures, 247, 113163. https://doi.org/10.1016/j.engstruct.2021.113163.
21. Higgins, C., Farrow III, W.C., Turan, O.T. (2012). Analysis of reinforced concrete beams with corrosion damaged stirrups for shear capacity. Structural Infrastructure Engineering, 8(11), 1080-1092. https://doi.org/10.1080/15732479.2010.504213.
22. Campione, G., Cannella, F., Cavaleri, L. (2017). Shear and flexural strength prediction of corroded RC beams. Construction and Building Materials, 149, 395-405. https://doi.org/10.1016/j.conbuildmat.2017.05.125.
23. Potisuk, T., Higgins, C.C., Miller, T.H., Yim, S.C. (2011). Finite element analysis of reinforced concrete beams with corrosion subjected to shear. Advances in Civil Engineering, 2011, 706803. https://doi.org/10.1155/2011/706803.
24. Coronelli, D., Gambarova, P. (2004). Structural assessment of corroded reinforced concrete beams: modeling guidelines. Journal of Structural Engineering, 130(8), 1214-1224. https://doi.org/10.1061/(ASCE)0733-9445(2004)130:8(1214).
25. Huang, L., Ye, H., Jin, X., Jin, N., Xu, Z. (2020). Corrosion-induced shear performance degradation of reinforced concrete beams. Construction and Building Materials, 248, 118668.

https://doi.org/10.1016/j.conbuildmat.2020.118668.
26. Imam, A., Azad, A.K. (2016). Prediction of residual shear strength of corroded reinforced concrete beams. International Journal of Advanced Structural Engineering, 8(3), 307-318. https://doi.org/10.1007/s40091-016-0133-x.
27. Lu, Z.H., Li, H., Li, W., Zhao, Y.G., Dong, W. (2018). An empirical model for the shear strength of corroded reinforced concrete beam. Construction and Building Materials, 188, 1234-1248. https://doi.org/10.1016/j.conbuildmat.2018.08.123.
28. Xu, S.H., Niu, D.T. (2004). The shear behaviour of corroded simply supported reinforced concrete beam. Chinese Journal of Building Structures, 25(5), 98-104.
29. Zhao, Y.X., Jin, W.L. (2008). Analysis on shearing capacity of concrete beams with corroded stirrups. Journal of Zhejiang University (Engineering Science), 42(1), 19-24.
30. ACI Committee 318 (2008). Building code requirements for structural concrete and commentary (ACI 318-08). American Concrete Institute, Farmington Hills, MI, USA.
31. GB 50010-2010 (2010). Code for design of concrete structures. China Building Industry Press, Beijing, China.
32. EN 1992-1-1 (2004). Eurocode 2: Design of concrete structures. Part 1-1: General rules and rules for buildings. British Standards, Brussels, Belgium.
33. Dinh, N.H., Nguyen, C.V., Bui, Q.H. (2025). Mechanical model for shear strength prediction of corrosion-damaged reinforced concrete beams under various corrosion schemes. Engineering Structures, 335, 120357. https://doi.org/10.1016/j.engstruct.2025.120357.
34. Xue, X., Seki, H. (2010). Influence of longitudinal bar corrosion on shear behavior of RC beams. Journal of Advanced Concrete Technology, 8(2), 145-156. https://doi.org/10.3151/jact.8.145.
35. Alemu, Y., Onkareswara, N.B., Alemayehu, H. (2025). Comparative analysis of machine learning techniques for predicting flexural behavior in RC beams. Asian Journal of Civil Engineering,27(1), 355-365. https://doi.org/10.1007/s42107-025-01507-4.
36. Wakjira, T.G., Ibrahim, M., Ebead, U., Alam, M.S. (2022). Explainable machine learning model and reliability analysis for flexural capacity prediction of RC beams strengthened in flexure with FRCM. Engineering Structures, 255, 113903. https://doi.org/10.1016/j.engstruct.2022.113903.
37. Sudheer, M., Rao, B. (2025). Machine learning based boosting models for predicting flexural strength of steel fiber reinforced concrete. Asian Journal of Civil Engineering, 26(8), 3507-3517. https://doi.org/10.1007/s42107-025-01384-x.
38. Taffese, W.Z., Zhu, Y. (2025). Explainable machine learning for predicting flexural capacity of reinforced UHPC beams. Engineering Structures, 343, 121188. https://doi.org/10.1016/j.engstruct.2025.121188.
39. Wakjira, T.G., Abushanab, A., Ebead, U., Alnahhal, W. (2022). FAI: Fast, accurate, and intelligent approach and prediction tool for flexural capacity of FRP-RC beams based on super-learner machine learning model. Materials Today Communications, 33, 104461. https://doi.org/10.1016/j.mtcomm.2022.104461.
40. Nguyen, T.A., Ly, H.B. (2024). Predicting axial compression capacity of CFDST columns and design optimization using advanced machine learning techniques. Structures, 59, 105724. https://doi.org/10.1016/j.istruc.2023.105724.
41. Deifalla, A., Salem, N.M. (2022). A machine learning model for torsion strength of externally bonded FRP-reinforced concrete beams. Polymers, 14(9), 1824. https://doi.org/10.3390/polym14091824.
42. Worden, K., Manson, G. (2007). The application of machine learning to structural health monitoring. Philosophical Transactions of the Royal Society A, 365(1851), 515-537. https://doi.org/10.1098/rsta.2006.1938.
43. Gui, G., Pan, H., Lin, Z., Li, Y., Yuan, Z. (2017). Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection. KSCE Journal of Civil Engineering, 21(2), 523-534. https://doi.org/10.1007/s12205-017-1518-5.
44. Zhang, Y., Burton, H. V., Sun, H., Shokrabadi, M. (2018). A machine learning framework for assessing post-earthquake structural safety. Structural Safety, 72, 1-16. https://doi.org/10.1016/j.strusafe.2017.12.

001.
45. Mangalathu, S., Hwang, S.H., Choi, E., Jeon, J.S. (2019). Rapid seismic damage evaluation of bridge portfolios using machine learning techniques. Engineering Structures, 201, 109785. https://doi.org/10.1016/j.engstruct.2019.109785.
46. Tang, Y., Wang, W., Huang, Q., Cheng, Y., Zhou, C. (2025). Shear performance prediction of RC beams shear-strengthened with FRP sheet: A machine learning driven design-oriented method. Engineering Structures, 334, 120240. https://doi.org/10.1016/j.engstruct.2025.120240.
47. Shu, J., Yu, H., Liu, G., Yang, H., Guo, W., Phoon, C., ..., Hu, H. (2024). Proposing an inherently interpretable machine learning model for shear strength prediction of reinforced concrete beams with stirrups. Case Studies in Construction Materials, 20, e03350. https://doi.org/10.1016/j.cscm.2024.e03350.
48. Alshboul, O., Almasabha, G., Al-Shboul, K.F., Shehadeh, A. (2023). A comparative study of shear strength prediction models for SFRC deep beams without stirrups using machine learning algorithms. Structures, 55,97-111. https://doi.org/10.1016/j.istruc.2023.06.026.
49. Ma, K., Qiao, L., Lin, G., Xing, G. (2024). Prediction of the shear strength of lightweight concrete beams without web reinforcement based on a machine learning model optimized by a genetic algorithm. Structures, 65, 106738. https://doi.org/10.1016/j.istruc.2024.106738.
50. Ergen, F., Katlav, M. (2024). Investigation of optimized machine learning models with PSO for forecasting the shear capacity of steel fiber-reinforced SCC beams with/out stirrups. Journal of Building Engineering, 83, 108455. https://doi.org/10.1016/j.jobe.2024.108455.
51. Arif, M., Luqman, M., Khawaja, U.I., Allah, N. (2025). Improved machine learning modeling for predicting shear capacity of RC deep beams. Civil Engineering and Environmental Engineering Resilient Smart Sustainable Solutions, 48, 252. https://doi.org/10.21741/9781644903414-28.
52. Yang, Y., Liu, G. (2023). Data-driven shear strength prediction of FRP-reinforced concrete beams without stirrups based on machine learning methods. Buildings, 13(2), 313. https://doi.org/10.3390/buildings13020313.
53. David, A.B., Olalusi, O.B., Awoyera, P.O., Simwanda, L. (2024). Suitability of mechanics-based and optimized machine learning-based models in the shear strength prediction of slender beams without stirrups. Buildings, 14(12), 3946. https://doi.org/10.3390/buildings14123946.
54. Zhang, J., Sun, Y., Li, G., Wang, Y., Sun, J., Li, J. (2022). Machine-learning-assisted shear strength prediction of reinforced concrete beams with and without stirrups. Engineering Computations, 38(2), 1293-1307. https://doi.org/10.1007/s00366-020-01076-x.
55. Mohammed, H.R.M., Ismail, S. (2022). Proposition of new computer artificial intelligence models for shear strength prediction of reinforced concrete beams. Engineering Computations, 38(4), 3739-3757. https://doi.org/10.1007/s00366-021-01400-z.
56. Yehia, S.A., Fayed, S., Zakaria, M.H., Shahin, R.I. (2024). Prediction of RC T-beams shear strength based on machine learning. International Journal of Concrete Structures and Materials, 18(1), 52. https://doi.org/10.1186/s40069-024-00690-z.
57. Uddin, M.N., Yu, K., Li, L. Z., Ye, J., Tafsirojjaman, T., Alhaddad, W. (2022). Developing machine learning model to estimate the shear capacity for RC beams with stirrups using standard building codes. Innovative Infrastructure Solutions, 7(3), 227. https://doi.org/10.1007/s41062-022-00826-8.
58. Megahed, K. (2024). Prediction and reliability analysis of shear strength of RC deep beams. Scientific Reports, 14(1), 14590. https://doi.org/10.1038/s41598-024-64386-w.
59. Wakjira, T.G., Ebead, U., Alam, M.S. (2022). Machine learning-based shear capacity prediction and reliability analysis of shear-critical RC beams strengthened with inorganic composites. Case Studies in Construction Materials, 16, e01008. https://doi.org/10.1016/j.cscm.2022.e01008.
60. Alacalı, S., Arslan, G. (2025). Assessment of shear strength of steel fiber reinforced concrete (SFRC) beams using gene expression programming. Computers and Concrete, 35(5), 537-555. https://doi.org/10.12989/cac.2025.35.5.537.
61. Alacali, S., Arslan, G. (2024). A prediction model for plastic hinge length of rectangular RC columns using gene expression programming. Neural Computing and Applications, 36(16), 9481-9501.

https://doi.org/10.1007/s00521-024-09578-1.
62. Alacalı, S., Akkaya, H.C., Sengun, K., Arslan, G. (2024). Proposal and evaluation of new models for predicting the FRP contribution to shear strength in reinforced concrete beams using gene expression programming. Neural Computing and Applications, 36(25), 15515-15544. https://doi.org/10.1007/s00521-024-09892-8.
63. Alacalı, S. (2022). A prediction model for strength and strain of CFRP-confined concrete cylinders using gene expression programming. Computers and Concrete, 30(6), 377-391. https://doi.org/10.12989/cac.2022.30.6.377.
64. Akkaya, H.C., Alacalı, S. (2025). Developing a new model using gene expression programming on the prediction of moment capacity of ferrocement elements. Bitlis Eren University Journal of Science, 14(2), 713-731. https://doi.org/10.17798/bitlisfen.1583956.
65. Rodriguez, J., Ortega, L.M., Casal, J. (1997). Load carrying capacity of concrete structures with corroded reinforcement. Construction and Building Materials, 11(4), 239-248. https://doi.org/10.1016/S0950-0618(97)00043-3.
66. Higgins, C., Farrow, W.C. III (2006). Test of reinforced concrete beams with corrosion damaged stirrups. ACI Structural Journal, 103(1), 133-141.
67. Li, X.T., Yin, H.G. (2010). Degradation mechanism and predicting models of shearing capacity for corroded reinforced concrete beams. Journal of Xuzhou Institute of Technology (Natural Science Edition), 25(4), 58-63.
68. El-Sayed, A.K., Hussain, R.R., Shuraim, A.B. (2016). Influence of stirrup corrosion on shear strength of reinforced concrete slender beams. ACI Structural Journal, 113(6), 1223-1234.
69. Juarez, C.A., Guevara, B., Fajardo, G., Castro-Borges, P. (2011). Ultimate and nominal shear strength in reinforced concrete beams deteriorated by corrosion. Engineering Structures, 33(12), 3189-3196. https://doi.org/10.1016/j.engstruct.2011.08.014.
70. Liu, S.T. (2013). The research on shear capacity of corroded RC beams. M.S. Thesis, Central South University, Changsha, Hunan, China.
71. Fu, C., Huang J., Dong Z. (2023). Shear behavior of reinforced concrete beams subjected to accelerated non-uniform corrosion. Engineering Structures, 286, 116081. https://doi.org/10.1016/j.engstruct.2023.116081.
72. Wang, L., Zhang, X., Zhang, J., Ma, Y., Liu, Y. (2015). Effects of stirrup and inclined bar corrosion on shear behavior of RC beams. Construction and Building Materials, 98, 537-546. https://doi.org/10.1016/j.conbuildmat.2015.07.077.
73. Alaskar, A., Alqarni, A.S., Alfalah, G., El-Sayed, A.K., Mohammadhosseini, H., Alyousef, R. (2021). Performance evaluation of reinforced concrete beams with corroded web reinforcement: experimental and theoretical study. Journal of Building Engineering, 35, 102038. https://doi.org/10.1016/j.jobe.2020.102038.
74. Han, S.J., Joo, H.E., Choi, S.H., Heo, I., Kim, K.S., Seo, S.Y. (2019). Experimental study on shear capacity of reinforced concrete beams with corroded longitudinal reinforcement. Materials, 12, 837. https://doi.org/10.3390/ma12050837.
75. Tang, Z., He, Z., Chen, Z., Chen, L., Xue, H., Zhuge, H. (2021). Experimental and numerical study on ultimate shear load carrying capacity of corroded RC beams. Advances in Civil Engineering, 2021(1), 4848483. https://doi.org/10.1155/2021/484848.
76. Ming, D., Qiu, M., Yuan, L., Shui, Y., Shi, G. (2025). Study on shear capacity of corroded reinforced concrete. Iranian Journal of Science and Technology, Transactions of Civil Engineering, 49(4), 3759-3777. https://doi.org/10.1007/s40996-024-01700-6.
77. Xue, X., Cheng, J. (2025). An improved method for predicting the shear strength of corroded reinforced concrete beams with experimental and numerical validation. Structure and Infrastructure Engineering, 21, 814-825. https://doi.org/10.1080/15732479.2023.2242822.
78. Sneed, L.H., Ramirez, J.A. (2010). Influence of effective depth on the shear strength of concrete beams-Experimental study. ACI Structural Journal, 107, 554-562.
79. Chen, H., Yi, W.J., Ma, Z.J. (2019). Shear size effect in simply supported RC deep beams. Engineering

Structures, 2019, 268-278. https://doi.org/10.1016/j.engstruct.2018.12.062.
80. Akkaya, H.C., Aydemir, C., Arslan, G. (2022). An experimental research on reinforced concrete deep beams fully wrapped with fiber reinforced polymers against shear. Case Studies in Construction Materials, 17, e01198. https://doi.org/10.1016/j.cscm.2022.e01198.
81. Tompos, E.J., Frosch, R.J. (2002). Influence of beam size, longitudinal reinforcement, and stirrup effectiveness on concrete shear strength. ACI Structural Journal, 99, 559-567. 10.14359/12295.
82. Li, S.S., Jin, T.C., Zheng, L.A., Zhang, G.Y., Li, H.M., Chen, A.J., Xie, W. (2023). Experimental study of shear performance of high-strength concrete deep beams with longitudinal reinforcement with anchor plate. Materials, 16, 6023. https://doi.org/10.3390/ma16176023.
83. Russo, G., Mitri, D., Pauletta, M. (2013). Shear strength design formula for RC beams with stirrups. Engineering Structures, 51, 226-235. https://doi.org/10.1016/j.engstruct.2013.01.024.
84. Akkaya, H.C., Aydemir, C., Arslan, G. (2024). Evaluation of shear behavior of short-span reinforced concrete deep beams strengthened with fiber reinforced polymer strips. Engineering Structures, 299, 117145. https://doi.org/10.1016/j.engstruct.2023.117145.
85. Cladera, A., Mari, A.R. (2005). Experimental study on high-strength concrete beams failing in shear. Engineering Structures, 27 (10), 1519-1527. https://doi.org/10.1016/j.engstruct.2005.04.010.
86. Higgins, C., Farrow, W.C. (2006). Tests of reinforced concrete beams with corrosion-damaged stirrups. ACI Structural Journal, 103 (1), 133-141.
87. Jeppsson, J., Thelandersson, S. (2003). Behavior of reinforced concrete beams with loss of bond at longitudinal reinforcement. Journal of Structural Engineering, 129 (10), 1376-1383. https://doi.org/10.1061/(ASCE)0733-9445(2003)129:10(1376).
88. Azam, R., Soudki, K. (2013). Structural behavior of shear-critical RC slender beams with corroded properly anchored longitudinal steel reinforcement. Journal of Structural Engineering, 139(12), 04013011. https://doi.org/10.1061/(ASCE)ST.1943-541X.000079.
89. Azam, R., Soudki, K. (2012). Structural performance of shear-critical RC deep beams with corroded longitudinal steel reinforcement. Cement and Concrete Composites, 34(8), 946-957. https://doi.org/10.1016/j.cemconcomp.2012.05.003.
90. Li, S.B., Zhang, X., Jia, L.D. (2011). Analysis for shear capacity of reinforced concrete beams with corrosion stirrups. Chinese Journal of Engineering Mechanics, 28, 60-63.
91. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J. (1984). Classification and regression trees. Wadsworth International Group, Monterey, CA, USA.
92. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324.
93. Geurts, P., Ernst, D., Wehenkel, L. (2006). Extremely randomized trees. Machine Learning, 63(1), 3-42. https://doi.org/10.1007/s10994-006-6226-1.
94. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y. (2017). Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems 30 (NIPS 2017), Long Beach, CA, USA, December.
95. Freund, Y., Schapire, R.E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55 (1), 119-139. https://doi.org/10.1006/jcss.1997.1504.
96. Chen, T., Guestrin, C. (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, San Francisco, CA, USA, August.
97. Dorogush, A.V., Ershov, V., Gulin, A. (2018). Catboost: Gradient boosting with categorical features support. arXiv preprint, arXiv:1810.11363. https://doi.org/10.48550/arXiv.1810.11363.
98. Smola, A.J., Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and Computing, 14(3), 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.
99. Williams, C.K., Rasmussen, C.E. (2006). Gaussian processes for machine learning. MIT Press, Cambridge, MA, USA.
100. Ferreira, C. (2001). Gene expression programming: A new adaptive algorithm for solving problems.

Complex Systems, 13(2), 87-129. https://doi.org/10.48550/arXiv.cs/0102027.
101. Safavian, S.R., Landgrebe, D. (1991). A survey of decision tree classifier methodology. IEEE Transactions on Systems, Man, and Cybernetics, 21(3), 660-674. https://doi.org/10.1109/21.97458.
102. Loh, W.Y. (2011). Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(1), 14-23. https://doi.org/10.1002/widm.8.
103. Liaw, A., Wiener, M. (2002). Classification and regression by randomforest. R News, 2(3), 18-22.
104. Louppe, G. (2014). Understanding random forests: From theory to practice. arXiv preprint, arXiv:1407.7502.
105. Schapire, R.E., Freund, Y. (2012). Boosting: Foundations and algorithms. MIT Press, Cambridge, MA, USA.
106. Drucker, H. (1997). Improving regressors using boosting techniques. Proceedings of the 14th international conference on machine learning, Nashville, TN, USA, July.
107. Dietterich, T.G. (2000). Ensemble methods in machine learning. Multiple classifier systems, first international workshop, MCS 2000, Cagliari, Italy, June.
108. Bentéjac, C., Csörgő, A., Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54(3), 1937-1967. https://doi.org/10.1007/s10462-020-09896-5.
109. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A.V., Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. Advances in neural information processing systems 31 (NeurIPS 2018), Montréal, Quebec, Canada, December.
110. Cortes, C., Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. https://doi.org/10.1007/BF00994018.
111. Vapnik, V.N. (1998). Statistical learning theory. John Wiley & Sons Inc., Hoboken, NJ, USA.
112. Schölkopf, B., Smola, A.J. (2002). Learning with kernels: Support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge, MA, USA.
113. Bishop, C.M. (2006). Pattern recognition and machine learning. Springer, New York, NY, USA.
114. Quiñonero-Candela, J., Rasmussen, C.E. (2005). A unifying view of sparse approximate gaussian process regression. Journal of Machine Learning Research, 6, 1939-1959.
115. Ferreira, C. (2006). Gene expression programming: Mathematical modeling by an artificial intelligence (2nd ed.), Springer-Verlag, Berlin, Germany.
116. Ferreira, C. (2025). Gepsoft: Gene expression programming software. Version 5.0, Gepsoft Ltd, Bristol, UK. https://www.gepsoft.com.
117. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ..., Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. The Journal of Machine Learning Research, 12, 2825-2830.
118. Chai, T., Draxler, R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 7(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014.
119. Akkaya, H.C., Alacalı, S. (2026). Predicting the punching shear capacity of RC slab-column connections with FRP bars using machine learning based algorithms. Latin American Journal of Solids and Structures, 23(2), e8817. https://doi.org/10.1590/1679-7825/e8817.
120. Lundberg, S.M., Lee, S.I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems 30 (NIPS 2017), Long Beach, CA, USA, December.
121. Williams, C.K.I., Rasmussen, C.E. (1996). Gaussian processes for regression. Advances in neural information processing systems 8 (NIPS 1995), Denver, CO, USA, November.
122. Maleki, F., Ovens, K., Gupta, R., Reinhold, C., Spatz, A., Forghani, R. (2022). Generalizability of machine learning models: Quantitative evaluation of three methodological pitfalls. Radiology: Artificial Intelligence, 5(1), e220028. https://doi.org/10.1148/ryai.220028.
123. Salehinejad, H., Kitamura, J., Ditkofsky, N., Lin, A., Bharatha, A., Suthiphosuwan, S., ..., Colak, E. (2021). A real-world demonstration of machine learning generalizability: Intracranial hemorrhage detection on head CT. arXiv preprint, arXiv:2102.04869. https://doi.org/10.48550/arXiv.2102.04869.

124. Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th international joint conference on artificial intelligence, Montreal, Quebec, Canada, August.
125. Project Jupyter (2025). Jupyterlab. Version 4.4.4. Project Jupyter. https://jupyter.org.
126. Anaconda, Inc (2025). Anaconda navigator. Version 24.3. Anaconda, Inc, Austin, TX, USA. https://www.anaconda.com.
127. Asamoto, S., Sato, J., Okazaki, S., Chun, P.J., Sahamitmongkol, R., Nguyen, G.H. (2021). The cover depth effect on corrosion-induced deterioration of reinforced concrete focusing on water penetration: Field survey and laboratory study. Materials, 14(13), 3478. https://doi.org/10.3390/ma14133478.
128. Khalid, M.A., Shinichi, M. (2024). Influence of partial cover thickness and mortar quality on steel corrosion in a chloride environment. Cogent Engineering, 11(1), 2297474. https://doi.org/10.1080/23311916.2023.2297474.
129. Jiang, H., Wang, S., Ren, Y., Li, Y., Song, S., Qin, Y. (2023). Safety analysis of rebar corrosion depth at the moment of corrosion-induced cover cracking. Sustainability, 15(3), 2491. https://doi.org/10.3390/su15032491.