

RESEARCH

Open Access



Prediction of RC T-Beams Shear Strength Based on Machine Learning

Saad A. Yehia^{1*} , Sabry Fayed², Mohamed H. Zakaria² and Ramy I. Shahin¹

Abstract

The contribution of shear resisted by flanges of T-beams is usually ignored in the shear design models even though it was proven by many experimental studies that the shear strength of T-beams is higher than that of equivalent rectangular cross-sections. Ignoring such a contribution result in a very conservative and uneconomical design. Therefore, the aim of this research is to investigate the capability of machine learning (ML) techniques to predict the shear capacity of reinforced concrete T-beams (RCTBs) by incorporating the contribution of the flange. Five machine learning (ML) techniques, which are the Decision Tree (DT), Random Forest (RF), Gradient Boosting Regression Tree (GBRT), Light Gradient Boosting Machine (LightGBM), and Extreme Gradient Boosting (XGBoost), are trained and tested using 360 sets of data collected from experimental studies. Among the various machine learning models evaluated, the XGBoost model demonstrated exceptional reliability and precision, achieving an R-squared value of 99.10%. The SHapley Additive exPlanations (SHAP) approach is utilized to identify the most influential input features affecting the predicted shear capacity of RCTBs. The SHAP results indicate that the shear span-to-depth ratio (a/d) has the most significant effect on the shear capacity of RCTBs, followed by the ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$), flange thickness (h_f), and flange width (b_f). The accuracy of the XGBoost model in predicting the shear capacity of RCTBs is compared with established codes of practice (ACI 318-19, BS 8110-1:1997, EN 1992-1-2, CSA23.3-04) and existing formulas from researchers. This comparison reinforces the superior reliability and accuracy of the machine learning approach compared to traditional methods. Furthermore, a user-friendly interface platform is developed, effectively simplifying the implementation of the proposed machine-learning model. The reliability analysis is performed to determine the value of the resistance reduction factor (ϕ) that will achieve a target reliability index ($\beta_T = 3.5$).

Keywords Reinforced concrete T-beams, Shear strength, Machine learning (ML), Extreme Gradient Boosting model (XGBoost), Decision tree model (DT), Random forest model (RF)

1 Introduction

RC beams with T-cross sections are utilized in a variety of construction applications, including bridge decks, building floors/slabs, retaining walls, and parking garages (Ayensa et al., 2019; Ramadan et al., 2022; Ribas González et al., 2017). The advantage of T-sections lies in the fact that a portion of the slab acts integrally with the beam, bending along with it under the loads to effectively resist straining actions. On the contrary, the shear resistance of RCTBs is typically assessed solely based on the area of the beam web. The contribution of the flanges to the

Journal information: ISSN 1976-0485 / eISSN 2234-1315.

*Correspondence:

Saad A. Yehia
saad_yehia@kfs-hiet.edu.eg

¹ Civil Engineering Department, Higher Institute of Engineering and Technology, Kafrelsheikh 33511, Egypt

² Civil Engineering Department, Faculty of Engineering, Kafrelsheikh University, Kafrelsheikh 33511, Egypt



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

shear strength is commonly overlooked in shear design models (Zararis IP et al., 2006). Nevertheless, recent studies have demonstrated a notable contribution of the flanges to the shear strength of T-beams (Amna & Monstaser, 2019; Ayensa et al., 2019; Bresler & MacGregor, 1967; Elgohary et al., 2019; Giaccio et al., 2011; Hawileh et al., 2022; Mhanna et al., 2020, 2021a, 2021b; Ramadan et al., 2022; Samad et al., 2016; Sarsam et al., 2018; Swamy et al., 1974; Thamrin et al., 2016).

In the study by Ramadan et al., (2022), the authors conducted a study on the shear behavior of T-beams, highlighting the impact of the beam flange area on shear strength. The results of the experiment showed that T-beams had a higher shear strength compared to rectangular beams with the same web size. Furthermore, the research discovered that elevating the ratio of flange thickness-to-beam depth from 0.3 to 0.5 within the examined range of variables caused an increase in shear strength of up to 54%. Likewise, increasing the ratio of flange width-to-web width from 3 to 5 led to an escalation in shear strength of up to 19%. Ayensa et al., (2019) displayed that the significance of the flange can result in considerable cost savings during the construction of new structures. It can also be a critical factor when assessing the shear strength of existing systems. Test results of a study by Bresler and MacGregor, (1967) concluded that the geometry of the beam, particularly for I and T-sections, has a significant impact on the shear strength and behavior due to the varying magnitudes of shearing stress developed in the web. This variation also affects the propagation of diagonal cracking. Also, Amna et al., (2019) presented an experimental and numerical model to investigate the shear behavior from using lightweight concrete beams. The research findings indicated a notable rise in the ultimate load as the shear span to depth ratio decreased. Conversely, there was an inverse relationship between the failure load and the shear span to depth ratio. Sarsam et al., (2018) conducted a literature review to investigate how flanges impact the shear capacity of RCTBs. Based on their analysis, they developed an equation to predict the shear capacity of RCTBs. Swamy and Qureshi, (1974) developed a procedure for determining the maximum shear strength of the compression zone in T-beams. This procedure was created based on Mohr's theory of failure and the biaxial stress criteria concept. However, this strategy requires a series of calculations and is complicated to implement into practice. The shear mechanism of RCTBs with simply supported subjected to a 4-point bending test was examined by Samad et al., (2016). The study encompassed similar variables and parameters, such as a shear span-to-effective depth ratio (a/d) of 3.5 and a longitudinal reinforcement of $\rho=2.15\%$. Results indicate significant differences between EC2

and ACI318-08 design codes when calculating the shear capacity V_m , and concrete shear resistance V_c of T-beams. In their study, Thamrin et al., (2016) found that the shear capacity of the T-beams tested was notably influenced by both the ratio of longitudinal reinforcement, and the size of the flange. Where the shear capacity of T-beams is generally reported to be 5–20% higher than that of beams with a rectangular cross-section, and current codes are using conservative empirical equations to calculate the shear capacity of beams. Elgohary et al., (2019) presented an experimental test of RCTBs without stirrups, where the flange width-to-web width ratio was varied. The test results indicated that the shear strength was improved by 10–30% owing to the involvement of the flange. Furthermore, some theoretical models, as demonstrated in the references (Bairan Garcia & Mari Bernat, 2006; Cladera et al., 2015; Kotsovos et al., 1987; Moayer & Regan, 1974; Tureyen & Frosch, 2003; Wolf & Frosch, 2007; Zararis IP et al., 2006), have considered the contribution of flanges to shear strength. Based on prior literature, it is evident that existing design formulas cannot accurately forecast the shear capacity of flanged sections, primarily due to the intricate nature of shear failure in T-sections. This underscores the growing necessity for more accurate prediction methodologies that account for the augmented shear capacity resulting from the inclusion of flanges in T-sections.

Modern machine learning (ML) algorithms present a promising solution to address these concerns, as they excel at handling complex problems with multiple variables without relying on assumptions. In the past decade, data-driven approaches including various ML techniques have had an increasing presence and importance in a wide range of civil engineering applications, including geotechnical (T. Q. Huynh et al., 2022; Vadyala et al., 2022), structural (Le-Nguyen et al., 2022; H. D. Nguyen et al., 2021a, 2021b, 2022; Shahin et al., 2023), material perspectives (J. Schmidt et al., 2019; J. Wei et al., 2019; Nguyen-Sy et al., 2020; Wang et al., 2020), etc. Moreover, ML techniques have demonstrated significant success in estimating the shear strength of RC beams. Several studies have utilized ML methods to predict the shear strength of RC beams, yielding promising results. For example, Goh (Goh, 1995) was among the pioneers in utilizing Artificial Neural Networks (ANNs) to predict the ultimate shear strength of deep beams. Similarly, Zhang et al., (2022) applied random forest (RF) in conjunction with an optimization technique to predict the RC beams shear strength, including the stirrup's effect on beam strength. Cladera and Mari, (2004) and Jung and Kim, (2008) utilized experimental data to train and test an artificial neural network (ANN)-based model for estimating the shear strength of RC beams. Additionally, Ashour

et al., (2003) and Alshboul et al., (2022) introduced a gene expression (GEP)-based model designed to calculate the shear strength of RC beams. Also, a variety of ML algorithms have been employed to predict the shear strength of steel fiber-reinforced concrete and ultra-high-performance fiber-reinforced concrete beams. Support Vector Machine (SVM), Artificial Neural Networks (ANN), k-Nearest Neighbor (k-NN), Gene Expression Programming (GEP), Decision Tree (DT), Random Forest (RF), XGBoost, AdaBoost, Gradient Boosting Regression Tree (GBRT), and CatBoost models (Abambres & Lantsoght, 2019; Chaabene & Nehdi, 2020; Jiang & Liang, 2021; Keshtegar et al., 2019; Rahman et al., 2021; Sarveghadi et al., 2019; Shahnewaz & Alam, 2020; Shatnawi et al., 2022; Solhmirzaei et al., 2020; Xiangyong Ni & Kang-kang Duan, 2022; Yaseen et al., 2018; Ye et al., 2023) have been extensively utilized for this purpose. These models offer diverse approaches to accurately predict the shear strength of RC beams, contributing to advancements in structural engineering research and design.

While machine learning (ML) proves highly beneficial in enhancing metamodels and enjoys widespread popularity among researchers, its implementation in structural engineering applications presents two significant challenges. The primary hurdle lies in the crucial initial step of converting input features into numerical data suitable for the algorithm's consumption. Meanwhile, the secondary challenge involves determining the optimal technique to address the prediction problem (Keshtegar et al., 2019). The current study investigated five prominent ML techniques, which are DT, RF, GBRT, LightGBM, and XGBoost. The primary objective was to ascertain the most suitable technique for predicting the shear capacity of RCTBs. DT algorithm provide inherent interpretability, enabling visualization of decision-making processes and identification of key influential features impacting predicted shear capacity (Sutton, 2005). By strategically combining multiple DT, RF algorithm achieve enhanced prediction accuracy compared to individual DT. This ensemble learning approach reduces prediction errors and fosters the development of more generalizable models that are less susceptible to overfitting the training data (Breiman L, 2001). GBRT algorithm exhibit remarkable proficiency in managing complex relationships between input features and the target variable (such as shear capacity). Their adeptness enables effective capture of non-linear effects within the data. Moreover, GBRTs often achieve high accuracy across diverse datasets (Jerome H. Friedman, 2001). LightGBM (Ke G et al., 2017) and XGBoost (Chen & Guestrin, 2016) algorithms are efficiency in computational speed and memory usage, making them ideal for handling large datasets common in structural engineering. They integrate robust

regularization techniques to prevent overfitting, enhancing model generalizability to new data. Additionally, both algorithms can handle missing values and categorical data.

2 Research significance

Despite the promising outcomes demonstrated by ML algorithms in diverse domains, there is a notable lack of literature focused on their application in estimating the shear capacity of RCTBs. This highlights the need for further research and improvement of ML techniques specifically within this area of study. Furthermore, from the previous literature, it can be determined that current design formulas cannot accurately predict the shear capacity of RCTBs, due to the complex nature of shear failure in T-sections. Hence, this study explores the potential of five ML algorithms—Decision Trees (DT), Random Forest (RF), Gradient Boosted Regression Trees (GBRT), LightGBM, and XGBoost—to predict the shear capacity of RCTBs. The models are trained and tested using experimental data consisting of 360 data points. The SHAP method was utilized to determine the significance of input features and illustrate their positive or negative impacts on the predicted outcomes. Furthermore, sensitivity analysis is conducted to assess the ML models' capacity to capture the influences of geometric and material parameters on their predictions. The developed ML model is compared against established codes of practice and existing empirical formulas. Furthermore, a user-friendly interface platform is developed, effectively simplifying the implementation of the proposed machine-learning model. Reliability analysis is performed to determine the value of the resistance reduction factor (ϕ) that will achieve the specified target reliability index.

3 Experimental Database

3.1 Database Preparation and Description of Parameters

A literature survey was conducted to gather experimental data on the shear capacity of RCTBs with and without shear reinforcement. The survey yielded 360 experimental results from the reviewed literature, as illustrated in Table 1. The input parameters that affect the shear capacity of RCTBs are web width of beam (b_w), effective depth of beam (d) (the distance between the steel rebar and the compression fiber's edge), flange width (b_f), flange thickness (h_f), compressive strength of concrete (f'_c) (conventional normal concrete), shear span-to-depth ratio (a/d), longitudinal steel ratio (ρ), and ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$). The output variable is maximum shear capacity (V_{exp}). Moreover, Table 2 displays the descriptive statistics for every parameter contained within the dataset, where descriptive statistics refer to the numerical

Table 1 Geometrical dimensions and material properties of the experimental tests

References	No. of tests	Geometrical dimensions				Concrete properties		Internal reinforcement		Shear capacity
		b_w (mm)	d (mm)	b_f (mm)	h_f (mm)	a/d	f'_c (MPa)	$\rho\%$	$\rho_v f_{yv}$ (MPa)	
(Abdul Aziz Abdul Samad et al., 2016,)	2	200	231	400	100	3.5	40.4–43.5	2.15	0	204–228
(Thamrin et al., 2016)	6	125	212–219	250	70	3.7–3.8	32	1–2.5	0	36.6–47.5
(KOTSOVOS et al., 1987)	4	50	240	200	50–65	3.3–10.4	40	5.2	0	19–37
(Elgohary et al., 2019)	5	150	470	300–900	50	1.27	19.4–20.2	1.71	0	165–185
(Ayensa et al., 2019)	21	200	494	200–800	50–200	2.96	40.2–42.6	1.63–2.10	0.76	260–326.6
(Ramadan et al., 2022)	48	100	280–580	300–600	45–300	0.69–1.42	27–29.1	1.39–2.87	0–1.81	76.5–392.7
(Kadr et al., 2019)	27	120	260	360	60	1–3	25.4–35	2.58	0–4.17	155–325
(Withit Pansuk et al., 2004)	6	150	300	150–600	75	3.5	23	3	1.07–1.84	180–313
(Bach et al., 1980)	39	200–380	335–358	800	90	2.92–3.13	10–35	0.90–5.61	0.40–3.34	184–408
(Hans Christian, 1974)	12	110	298	400	80	3.5	22.9–34.2	3.83	0.76–1.60	107–139
(Braestrup et al., 1976)	26	200	359	800	90	2.92	7.94–12.7	1.68	0.35–4.83	157–280
(Swamy & Qureshi, 2015)	45	76.2	95.25	228.6	25.4	3–8	19–32.3	3.52	0–1.84	11.25–36.7
(Palaskas et al., 1981)	18	190.5	370–399	610	102	3.88–4.18	26–38	0.49–0.95	0–0.76	58–140
(Bousselham A & Chaallal O, 2006)	3	152	350	508	102	3	25	3.76	0–4.9	81.3–295
(Ferguson & Thompson, 1953)	11	102–178	178–210	432–483	38	3.4–4	17.5–45.4	2.73–4.78	0	21.5–52.1
(Placas A et al., 1971)	32	152	254	610	76	3.36–7.20	12–36.8	1.25–4.16	0–2.25	54.7–239
(Sahoo et al., 2016)	2	150	217	300	50	2.5–3	23	1.85	0	40–43.5
(Wehr KE, 1967)	3	152	279	610–914	76	3.88	27.7–29.6	1.33	0	44.5–51
(Leonhardt F & Walther R., 1962)	10	50–160	300–375	300–960	75–80	3.33–3.50	15.4–25	1.39–8.34	0.57–3.5	101–375
(Kotsovou & Cotsivos, 2018)	2	150	360	400	80	2.4	50	2.23	0.78–1.56	151–203.6
(Abd, 2019)	38	50–190	240–395	200–610	65–102	3.3–10.4	23.4–43	0.49–5.23	0–1.64	19–160

Table 2 Input and output parameters distribution

Description	Input parameter	Mean	STD	Min	Max	Q1	Q2	Q3
Geometrical dimensions	b_w (mm)	144.11	55.85	50	380	100	150	190.5
	d (mm)	297.88	108.17	95.25	580	254	298	359
	b_f (mm)	505.21	210.82	150	960	300	500	610
	h_f (mm)	85.69	46.37	25.40	300	60	80	97.5
	a/d	3.25	1.55	0.69	10.4	2.92	3.06	3.6
Concrete properties	f'_c (MPa)	27.59	8.92	7.94	57	23	28.8	32.25
Internal reinforcement	$\rho\%$	2.57	1.23	0.49	8.34	1.64	2.58	3.52
	$\rho_v f_{yv}$ (MPa)	1.03	1.17	0	5.56	0	0.76	1.55
Shear capacity	V_{exp} (kN)	167	102.59	11.25	408	76.8	165	252.45

STD standard deviation, *Min* minimum, *Max* maximum, and Q1, Q2, Q3: 25th, 50th, and 75th percentiles

values that can be used to describe a dataset, such as its measures of central tendency (mean) and variability (standard deviation).

Figure 1 displays a histogram of the data distribution and features. It displays the frequency of both input and output features within a specific range. The outcomes indicate that the web width (b_w) spans from 50 to 200 mm, and the effective depth (d) of the beams varies between 200 and 400 mm, with a notable concentration around 300 mm. Additionally, the flange width (b_f) lies within 200–600 mm, accompanied by a flange thickness (h_f) around 85 mm. The a/d ratio predominantly centers on a value of 3. The concrete compressive strength stands approximately at 30 MPa, while the longitudinal steel ratio (ρ) hovers around 0.025%. Moreover, the product of the shear reinforcement ratio and the yield strength of shear reinforcement ($\rho_v f_{yv}$) approximates 1 MPa. The maximum shear capacity (V_{exp}) of the beam encompasses a range of 10–400 kN, with a primary peak around 150 kN. Finally, the histogram highlights the need for future experimental programs to address the gaps in the dataset.

3.2 Pearson Correlation Analysis

The Pearson correlation coefficient (PCC) is a statistical metric used in sensitivity analysis to measure of the significance of a linear relationship between two variables. It is also known as Pearson's r or the Pearson coefficient for product–moment correlation (PPMCC) (Schober & Schwarte, 2018). The PCC has a value between -1 and $+1$, with 1 representing a perfect positive linear relationship between the variables, -1 denotes a perfect negative linear relationship, while 0 denotes no linear relationship. In addition, the PCC evaluates the strength and directionality of a relationship between two variables. Figure 2 shows the results of the Pearson

correlation analysis of nine parameters (b_w , d , b_f , h_f , a/d , f'_c , ρ , $\rho_v f_{yv}$, V_{exp}). The degree of correlation between b_w and b_f reaches 0.75, while the degree of correlation between d and h_f reaches 0.70, which indicated that features have a strong relationship. Among all the input features, it appears that the effective depth of the beam (d) has the greatest effect on the output, with a correlation coefficient of 0.63. Furthermore, it can be observed that the flange dimensions (thickness (h_f) and width (b_f)) positively affect the shear capacity of RCTBs, with correlation coefficients of 0.57 and 0.50, respectively. This highlights the significance of the flange in predicting the shear capacity of RCTBs. These results indicate that variations in these variables may have a significant impact on the output, while changes in the other input parameters may have a comparatively smaller effect.

4 Machine Learning Models

4.1 Overview of Machine Learning Models

The utilization of machine learning has gained prominence within the field of structural engineering, enabling the analysis and prediction of complex structural behaviors with improved accuracy and efficiency. Machine learning algorithms can identify patterns and relationships within the data, aiding in the modeling and simulation of structural systems. These algorithms can be applied to various aspects of structural engineering, including structural optimization, damage detection, and structural health monitoring. For instance, researchers have successfully utilized machine learning techniques to predict the behavior of structures under different loading conditions, optimize structural design parameters, and detect and classify structural damage. Figure 3 demonstrates a standard ML workflow employed in predictive modeling. Using a learning algorithm and an initial

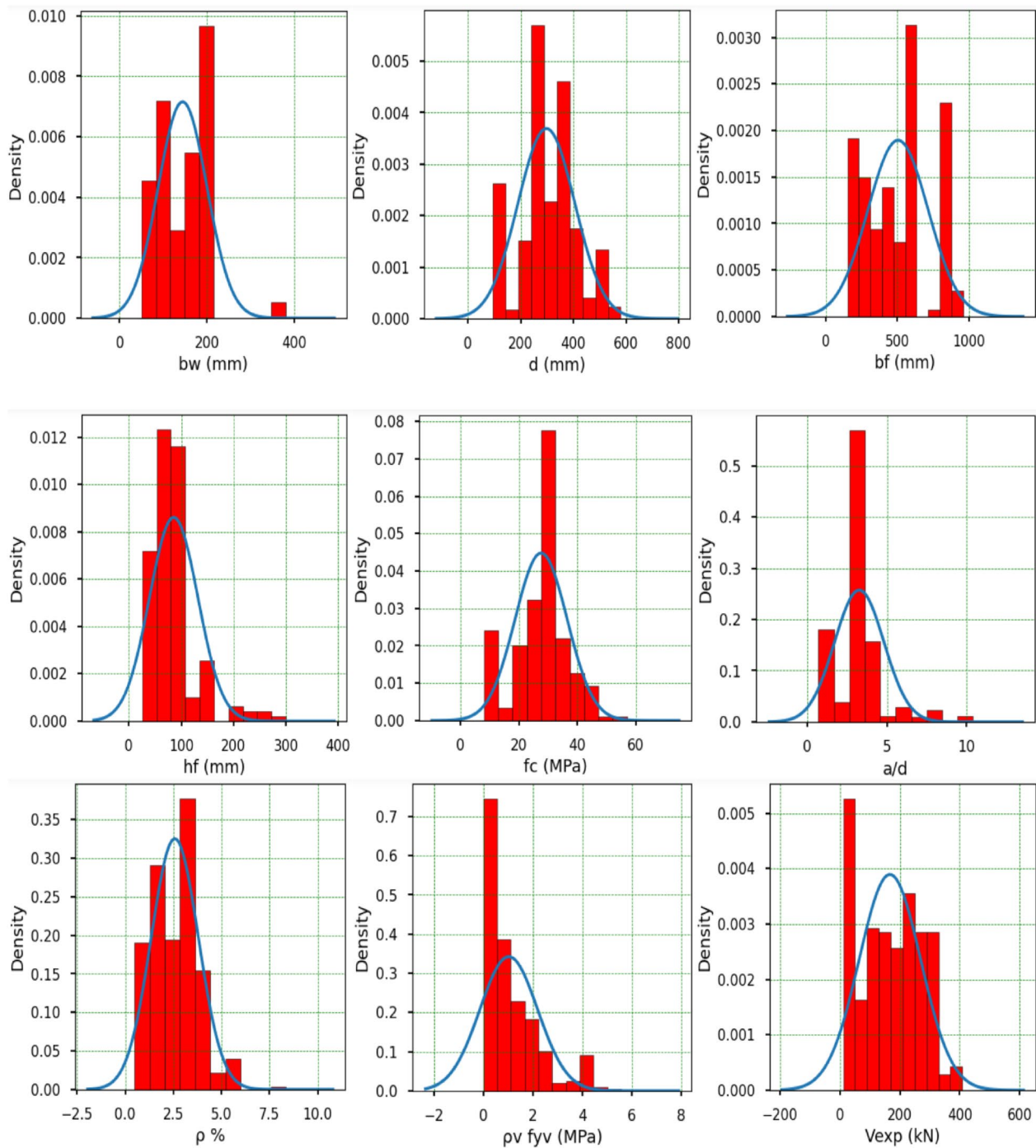


Fig. 1 Histogram with a normal distribution fit for both input and output features

dataset, computer systems can be trained to continually improve and learn until they reach the desired performance level. Therefore, the ML model accuracy is dependent on the nature and characteristics of the initial data, as well as the effectiveness and efficiency of the learning algorithm employed.

4.2 ML Algorithms

The current study investigated five ML algorithms commonly used in the field of structural engineering (Thai, 2022). As these algorithms have been previously explained in detail in other publications, the subsequent sections will simply emphasize their key features.

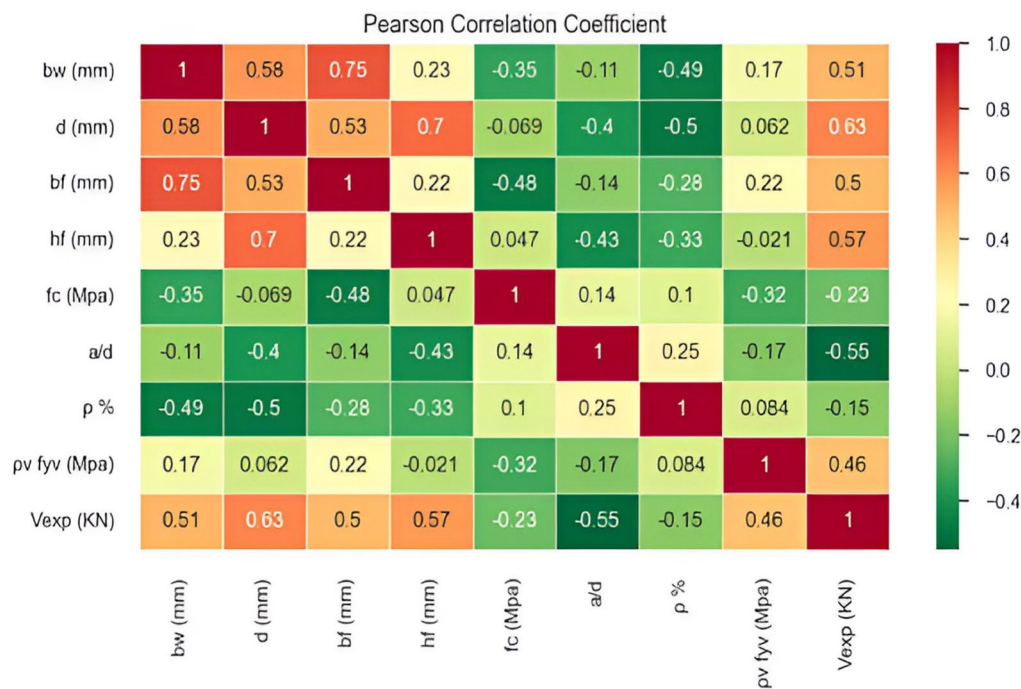


Fig. 2 Pearson correlation analysis of nine parameters

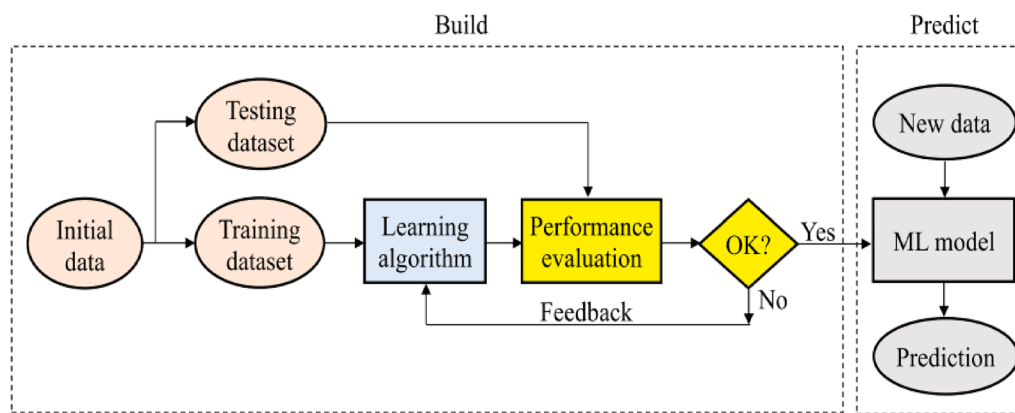


Fig. 3 Typical workflow of ML

4.2.1 Decision Tree (DT)

DT (Sutton, 2005) is a type of tree-based model used for visualizing the decision-making process. It is also commonly known as CART (Classification and Regression Tree), which refers to the model's capability to perform both classification and regression tasks. Figure 4 illustrates that a decision tree is composed of four essential components: a root node, decision nodes, leaf (terminal) nodes, and two or more branches. The root node is positioned at the top of the tree and serves as the chief decision node, reflecting the goal of the decision tree.

The decision node, on the other hand, is where a condition is established that splits the dataset, while the leaf node denotes the conclusion of a branch, demonstrating a decision to be performed. In a decision tree, the root node represents the source data, and this data is recursively divided into smaller subsets using splitting conditions according to several metrics such as MSE (in the case of regression problems). The process of splitting is repeated for each subset that is derived until no further splits can be made that reduce the chosen metric or until the tree reaches its maximum allowable depth.

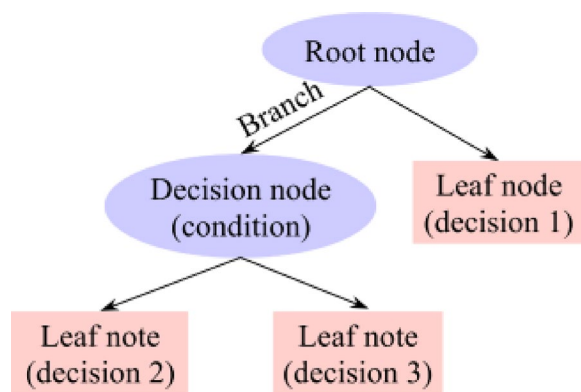


Fig. 4 Example of a Decision Tree (DT)

This iterative process continues until the resulting tree is deemed satisfactory for the problem at hand.

4.2.2 Random Forest (RF)

Leo Breiman, (2001) developed RF algorithm, which used decision trees (DTs) as its base or weak learners and combines them to enhance the model's performance. DTs are prone to overfitting and instability but using ensemble learning with multiple DTs can help overcome these issues. In RF, several decision trees are trained on distinct random subsets of the dataset and features to improve the model's performance and robustness, and the final prediction is generated by voting on the outputs of these individual trees, as shown in Fig. 5. Random Forest can be advantageous over DT when dealing with large databases with many input variables, as it can handle them

more efficiently by employing a large number of trees. While RF may train faster than a single DT, it can take longer to develop predictions due to the need to combine results from multiple trees. The default parameters of RF are frequently sufficient to generate acceptable outcomes, but hyperparameter tuning can be used to enhance the model's accuracy or speed. Overall, RF is a strong technique that inherits many of the merits of decision trees while also being more robust and adaptable for certain types of datasets and problems.

4.2.3 Gradient Boosting Regression Tree (GBRT)

Gradient Boosting Regression Tree (GBRT) was developed by Friedman, (2001). It is based on the AdaBoost algorithm with two modifications to the weak learners. In the first modification, decision trees (DTs) were used as weak learners. The second modification is that weights are updated based on the residual errors of the previous weak learner, rather than its classification errors. GBRT has become a popular ML technique for regression tasks due to their ability to handle complex interactions between variables and the flexibility to learn non-linear relationships. Figure 6 illustrates that the training of the next tree considers the residual error from the previous tree. After the initial weak learner, each subsequent tree is trained. The final model can capture the residual errors of the weak learners and improve the accuracy of predictions. GBRT also offer flexibility by enabling the tuning of hyperparameters, including the number and depth of trees, and the learning rate to control the model's convergence and training speed.

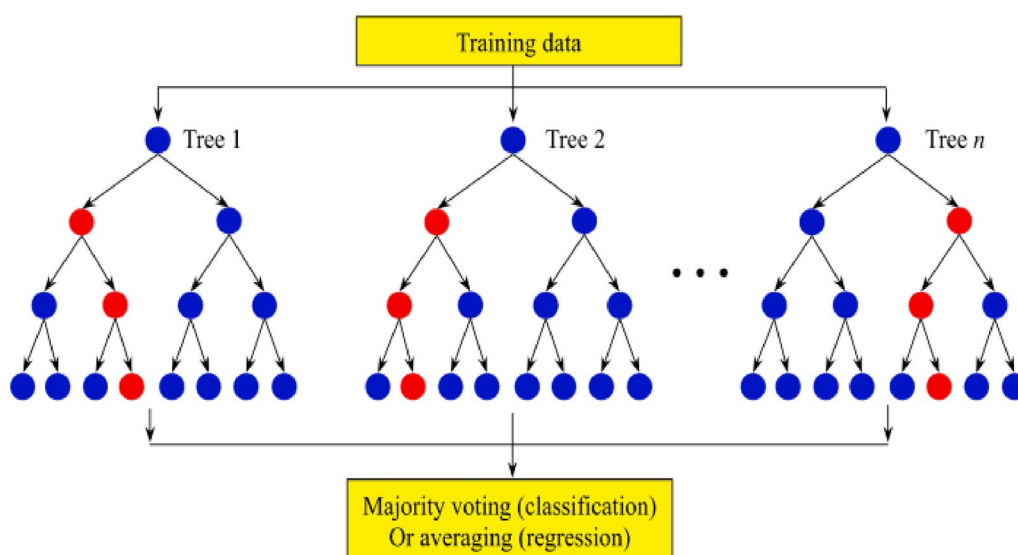


Fig. 5 Flowchart of RF (parallel training)

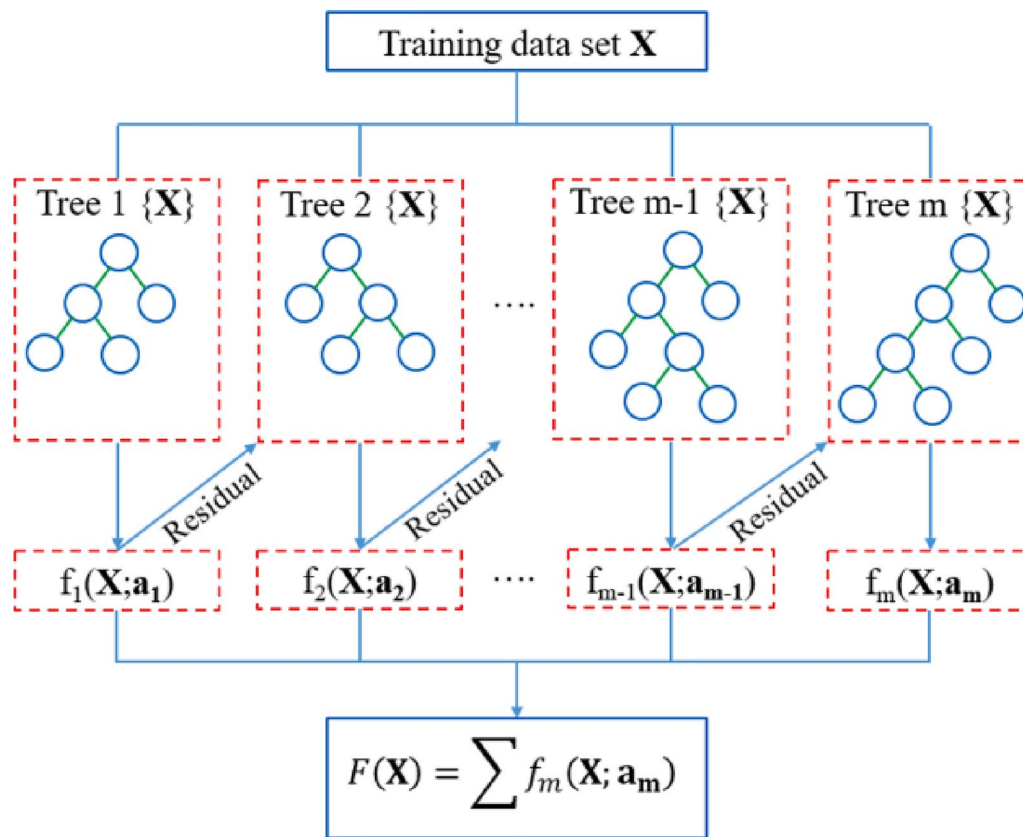


Fig. 6 Illustration of the GBRT model

4.2.4 Light Gradient Boosting Machine (LightGBM)

Ke et al., (2017) developed Light Gradient Boosting Machine algorithm with a focus on computational efficiency without sacrificing accuracy, boasting a speed that can up to 20 times more rapid than Gradient Boosting Machine. Light word refers to its impressive speed compared to other boosting algorithms, including XGBoost, which can be slow to train with large datasets. The primary distinction between LightGBM and other boosting algorithms lies in the method of its expanded tree. LightGBM uses a leaf-wise tree growth strategy to select the leaf with the largest loss (see Fig. 7) to reduce additional loss and improve accuracy than depth-first strategies used by other boosting algorithms when developing trees on the same leaf. This is in contrast to level-wise tree growth strategies used by many other gradient boosting algorithms, which may be less effective in reducing loss on individual leaves. LightGBM relies on two key features to achieve exceptional speed when compared to other boosting frameworks: (1) gradient based one-side sampling (GOSS) and (2) exclusive feature bundling (EFB). The use of GOSS and EFB can speed up training time for

gradient boosting decision trees by reducing computational complexity. While the search results include information on LightGBM's design and performance.

4.2.5 Extreme Gradient Boosting (XGBoost)

XGBoost was developed by Chen et al., (2016), where XGBoost is considered to be one of the most efficient machine learning methods as it enhances the efficiency of Gradient Boosting Machines. XGBoost is a machine learning algorithm that employs several advanced techniques to train a high-dimensional data in a timely and precise manner. The algorithm uses a compressed column to store input dataset, which reduces the sorting cost and speeds up the training process. Additionally, it leverages a randomization technique to minimize overfitting and increase generalization. Lastly, XGBoost uses parallel and distributed computing techniques to fully utilize all available CPU cores during training and split finding. This enhances the scalability and performance of the algorithm, making it possible to train models on very large datasets with a large number of features.

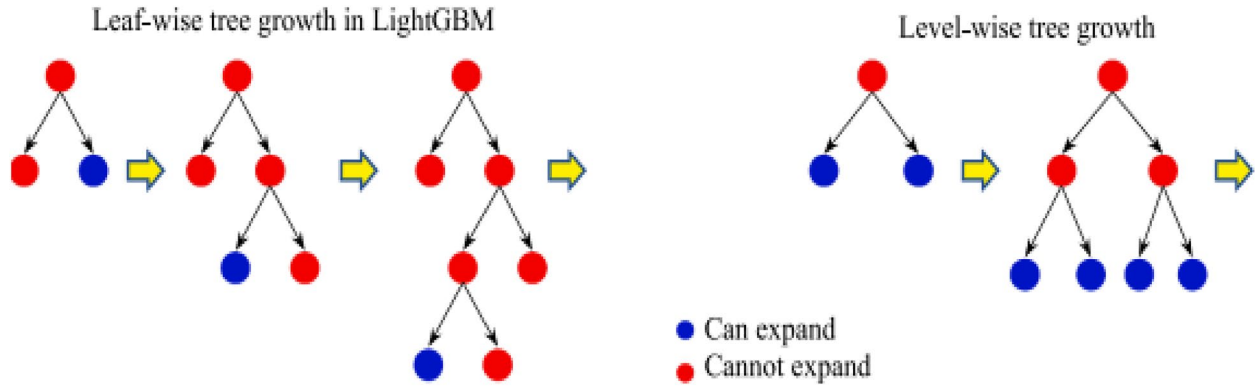


Fig. 7 Outline representation of LightGBM and another Bas

4.3 Development of ML Models

The ML models proposed in this paper, comprising eight input parameters, are selected to identify the optimal model with the most influential parameters for predicting the target variable (shear capacity). To address challenges associated with low learning rates at extreme parameter values and enhance the accuracy and speed of modeling, the collected database is normalized to a range of $[-1, 1]$. The four common statistical metrics, specifically mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and correlation coefficient (R^2), are frequently utilized to assess the efficacy of prediction models. A reduced value for MSE, RMSE, and MAE corresponds to improved model performance, whereas R^2 values within the range of 0–1 quantify the concurrence between projected and actual values. Consequently, elevated R^2 values signify enhanced model effectiveness. The formulations for these four parameters are as follows (Wakjira et al., 2022):

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \quad (1)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (4)$$

where $\hat{y}$ represents the predicted value, y is the corresponding actual value, $\bar{y}$ is the mean of all y values in the data set, and N is the sample size.

4.3.1 Hyperparameter Tuning and Cross-Validation

There are multiple techniques that can improve the performance of ML models, and one effective strategy is to fine-tune the hyperparameter. Performance of a model can be greatly affected by the choice of hyperparameter values, and a systematic approach to varying the values and evaluating performance for each combination is called hyperparameter optimization. The grid search method is widely employed for hyperparameter tuning purposes. In order to address concerns related to overfitting during the hyperparameter optimization process, the approach of K-fold cross-validation is employed. The dataset is first partitioned into a training set and a testing set. Specifically, 80% of the entire dataset is for the training set, and the remaining 20% is for the testing set. The K-fold cross-validation involves carrying out the following steps:

- For the intention of cross-validation, it is customary to divide the training dataset into K-equivalent groups or folds.
- In each iteration, use $K-1$ folds for training and the remaining fold for validation.
- Repeat this process until every fold has been used for validation at least once. This approach allows for comprehensive validation of the model and reduces the risk of overfitting.

For hyperparameter optimization in this study, grid search technique was used together with tenfold

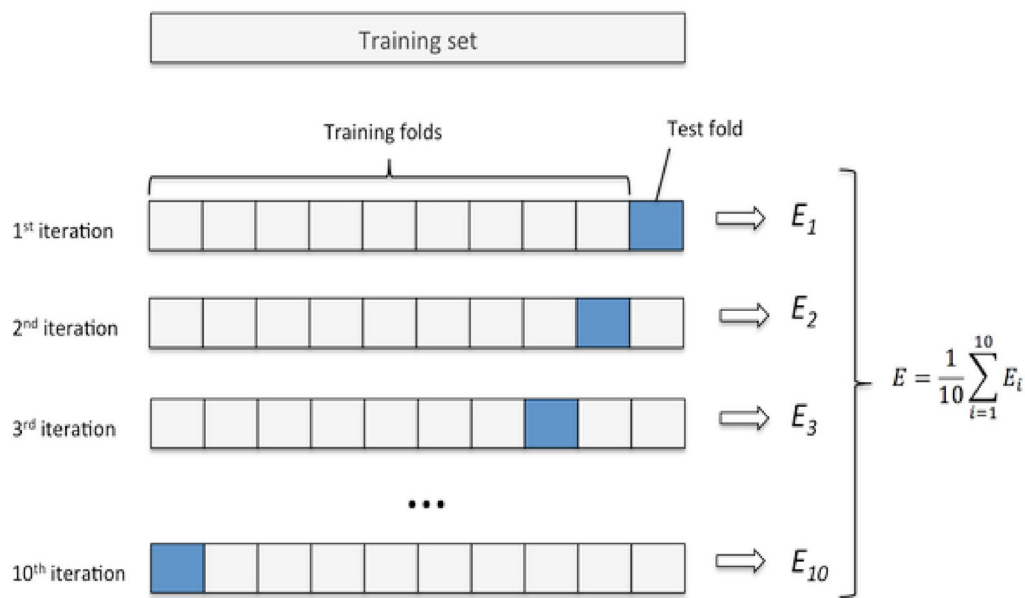


Fig. 8 Hyperparameter tuning using K-fold cross-validation

cross-validation ($K=10$). During each iteration of the tenfold cross-validation, 90% of the training dataset is used to train the model while the remainder 10% is used as a validation dataset. Figure 8 depicts the K-fold cross-validation used in this work for training and for the hyperparameter selection of the model.

For example, the effectiveness of the XGBoost algorithm depends on the initial setup of hyperparameters like the number of trees ($n_estimators$) and the learning rate (H. Nguyen et al., 2021a, 2021b). The optimization of XGBoost hyperparameters using tenfold cross-validation can be represented through four charts, each presenting the R^2 score for a different number of estimators (100, 200, 300, 400). In each chart, there are four lines indicating various maximum depths (4, 8, 12, 16). According to Fig. 9, the R^2 score exhibits its optimal fit when the learning rate reaches 0.5, the number of estimators equals 300, and the maximum depth equals 4. Moreover, with a high learning rate, the number of estimators shows no influence on the R^2 score. Conversely, when the learning rate drops below 0.2, a greater number of estimators is necessary for the R^2 score to achieve a stable maximum value. Furthermore, Table 3 synthesizes the optimal hyperparameter values for the DT, RF, GBRT, LightGBM and XGBoost models, which were all subjected to hyperparameter optimization.

5 Results and Discussion

5.1 Model Prediction Results

This section explores the predictive capacity of the developed ML models. After identifying the most effective hyperparameters for each ML algorithm. Figure 10 presents scatter plots illustrating the relationship between predicted and actual (experimental) shear capacity values of RCTBs using various ML models. The developed ML models demonstrate a strong alignment between predicted and experimental shear capacity, with R^2 values exceeding or equaling 98%. Among these models, the XGBoost model demonstrated superior predictive performance, yielding an R^2 value of 0.9914 and minimal values for RMSE, MAE, and MAPE, which were 9.53, 3.75, and 0.03, respectively. In contrast, the RF model recorded the lowest R^2 value of 98.20 and higher values for RMSE (13.80), MAE (5.80), and MAPE (0.0422), as depicted in Table 4. Importantly, across all ML models, the predicted shear capacity values cluster closely around the 45° diagonal line, indicating a precise correlation between the experimental and forecasted shear capacity values. Additionally, Fig. 11 illustrates the residuals of predicted shear capacity of RCTBs, depicted as the difference between V_{exp} and V_{pred} on both training and testing datasets. It is noteworthy that the residuals for all models are centered around zero. Therefore, it can be concluded

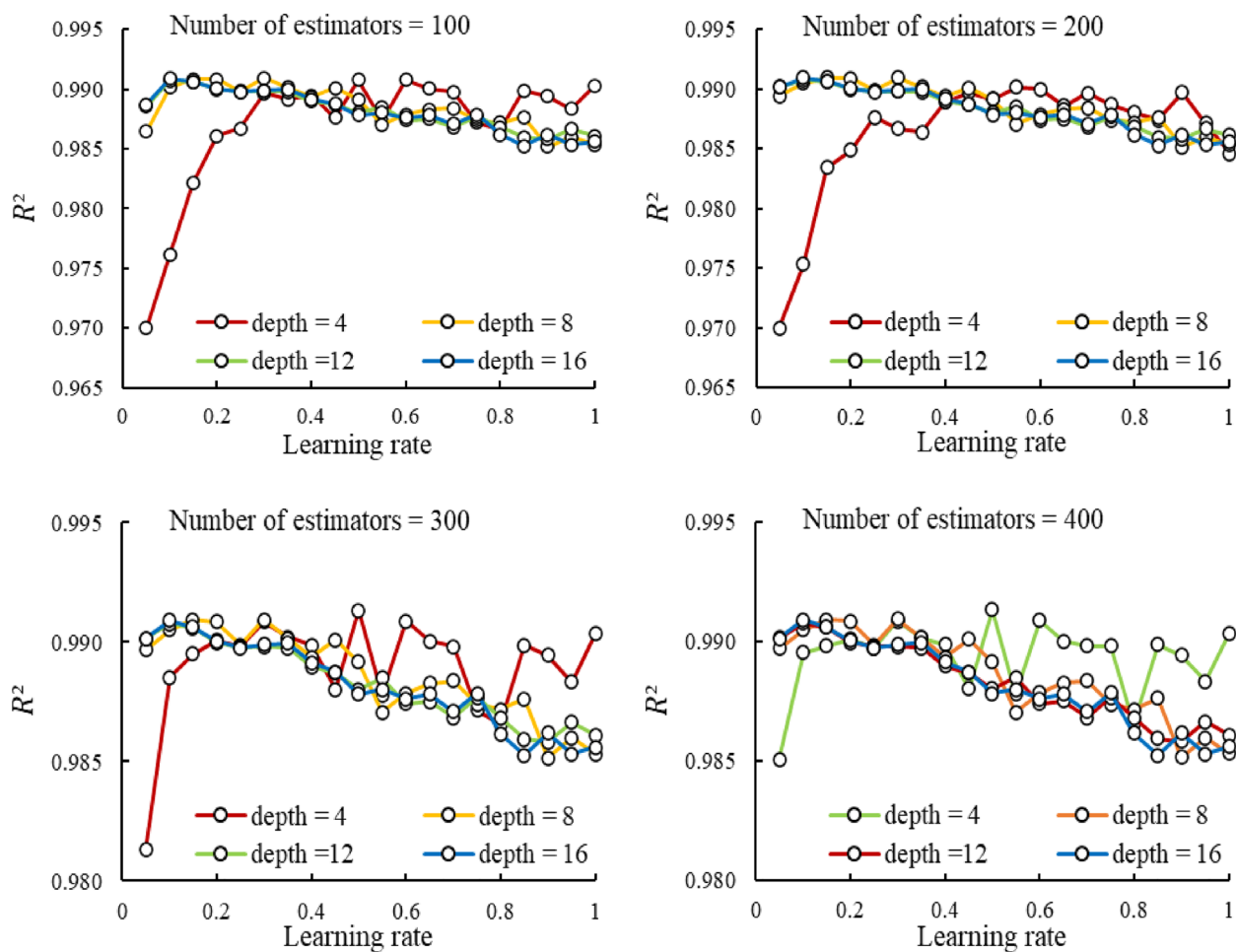


Fig. 9 XGBoost hyperparameters optimization

Table 3 Optimal hyperparameter values for each ML model

Model	Parameters
DT	criterion='squared_error', splitter='best', max_depth=10, max_features=None, max_leaf_nodes=None, min_impurity_decrease=0, ccp_alpha=0.08, min_samples_leaf=1, min_samples_split=2, min_weight_fraction_leaf=0.0, random_state=27
RF	bootstrap='False', max_depth=80, max_features=3, min_samples_leaf=1, min_samples_split=2, n_estimators=1000
GBRT	learning_rate=0.2, max_depth=10, max_features='log2', min_samples_leaf=4, min_samples_split=3, n_estimators=100
LightGBM	boosting_type='gbdt', learning_rate=0.9, max_depth=-1, n_estimators=300, num_leaves=31, objective='regression'
XGBoost	learning_rate=0.5, max_depth=4, min_child_weight=1, subsample=1, colsample_bytree=0.5, n_estimators=300, gamma=25, objective='reg:squarederror'

that the models demonstrate high efficiency in predicting the shear capacity.

To facilitate comparison of the performance of the presented ML models, Fig. 12 displays the Taylor diagram (Taylor, 2001) for the ML models. Taylor diagram represents predicted model values in comparison to the original data by plotting correlation (r) and standard

deviation against a reference dataset. In Fig. 12 the XGBoost model directly lies near the reference line of the original dataset. The Taylor diagram can also explain the reliability and precision of the ML models. Moreover, to visualize the outcomes of the models, graphical representations of the experimental and predicted shear capacity values versus the experiment

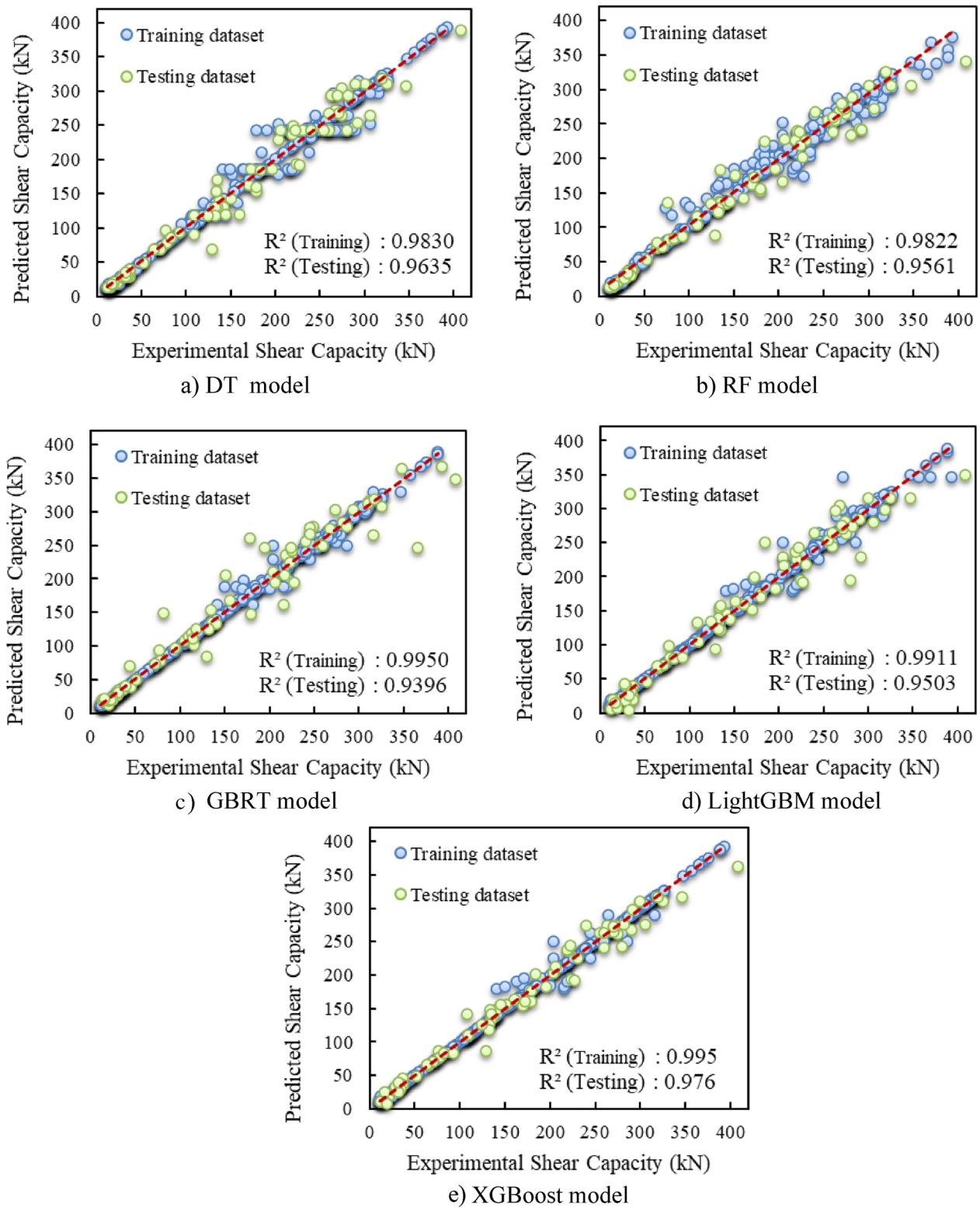


Fig. 10 Experimental versus predicted shear capacity of RCTBs for **a** DT, **b** RF, **c** GBRT, **d** LightGBM, and **e** XGBoost models

Table 4 Performance metrics of ML models

Model	Data set	Performance measures			
		RMSE (kN)	MAE (kN)	MAPE (%)	R^2 (%)
DT	Training dataset	13.38	6.71	4.76	98.30
	Test dataset	19.41	13.91	8.96	96.35
	All	13.70	5.80	4.22	98.20
RF	Training dataset	13.68	8.35	6.00	98.22
	Test dataset	21.27	13.65	8.65	95.61
	All	13.80	5.80	4.22	98.20
GBRT	Training dataset	7.12	3.12	2.00	99.50
	Test dataset	27.31	16.45	13.11	93.96
	All	13.79	5.80	4.22	98.20
LightGBM	Training dataset	9.68	3.96	2.78	99.11
	Test dataset	22.64	15.82	14.21	95.03
	All	13.33	6.33	5.08	98.30
XGBoost	Training dataset	7.23	1.97	1.52	99.50
	Test dataset	15.63	10.84	9.03	97.63
	All	9.53	3.75	3.025	99.14

number are provided, as depicted in Fig. 13. Additionally, error data are employed to enhance the visibility of the generated models. It can be observed that, among all models, the XGBoost model demonstrates that most of the predicted points are closer to the experimental points. It also indicates that the predicted errors in the XGBoost model are lower than those in other models. However, one potential reason for the occurrence of predicted errors is the scarcity of experimental data points available in this region, emphasizing the necessity for future experimental programs to address this gap. Finally, the superior performance of the XGBoost model in predicting the shear capacity of RCTBs leads to a focus on this methodology in subsequent investigations.

5.2 Comparison with Shear Design Models

In this section, the accuracy of the XGBoost model is evaluated by comparing its predictions of the shear capacities of RCTBs with those calculated using the American Concrete Institute (ACI 318-19) (2019), British Standard (BS 8110-1:1997) (1997), European Standard (Eurocode EN 1992-1-2), (2004), Canadian Standards Association (CSA23.3-04) (2004) design guidelines, and existing formulas from researchers. The existing formulas suggested by Sarsam et al., (2018), Thamrin et al., (2016), and Zararis et al., (2006) are presented in Table 5.

The shear capacity of RCTBs was determined using shear design models. It should be noted that all shear design models were applied to all specimens, except for the Thamrin et al. model, which predicted the capacity of specimens without steel stirrups. Figure 14 illustrates the experimental versus predicted shear capacity based on shear design models and the XGBoost model. The solid line denotes the best fit between the experimental and predicted responses. Furthermore, the assessment of both the six models and XGBoost model based on the mean, standard deviation (STD), and coefficient of variation (COV) of the V_{Exp} to V_{Pred} ratio is demonstrated in Table 6. It can be clearly indicated from Table 6 that all shear design provisions significantly underestimated the shear capacity of RCTBs, where the average ratio of experimental to predicted shear values for all the codes was above 1.0 (in the range of 1.87–2.94). Furthermore, the CSA23.3-04 (2004) code provided the most accurate predictions for the shear capacity of RCTBs, with an average ratio of V_{Exp} to V_{Pred} shear capacity of 1.87. Among the three existing formulas developed to predict the shear capacity of RCTBs, Table 6 indicates that the Zararis et al.'s model (2006) offered the most accurate predictions for the shear capacity of RCTBs overall. Particularly, the average ratio of V_{Exp} to V_{Pred} shear capacity is 1.71. The accuracy of the Zararis et al. model comes from its incorporation of the whole area of the T-beam and its determination of an effective width that is suitable for predicting the shear capacity. Following that, the predictions of the Sarsam et al., model (2018) also explore the influence of flanges on the shear capacity of RCTBs (average ratio of V_{Exp} to V_{Pred} = 1.78). Finally, despite Thamrin et al.'s (2016) equation being developed to predict the shear capacity of RCTBs without internal stirrups, it failed to accurately estimate the true capacity and provided conservative results. In particular, the mean ratio of V_{Exp} to V_{Pred} was 1.65 with COV of 32%.

In conclusion, the contribution of the flange to shear capacity is significant and should not be neglected when determining the shear capacity of RCTBs. Unfortunately, most existing design guidelines overlook this aspect, resulting in inaccurate estimations of the true shear capacity. With respect to the models proposed in the literature, the model presented by Zararis et al., (2006) offered accurate estimations of the shear capacity of RCTBs and could be safely utilized for designing RCTBs in shear. Furthermore, the XGBoost model demonstrated better predictive capability in contrast with the shear design models, with an average ratio of V_{Exp} to V_{Pred} was 1.0045 ± 0.0596 and COV of 5.93%, as observed in Fig. 14h and Table 5.

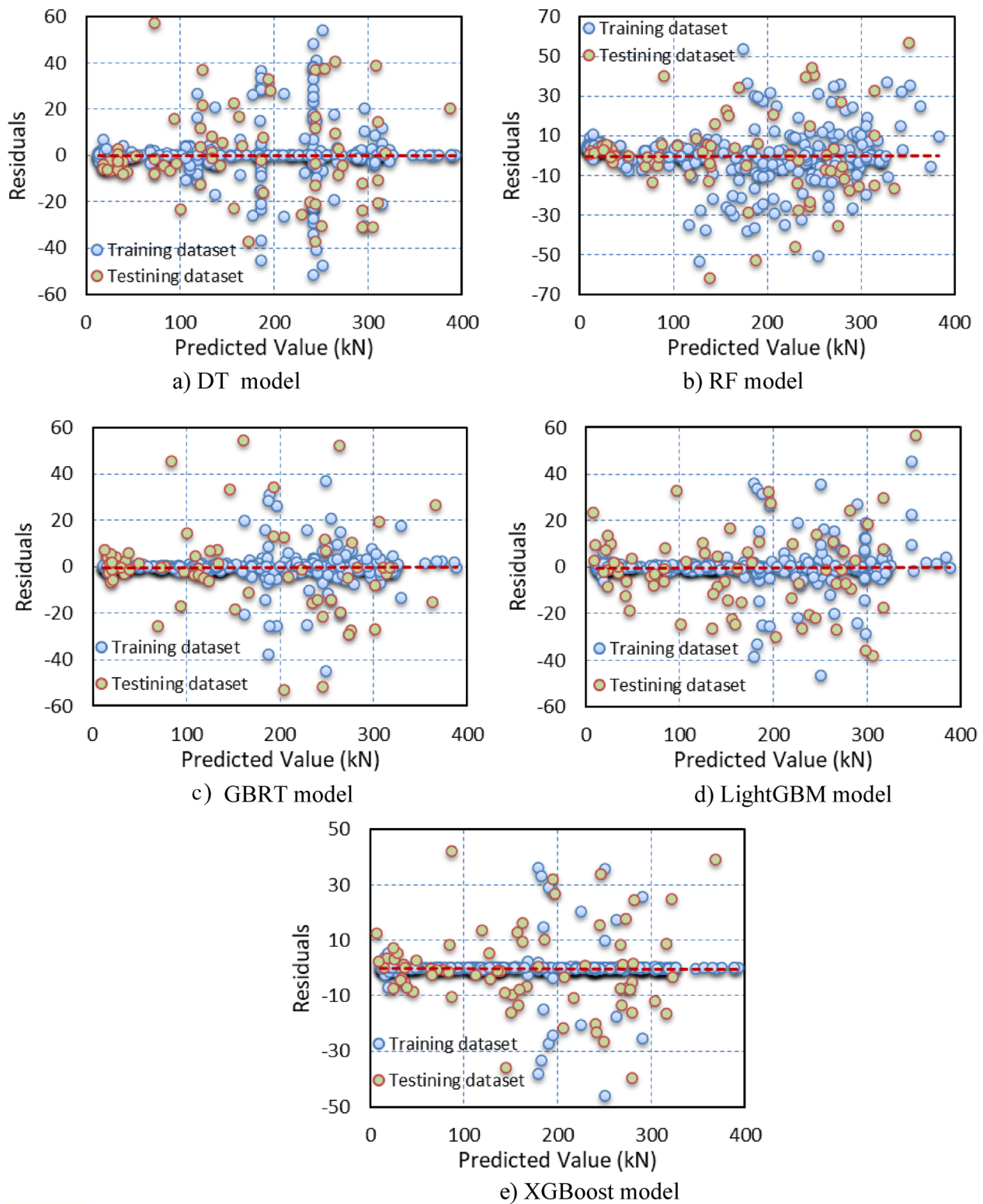


Fig. 11 Residual of the predicted shear strength in both the training and test datasets for **a** DT, **b** RF, **c** GBRT, **d** LightGBM, and **e** XGBoost models

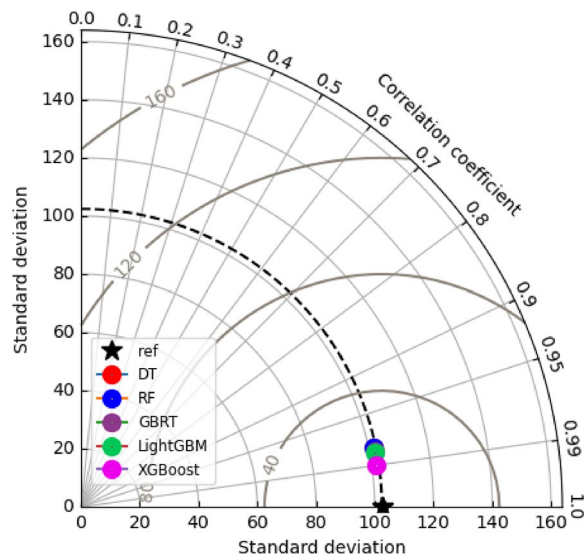


Fig. 12 Taylor graph of DT, RF, GBRT, LightGBM, and XGBoost models

5.3 Explainability of XGBoost Model Using SHAP Approach

In this study, a unified SHAP method was utilized to explain the output of an XGBoost model and to identify the most variables and their interactions that impact the shear capacity of RCTBs. The most significant factor is determined through the absolute SHAP value. The x-axis position of each point represents the Shapley value for the corresponding factor, indicating its impact on the shear capacity. The y-axis displays a list of factors arranged according to their level of importance, as shown in Fig. 15a. The SHAP results indicate that the shear span-to-depth ratio (a/d) has the most significant effect on the shear capacity of RCTBs. This is followed by the ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$), flange thickness (h_f), and flange width (b_f). The SHAP summary plot, shown in Fig. 15b, was used to evaluate the influence of input features on prediction results. Each point on the plot represents a prediction instance, with positive and negative SHAP values indicating the correlation between input features and outputs. Furthermore, the color of each point indicates the value of the input feature, varying from low (blue) to high (red). In general, it is observed that all input features show positive Shapley values (red points) except a/d , suggesting that higher values of these parameters result in an increase in the ultimate shear capacity of RCTBs.

5.4 Sensitivity Analysis

The sensitivity analysis of the XGBoost model involves varying one parameter at a time while keeping the rest

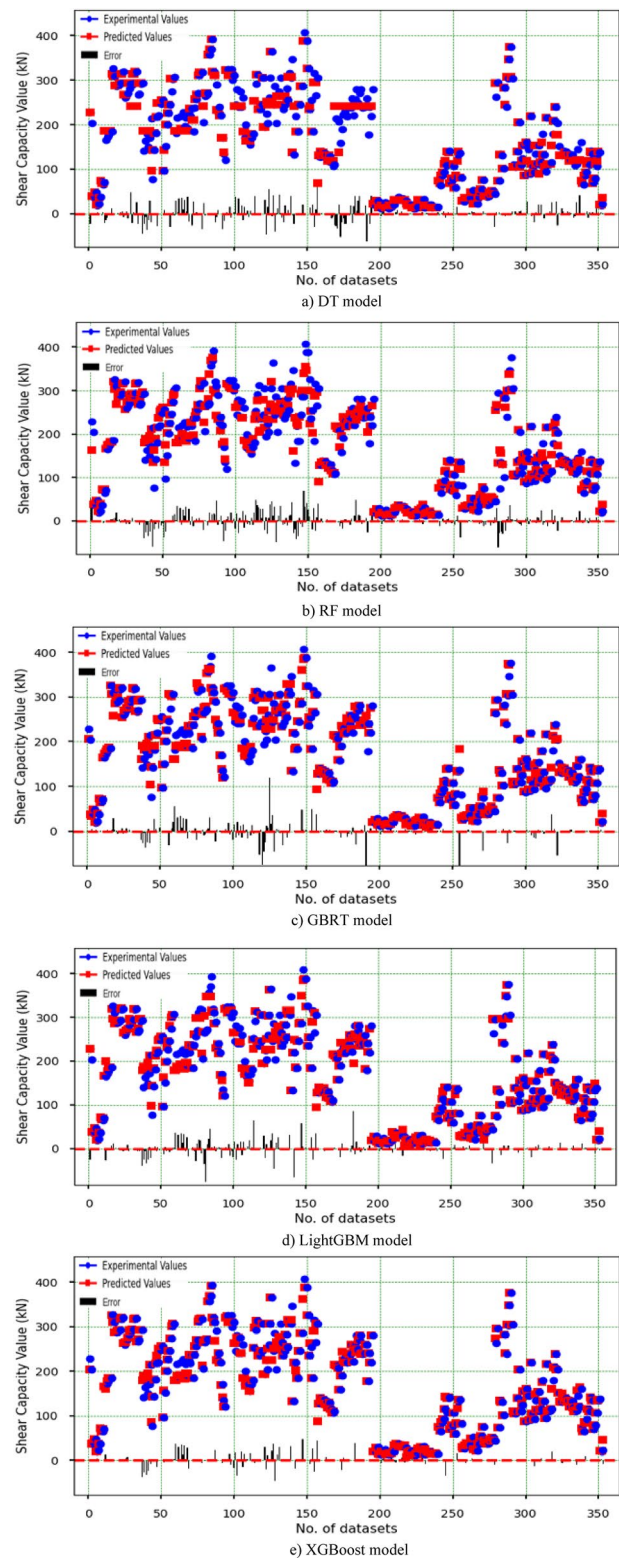


Fig. 13 Experimental versus predicted shear capacity of RCTBs with errors of **a** DT, **b** RF, **c** GBRT, **d** LightGBM, and **e** XGBoost models

Table 5 Existing design models for estimating the shear capacity of RCTBs

Models	Formulae
ACI 318-19, (2019)	If $A_V \geq A_{V,min}$ $V_n = \text{Either of } \begin{cases} \left[0.17\lambda\sqrt{f'_c} + \frac{N_u}{6A_g} \right] b_w d + \frac{A_v f_{yv} d}{s} \\ \left[0.66\lambda\rho_w^{1/3}\sqrt{f'_c} + \frac{N_u}{6A_g} \right] b_w d + \frac{A_v f_{yv} d}{s} \end{cases}$ If $A_V < A_{V,min}$ $V_n = \left[0.66\lambda_s\lambda\rho_w^{1/3}\sqrt{f'_c} + \frac{N_u}{6A_g} \right] b_w d + \frac{A_v f_{yv} d}{s}$, where $\lambda_s = \sqrt{\frac{2}{1 + \frac{d}{250}}} \leq 1.0$; $\frac{N_u}{6A_g} \leq 0.05f'_c$
BS 8110-1:1997, (1997)	$V_n = \frac{0.79\left(\frac{100A_s}{b_v d}\right)^{\frac{1}{3}}\left(\frac{400}{d}\right)^{\frac{1}{4}}\left(\frac{f_{cu}}{25}\right)^{\frac{1}{3}}}{\gamma_m} b_v d + \frac{0.95A_{sv}f_{yv}d}{s_v}$ where $f_{yv} = f_y$; $A_{sv} = A_s$; $s_v = s$; $b_v = b_w$; $f_{cu} = \frac{f'_c}{0.82}$; $\gamma_m = 1.25$ the limitations provided by BS 8110-1:1997 are as follows: $0.15 \leq \frac{100A_s}{b_v d} \leq 3.0$, $\frac{400}{d} \geq 1.00$ and $\frac{f_{cu}}{25} \leq 1.6$
Eurocode EN 1992-1-2, (2004)	$V_n = \left[C_{Rd,c}k(100\rho_l f_{ck})^{\frac{1}{3}} \right] b_w d + \frac{A_{sw}z f_{yw} \cot\theta}{s}$ where $C_{Rd,c} = 0.12$; $k = 1 + \sqrt{\frac{200}{d}}$; $f_{ck} = f'_c$; $\rho_l = \rho_w \leq 0.02$; $A_{sw} = A_s$; $z = \text{lever arm taken as } 0.9d$ $f_{yw} = f_y$; $\theta = \text{compression strut angle } 45^\circ \text{ or } 21.8^\circ$
CSA23.3-04 model, (2004)	$V_n = \left[\lambda\beta\sqrt{f'_c} \right] b_w d_v + \frac{A_v f_{yv} \cot\theta}{s}$ where $d_v = \max(0.9d, 0.72h)$ If $h < 250$ mm, $\beta = 0.21$, $\theta = 42^\circ$ If $f_y < 400$ MPa and $f'_c > 35$ MPa; $\theta = 35^\circ$ and $\beta = 0.18$ if $A_V \geq A_{V,min}$, Otherwise $\beta = 230/(1 + d_v)$
Sarsam et al., (2018)	$V_n = 1.7 \left(f'_c \rho_w \frac{V_{UD}}{M_U} \right)^{0.5} \alpha + \frac{\phi_s A_v f_{yt} d}{s}$ where $\frac{V_{UD}}{M_U} \leq 0.5$; $\alpha = b_w d + 0.7(b - b_w)h_f$
Thamrin et al., (2016)	$V_n = \frac{1}{2} \left[\alpha \sqrt{f'_c} + 120\rho_w \frac{d}{a} \right] b_w d$ where $\alpha = 1.0 + \left(\frac{b_f h_f}{4d^2} \right)$
Zararis et al., (2006)	$V_n = \left[1.2 - 0.2 \frac{a}{d} \right] \frac{\zeta}{d} f_{ct} b_{ef} d + \left[0.5 + 0.25 \frac{a}{d} \right] \rho_v f_{yv} b_w d$ where $\left[1.2 - 0.2 \frac{a}{d} \right] \geq 0.65$ (d in meters) $f_{ct} = 0.3f'_c^{\frac{2}{3}}$ $\rho_v = \frac{A_v}{b_w s}$ $b_{ef} = b_w \left[1 + \frac{0.5 \frac{h_f}{d} \left(\frac{b_f}{b_w} - 1 \right)}{\zeta/d} \right]$ $\left[\frac{\zeta}{d} \right]^2 + \left[1.5 \frac{h_f}{d} \left[\frac{b_f}{b_w} - 1 \right] + \left(600 \frac{\rho + \rho'}{f'_c} \right) \right] \frac{\zeta}{d} - 600 \frac{\rho + \rho'}{f'_c} = 0.0$

constant to identify the effects of parameters in modeling. the sensitivity analysis in this study is focused on parameters: web width (b_w), effective depth of beam (d), flange width (b_f), flange thickness (h_f), compressive strength of concrete (f'_c), shear span-to-depth ratio (a/d), Longitudinal steel ratio (ρ), and ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$). The result of the sensitivity analysis is presented in Fig. 16. In general, an increase in all input features results in an increase in the ultimate shear capacity of RCTBs, except for the increase in a/d , which decreases the ultimate shear capacity of RCTBs, as shown in Fig. 16f. This also indicates that increasing the flange dimensions enhances the shear capacity RCTBs, as shown in Fig. 16c, d. Furthermore,

the predicted outcomes from sensitivity analysis of the XGBoost model for the shear capacity of RCTBs with diverse geometric and material properties align with the findings reported in references (Ayensa et al., 2019; Hawileh et al., 2022; Kadr et al., 2019; Ramadan et al., 2022) concerning the structural behavior of RCTBs.

5.5 Noise sensitivity analysis

Noise, defined as unwanted or irrelevant information in data, can significantly impact the performance of machine learning algorithms. Understanding the effects of noise on model performance is crucial for ensuring the robustness and reliability of predictive models. The impact of noise on XGBoost model performance was analyzed by comparing the performance metrics

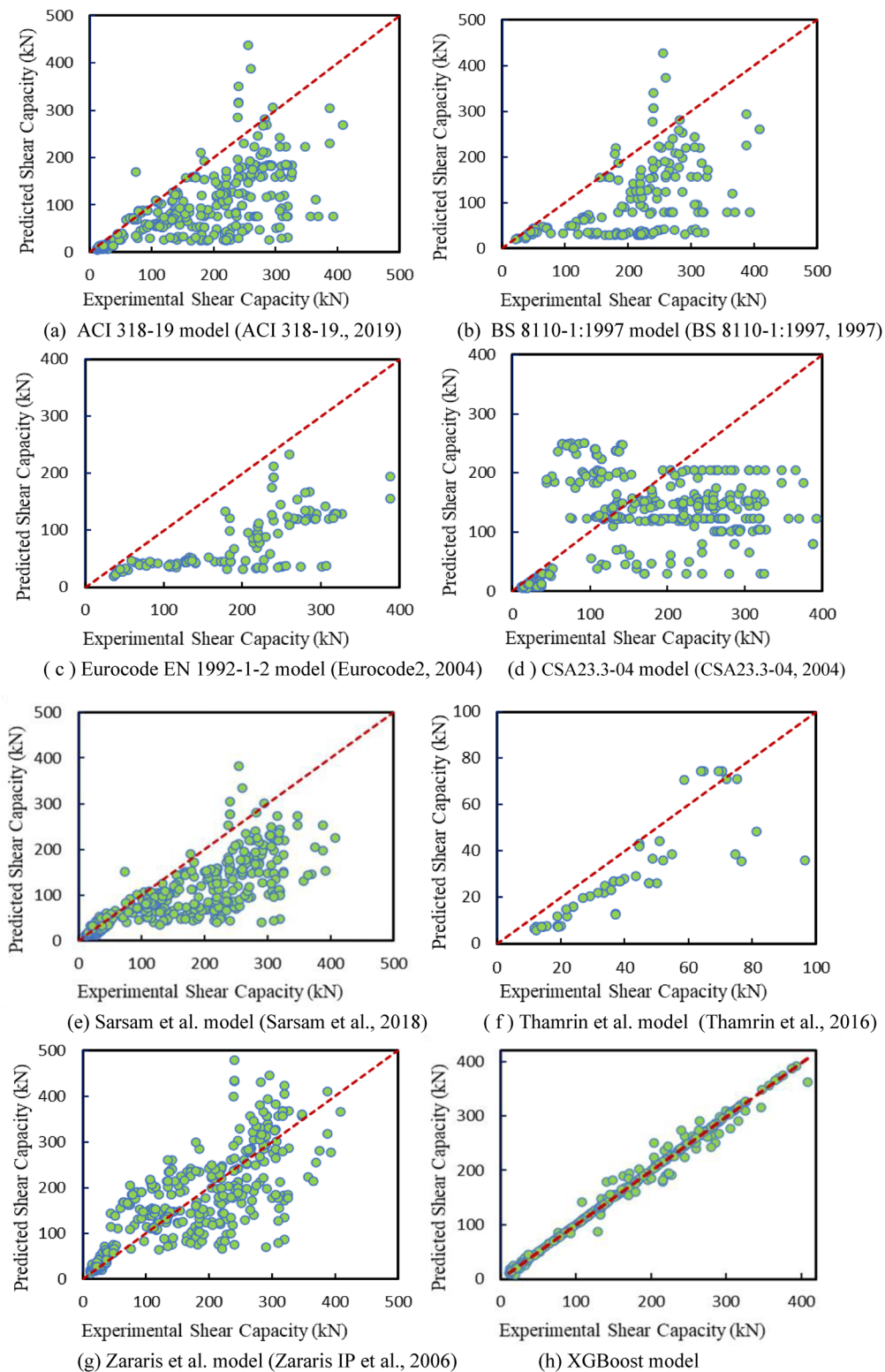


Fig. 14 Comparison of predicted shear capacities by shear design models and XGBoost model. **a** ACI 318–19, model (2019). **b** BS 8110–1:1997, model (1997). **c** Eurocode EN 1992–1-2, model (2004). **d** CSA23.3–04, model (2004). **e** Sarsam et al., model (2018). **f** Thamrin et al. model (2016). **g** Zararis et al., model (2006). **h** XGBoost model

Table 6 Assessment of shear design models and XGBoost model based on V_{Exp} to V_{pred} ratio

Model	Mean	STD	COV
ACI 318-19, model (2019)	2.31	1.67	0.73
BS 8110-1:1997, model (1997)	2.94	2.17	0.74
Eurocode EN 1992-1-2, model (2004)	2.82	1.55	0.55
CSA23.3-04 model (2004)	1.87	1.40	0.75
Sarsam et al., model (2018)	1.78	0.95	0.54
Thamrin et al., model (2016)	1.65	0.53	0.32
Zararis et al., model (2006)	1.71	1.10	0.64
XGBoost model	1.0045	0.0596	0.0593

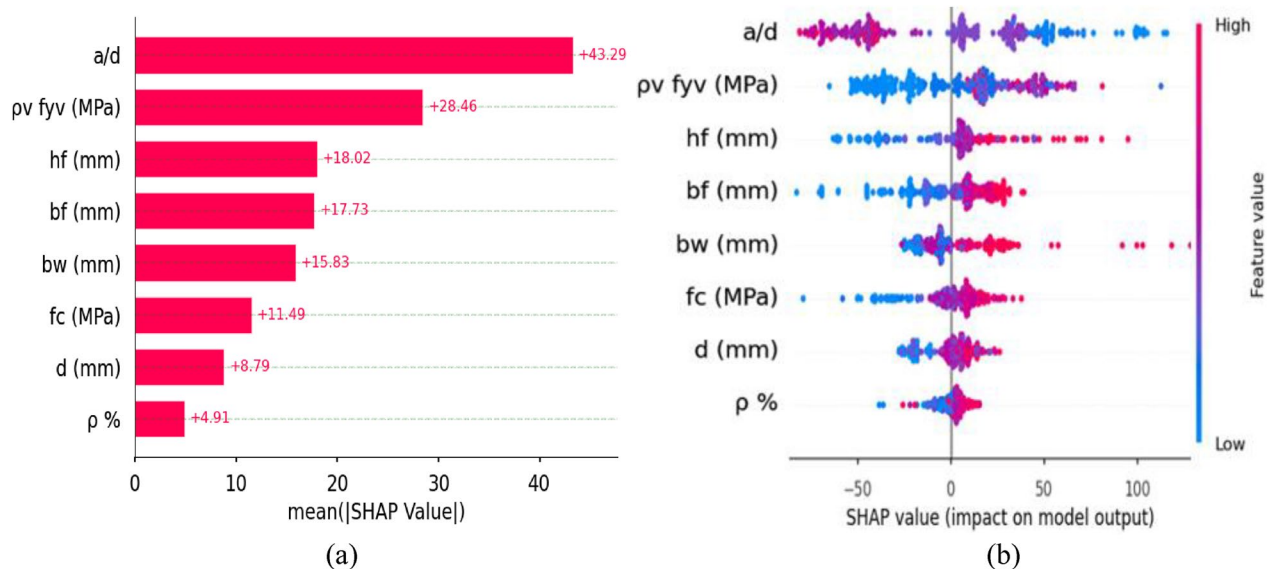
STD standard deviation, COV coefficient of variation

obtained under different noise levels. Trends and patterns in model behavior were identified to understand how noise influences predictive accuracy and generalization capabilities. The XGBoost model demonstrates robust performance with an R-squared value of 98% in the absence of noise. However, as sample noise increases by 100%, the model exhibits a notable decline in R-squared by 23.50%, resulting in a value of 75%. This decreasing trend persists with each successive 10% increase in sample noise from the noise-free condition, as illustrated in Fig. 17a. Similarly, concerning mean absolute error (MAE) and root mean squared error (RMSE), the XGBoost model achieves 10.9 kN and 15.6 kN, respectively, in the absence of noise. Nevertheless, with a 100% increase in sample noise, both MAE and RMSE experience a substantial uptrend, increasing by 60%. This increasing trend continues steadily with each 10% increment in sample noise from the noise-free condition, as

depicted in Fig. 17b, c. Moreover, concerning the mean absolute percentage error (MAPE), the XGBoost model exhibits a similar trend with an increase in sample noise. With a 100% increase in noise, the XGBoost experiences a 55% MAPE rise. The graphical illustration of the MAPE trend is presented in Fig. 17d. As a result, the XGBoost model exhibited a decrease in predictive accuracy with increasing levels of noise, indicating its sensitivity to noisy inputs.

5.6 Graphical User Interface

The study developed a graphical user interface (GUI) platform utilizing the most precise model, namely XGBoost, to enhance accessibility for both practical engineers and the research community. This platform provides an easy-to-use interface to input data and obtain predictions of the shear capacity of RCTBs. The use of this GUI platform eliminates the need for users to have extensive knowledge of machine learning techniques, making the research more accessible to a wider audience. Python library allowed the research team to build an interactive UI for the machine learning model. Users can provide input feature values and quickly obtain the corresponding shear strength value through this interface. The GUI platform is accessible via reference (Yehia, 2024). It is important to note that this GUI is exclusively applicable for RCTBs with geometric and material properties outlined in Table 1, as the XGBoost algorithm was trained using these specific ranges.

**Fig. 15** **a** Overall importance of the input features, and **b** summary plot of the input feature effects

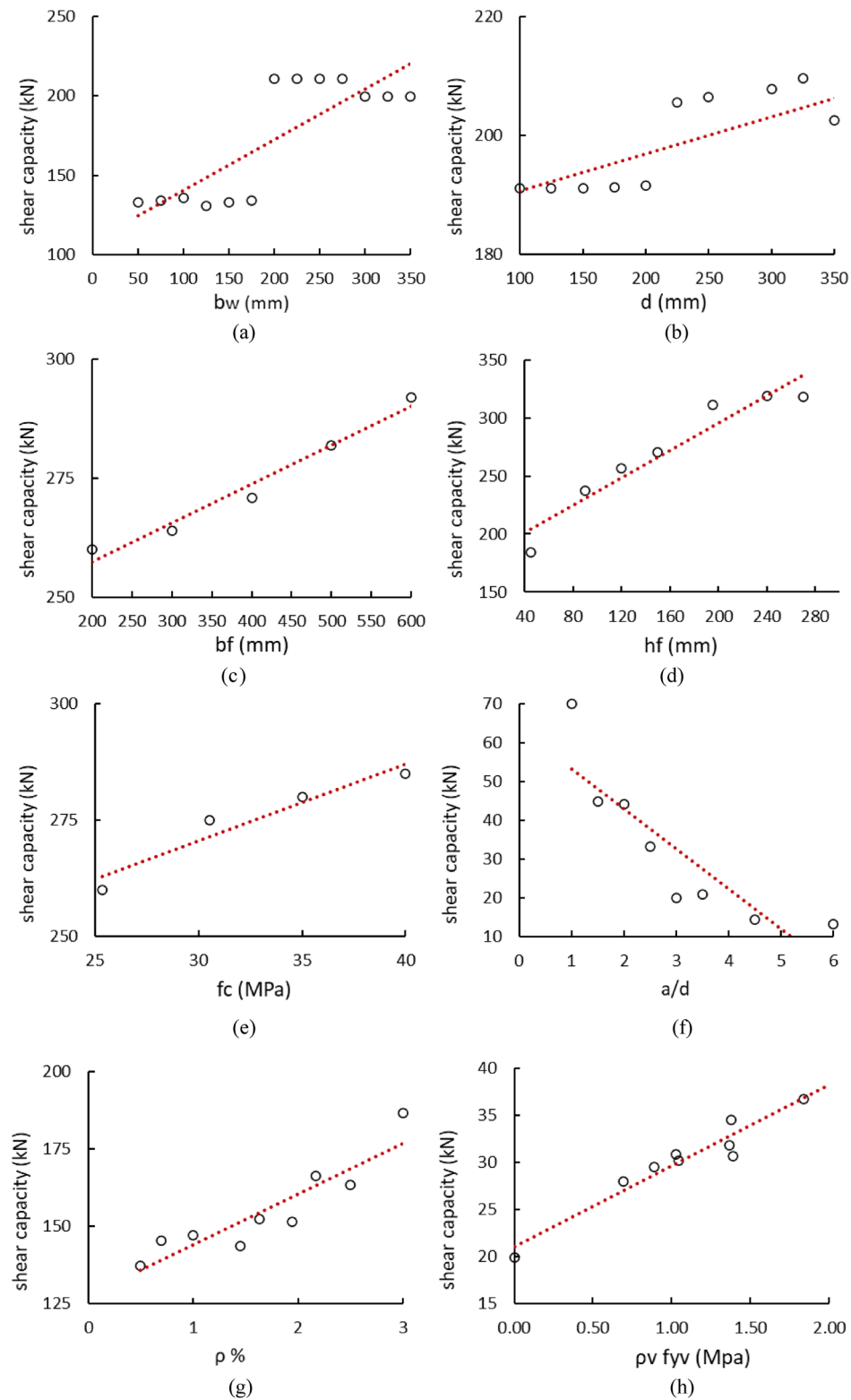


Fig. 16 Sensitivity analysis results for: **a)** web width (bw), **b)** effective depth of beam (d), **c)** flange width (bf), **d)** flange thickness (hf), **e)** compressive strength of concrete (f_c), **f)** shear span-to-depth ratio (a/d), **g)** longitudinal steel ratio (ρ), and **h)** ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$)

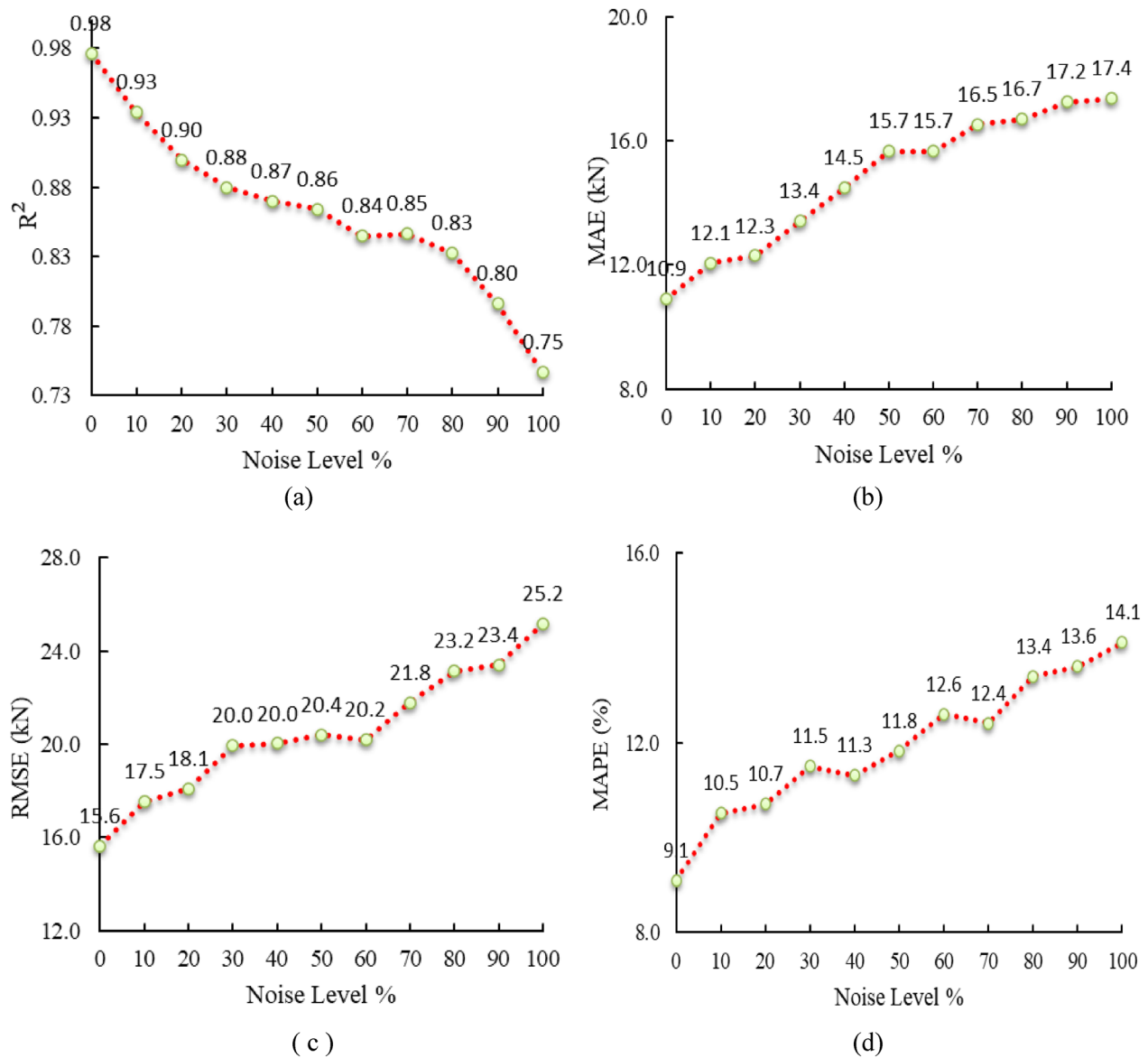


Fig. 17 Noise level versus **a** R^2 , **b** MAPE, **c** RMSE, and **d** MAPE of XGBoost model

6 Reliability Analysis for Shear Capacity of RCTBs

Reliability analysis in structural engineering is a process of evaluating the probability of failure of a structure. It is used to ensure that structures are designed to meet their intended function and performance requirements, and that they are safe under all anticipated loading conditions. To ensure safety, structures are designed such that their capacity (R) exceeds the demand (Q). The load and resistance factor design (LRFD) method is a common approach for calculating the limit state of a structure. In LRFD, the limit state is defined as the point at which the structure fails to meet its intended

function. The LRFD equation for the limit state in the resistance factor format is:

$$\phi R_n \geq \sum \gamma_i Q_i, \quad (5)$$

where the nominal resistance (R_n) was reduced by the capacity reduction factor (ϕ) to account for uncertainties in the material properties and loading conditions. The load effect due to each type of load (Q_i) was multiplied by the load partial safety factor (γ_i) to account for the variability of the loads. The load combination of 1.2DL + 1.6LL as specified in ACI 318 (ACI Committee 318, 2014) was

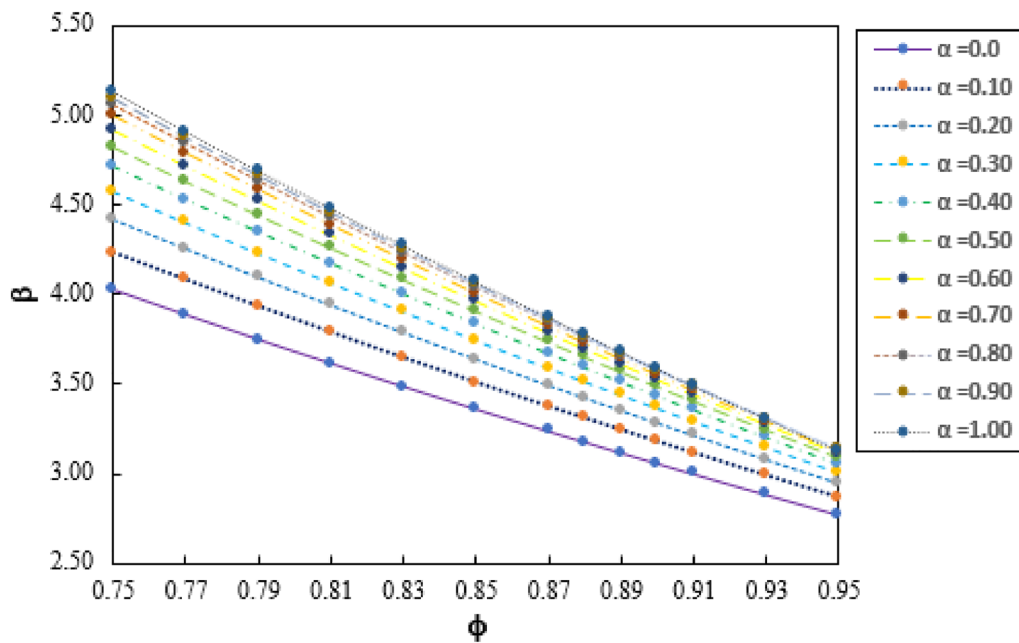


Fig. 18 α versus β response

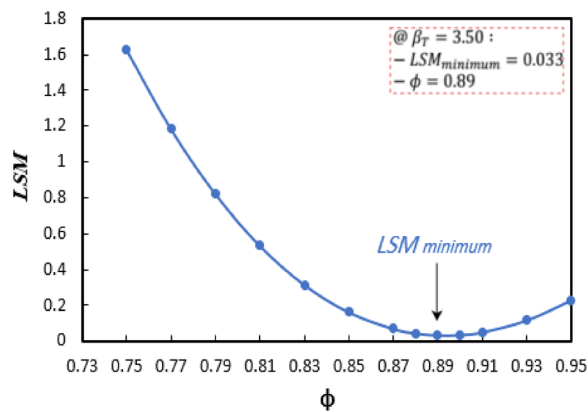


Fig. 19 Calibration of the shear strength reduction factor

used in the reliability analysis, where DL is the dead load and LL is the live load.

The target reliability index is selected based on the consequences of failure (Wight JK, 2016). A higher target reliability index corresponds to a lower probability of failure. According to (Wight JK, 2016), the target reliability index is typically between 3.0 and 3.5. Sze-rszen and Nowak, (2003) assumed a normal distribution for both the dead load and live load distributions. The bias of the dead load for cast-in situ concrete is 1.05, and the coefficient of variation is 0.1. The bias of the 50 year live load is 1.0, and the coefficient of variation is 0.18.

The reliability index β was calculated for a range of capacity reduction factors $\phi = 0.75:0.95$ and for each

case of $\alpha = 0.0:0.10:1.00$, where α is the ratio between live load to dead load. The results are plotted in Fig. 18. As expected, the value of β increased as ϕ decreased.

The resistance reduction factor was calibrated to achieve a target reliability index of 3.50. This was done using the least square method in Eq. (6):

$$LSM = \frac{1}{n} \sum_{i=1}^n (\beta_i - \beta_T)^2 \quad (6)$$

A larger safety margin is associated with a smaller resistance reduction factor. Figure 19 shows how the LSM varies as the resistance reduction factor (ϕ) changes for a target reliability index of 3.5. Furthermore, Fig. 19 shows that the minimum LSM corresponds to a capacity reduction factor of 0.89, which is recommended for RCTBs to achieve a target reliability index of 3.5 based on the proposed XGBoost model.

7 Conclusion

This paper has presented five machine learning (ML) models—Decision Trees (DT), Random Forest (RF), Gradient Boosted Regression Trees (GBRT), Light Gradient Boosting Machine (LightGBM), and Extreme Gradient Boosting (XGBoost)—for predicting the shear capacity of RCTBs. ML models were trained using extensive datasets of experimental data to comprehend the intricate relationships between input parameters and the corresponding shear capacity. The accuracy of shear design models in predicting the shear capacity of RCTBs was

also assessed. A user-friendly GUI tool has been developed, enabling practicing engineers to calculate the shear capacity of RTCBs. A reliability analysis was performed to determine the value of the resistance reduction factor that achieves a target reliability index.

The study yielded the following outcomes:

1. The XGBoost model surpasses other ML models in predicting the shear capacity of RTCBs, achieving an R^2 of 0.9914 and exhibiting the lowest values of RMSE (9.53 kN), MAE (3.75 kN), and MAPE (3.025%).
2. Among the existing design guidelines, the CSA23.3–04 (2004) code provided the most accurate predictions for the shear capacity of RCTBs, with an average ratio of V_{Exp} to V_{Pred} shear capacity of 1.87.
3. The model presented by Zararis et al. offered a higher degree of accuracy than other analysis models in determining the shear capacity of RCTBs, achieving an average ratio of V_{Exp} to V_{Pred} of 1.71 and a COV of 0.64.
4. The relative importance of input parameters on the shear capacity of RCTBs, as ranked by SHAP, is as follows: the shear span-to-depth ratio (a/d), ratio of shear reinforcement multiplied by the yield strength of shear reinforcement ($\rho_v f_{yv}$), flange thickness (h_f), and flange width (b_f).
5. A graphical user interface (GUI) platform based on the XGBoost model is developed and can be accessed through the following reference (Yehia, 2024).
6. A capacity reduction factor $\phi=0.89$ is calibrated to achieve a reliability index ($\beta_T=3.5$).

Furthermore, this study has potential limitations as the ML models provided were trained exclusively on the geometric and material properties of the experimental data gathered. Therefore, future research should expand beyond these specific ranges by conducting additional experiments. This approach will not only help address gaps in the dataset, but also enable multi-dimensional validation and refinement of the methodology employed in this study. Additionally, the current study neglected to consider the type of concrete when predicting the shear capacity of RCTBs. Therefore, future research should explore how different types of concrete affect shear capacity and focus on enhancing the existing ML models to incorporate this variable.

Abbreviations

A_g	Gross area of the section (mm ²)
A_s	Longitudinal steel reinforcement area (mm ²)
A'_s	Top longitudinal reinforcement (mm ²)
A_v	Total cross-sectional area of vertical stirrups reinforcement (mm ²)
b	Flange width (mm)
b_{ef}	Effective width (mm)

b_f	Width of the flange (mm)
h_f	Flange height (mm)
N_u	Factored axial load (N)
ρ'	Ratio of top reinforcement = $A'_s/b_w d$
β	Factor indicating ability of cracked concrete to transmit tension and shear
ρ_w	Ratio of longitudinal reinforcement = $A_s/b_w d$
ρ_t	Depth ratio = h_f/h
d'	Distance from extreme compression fiber to the top reinforcement (mm)
λ	Reduction factor for light weight concrete, taken as 1 for normal weight concrete
z	Lever arm taken as 0.9 times the effective depth
$C_{Rd,c}$	Coefficient for design shear resistance of concrete
M_U	Ultimate moment
b_w	Width of the web (mm)
c	Distance from the extreme compression fiber to neutral axis (mm)
d	Distance from extreme compression fiber to the bottom reinforcement (mm)
f'_c	Concrete compressive strength (MPa)
f_{ct}	Tensile strength of concrete (MPa)
f_y	Yield strength of longitudinal reinforcement (MPa)
f_{yv}	Yield strength of stirrups (MPa)
V_c	Shear contribution of concrete (N)
V_n	Nominal shear capacity of the member (N)
V_p	Component in the direction of the applied shear of the effective pre-stressing force
θ	Angle of inclination of diagonal compressive stresses
ρ_b	Width ratio = b_f/b_w
ρ_v	Ratio of shear reinforcement
s	Spacing of stirrups (mm)
γ_m	Material partial safety factor
f_{ck}	Characteristic cube strength of concrete
k	Factor considering shear size effect
ϕ_s	Resistance factor for shear reinforcement

Acknowledgements

The authors wish to acknowledge the financial support provided by The Science, Technology & Innovation Funding Authority (STDF) and The Egyptian Knowledge Bank (EKB). Many thanks also go to reviewers for their constructive suggestions.

Author contributions

Saad A. Yehia: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, visualization, project administration, writing—original draft, project administration. Sabry Fayed: software, investigation, data curation, data curation, writing—original draft. Mohamed H. Zakaria: software, investigation, data curation, data curation, writing—original draft. Ramy I. Shahin: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, visualization, writing—original draft.

Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Declarations

Informed consent

Informed consent was obtained from all individual participants included in this study.

Competing interests

The authors declare that they have no competing interests.

Received: 10 September 2023 Accepted: 14 April 2024
Published online: 13 August 2024

References

- Abambres, M., & Lantsoght, E. O. L. (2019). ANN-based shear capacity of steel fiber-reinforced concrete beams without stirrups. *Fibers*, 7(10), 88. <https://doi.org/10.3390/FIB7100088>
- Abd, R. T. (2019). Prediction of the shear strength of concrete T-beams using artificial neural networks model. *University of Thi-Qar Journal for Engineering Sciences*, 10(1), 168–181. [https://doi.org/10.3166/tqujes.10.1.354\(2019\)](https://doi.org/10.3166/tqujes.10.1.354(2019))
- ACI Committee 318. (2014). American Concrete Institute. Building code requirements for structural concrete (ACI 318–14): an ACI standard : commentary on building code requirements for structural concrete (ACI 318R-14). *An ACI Report*.
- ACI 318–19. (2019). *Building code requirements for structural concrete*. Farmington Hills: American Concrete Institute.
- Alshboul, O., Almasabha, G., Shehadeh, A., Mamlook, R. E. A., Almuflih, A. S., & Almakayee, N. (2022). Machine learning-based model for predicting the shear strength of slender reinforced concrete beams without stirrups. *Buildings*, 12(8), 1166.
- Amna, H. A., & Monstaser, W. M. (2019). Shear behaviour of reinforced light-weight concrete T-beams. *Life Sci J*, 16(8), 11–27. <https://doi.org/10.7537/marslsj160819.02>
- Ashour, A. F., Alvarez, L. F., & Toropov, V. V. (2003). Empirical modelling of shear strength of RC deep beams by genetic programming. *Computers & Structures*, 81(5), 331–338. [https://doi.org/10.1016/S0045-7949\(02\)00437-6](https://doi.org/10.1016/S0045-7949(02)00437-6)
- Ayensa, A., Oller, E., Beltrán, B., Ibarz, E., Mari, A., & Gracia, L. (2019). Influence of the flange's width and thickness on the shear strength of reinforced concrete beams with T-shaped cross section. *Engineering Structures*, 188, 506–518. <https://doi.org/10.1016/J.ENGSTRUCT.2019.03.057>
- Bairan Garcia, J. M., & Mari Bernat, A. R. (2006). Coupled model for the non-linear analysis of anisotropic sections subjected to general 3D loading. Part 1: theoretical formulation. *Computers & Structures*, 84(31–32), 2254–2263. <https://doi.org/10.1016/J.COMPSTRUCT.2006.08.036>
- Bousselham, A., & Chaallal, O. (2006). Behavior of reinforced concrete T-beams strengthened in shear with carbon fiber-reinforced polymer-an experimental study. *ACI Structural Journal*, 103(3), 339.
- Braestrup, M. W., Nielson, M. P., Bach, F., & Jensen, B. C. (1976). *Shear tests on reinforced concrete T-beams: series T*. Technical University of Denmark.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Bresler, B., & MacGregor, J. G. (1967). Review of concrete beams failing in shear. *Journal of the Structural Division*, 93(1), 343–372. <https://doi.org/10.1061/JSDAEG.0001586>
- BS 8110–1:1997. (1997). *British Standard. "Structural Use of Concrete, Part 2: Code of practice for design and construction"*. British Standard Institution, London.
- Chaabene, W. B., & Nehdi, M. L. (2020). Novel soft computing hybrid model for predicting shear strength and failure mode of SFRC beams with superior accuracy. *Composites Part C: Open Access*, 3, 100070. <https://doi.org/10.1016/J.JCOMC.2020.100070>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 13–17 August 2016, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Cladera, A., & Mari, A. R. (2004). Shear design procedure for reinforced normal and high-strength concrete beams using artificial neural networks. Part I: beams without stirrups. *Engineering Structures*, 26(7), 917–926. <https://doi.org/10.1016/J.ENGSTRUCT.2004.02.010>
- Cladera, A., Mari, A., Ribas, C., Bairán, J., & Oller, E. (2015). Predicting the shear-flexural strength of slender reinforced concrete T and I shaped beams. *Engineering Structures*, 101, 386–398. <https://doi.org/10.1016/J.ENGSTRUCT.2015.07.025>
- CSA23.3–04. (2004). Design of concrete structures. *Mississauga, Ontario, Canada*.
- Elgohary, A., Abdelhazef, A., & Asran, A. (2019). Effect of flange width on shear strength of RC T-beams. *Journal of Al Azhar University Engineering Sector*, 14(52), 875–882.
- Eurocode2. (2004). European Committee for Standardization. "Design of concrete structures - Part 1–2". Brussels.
- Ferguson, P. M., & Thompson, J. N. (1953). Diagonal tension in T-beams without stirrups. *Journal Proceedings*, 49(3), 665–675. <https://doi.org/10.1435/11844>
- Bach, F., Nielsen, M. P., & Bræstrup, M. W. (1980). *Shear tests on reinforced concrete T-beams series V, U, X, B and S*. Technical University of Denmark, Department of Civil Engineering.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.
- Giaccio, C., Al-Mahaidi, R., & Taplin, G. (2011). Experimental study on the effect of flange geometry on the shear strength of reinforced concrete T-beams subjected to concentrated loads. *Canadian Journal of Civil Engineering*, 29(6), 911–918. <https://doi.org/10.1139/L02-099>
- Goh, A. T. C. (1995). Prediction of ultimate shear strength of deep beams using neural networks. *Structural Journal*, 92(1), 28–32. <https://doi.org/10.1435/1470>
- Hans Christian. (1974). *Shear Tests on 12 Reinforced Concrete T-Beams*. Technical University of Denmark Department of Civil Engineering.
- Hawileh, R. A., Mhanna, H. H., & Abdalla, J. A. (2022). Effect of flange geometry on the shear capacity of RC T-beams. *Procedia Structural Integrity*, 42, 1198–1205. <https://doi.org/10.1016/J.PROSTR.2022.12.153>
- Huynh, T. Q., Nguyen, T. T., & Nguyen, H. (2022). Base resistance of super-large and long piles in soft soil: performance of artificial neural network model and field implications. *Acta Geotechnica*. <https://doi.org/10.1007/S11440-022-01736-w>
- Jiang, C. S., & Liang, G. Q. (2021). Modeling shear strength of medium- to ultra-high-strength concrete beams with stirrups using SVR and genetic algorithm. *Soft Computing*, 25(16), 10661–10675. <https://doi.org/10.1007/S00500-021-06027-2/METRICS>
- Jung, S., & Kim, K. S. (2008). Knowledge-based prediction of shear strength of concrete beams without shear reinforcement. *Engineering Structures*, 30(6), 1515–1525. <https://doi.org/10.1016/J.ENGSTRUCT.2007.10.008>
- Kadr, A., Asst, K., Abdul, J., & Khudhair, S. (2019). Experimental study of the shear strength of self compacting concrete T-beams made with recycled concrete as coarse aggregate. *University of Thi-Qar Journal*, 14(1), 31–52. <https://doi.org/10.3279/utq/utj/vol14/1/4>
- Ke G, Meng Q, Finley T, & Wang T. (2017). LightGBM: a highly efficient gradient boosting decision tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 3149–3157.
- Keshtegar, B., Bagheri, M., & Yaseen, Z. M. (2019). Shear strength of steel fiber-unconfined reinforced concrete beam simulation: application of novel intelligent model. *Composite Structures*, 212, 230–242. <https://doi.org/10.1016/J.COMPSTRUCT.2019.01.004>
- Kotsovos, M. D., Bobrowski, J., & Eibl, J. (1987). Behaviour of reinforced concrete T-beams in shear. *Structural (The) Engineer. Part b: the Journal of the Institution of Structural Engineers-Quarterly*, 65(1), 1–10.
- Kotsovos, G. M., & Cotsovos, D. M. (2018). Shear failure criterion for RC T-beams. *Engineering Structures*, 160, 44–55. <https://doi.org/10.1016/J.ENGSTRUCT.2017.12.044>
- Le-Nguyen, K., Minh, Q. C., Ahmad, A., & Ho, L. S. (2022). Development of deep neural network model to predict the compressive strength of FRCM confined columns. *Frontiers of Structural and Civil Engineering*, 16(10), 1213–1232. <https://doi.org/10.1007/S11709-022-0880-7/METRICS>
- Leonhardt F, & Walther R. (1962). Schubversuche an Einfeldrigen Stahlbetonbalken mit und ohne Schubbewehrung zur Ermittlung der Schubtragfähigkeit und der Oberen Schubspannungsgrenze. Heft 151, Deutscher Ausschuss für Stahlbeton. Berlin:W. Ernst & Sohn, 68P.
- Mhanna, H. H., Hawileh, R. A., & Abdalla, J. A. (2020). Shear strengthening of reinforced concrete T-beams using CFRP laminates anchored with bent CFRP splay anchors. *Procedia Structural Integrity*, 28, 811–819. <https://doi.org/10.1016/J.PROSTR.2020.10.095>
- Mhanna, H. H., Hawileh, R. A., & Abdalla, J. A. (2021a). Shear behavior of RC T-beams externally strengthened with anchored high modulus carbon fiber-reinforced polymer (CFRP) laminates. *Composite Structures*, 272, 114198. <https://doi.org/10.1016/J.COMPSTRUCT.2021.114198>
- Mhanna, H. H., Hawileh, R. A., Abdalla, J. A., Salama, A. S. D., & Alkhrdaji, T. (2021b). Shear strengthening of reinforced concrete T-beams with anchored CFRP laminates. *Journal of Composites for Construction*, 25(4), 04021030. [https://doi.org/10.1061/\(ASCE\)CC.1943-5614.0001141](https://doi.org/10.1061/(ASCE)CC.1943-5614.0001141)
- Moayer, M., & Regan, P. E. (1974). Shear strength of prestressed and reinforced concrete T-beams. *Special Publication*, 42, 183–214. <https://doi.org/10.1435/17284>

- Nguyen, H. D., LaFave, J. M., Lee, Y. J., & Shin, M. (2022). Rapid seismic damage-state assessment of steel moment frames using machine learning. *Engineering Structures*, 252, 113737. <https://doi.org/10.1016/J.ENGSTRUCT.2021.113737>
- Nguyen, H. D., Truong, G. T., & Shin, M. (2021a). Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. *Engineering Structures*, 235, 112067. <https://doi.org/10.1016/J.ENGSTRUCT.2021.112067>
- Nguyen, H., Vu, T., Vo, T. P., & Thai, H. T. (2021b). Efficient machine learning models for prediction of concrete strengths. *Construction and Building Materials*, 266, 120950. <https://doi.org/10.1016/J.CONBUILDMAT.2020.120950>
- Nguyen-Sy, T., Wakim, J., To, Q. D., Vu, M. N., Nguyen, T. D., & Nguyen, T. T. (2020). Predicting the compressive strength of concrete from its compositions and age using the extreme gradient boosting method. *Construction and Building Materials*, 260, 119757. <https://doi.org/10.1016/J.CONBUILDMAT.2020.119757>
- Ni, X., & Duan, K. (2022). Machine learning-based models for shear strength prediction of UHPFRC beams. *Mathematics*, 10(16), 2918.
- Palaskas, M. N., Attiogbe, E. K., & Darwin, D. (1981). Shear strength of lightly reinforced T-beams. *Journal of the American Concrete Institute*, 78(6), 447–455.
- Pansuk, W., Ueda, T., & Takahashi, R. (2004). Investigation on shear capacity of reinforced concrete T-Beams using 3d nonlinear finite element analysis. *JSCCE*, 50A(3), 991–998.
- Placas, A., Regan, P. E., & Baker, A. L. L. (1971). Shear failure of reinforced concrete beams. *ACI Journal*, 68(10), 763–773. <https://doi.org/10.1435/15237>
- Rahman, J., Ahmed, K. S., Khan, N. I., Islam, K., & Mangalathu, S. (2021). Data-driven shear strength prediction of steel fiber reinforced concrete beams using machine learning approach. *Engineering Structures*, 233, 111743. <https://doi.org/10.1016/J.ENGSTRUCT.2020.111743>
- Ramadan, O. M., Abdel-Kareem, A. H., El-Azab, I. A., & Abousafa, H. R. (2022). Flange contribution to the shear strength of RC T-beams with flange in compression. *Buildings*, 12(6), 803. <https://doi.org/10.3390/BUILDINGS12060803>
- Ribas González, C. R., Fernández Ruiz, M., Rodrigo, C., & González, R. (2017). Influence of flanges on the shear-carrying capacity of reinforced concrete beams without web reinforcement. *Structural Concrete*, 18(5), 720–732. <https://doi.org/10.1002/SUCO.201600172>
- Sahoo, D. R., Bhagat, S., & Reddy, T. C. V. (2016). Experimental study on shear span to effective-depth ratio of steel fiber reinforced concrete T-beams. *Materials and Structures*, 49(9), 3815–3830. <https://doi.org/10.1617/S11527-015-0756-6/TABLES/5>
- Samad, A. A. A., Mohamad, N., Al-Qershi, M. A. H., Jayaprakash, J., & Mendis, P. (2016). Shear mechanism and shear strength prediction of reinforced concrete Tbeams. *Jurnal Teknologi*, 78(5), 471–476.
- Sarsam, K., Khalel, R., & Mohammed, N. (2018). Influence of flange on the shear capacity of reinforced concrete beams. *MATEC Web of Conferences*, 162, 04003. <https://doi.org/10.1051/MATECONF/201816204003>
- Sarveghadi, M., Gandomi, A. H., Bolandi, H., & Alavi, A. H. (2019). Development of prediction models for shear strength of SFRCB using a machine learning approach. *Neural Computing and Applications*, 31(7), 2085–2094. <https://doi.org/10.1007/S00521-015-1997-6/METRICS>
- Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *NPJ Computational Materials*. <https://doi.org/10.1038/S41524-019-0221-0>
- Schober, P., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia and Analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Shahin, R. I., Ahmed, M., Yehia, S. A., & Liang, Q. Q. (2023). ANN model for predicting the elastic critical buckling coefficients of prismatic tapered steel web plates under stress gradients. *Engineering Structures*, 294, 116794. <https://doi.org/10.1016/J.ENGSTRUCT.2023.116794>
- Shahnewaz, M., & Alam, M. S. (2020). Genetic algorithm for predicting shear strength of steel fiber reinforced concrete beam with parameter identification and sensitivity analysis. *Journal of Building Engineering*, 29, 101205. <https://doi.org/10.1016/J.JOBE.2020.101205>
- Shatnawi, A., Alkassar, H. M., Al-Abdaly, N. M., Al-Hamdany, E. A., Bernardo, L. F. A., & Imran, H. (2022). Shear strength prediction of slender steel fiber reinforced concrete beams using a gradient boosting regression tree method. *Buildings*, 12(5), 550.
- Solhmirzaei, R., Salehi, H., Kodur, V., & Naser, M. Z. (2020). Machine learning framework for predicting failure mode and shear capacity of ultra- high performance concrete beams. *Engineering Structures*, 224, 111221. <https://doi.org/10.1016/J.ENGSTRUCT.2020.111221>
- Sutton, C. D. (2005). Classification and regression trees, bagging, and boosting. *Handbook of statistics*, 24:303–329.
- Swamy, R. N., & Qureshi, S. A. (2015). Shear behaviour of reinforced concrete t beams with web reinforcement. *University of Melbourne*, 57(Part 2), 35–49. <https://doi.org/10.1680/IICEP.1974.4097>
- Swamy, R. N., Qureshi, S. A., Eng, C., Struct, M. I., Eng, B., & Eng, M. (1974). An ultimate shear strength theory for reinforced concrete T-beams without web reinforcement. *Matériaux Et Construction*, 7(3), 181–189. <https://doi.org/10.1007/BF02473833>
- Szerszen, M. M., & Nowak, A. S. (2003). Calibration of design code for buildings (ACI 318): part 2—reliability analysis and resistance factors. *Structural Journal*, 100(3), 383–391. <https://doi.org/10.1435/12614>
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192. <https://doi.org/10.1029/2000JD900719>
- Thai, H. T. (2022). Machine learning for structural engineering: a state-of-the-art review. *Structures*, 38, 448–491. <https://doi.org/10.1016/J.JISTRUC.2022.02.003>
- Thamrin, R., Tanjung, J., Aryanti, R., Fitrah Nur, O., & Devinus, A. (2016). Shear strength of reinforced concrete t-beams without stirrups. *Journal of Engineering Science and Technology*, 11(4), 548–562.
- Tureyen, A. K., & Frosch, R. J. (2003). Concrete shear strength: another perspective. *Structural Journal*, 100(5), 609–615. <https://doi.org/10.1435/12802>
- Vadyala, S. R., Betgeri, S. N., Matthews, J. C., & Matthews, E. (2022). A review of physics-based machine learning in civil engineering. *Results in Engineering*, 13, 100316. <https://doi.org/10.1016/J.RINENG.2021.100316>
- 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 FRM. *Engineering Structures*, 255, 113903. <https://doi.org/10.1016/J.ENGSTRUCT.2022.113903>
- Wang, A. Y. T., Murdock, R. J., Kauwe, S. K., Oliynyk, A. O., Gurlo, A., Brgoch, J., Persson, K. A., & Sparks, T. D. (2020). Machine learning for materials scientists: an introductory guide toward best practices. *Chemistry of Materials*, 32(12), 4954–4965. https://doi.org/10.1021/ACS.CHEMMATER.0C01907/SUPPL_FILE/CMOC01907_SI_001.PDF
- Wehr, K. E. (1967). *Shear strength of reinforced concrete T-beams*. Lafayette: Research Report No.4 on Joint Research Highway Project, Purdue University.
- Wei, J., et al. (2019). Machine learning in materials science. *Infomat*, 1(3), 338–358. <https://doi.org/10.1002/Inf2.1.2028>
- Wight, J. K. (2016). *Reinforced concrete mechanics and design* (7th ed.). Pearson Education Inc.
- Wolf, T. S., & Frosch, R. J. (2007). Shear design of prestressed concrete: a unified approach. *Journal of Structural Engineering*, 133(11), 1512–1519. [https://doi.org/10.1061/\(ASCE\)0733-9445\(2007\)133:11\(1512\)](https://doi.org/10.1061/(ASCE)0733-9445(2007)133:11(1512))
- Yaseen, Z. M., Tran, M. T., Kim, S., Bakhshpoori, T., & Deo, R. C. (2018). Shear strength prediction of steel fiber reinforced concrete beam using hybrid intelligence models: a new approach. *Engineering Structures*, 177, 244–255. <https://doi.org/10.1016/J.ENGSTRUCT.2018.09.074>
- Ye, M., Li, L., Yoo, D. Y., Li, H., Zhou, C., & Shao, X. (2023). Prediction of shear strength in UHPC beams using machine learning-based models and SHAP interpretation. *Construction and Building Materials*, 408, 133752. <https://doi.org/10.1016/J.CONBUILDMAT.2023.133752>
- Yehia, S. A. (2024). *GUI of XGBoost-model*. <https://huggingface.co/spaces/saad-yehia/xgboost-model>
- Zararis, I. P., Karaveziroglou, M. K., & Zararis, P. D. (2006). Shear strength of reinforced concrete T-beams. *ACI Structural Journal*, 103, 693–700.
- 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 with Computers*, 38(2), 1293–1307. <https://doi.org/10.1007/S00366-020-01076-X/FIGURES/12>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Saad A. Yehia Is an Assistant Professor of the Civil Engineering Department, Higher Institute of Engineering and Technology, Kafrelsheikh 33511 Egypt.

Sabry Fayed Is an Associate Professor of the Civil Engineering Department, Faculty of Engineering Kafrelsheikh University, Kafrelsheikh 33511 Egypt.

Mohamed H. Zakaria Is an Assistant Professor of the Civil Engineering Department, Faculty of Engineering, Kafrelsheikh University, Kafrelsheikh 33511 Egypt.

Ramy I. Shahin Is an Assistant Professor of the Civil Engineering Department, Higher Institute of Engineering and Technology, Kafrelsheikh 33511 Egypt.