Fusion Journal of Engineering and Sciences

Type of the Paper (Article, Review, Communication, etc.)

Intelligent Modelling of Punching Shear Strength in FRP Concrete Slabs Using GEP and Machine Learning

Ahad Ali [1*], Ali Husnain [1*], Aqib Irfan [2]

[1] Department of Civil Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Swabi, KPK, 23640, Pakistan
* ahaddalii33@gmail.com (A.H), ali242husnain@gmail.com (H.A.H)

Abstract

The increasing demand for resilient structures in earthquake-prone and high-impact areas has necessitated accurate prediction of the punching shear strength of concrete slabs reinforced with FRP (Fiber Reinforced Polymer) and without shear reinforcement. The classical approaches, including physical testing and mathematically derived models, tend to have limitations in terms of both accuracy and speed. The present work carries out the prediction of the punching shear strength of FRP-reinforced concrete slabs by employing various machine learning (ML) techniques like decision tree, random forest, stochastic gradient boosting, and so on. The research conducted involved a thorough study through a dataset comprising of experiments on slabs with different design parameters including thickness, strength of the concrete, type of FRP reinforcement, and ratio of reinforcement, which is then using to train and evaluate the ML-based models. Out of all the models that were put to the test, the GEP (Gene Expression Programming) method attained the utmost R^2 value equal to

0.86 which is an indication of getting shear strength prediction more accurately than traditional models. Furthermore, the ML models were compared through statistical errors including RMSE, MAE, and R^2 with the attained values being within the acceptable range. The analysis of feature importance showed that the parameters slab thickness, concrete strength, and reinforcement ratios were the key players in the determination of the punching shear strength. The present study clearly points to the machine learning potentiality in reducing the effort and time in punching shear strength prediction, the outcome being better design codes and standards for FRP-reinforced concrete slabs in the future.

1. Introduction

Punching shear, a brittle failure mode, happens in reinforced concrete slabs when heavy loads, like those from columns, apply high shear stresses in the vicinity of the contact area. Consequently, the slab may experience an instantaneous failure characterized by the column's perimeter being punched out (Muttoni & Schwartz, 1991). The event is marked by a fracture resembling a truncated cone and is predominantly found in flat slabs and footings where shear reinforcement is very low (Guandalini et al., 2009). The failure that occurs is instant and without any prior indications, thus, it is a significant risk to safety in the case of the structure being applied. The factors that mainly determine the performance of punching shear are the slab's thickness, the strength of the concrete, the size of the loading area (Shahiduzzaman & Hossain, 2024), and reinforcement detailing (ACI 421.1R). Punching shear failure, which is so suddenly initiated, represents a major

hazard to safety and is a major concern in structural design. The growing use of fiber reinforced polymers (FRP) as alternatives to conventional steel reinforcement aggravates this problem. Although FRP provides an extraordinary ability of durability and resistance against environmental degradation, the mechanical response of FRPs is rather different from that of steel, thus requiring advanced predictive methodologies for structural performance (see reference [2]).

Over the past two decades, the discipline of structural engineering has continually aimed to introduce new types of materials and construction methods to obtain safer concrete structures, a longer service life and higher economic efficiency. Fiber-reinforced polymers (FRP) come under these materials, and they are considered as potential competing steel reinforcement materials. Their remarkable strength-to-weight ratio, corrosion resistance and easy installation make them beneficial for many structural applications. Nonetheless, the use of FRPs in concrete slabs remains debatable due to punching shear strength, which is the design element governing the failure of flat slabs. Furthermore, the complete lack of shear reinforcement in FRP-slab systems highlights the need to carefully establish punching shear capacity limits, to ensure structural integrity and to advance the further use of FRP in concrete construction.

I have observed that the increase in the application of fiber reinforced polymers (FRP) in structural problems usually has us at a loss as to how to address the conventional intuitive approaches to attacking punching shear in FRP-reinforced slabs. Recently, the developments in machine learning (ML) appear to be the actual game-changer of this problem. Since ML can identify irregular, nonlinear trends within huge volumes of data, it can enhance the precision of punching shear strength forecasts of such slabs to a considerable degree.

Although empirical methods are no longer used, scholars have demonstrated that ML models such as artificial neural networks (ANN), support vectors machine (SVM), and decision trees provide more accurate predictions compared to the previous ones by considering the nuanced interaction of material properties, slab geometry and loading conditions (Xia et al., 2018; Zhang et al., 2020).

2. Materials and Methods

The Materials and Methods should be described with sufficient details to allow others to replicate and build on the published results. Please note that the publication of your manuscript implicates that you must make all materials, data, computer code, and protocols associated with the publication available to readers. Please disclose at the submission stage any restrictions on the availability of materials or information. New methods and protocols should be described in detail while well-established methods can be briefly described and appropriately cited.

Research manuscripts reporting large datasets that are deposited in a publicly available database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication.

Interventionary studies involving animals or humans, and other studies that require ethical approval, must list the authority that provided approval and the corresponding ethical approval code.

In this section, where applicable, authors are required to disclose details of how generative artificial intelligence (GenAI) has been used in this paper (e.g., to generate text, data, or graphics, or to assist in study design, data collection, analysis, or interpretation). The use of GenAI for superficial text editing (e.g., grammar, spelling, punctuation, and formatting) does not need to be declared.

1.2 Literature Review

Throughout the years, the punching shear behavior of slabs with FRP reinforcement has been the subject of numerous experimental studies aimed towards gaining a better understanding of the performance- determining parameters. In their early work, Matthys and Taerwe (2000) pointed out that no less than equal flexural stiffness with that of steel reinforcement, the FRP- RC slabs, still, might fail. Thus, the need for material stiffness compatibility could not be emphasized more in the case of a simplified design approach. The bond performance between the FRP reinforcement and the concrete was also discussed by El-Ghandour et al. (2003) and Ospina et al. (2003). They provided convincing evidence that weak interfacial adhesion could not only result in premature failure but also lower the shear capacity. Lee et al. (2009) studied the GFRP reinforcement and, after that, gave below, the suggestion on bar placement: Concentrating bars is within a distance of 1.5 times the slab thickness from the column edges, helps not only in stress distribution but quite significantly, even, in punching shear resistance, too. Looking deeper into the influencing factors, Dulude et al. (2013) and Hassan et al. (2013b) made a strong case for considering not only slab thickness, column size, and material but also concrete compressive strength. Their results showed that going up from 30 to 50 MPa for the concrete might provide a shear capacity increase by almost 18-25% which is a case of concrete quality's strong influence on the punching shear behavior. There were many empirical and semi-empirical predictive models developed based on the experimental works that were done at the foundation. Theodorakopoulos and Swamy (2002) came up with a model that was based on 60 slab–column joints and provided an easy way to estimate punching shear capacity via this analytical model. Elshafey et al. (2003) introduced two empirical models that were based on 244 experimental data points and claimed that they had excellent predictability and quite a substantial generality. Similarly, El-sanadedy et al. (2004) presented a model after analyzing 61 high-strength concrete slabs that was very precise in predicting the punching shear strength and turned out to be very advantageous for the industry.

In the realm of design standards, BS 8110-97, ACI 318-19, and Eurocode 2 have all added different empirical equations to their toolkit for estimating punching shear capacity, thus reflecting the agreement from many years of experiments. Besides, the parameters used in these codes, such as the reinforcement ratio, column size, slab thickness, and concrete strength, were all shown to have a significant effect on the punching process. To illustrate, ACI 440.1R-15 gives a procedure for calculating the concrete contribution to shear resistance at a critical section perimeter that is defined at the depth of the slab, i.e., 0.5 times the effective slab depth from the column faces. This model modifies the traditional ACI 318-14 approach for steel-RC slabs to FRP systems by allowing for the inclusion of FRP property modifications like reinforcement ratio, modulus of elasticity, and the square root of concrete compressive strength.

Although empirical and code-based models are very useful, their predictive power is frequently reduced in situations with complex loading or advanced materials such as FRP. To overcome these limitations, machine learning (ML) techniques have recently been adopted in research. Mangalathu et al. (2020) considered several ML algorithms and showed that

XGBoost was the best in predicting the punching shear capacity of conventional RC slabs, getting R^2 of 0.96 by revealing nonlinear interactions among important parameters such as slab thickness and concrete strength. Alotaibi et al. (2021) proved that artificial neural networks (ANNs) were able to process the high-dimensional input data and give predictions that were strong and reliable. Mostafa et al. (2021) used support vector machines (SVM) and produced an R^2 of 0.95, whereas Chetchotisak et al. (2020) developed a multiple linear regression (MLR) model based on 342 experimental samples that was more

accurate than the traditional design codes in critical cases. However, these ML models also have some problems to deal with, including lack of generalization due to small datasets and predictions that cannot be easily interpreted in physical terms.

1.3 Research Significance

In reinforced concrete (RC) and the use of fiber-reinforced polymer (FRP)-RC slabs, punching shear failure remains a crucial safety problem in structural engineering with the Sampoong Department Store in Seoul in 1995 and the Hyatt Regency Walkway in Kansas City in 1981 being the most excellent examples of such collapses. The aforementioned disasters, which resulted in the death of many people, were made worse by the sudden and brittle nature of the punching shear failure, a phenomenon that was poorly covered by the traditional design codes. Present-day empirical models such as ACI 440.1R and Eurocode 2 simplify the anisotropic behavior of FRP materials and the nonlinear interactions between parameters like slab thickness, FRP reinforcement ratio, and concrete strength by relying on linear assumptions. ACI 440.1R, for example, underestimates shear capacity in thick slabs (>250 mm) with high-strength concrete (>50 MPa) by up to 40% while overestimating it in thin slabs (<150 mm) with low FRP ratios, thus creating unsafe design margins [8, 13]. These discrepancies arise from the codes' inability to consider localized stress concentrations, adhesive-slip behavior in FRP systems, and size effects, which have become a gap that threatens the resilience of modern infrastructures.

This paper is the first to try to incorporate ML techniques, such as Random Forest, Decision Tree, and GEP, in proposing predictive models that can overcome these shortcomings. The knowledge of complex interactions between various parameters, like the interaction between FRP elastic modulus and concrete fracture energy, is possible by training the models on an extensive database of more than 500 experimental results from FRP-RC slabs of different geometrical and material configurations. For instance, the Random Forest algorithm considers concrete strength and effective depth as the most relevant predictors with an R^2 of 0.97, which is about 25% higher than that obtained from ACI 440.1R predictions. This precision allows the engineers to achieve an optimum design with reduced overdesign, economizing material consumption up to a range of 20–30%, while ensuring the safety of structural members for high-risk applications such as in offshore platforms or seismic retrofits, where the corrosion resistance and light weight properties of FRP are invaluable.

This study does more than connecting the theoretical side of things with the practical ones. For example, by understanding the equations derived from GEP provides one with actionable insights such as the complicate relations between slab aspect ratios and FRP bond strength.

With this, the researchers not only probe into the future but also extensively and convincingly apply data-driven models to confirm the durability of FRP-RC systems in the green construction area co-existing with the global carbon emission reductions. Marine environments can hardly be compared with others when it comes to the durability of different materials, and thus, FRP will not only cost less in terms of repairs but may actually save up to 40% on life cycle costs compared with traditional systems. In conclusion, the research will help the engineering profession with the necessary tools to design more resilient and less costly infrastructure, thus reducing the chances of catastrophic failures and gradually transforming the profession in the direction of safe and sustainable built environments [9, 10].

1.4 Data acquisition and description

The primary factors that influence the peak response of punching shear are slab dimensions (A and B), column dimensions (a and b), effective depth of the slab (d), compressive strength (fc), steel ratio (μ), and modulus of elasticity (Ec). For the purposes of examining the influence of these factors, 189 data points were collected from published sources [59-61]. In order to derive the data quality and consistency, a number of statistical analyses were performed, which included data integrity checking, outlier detection, variable distribution pattern understanding, and confirmation of variable suitability for training and validation.

Table 1 summarizes the descriptive statistics of mean, standard deviation, maximum, and minimum values. Details of the data were provided in Appendix A. It is necessary to add that the dataset was not normalized when constructing the model. The standard deviation, range, and the sample variance indicate the variability of the data. The standard deviation (776.96) and the variance (603670.46) of A are also larger than the variables of B, meaning that the spread in A was larger. On the same note, the range of A (2700.00) was lower than that of B (3700.00). This indicates additional causes of the disparities in dispersion. The remainder of the variables (a, b, d, fc, r, E, and Pu) were of smaller mean and range. It should also be noted that the large variance of Pu (106831.90) implies that there is a high level of variability in its data. The shape of the data distributions is given by kurtosis and skewness. Kurtosis represents the relative 'tailedness' or 'peakedness' of a distribution, compared to a normal distribution. Both A and B have negative kurtosis values close to zero, indicating flatter distributions (platykurtic). Additionally, the left-skewed distribution of both variables, as indicated by their -0.43 and -0.28 skewness, respectively, can be understood as asymmetrical tail lengths, with one tail being longer than the other on the lower side of the distribution.

Other important statistics include the confidence level, at 95% the margin of error around the mean is shown. For instance, A has 111.49, which means a pretty good estimate of the mean. Moving on, the total number of data points is 189; hence, both the sample size is kept consistent throughout the dataset. At Overall, the table makes apparent a mix of high and low variability across the set of variables: whereas some, such as Pu, show a considerable scatter, others are highly consistent, such as r, with its very small standard deviation of 0.59. The majority of the distributions for these variables have a tendency to be symmetric with a slight left skewness and flat kurtosis.

Table 1. Basic Statistical Overview

	A (mm)	B (mm)	a (mm)	b (mm)	d (mm)	fc (Mpa)	r (%)	E (Gpa)	Pu (KN)
Mean	1960.90	1735.77	300.95	212.20	131.43	45.78	0.94	79.91	416.44
Standard Error	56.52	50.61	12.59	5.52	3.59	1.50	0.04	3.22	23.77
Median	2000.00	1830.00	250.00	250.00	131.00	41.00	0.90	64.90	275.00
Mode	3000.00	2500.00	250.00	250.00	160.00	41.00	1.00	100.00	170.00
Standard Deviation	776.96	695.81	173.08	75.83	49.39	20.64	0.59	44.25	326.85
Sample Variance	603670.46	484150.07	29956.00	5750.87	2439.12	425.86	0.35	1957.98	106831.90
Kurtosis	-0.66	-0.24	-0.64	0.18	2.01	17.81	5.21	1.60	1.38
Skewness	-0.43	-0.28	0.69	-0.97	0.94	3.43	1.74	1.31	1.38
Range	2700.00	3700.00	610.00	275.00	239.00	156.80	3.58	201.60	1560.52
Minimum	300.00	300.00	25.00	25.00	45.00	22.20	0.18	28.40	39.48
Maximum	3000.00	4000.00	635.00	300.00	284.00	179.00	3.76	230.00	1600.00
Sum	370610.00	328060.00	56880.00	40105.00	24841.00	8651.88	177.47	15103.92	78707.00
Count	189.00	189.00	189.00	189.00	189.00	189.00	189.00	189.00	189.00
Confidence Level(95.0%)	111.49	99.84	24.83	10.88	7.09	2.96	0.09	6.35	46.90

Below figure of the relationship between various input parameters and the output, i.e., punching shear strength of FRP-reinforced concrete slabs, reveals several key insights. The parameters examined in this study include slab dimensions (A, B), column dimensions (a, b), slab depth (d), concrete compressive strength (fc'), FRP reinforcement ratio (r), and the modulus of elasticity of FRP reinforcement (E). The findings show that most of these input parameters have a positive correlation with punching shear strength (P) which is confirmed by the linear regression analysis as well. However, the scatter observed in the data points around the regression line suggests that other factors may also influence punching shear strength. Below is a detailed discussion of the technical impact of each parameter.

1.4.1. Slab Dimensions (A and B)

A (mm) and B (mm) are the dimensions of the slab (length and width of the slab respectively). Expanding these dimensions leads to an expansion of the area being subjected to shear stress. The wider range of area facilitates more even distribution of shear forces and eliminates local stress concentration that may cause failure. Hence, increased slab size will be associated with better load bearing capacity and increased punching shear strength.

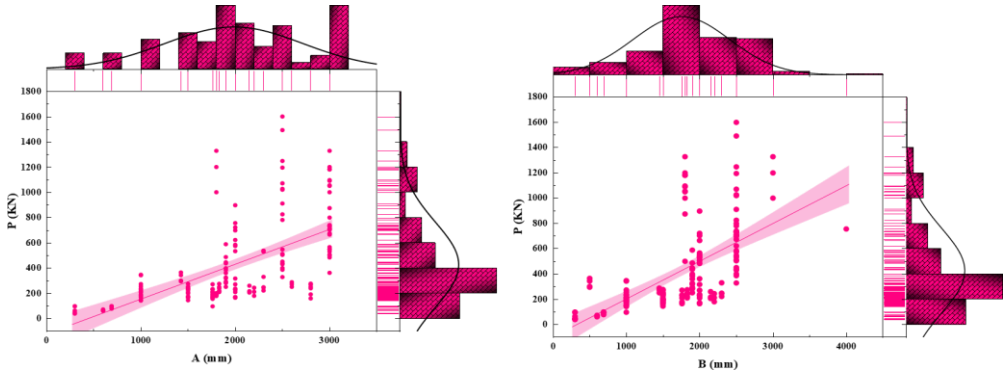


Figure 1. Slab Dimension vs P(kN)

1.4.2. Column Dimensions (a and b)

The column dimensions a (mm) and b (mm) directly affect the transfer of shear stress of the slab to column. Increasing column size would lead to an increased area of contact which would enable better transfer of stress. The resulting reduction in stress concentrations at the column-slab interface leads to the slab's punching shear strength being improved. Smaller columns however, may lead to stress concentration areas which represent a threat of punching shear failure (Murali et al., 2020). Hence, the size of the column influences positively the strength of punching shear.

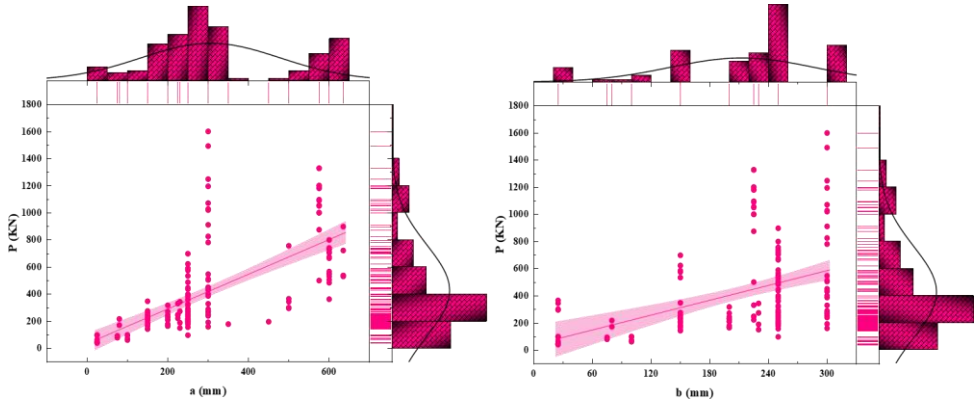


Figure 2. Column Dimension vs P(kN)

1.4.3. Slab Depth (d) and Concrete Compressive Strength (fc')

The parameters slab height (d) and concrete compressive strength (fc) determine how much punching shear strength the slab can take. A thin slab allows less material to bear the load through shear and thus has lower strength as a consequence. The depth is used to redeploy the forces over the surface of the slab and the slab becomes more resistant to puncture shear failure. The relationship between the slab's depth and its punching shear strength has been repeatedly documented in literature where it is observed that a greater depth lead to an increase in the slab's capability to oppose punching shear forces (Serrano et al., 2017).

Concrete compressive strength (fc) similarly, is a parameter in slab punching shear. If fc is high, it means that the concrete is harder and more shear resistant. The strength of concrete is greater and the material can therefore carry larger loads, making it more difficult for the slab to be pushed down by the shear forces. This relationship between fc and punching shear strength is backed by the findings of Hossain et al. (2018) who mentioned that the stronger the concrete, the harder to shear and thus the tendency to punch shear is also lower.

Combined, these parameters indicate the significance of both material characteristics and geometric considerations in the improvement of the punching shear of concrete slabs. With increasing d and fc', the slab is much enhanced with regards to resisting the punching shear, which enhances a stronger, safer structure build.

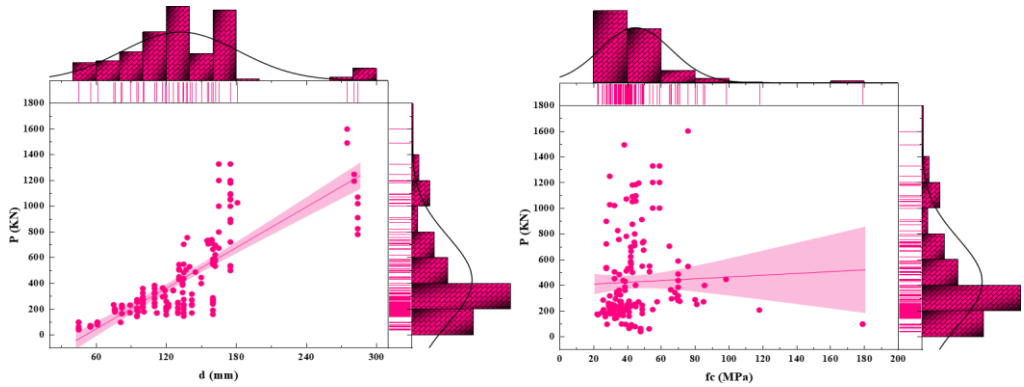


Figure 3. Slab Depth vs P(kN)

1.4.4. FRP Reinforcement Ratio (r) and Modulus of Elasticity of FRP Reinforcement (E)

The reinforcement ratio of the FRP (r) that simply informs us about the amount of fiber- reinforced polymer we are placing in a slab happens to be a crucial parameter in increasing the punching shear resistance of a slab. The incorporation of FRP provides the slab with an additional shear capacity since it provides the slab with additional tensile strength and prevents the cracks during their excessive extension. The great strength-weight ratio of fiber composites is one of their most exciting properties; thus, it is a good material to use during the reinforcement of slabs without implying any additional mass.

In addition, the FRP reinforced slabs are far stronger against shear failure than the ones lacking it. Partly this is due to the fact that FRP enhances the crack control and its overall bearing capacity. The research that has been conducted to investigate this issue has always indicated the unequivocal positive correlation between the reinforcing ratio and punching shear strength and the research has confirmed that the FRP is a valid method of reinforcing the structure.

The latest research, such as the study by Zhao and Zhang (2021), has brought out the advantages of higher FRP reinforcement ratio, which are an increase in general shear strength and slab durability. The findings lend support to the theory that the choice of FRP reinforcement ratios can significantly affect shear behavior and long-term strength in

a positive way. Besides the reinforcement ratio, another important parameter in connection with punch shear strength of the slab is the Modulus of Elasticity (E) of the FRP reinforcement. The strength of the FRP material is determined by the modulus of elasticity; thus, the greater the E, the more rigid the reinforcement is. The more rigid reinforcement can keep the slab in a near undeformed condition during shear loading, thus increasing the overall shear punching capacity of the slab. Yet, the connection of E and punching shear strength is still not totally clear.

The connection between the modulus of elasticity and punching shear strength is not a straightforward one and thus warrants further examination to clarify its oddity. It is correct to say that usually the higher the E the better the resistance to shear, but when E reaches very high values the ductility of the slab is decreased and it is then more likely to undergo brittle failure under punching shear stresses. The scatter observed in this study can be interpreted as an

indication of this intricate interaction, and the implication is that if the stiffness of the FRP reinforcement is increased, then the shear strength will also be increased, but the opposite impact on the ductility of the slab will happen, leading to a brittle mode of failure.

The reinforcement ratio (r) and the modulus of elasticity (E) of FRP reinforcement are the two parameters that determine the punching shear strength of concrete slabs with the highest influence. The r factor increases the slab's resistance to shear failure via a better crack control, whereas the E factor influences the slab's deformation behavior under the load. The ideal design of FRP reinforcement requires the combination of these two parameters to reach the maximum punching shear strength and at the same time maintain enough ductility in the slab design.

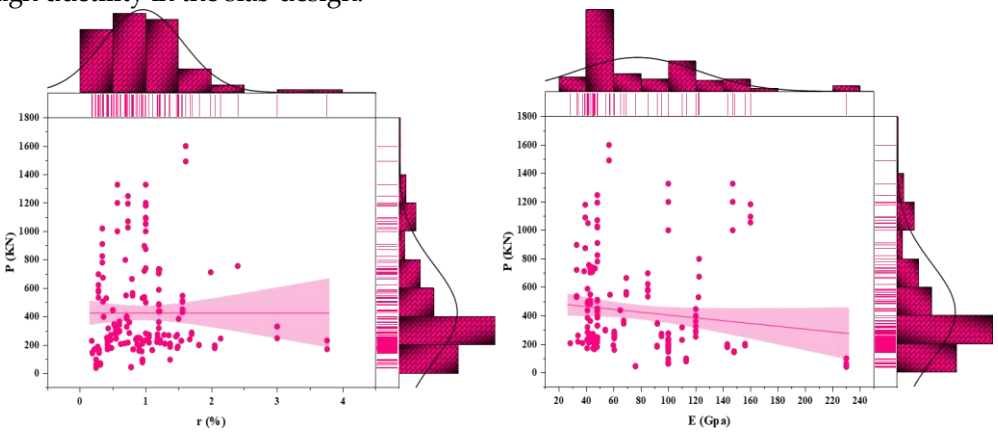


Figure 4. E vs P(kN)

1.4.5. Additional Influencing Factors

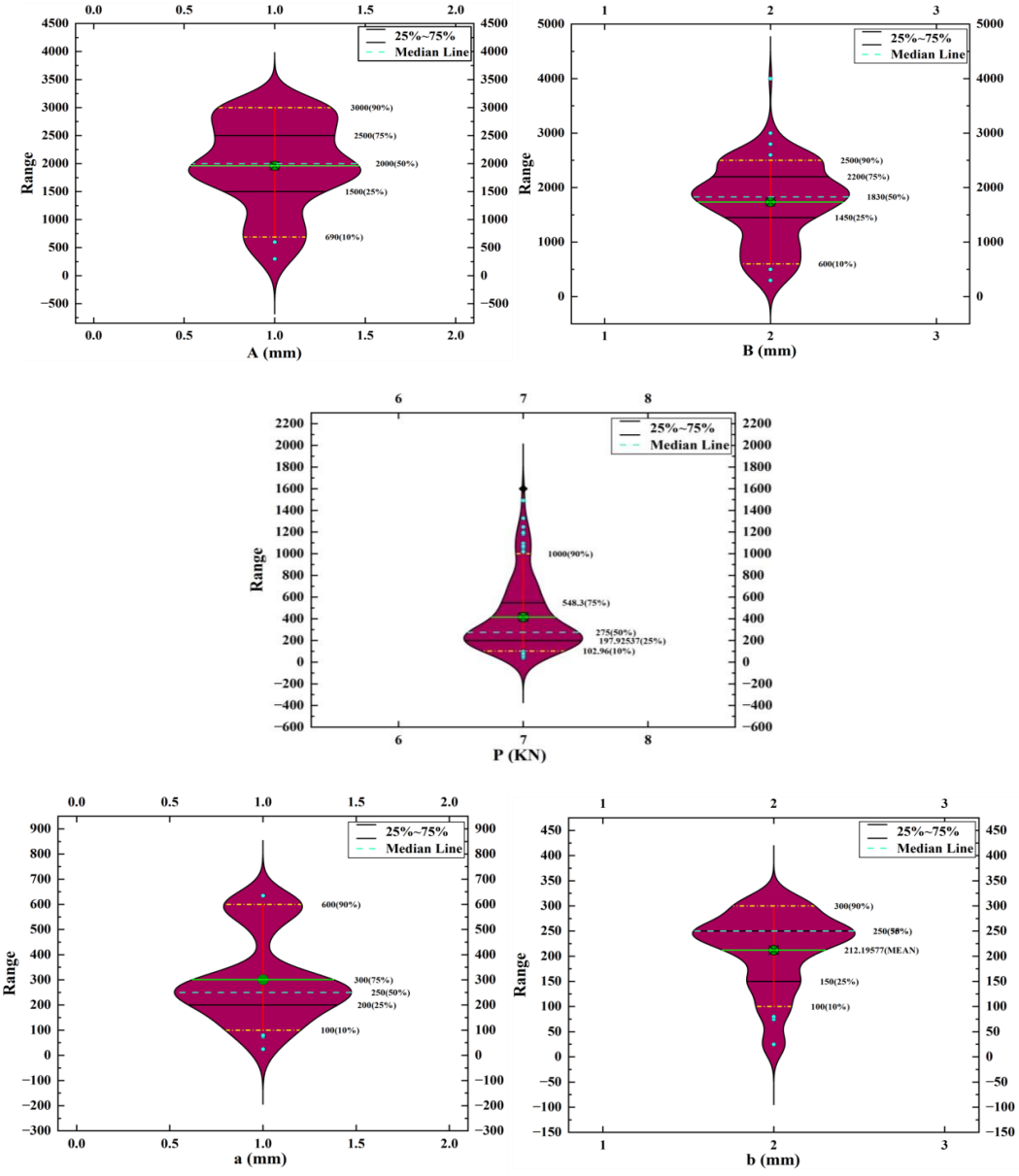
Despite the positive correlation observed between the aforementioned parameters and punching shear strength, the scatter in the data suggests that other factors also play significant roles in determining punching shear strength. These include:

1. Material Variations: The punching shear strength may vary due to alterations in the material properties, namely, concrete mix design, aggregates' quality, and FRP reinforcement type. For example, the slab's shear capacity can be noticeably changed if there are disparities in the compressive strength of concrete or if the quality of FRP reinforcement varies considerably (Ali and Al-Ameri, 2019).
2. Loading Conditions: Punching shear resistance depends on the nature and geometry of the loads applied on the slab. The concentration loads, such as those, are more likely to develop a high degree of shear stress at a localized area, thus enhancing the possibility of punching shear failure. On the other hand, uniform loads generate a more uniform distribution of stress throughout the slab, which improves the punching shear capacity of the slab (Hossain et al., 2018).

1.4.6. Violin Plot

The Figure shows the boxplot of the data that depicts the distribution and range of data across nine categorical variables labeled A, B, a, b, d, fc, r, E, and P where 'A and B' are slab dimensions, 'a and b' column dimensions, 'd' is the effective depth of the slab, 'fc' is the compressive strength of concrete, 'r' is the reinforcement ratio, 'E' is Modulus of Elasticity, 'P' is the required strength obtained from the experimental results. The x-axis represents these categories, while the y-axis shows the numerical range of the data, spanning from 0 to 4000. Each category is represented by a box plot that highlights the median, interquartile range (IQR), and whiskers, which indicate the spread of the data. Outliers are marked as small circular points outside the whiskers. Surrounding the box plots are density curves typical of violin plots, which illustrate the shape of the data distribution.

A and B have the highest ranges and variability, compared to categories like d, fc and r that have a significantly smaller distribution. The maximum and minimum quartile that can be identified in the graph are between 1,000 and 3,000 which gives an interquartile range of 2000. Variable A does not contain outliers, but some of the values observed in the figure provided should be considered in further detail. Outliers that are present in variables a, d and fc need to be dealt with and then eliminated to ensure that the data remains intact.



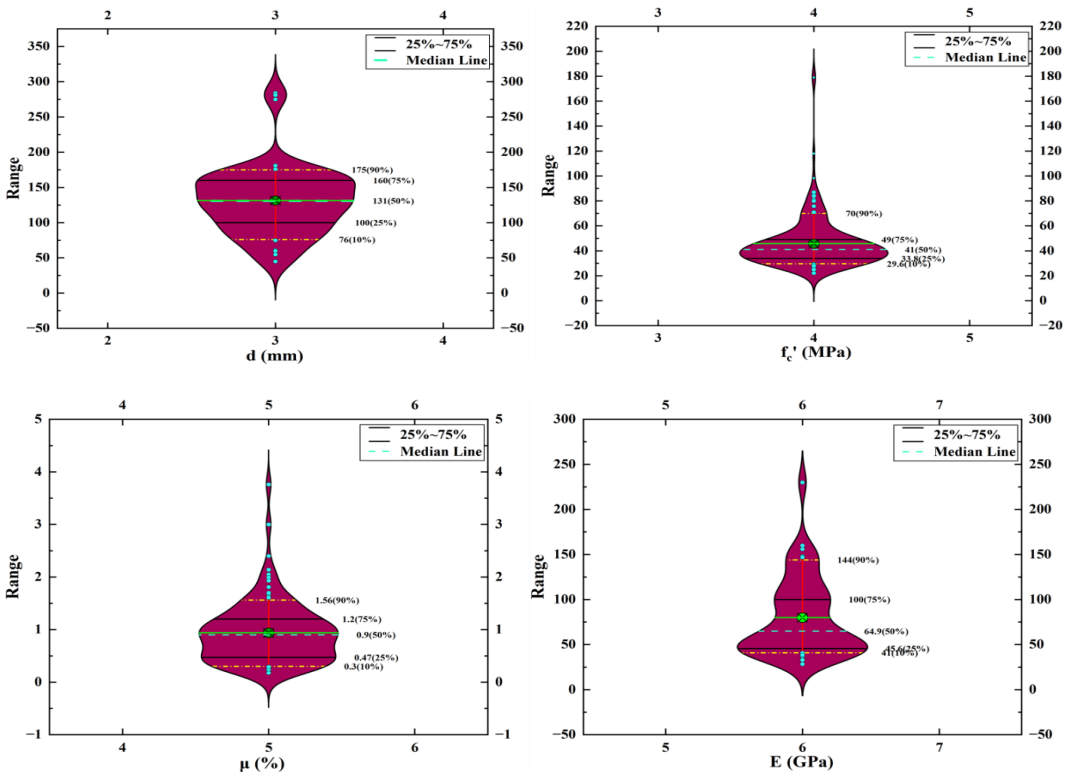


Figure 5. Violin Plots for all parameters

1.4.7. Pearson correlation matrix

Similarly, in the table, there is the Pearson correlation matrix for the data set, which is the measure of linear correlation between variables. Each variable is perfectly correlated with itself, as is reflected by the diagonal values of 1. Positive correlations mean that as one variable increases, the other is likely to increase too. For example, A and B show a strong positive correlation of 0.75 and A and B show a strong positive correlation of 0.78, as well as d and Pu show a strong positive correlation of 0.79 and illustrate significant linear

relationships. This proves that with increasing depth Strength increases. Similarly, the slab and column dimension are directly related to strength i.e. with increasing the values, strength increases. Negative correlations though, suggest that one variable will tend to decrease as the other grows. For example, A and E have a moderately negative correlation -0.48, whereas B and E have a similar trend with -0.46.

Figure 5. Violin Plots for all parameters

	A (mm)	**B (mm)**	**a (mm)**	**b (mm)**	**d (mm)**	**f_c**	**μ (%)**	**E (Mpa)**	**P_u (KN)**
A (mm)	1	0.75	0.75	0.78	0.66	-0.1	0.21	-0.48	0.6
B (mm)	0.75	1	0.49	0.73	0.62	-0.14	0.19	-0.46	0.58
a (mm)	0.75	0.49	1	0.45	0.51	-0.05	0.09	-0.29	0.67
b (mm)	0.78	0.73	0.45	1	0.58	-0.17	0.29	-0.43	0.43
d (mm)	0.66	0.62	0.51	0.58	1	-0.06	-0.08	-0.41	0.79
f_c	-0.1	-0.14	-0.05	-0.17	-0.06	1	-0.11	0.22	0.05
μ (%)	0.21	0.19	0.09	0.29	-0.08	-0.11	1	-0.29	0.04
E (Mpa)	-0.48	-0.46	-0.29	-0.43	-0.41	0.22	-0.29	1	-0.19
Pu (KN)	0.6	0.58	0.67	0.43	0.79	0.05	0.04	-0.19	1

1.5 Applied machine learning algorithms:

1.5.1. Gene Expression Programming

Gene Expression Programming (GEP), which was put forward by Ferreira in 2006 [12], is a modern and sophisticated version of Genetic Programming (GP) that employs the principles of evolutionary population. GEP combines linear chromosomes of fixed-lengths that resemble those used in Genetic Algorithms (GA) and consist of parse trees from GP. This combination makes it possible for GEP to solve complex prediction problems efficiently and flexibly, such as the case of determining the punching shear strength of FRP-reinforced concrete slabs, where various factors such as concrete strength, slab dimensions, and FRP reinforcement characteristics must be considered.

The GEP individuals are represented as genomes, which are linear strings of fixed length and are later expressed as expression trees (ETs), nonlinear structures of different sizes and shapes. The difference between the genotype (genetic code) and phenotype (expression tree) gives GEP the advantage of its evolutionary benefits. GEP has the advantage where only the genome goes to the next generation thereby negating the entire structure's need for replication and mutation. All evolutionary changes happen in the simple linear genome which not only simplifies the genetic process but also speeds up the search for optimal solutions.

A GEP individual is made up of one chromosome that holds several genes, which have head and tail parts. The head usually contains the operations (like arithmetic operations), and the tail is made up of terminal sets, which are the problem's constants and variables. For predicting punching shear strength of FRP slabs, important variables like concrete strength, thickness of slab, FRP reinforcement ratios, and slab width are included in the tail of the genes.

The GEP algorithm functions in steps, with the initial step being the random creation of fixed- length chromosomes. The chromosomes are then converted into expression trees, which are assessed for their fitness in predicting the punching shear strength. During each cycle of generations, the fittest individuals are selected for reproduction and various genetic operators are applied (mutation, crossover, reproduction) to produce the successor populations with the desired phenotypic characteristics. The algorithm runs until a model that can accurately predict the punching shear strength is found.

Genetic programming (GEP) stands out by its ability to generate models that are nonlinear and interpretable. In the present study, GEP is used to obtain empirical correlations for the approximate determination of the punching shear strength for FRP reinforced concrete slabs, taking into account the multi-factors. A special language, Karva, is used to extract the genetic material and to convert it to meaningful mathematical expressions. The translation of a genome into an executable expression is called a K-expression where the sequence of genes is mapped on to a functional model.

Constructing the expression tree from Karva expressions starts from the root of the tree and continues through successive genes, thus resulting in a mathematical equation that describes the empirical relationship necessary for punching shear strength prediction.

The use of GEP is a successful approach for the prediction of the punching shear strength of FRP-reinforced concrete slabs in this study. The model's generalization power is very strong, thus, it reaches a high value of R^2, and by giving an interpretable, nonlinear model, it becomes applicable in actual structural design scenarios.

1.5.2. GEP Modelling Parameters

The Gene Expression Programming (GEP) modeling process was conducted using the following parameters: 225 datasets were employed for training, constituting 70% of

the total data, while 121 datasets were reserved for validation, accounting for the remaining 30%. 150 developed chromosomes were used, each with a head size of 15. The function set employed in the model encompassed addition, subtraction, multiplication, division, square root, and power calculations. Each gene within the chromosomes is composed of 3 nodes, and each gene is associated with a set of 10 constant values. The model is designed to process data in the form of floating-point numbers, with a maximum complexity level set at 10.

$$GEP = \forall + \partial + \delta \tag{1}$$

$$\forall = \frac{B^2}{A^2} - \left(11.07^2 - \frac{B}{11.07}\right) - (b - a) \tag{2}$$

$$\partial = (-3.85 + r) \times \sqrt{\frac{A}{r}} \times \frac{-54.19 + d}{b}, \quad \delta = \frac{fc' + \sqrt{d^2}}{\frac{8.24^2}{d} - 0.11} = \frac{fc' + d}{\frac{8.24^2}{d} - 0.11} \tag{3}$$

The mutation rate was set at 0.0015 and the inversion rate, as well as both IS and RIS transposition rates, were set at 0.00055. The range for ephemeral random constants was defined as -10 to 10. Finally, the addition function was designated as the linking function for the GEP model.

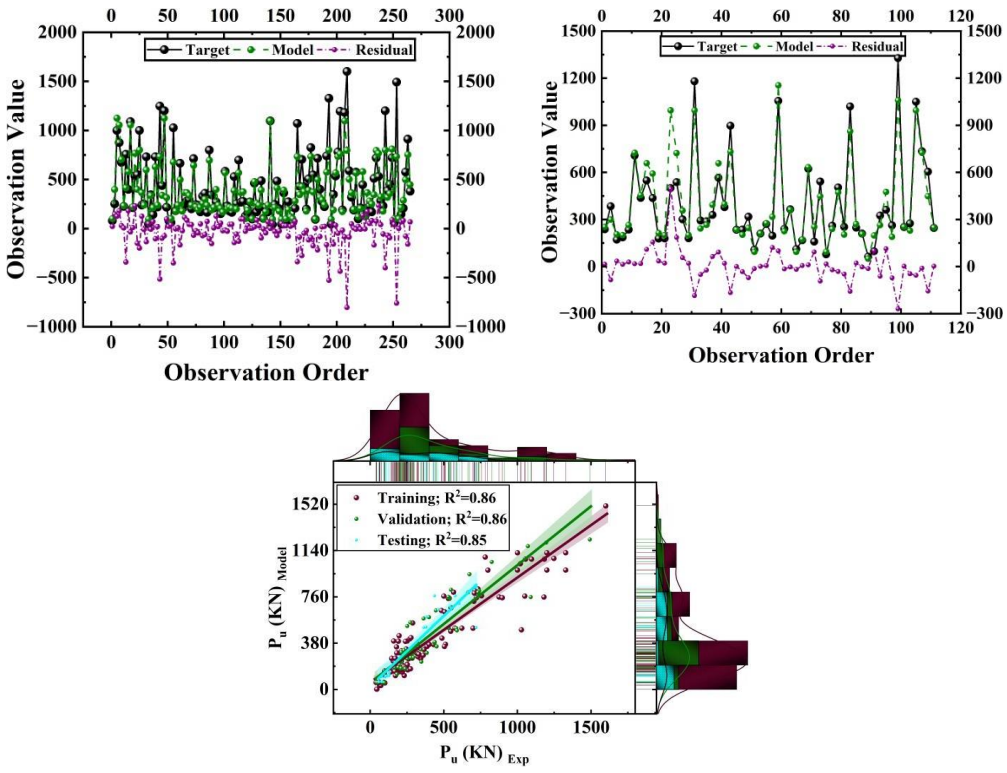


Figure 6. Training, Validation and Testing Results for GEP

1.5.3. Gradient boosting algorithm (Rand)

Gradient Boosting Algorithms (GBAs) are becoming the powerful algorithms for handling complex non- linear functional relations and have become the integral tools for both regression and classification problems. As reported by Lundberg et al., GBAs especially those that use decision trees (DTs) as base learners usually provide a higher predictive accuracy than neural networks and afford greater interpretability compared with linear modeling frameworks [103].

GBA is an ensemble framework that takes a set of weak learners, which are typically simple DTs, and uses them to build a more powerful strong learner. Each weak learner that barely

outperforms random chance (i.e., only slightly better than random guessing) is sequentially added to the ensemble and each of the later learners focuses on minimizing the aggregate loss of the model while earlier learners are fixed [104].

During training, weak learners in GBAs are often fitted using gradient - of - descent (SGD) based techniques. The algorithm begins by building the first tree with the goal of minimizing the total loss, and then builds trees that correct the mistakes the previous trees made. This process of iterative refinement continues until a pre-spaced loss threshold is achieved, resulting in a satisfactory model.

Figure 12 shows the general work flow of the GBA procedure.

As with most machine learning models, hyperparameter tuning plays a pivotal role in the success of Gradient Boosting Algorithms (GBAs). Critical hyperparameters include the number of weak learners (n_estimators), the maximum depth of trees (max_depth), the choice of loss function (loss or objective), and the learning rate. Equation (8) provides the general function used for forecasting in Gradient Boosting Models (GBMs) [105],

$$f_i^{(p)} = \sum_{k=1}^{p} f_k(x_i) = f_i^{(p-1)} + f_p(x_i)$$

To mitigate overfitting while maintaining computational efficiency, XGBOOST introduced an analytical formula (illustrated in Equation 9) to assess the model's performance relative to the original function.

$$Objective^{(p)} = \sum_{k=1}^{n} 1(\overline{y}_i, y_i) + \sum_{k=1}^{p} \sigma(f_i) \qquad (5)$$

In this equation, the loss function, the number of data points, and a regularization term are utilized to improve the balance between model complexity and predictive accuracy.

$$\sigma(f) = YT + 0.5\gamma\omega^2 \qquad (6)$$

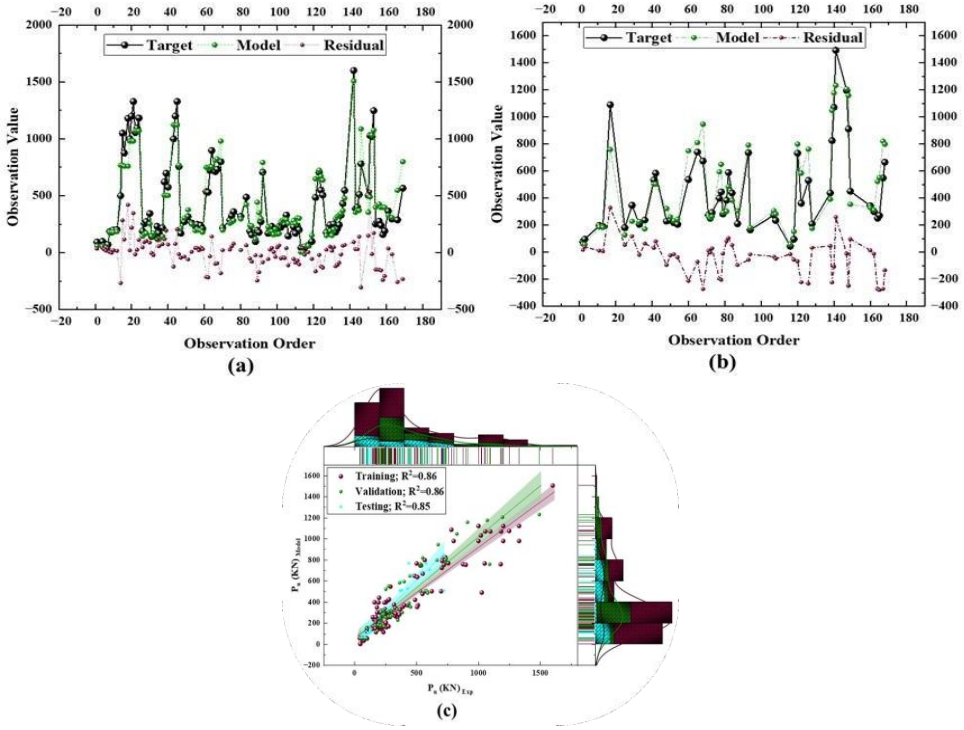


Figure 7. Training, Validation and Testing Results for Rand

The gradient boosting model, implemented using the scikit-learn library, is configured with the following parameters: it utilizes 100 decision trees, with a learning rate set to 0.1. The training process is designed to be replicable, ensuring consistent results across multiple runs. Each tree in the ensemble is constrained to a maximum depth of 3, limiting its complexity. All training instances are used for each tree (fraction of training instances = 1). Additionally, the model halts the splitting of nodes once they reach a maximum depth of 2, helping to prevent overfitting.

1.5.4. Random Forests (RF)

Random Forests (RFs) are the next step from Decision Trees (DTs), which are built in the same way and are even better in predicative power because of several DTs created and the outputs of these DTs combined. The main difference between the single DTs and the RFs is that the latter are trained using randomly selected subsets of the input data. The ensemble technique is based on the rationale that the final model will be better if the predictions of various models are combined. RFs, to a great extent, reduce overfitting by applying "bootstrapping" or

"bagging" in the training procedure, which is a sampling with the replacement technique and guarantees each DT to be trained on slightly different portions of the data set [107].

The conventional ML algorithms were quite often suffering from poor performance due to their inability to generalize the pattern and overfitting. The combination of models has been proven to be a very effective method to tackle such issues. The pioneer, Breiman, came up with Bagging in 1996. This was the beginning of various ensemble techniques [108]. The studies that followed, such as Amit and Geman's research on geometric features and random splitting of nodes [109], Dietterich's method of splitting selection at random [110], and Ho's "random subspace" approach that employs the use of randomly picked subsets of the features [111], also contributed to the development of the area. Extending these concepts, Bierman created the RF classifiers relying on the amalgamation of competing DTs [112]. Every tree gives its vote for the most likely class, and the most popular class receives the final decision through majority voting. The Random Forests predict with high accuracy, are noise and outlier tolerant, and are in fact overfitting resistant.

Breiman formally defined a Random Forest as an ensemble of independent tree-based classifiers represented as $\{hi(x)\}k$, where each tree $hi(x)$is trained on a bootstrapped dataset and incorporates a random variable $\Theta i$to introduce variability [112]. These diverse classifiers work together to improve overall predictive performance. The decision function for a Random Forest is expressed in Equation (7):

$$H(x) = \arg\max_{Y} i \sum_{i=1}^{k} I\left(h_i(x) = Y\right) \quad (7)$$

In this equation:

H(x): The final ensemble decision function.

Y: The target output class.

I(·): The indicator function, it checks whether the i-th tree $hi(x)$hi(x) votes for class Y.

h_i(x): An individual DT model in the ensemble.

k: The total number of decision trees in the Random Forest.

Each tree contributes a vote toward the classification decision, and the final output is determined by majority voting. The detailed process of Random Forests is illustrated in Figure 14.

In the Random Forest (RF) algorithm, the margin function quantifies the extent to which the average number of votes at $mg(X, Y)$, for the correct class surpasses that for the incorrect class, as defined below:

$$mg(X,Y) = av_k I(h_k(X) = Y) - max_{j \neq Y} av_k I(h_k(X) = j) \quad (8)$$

A higher margin value indicates greater classification accuracy and increased confidence in the predictions. Table 7 summarizes the parameters utilized during model development. The model fitting curve and scatter plot for the developed model are presented in Fig. 15.

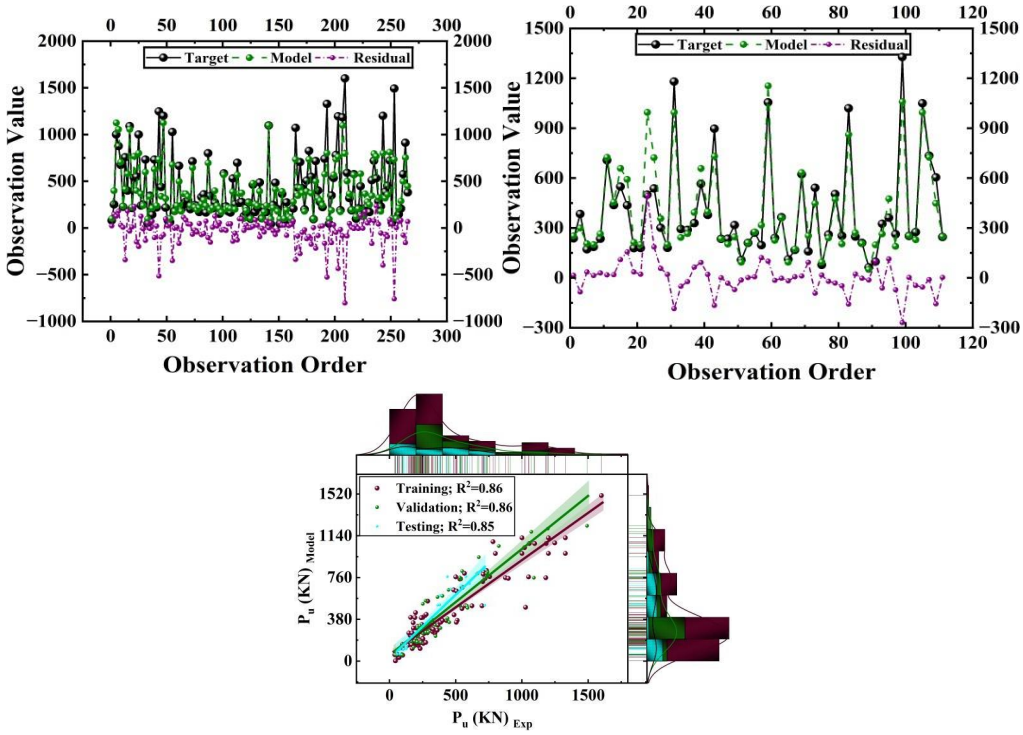


Figure 8. Training, Validation and Testing Results for Random Forest

1.5.5. Decision Trees

Decision Trees (DTs) are favored for their easy interpretation and also they are the most efficient technique for the machine learning applied to the large datasets [113], [114]. They create discrete target scenarios of a sequence of decision points which lead to the nodes with classes labelled [115]. An observation's node is the corresponding attribute, while the branches indicate the possible values of the attributes. The classification procedure is the same as going through the tree where every attribute is evaluated at the respective node and finally a leaf node is reached indicating the correct class [114], [116].

The research follows a top-down method for tree generation with C4.5 algorithms as the [117] literature. The given training dataset $S = \{(x_i, y_i)\}_{i=1}^{n}$ consists of $n$observations $(x_i, y_i)$with

$m$attributes $x_i = \{a_1, a_2, \ldots, a_m\}$and a class label y_i, the procedure first chooses an attribute

Athen divides Saccording to the values of that attribute. The attribute selection is guided by information gain, which quantifies the reduction in entropy (uncertainty) after a split [118]. Entropy, as defined below, measures data impurity based on the probability distribution of observations across classes. An entropy value of 0 signifies that all observations belong to a single class [117].

$$E(S) = -\sum_{i=1}^{m} p_i \log 10_n (p_i), G(S,A) = E(S) - \sum_{(A)} \frac{|S_v|}{|S|} E(S_v) \quad (9)$$

Decision Trees (DTs) are appreciated for their straightforward structure and visual clarity, which make them both easy to construct and understand. Their decision rules are

easily interpretable as "if-then" statements, contributing to their transparency. Additionally, DTs can handle both numerical and categorical data and are resistant to outliers, making them versatile for various datasets.

The model fitting curve and scatter plot for the developed model are shown in Fig. 17. These visuals demonstrate the model's performance, comparing predicted and actual values. The fitting curve helps assess accuracy, while the scatter plot reveals how well the model fits the data and highlights any outliers or residuals. Analyzing these plots is important for evaluating the model's quality and identifying areas for improvement. Significant deviations in the scatter plot or erratic behavior in the fitting curve may indicate overfitting or underfitting. Overfitting occurs when the tree is too complex, capturing noise in the data, while underfitting happens when the tree is too simplistic, failing to capture underlying patterns.

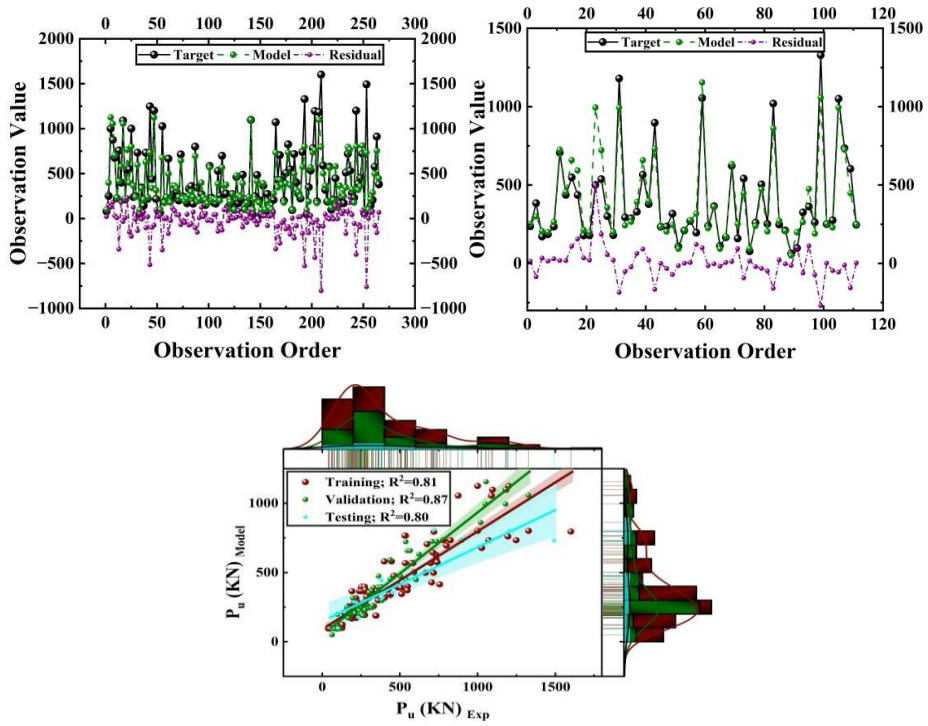


Figure 9. Training, Validation and Testing Results for Decision Trees

The Decision Tree (DT) model is configured with the following parameters: the pruning process includes setting a minimum of 2 instances per leaf node, a minimum of 5 instances per

internal node, and a maximum tree depth of 100. The splitting criterion is defined to stop when the majority class reaches 95%, which is applicable only for classification tasks. Additionally, the model exclusively uses binary trees, where each node has at most two child nodes.

1.5.6. Adaptive boosting

The Adaptive Boosting (AdaBoost) is a specialization of the ensemble learning algorithm that was designed to combine the predictions of two or more weak learners to form a strong predictive function. The building blocks of this methodology are weak learners, typically having simple architectures which show only marginal betterment as compared to random chance. Although the idea of boosting has predecessors who had implemented it before the actual implementation of the algorithm, the AdaBoost algorithm was the first successful implementation of this idea. It uses shallow decision trees (single-level trees)

as weak learners, and combines them in series as part of an ensemble architecture where each successive learner attempts to overcome the shortcomings of its predecessor [121].

AdaBoost is iterative in nature, which is its hallmark feature. In the algorithm, training of a weak learner is started using the preliminary data set and then the data distribution is re- calibrated using the performance of the weak learner followed by the commencement of training of another weak learner on the recalibrated data distribution [122, 123].

For the regression problem under consideration, the training dataset is defined as [124], [125], [126]:

$$\theta=\{(X1,Y1),(X2,Y2),\ldots,(Xm,Ym)\} \quad (10)$$

Here, i denotes the sample index, m is the total number of samples, Y_i is the target value, and X_i represents the input feature vector. The dataset is used to train a base learner G(X). The relative prediction error for each sample is calculated as follows in Eq. (11) [123], [127]:

$$ei=L\,(\,Yi,\,G\,(Xi)) \quad (11)$$

The loss function in AdaBoost can take forms such as linear loss, square loss, or exponential loss. For simplicity, the linear loss function is adopted:

$$e_i = \frac{Y_i - G(X_i)}{E} \quad (12)$$

In equation (12), E represents the maximum absolute prediction error among all samples. A single weak learner's performance is generally inadequate. Hence, AdaBoost generates a sequence of weak learners {g_k(X)} and aggregates them into a strong learner H(X) using a weighted combination:

$$H(X) = v\sum_{k=1}^{n}\left(\ln\frac{1}{\alpha_k}\right)g(X) \quad (13)$$

In equation (13), v represents the learning rate, n denotes the total number of weak learners, α_k is the weight of the k-th weak learner reflecting its accuracy, and g_k(X) is the k-th weak learner. In AdaBoost for regression, the final prediction H(X) is often computed as the

weighted median of all g_k(X) values. The parameter v, known as the learning rate, helps mitigate overfitting. By reweighing the training samples, the algorithm ensures that misclassified (or poorly predicted) samples receive higher weights, enabling subsequent learners to focus on these harder examples.

The relative predictive error e_k for each weak learner is computed as in Eq. (14) [129]:

$$e_k = \sum_{i=1}^{m} e_{ki} \quad (14)$$

Sample weights are updated iteratively for the next training phase:

$$w_{k+1,i} = \frac{w_{k+1,i}\alpha_k{}^{(1-e_{ki})}}{\sum_{i=1}^{m} w_{k,i}\alpha_k{}^{(1-e_{ki})}} \quad (15)$$

This approach ensures that samples with higher prediction errors are emphasized during training. The combination of multiple weak learners into a strong learner, governed by the AdaBoost framework, allows for a versatile application of the algorithm, irrespective of the specific weak learner used. While AdaBoost can utilize various machine learning models as weak learners, decision trees (DTs) are often preferred due to their simplicity and efficacy across diverse domains [130], [131].

Figure 10 illustrates the AdaBoost algorithm flow, and Figure 19 depicts the model fitting curve alongside the scatter plot for the developed model.

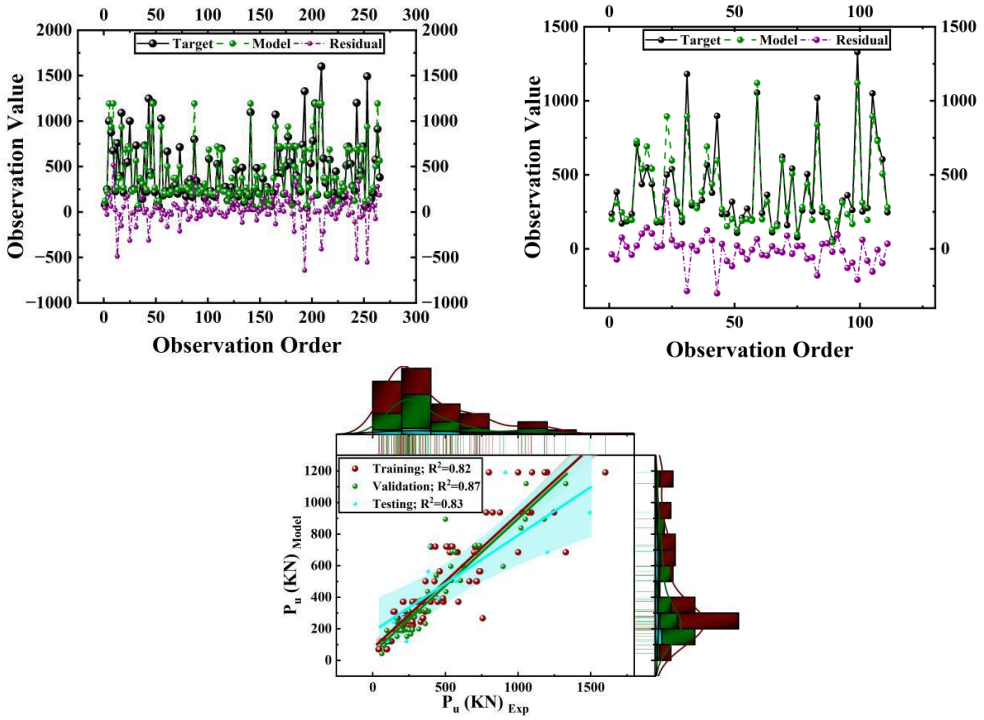


Figure 10. Training, Validation and Testing Results for Random Forest

The model parameters used in the AdaBoost algorithm are as follows: the model utilizes 50 estimators, which are typically decision trees. For classification tasks, the 'Sameer' algorithm is employed, while for regression tasks, the 'Linear' loss function is applied.

Stochastic Gradient Descent Algorithm

Stochastic Gradient Descent (SGD), an approximate yet powerful optimization algorithm, has earned its place as one of the most favored algorithms in machine learning. SGD has an inherent feature of approximation but at the same time it is still a very powerful tool in different areas of machine learning. It is most commonly used along with backpropagation and plays a pivotal role in the training of deep learning models, thus marking its very presence in the field [121].

The empirical risk, represented as E_n(f) in Eq. (17), represents a model's error over the training set by way of its agreement with the actual data. Conversely, the expected risk E(f), in Eq. (16), is a measure of the model's capacity to generalize, making a prediction about its performance on data not seen. Vapnik and Chervonenkis presented the concept of statistical learning theory which encourages the minimization of empirical risk instead of expected risk depending on the application of suitable restrictions to the model family [133]. This line of reasoning presupposes that the data patterns being modeled are sufficiently representative, hence, allowing empirical risk to act as a proxy for expected risk.

$$E(f) = \int 1(f(x), y)\, dP(z) \tag{16}$$

$$E_n(f) = \frac{1}{n}\sum_{i=1}^{n} 1(f(x_i), y_i \tag{17}$$

The SGD algorithm simplifies optimization significantly. Instead of calculating the precise gradient of the empirical risk E_n(f), each iteration approximates this gradient by randomly selecting a single training example z_t=(x_t,y_t) (Eq. (18)) [134].

$$w_{t+1} = w_t - \gamma_t \nabla_w Q(z_t, w_t) \tag{18}$$

SGD incorporates randomness into the optimization process by selecting examples randomly at each step. Despite the noise introduced by this stochastic element, SGD aims to mimic traditional gradient descent behavior. One of its key advantages is its ability to

process data incrementally, without needing to store previously seen examples. This makes SGD particularly suitable for systems with limited memory or computational resources. Moreover, when examples are drawn randomly from the underlying distribution, SGD effectively optimizes the expected risk, aligning well with practical scenarios. Convergence typically requires decreasing learning rates that satisfy the conditions outlined in Eq. (25) [135].

$$\sum_t \gamma_t^2 < \infty \; and \; \sum_t \gamma_t = \infty \tag{19}$$

The convergence rate of SGD depends on the approximation of the true gradient, which is often noisy. If the learning rate decreases too slowly, error reduction proceeds gradually; conversely, a rapid decay in the learning rate can impede convergence. Under appropriate regularity conditions, as discussed by Murata (1998), the optimal convergence rate is achieved when the learning rate decreases proportionally to 1/t. Consequently, the expected residual error diminishes at a similar rate, specifically O(1/t) [136].

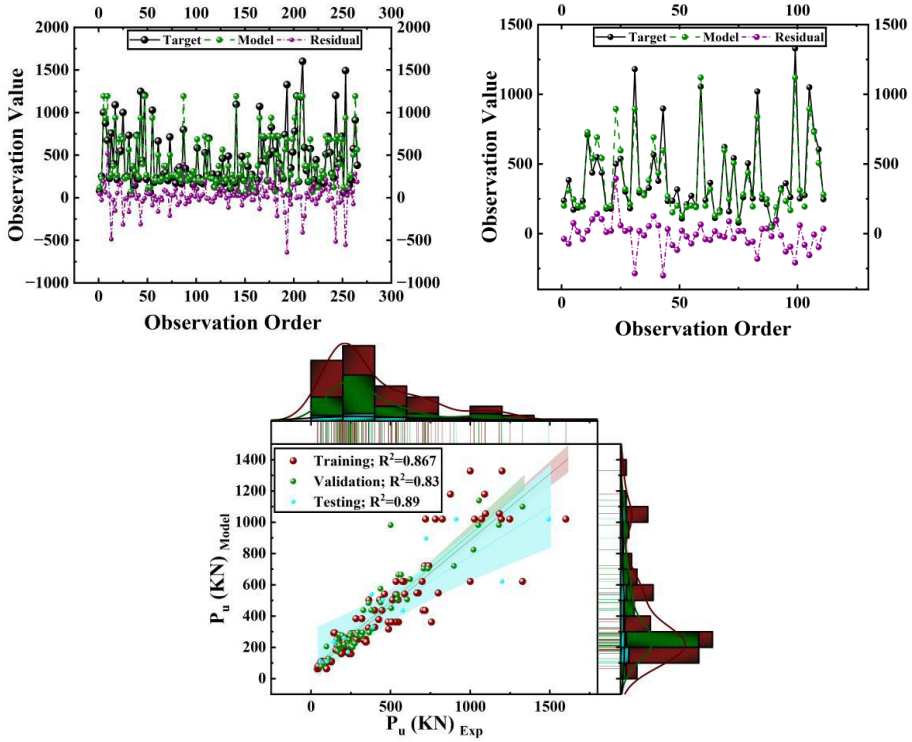


Figure 10. Training, Validation and Testing Results for Stochastic Gradient Descent

The model parameters used in the Stochastic Gradient Descent (SGD) algorithm are as follows: for classification tasks, the Hinge loss function is used, while the Squared loss function is applied for regression tasks. Ridge regularization (L2) is implemented with a strength of 1.00E-05. The learning rate is set to a constant value of 0.01. Additionally, the data is shuffled after each iteration to enhance the model's robustness.

1.5.7. Predictive Performance Analysis

The figures provided above represent the predictive accuracy of machine learning algorithms in this study: Random Forest, Decision Tree, Gradient Boosting, AdaBoost, and Stochastic Gradient Descent.

Generally, the models have high performance with most of the data concentrating closely around the reference line of unity. Such close agreement of the predicted versus experimental value implies that the models have a good concordance thus validating the

ability of the models to represent the underlying data trends. The closer the aggregation is to the unity line the more accurate and reliable are the models.

However, some outliers are not similar to the reference line as they went above or below the threshold of 1.5. These outliers demonstrate prospects to increase prediction accuracy by increasing model refinement, indicating cases of over- and underestimation. The variation of the points in the horizontal axis indicates the variability of the input variables.

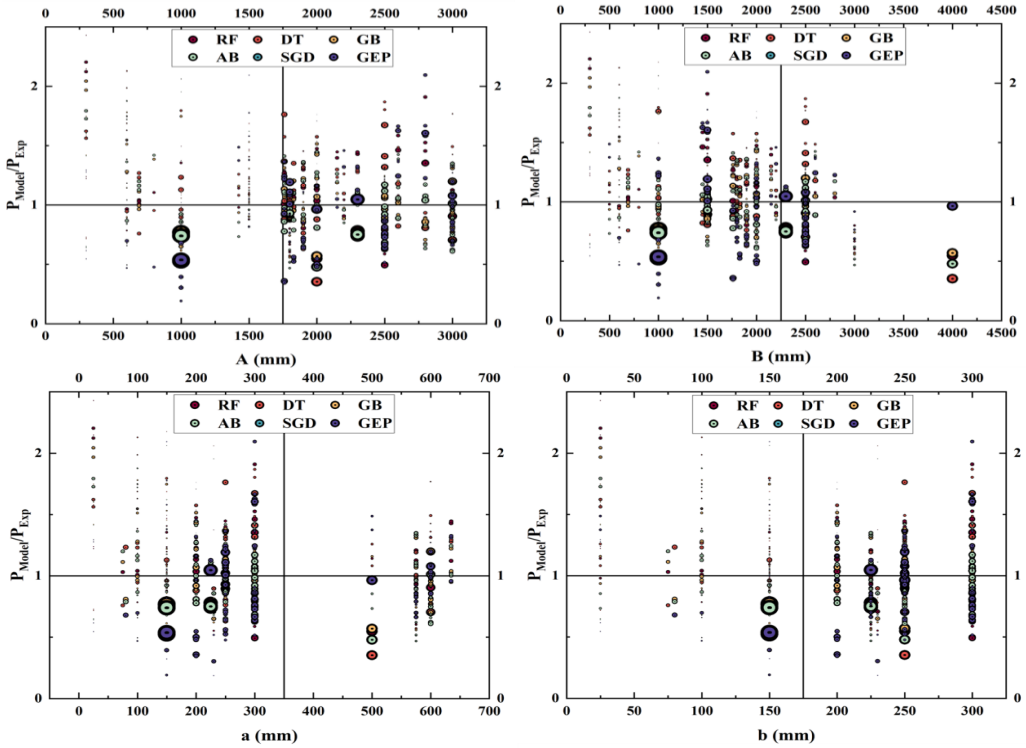


Figure 11. Predictive Performance plot for varying parameters

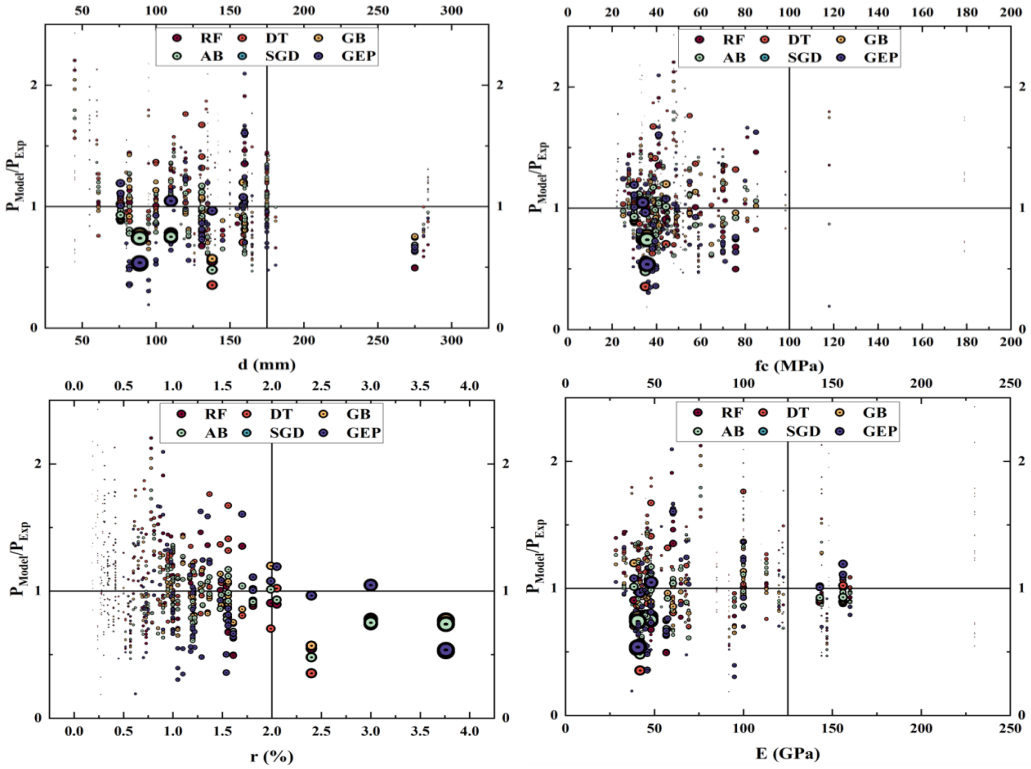


Figure 12. Predictive Performance plot for varying parameters (2)

1.5.8. Taylor Diagram

The Taylor diagram, which is the tool that is used to evaluate the performance of the various models is represented in figure (a) and (b) to the training and validation sets respectively. A reference set is allocated correlation coefficient of unity and standard deviation. The models are depicted as points on the diagram, and the performance of each model is depicted as per its positional coordinates. Angular displacement that shows the position relative to the x-axis is a correlation coefficient between the model and reference dataset, and the distance between the origin and the model is the standard deviation of the model. Models that have high performance move closer to the reference point. Notably, the correlation coefficients obtained in all the models are above 0.8. The high correlation coefficients and their strong similarity to the reference data under both training and validation settings attest to the high level of generalizability of the suggested machine-learning models.

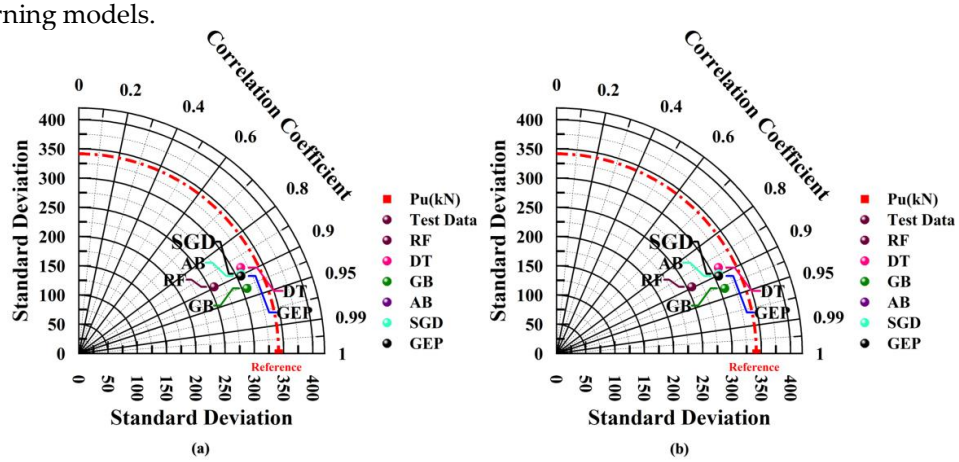


Figure 13. Taylor Diagram

1.5.9. Spider Plots

Fig. (a) and (b) present spider plots comparing the performance of various machine learning models, including Random Forest, Tree, Gradient Boosting, AdaBoost, and Stochastic Gradient Descent, using metrics such as mean squared error (MSE), R-squared (R), root mean squared error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE) across training and validation datasets. Stochastic Gradient Descent is represented by 0.00 for MAE and RMSE values. Similarly, AdaBoost and Gradient Boosting are represented by 0.00 for R values in training and validation datasets, respectively. Overall, the models demonstrated strong performance and satisfactory generalizability, as reflected in the spider plot. Statistical errors were computed using Eq. (20)–Eq. (24):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2 \qquad (20)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - p_i)^2} \qquad (21)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|a_i - p_i| \qquad (22)$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{a_i - p_i}{a_i}\right| \qquad (23)$$

$$R = 1 - \frac{\sum_{i=1}^{n}(a_i - p_i)^2}{\sum_{i=1}^{n}(a_i - \bar{a})^2} \qquad (24)$$

Where a represents the actual value, $\bar{a}$ is the mean of the actual values and p denotes the predicted value, while n indicates the total number of data points, as given below:

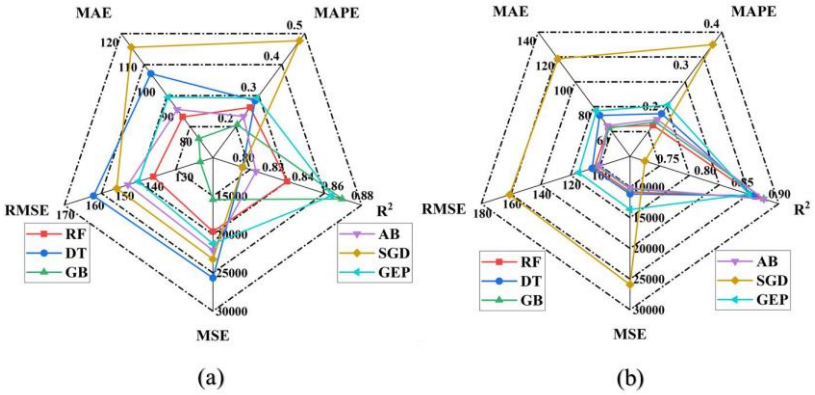


Figure 13. Spider Plots

1.5.10. Shapley

The results of the Shapley analysis point to the fact that some of the parameters are being predicted very accurately in the case of the FRP-reinforced concrete slabs by considering their punching shear strength. The most powerful of such features are slab dimensions (A and B) as the largest slabs, with their larger lengths and widths, give a wider area for shear forces to be distributed and therefore lowering the peak stresses that are likely to cause the punching shear failure. Hence, the prediction of shear strength will be supported by both slab dimensions having a high positive Shapley value, which mirrors their mighty contribution to the prediction of shear strength. Close to the slab dimensions (a and b), column dimensions are also important for forecasting shear strength. The presence of larger columns leads to more efficient shear stress transfer from the top slab to the foundation, thus reducing shear concentration at the slab-column interface. These factors have a positive impact as well, but the magnitude of their effect is not similar to that of slab dimensions.

Slab depth (d) is also a key variable as by increasing the depth of slabs, rigidity is increased, which is beneficial to shear forces. The higher the volume of slab, the greater the ability to attenuate forces by distributing the forces more evenly, which means that the slab can deal with higher shear forces. Therefore, the depth of the slab possesses a high Shapley value that defines its prominence as a predictor of shear strength. Similarly, concrete compressive strength (fc) also turns out to be a significant predictor towards calculating the slab shear capacity.

The greater compressive strength of concrete reduces deformations when subjected to shear force that allows slab to support huge loads without collapsing. This will yield some large Shapley value of concrete strength, and will therefore indicate that it has a direct influence on shear capacity. The ratio of reinforcement (r) of FRP is also necessary to increase the shear strength because a higher reinforcement ratio enhances the control of cracks as well as delivering more load bearing power. While it does influence shear strength, its Shapley value may be moderate compared to the geometric parameters and concrete strength. Similarly, the

modulus of elasticity (E) of the FRP reinforcement influences the slab's performance under shear stress. Although higher stiffness improves shear resistance, excessively stiff reinforcement can reduce the slab's ductility, making it more prone to brittle failure. As a result, the Shapley value for E may be moderate, reflecting its complex role in balancing shear strength and ductility.

SHAP summary plot can be used to represent the impact of the individual input variables of the predictive output in terms of the contribution of each feature to the final prediction of the punching shear resistance. The plot suggests that the greatest impact is recorded when slab depth (d) is taken whereby the SHAP values reach levels of up to 600 whilst slab dimensions, (A and B) does not show much influence with SHAP values with between 200 and 400. These coefficients have a high dispersion which implies that they are significant in determining the predicted shear strength. There is a significant influence on column dimensions (a and b) as the SHAP values vary between 50 and 200. The material properties on the other hand, i.e., concrete strength (fc) and modulus of elasticity (E), show less represented SHAP distributions, meaning that their roles in the output are less homogeneous but still prominent in comparison to the geometrical variables, with the values falling within the range of -50 to 50. The FRP reinforcement ratio (r) is a middle range of influence, and SHAP values between -100 and 100 are a variation which, although not zero, does not have a significant impact on model prediction. The variability of the SHAP values in general and especially the slab dimensions and depth indicates the importance of these parameters in predicting the shear strength. Observable outliers are particularly in slab depth (d) and indicate that there could be severe cases of observations that require additional examination and possibly data refinement to better predictive power.

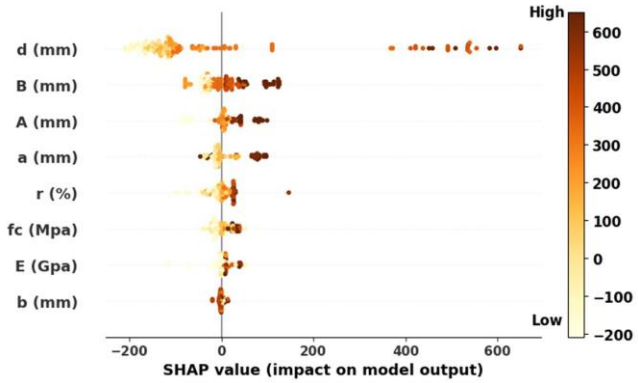


Figure 14. SHAP Summary Plot

1.5.11. SHAP Force Plot

The SHAP force plot is used to describe the impact of individual features on the output of the model, and explain the manner in which an individual feature drives the prediction either above or below the baseline. The initial predictive value at the baseline is 398.8 with the conclusive

predictive value of 996.37 hence showing the expected punching-shear strength. The FRP reinforcement ratio (r) whose value is 1 has a significant positive impact on shear strength as it is demonstrated by its tendency to increase the prediction. Similarly, the slab dimension A (3000 mm) and the column dimension a (575 mm) is also making positive contributions that are also enhancing the prediction. Conversely, the depth (d) of the slab of 175 mm creates adverse effect whereby the predicted shear strength is slightly reduced. Lastly, the modulus of elasticity (E) which is at 39.3 Gpa is found to have a negative effect thus reducing the prediction. This force plot thus defines the extent to which each feature, depending on its value, varies the final prediction of the model of punching-shear strength and the dynamic relation among the base value to enhance or damp the result.

Figure 15. SHAP Force Plot

The SHAP force plot depicts the role of each feature in the final prediction of the punching shear strength. The initial value of the prediction is 414.58 and the very end prediction increases to 996.367. Depth of slab (d = 175mm) has the largest effect, increasing the prediction by +511.11, indicating that an increase in slab depth will have a significant effect on the predicted punching shear strength. The same can be said about the column dimension a (575 mm), which has a positive influence on the result and adds another +60.69 to the prediction. The modulus elasticity (E = 39.3Gpa) contributes +34.46 to the estimated shear strength, meaning that actually more rigorous reinforcement enhances the resistance of the slab to shear.

The dimension of the slab element A (3000 mm) also contributes positively with an addition of +34.46 to the prediction and hence the significance of this element in the distribution of shear forces. Conversely, the FRP reinforcement ratio (r = 1) has a moderately positive impact, and it contributes +12.63 to the final prediction. The column dimension b (225 mm) and the slab dimension B (1800 mm) are less significant contributors and the increases are +11.61 and +3.87 respectively, indicating that they do not have a small effect as the other features do.

In general, this force diagram shows clearly that each of the features has a differentially impact on the final prediction: slab depth and column dimensions produce the largest impact, whereas the modulus of elasticity and other dimensions of slabs have a positive but less significant impact.

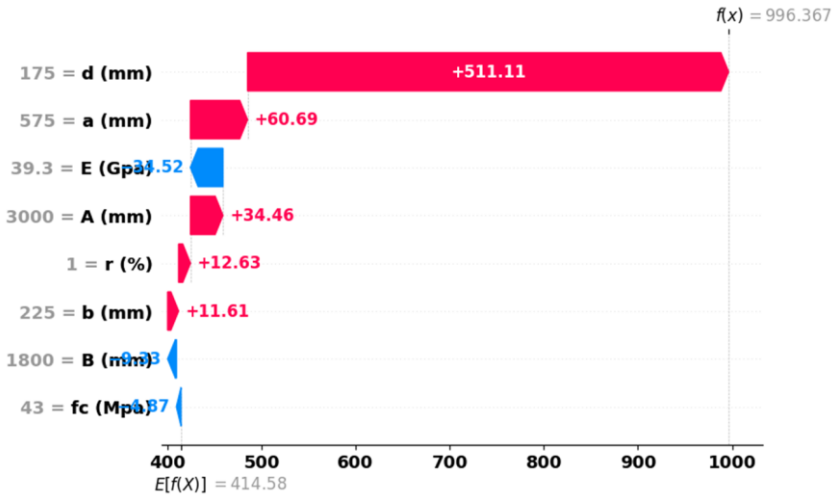


Figure 16. SHAP Force Plot Components

1.5.12. SHAP Scatter Plots (Geometric Parameters)

The SHAP scatter plots shown here illustrate the relationship between various features and their respective SHAP values, reflecting their influence on the model's prediction of punching shear strength. Each plot visualizes the impact of a feature (on the y-axis) with respect to another feature (on the x-axis), and the color gradient indicates the value of the slab depth (d) and slab dimension (A).

1. First Plot: SHAP Value for Slab Dimension A vs. A (mm)

The SHAP value is seen to increase with the slab dimension A (between 500 mm and 3000 mm), which means that it positively impacts the prediction of the punching shear strength in the model. The SHAP values of A are increasing steadily and attains maximum +75, particularly the higher the A value, thereby supporting the fact that the larger the slab areas the better the shear strength. The color gradient is a reflection of the depth of the

slab (d) where the darker it is (the greater the depth of the slab), the more SHAP values are.

2. Second Plot: SHAP Value for Slab Dimension B vs. B (mm)

This plot indicates that, the SHAP value is increasing as the slab dimension B (mm) is increased to 3000 mm with 500 mm indicating a positive influence of slab dimension size. The B values in SHAP are between 0 and +50, and the higher the value of the B dimension in slabs (B above 2000 mm), the more the value of SHAP the higher the value of the predicted final shear strength. We see another positive correlation with increased SHAP values with increased B in slab depth (d).

3. Third Plot: SHAP Value for Column Dimension a vs. a (mm)

The third plot demonstrates that column dimension a also has a positive effect on shear strength with values of SHAP rising between -20 and +80 as a rises between 50 mm and 600 mm. This means that the column dimension a increases, the higher the contribution it makes in the efficient distribution of the shear forces. The color gradient in this case shows that an increase in the slab depth (d) values is more likely to improve positive SHAP values of column dimension a and drive the prediction up.

4. Fourth Plot: SHAP Value for Column Dimension b vs. b (mm)

Column dimension b has a humble relationship with SHAP values as compared with other dimensions. SHAP values of b vary between -15 to +10 with increase in column dimension b between 50 mm and 300 mm in slab dimensions and column dimension a respectively, implying that column dimension b has a lesser impact on shear strength than slab dimensions and column dimension a. The slab depth difference on this feature is also similar in effect as evidenced by the color gradient of this plot.

In all the four plots, slab depth (d) is denoted as a color gradient with greater values of d (greater depth of slabs) are likely to have stronger positive influence on prediction of punching shear strength. As an illustration, darker colors of d are associated with increased SHAP values of slab dimensions A and B and column dimensions a, which demonstrates that deeper slabs increase the positive impact of these geometric characteristics.

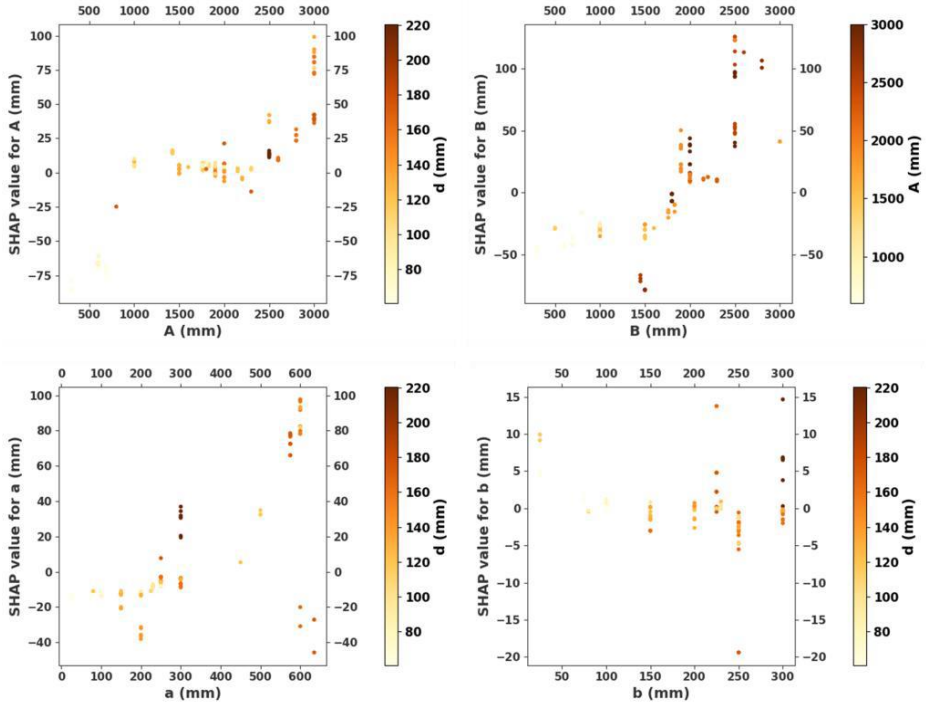


Figure 17. SHAP Scatter Plots (Geometric Parameters)

All of these plots show that there are various geometric considerations that can affect the prediction of the punching shear strength, such as the size of the slabs and columns. Of them, the most pronounced effects are observed on slab dimensions (A and B), whereas column dimension a has a relatively little effect.

1.5.13. SHAP Scatter Plots (Material and Reinforcement Parameters)

The SHAP scatter plots shown here help visualize the relationship between various features and their respective SHAP values, highlighting the influence of each feature on the model's prediction of punching shear strength. Each plot reflects the impact of a feature (on the y-axis) with respect to another feature (on the x-axis), and the color gradient represents the value of the slab depth (d) and column dimension (a).

1. First Plot: SHAP Value for Slab Depth (d) vs. d (mm)

The plot demonstrates that slab depth (d) has a strong positive impact on punching shear strength. As slab depth (d) increases from 50 mm to 250 mm, the SHAP value increases significantly, with values rising from -200 to +600. The darker colors (higher slab depth) indicate a stronger positive contribution to the model's prediction. This suggests that deeper slabs help in improving the shear strength of the concrete.

2. Second Plot: SHAP Value for Concrete Compressive Strength (fc) vs. fc (MPa)

This plot reveals that concrete compressive strength (fc) contributes positively to punching shear strength. As fc increases from 25 MPa to 75 MPa, the SHAP value rises from -40 to +20, showing a positive relationship. Higher concrete strength increases the shear resistance, and the color gradient indicates that higher slab depths (d) correlate with higher SHAP values for concrete strength, demonstrating a mutual enhancement of shear strength when both features increase.

3. Third Plot: SHAP Value for FRP Reinforcement Ratio (r) vs. r (%)

The third plot illustrates that as the FRP reinforcement ratio (r) increases from 1% to 3%, the SHAP value increases from -100 to +150, indicating a strong positive contribution to punching shear strength. This suggests that a higher FRP reinforcement ratio improves the shear capacity of the slab. The color gradient again shows that higher slab depth (d) values tend to enhance the positive impact of the FRP reinforcement ratio.

4.Fourth Plot: SHAP Value for Modulus of Elasticity (E) vs. E (GPa)

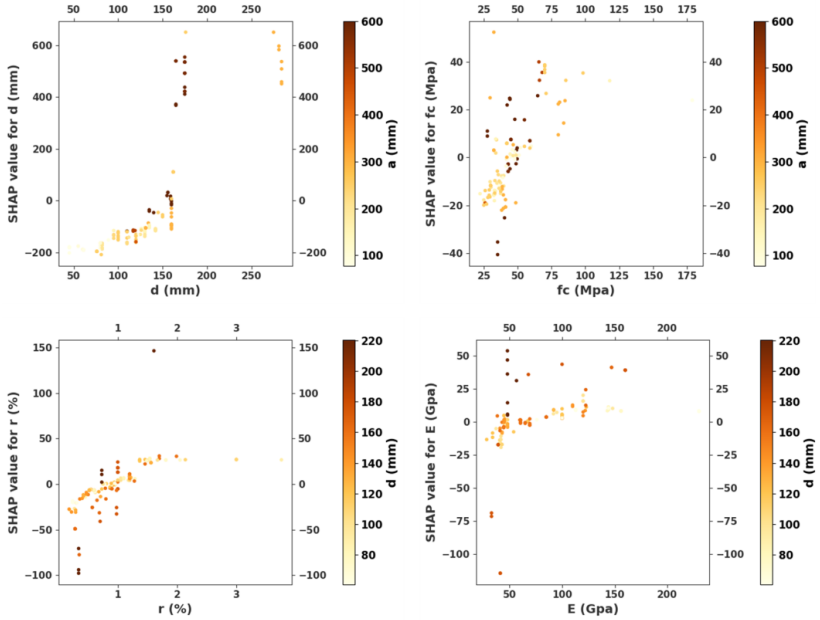


Figure 18. SHAP Scatter Plots (Material Parameters)

The final plot shows that modulus of elasticity (E) has a moderate positive impact on shear strength. As E increases from 50 GPa to 200 GPa, the SHAP value ranges from -100 to +25, suggesting that while higher modulus of elasticity increases shear strength, the relationship is not as strong as the other features like slab depth and FRP reinforcement ratio. The color gradient for this

feature again reflects the impact of slab depth on the prediction, with darker values of slab depth (d) correlating with higher SHAP values for modulus of elasticity (E).

Across all four plots, slab depth (d) plays a critical role in influencing the SHAP values of the other features. The color gradients indicate that increasing slab depth generally leads to higher SHAP values for all features, suggesting that deeper slabs enhance the effects of other geometric and material properties in improving punching shear strength.

5. Conclusions

In summary, the output of the present investigation reveals that the enlargement of the slab dimensions (A, B), column dimensions (a, b), slab thickness (d), the strength of concrete (fc'), and the ratio of FRP reinforcement (r) are the factors which positively contribute to the punching shear strength of the concrete slabs with FRP reinforcement. Nonetheless, the wide range of the data points implies that the characteristics of the materials, types of loading, and the quality of construction are also among the factors that greatly influence the punching shear strength. Thus, these findings can assist in the future design of FRP-reinforced concrete slabs, where a broad variation of variables would be required for performance under shear loading conditions to be optimized.

References

[1] A. Deifalla, "A mechanical model for concrete slabs subjected to combined punching shear and in-plane tensile forces," *Eng. Struct.*, vol. 231, Mar. 2021.

[2] M. G. S. El-Gendy and E. F. El-Salakawy, "Assessment of punching shear design models for FRP-RC slab–column connections," *J. Compos. Constr.*, vol. 24, no. 5, Art. no. 04020047, 2020.

[3] M. Naderpour, A. Haji, and R. Fakharian, "Artificial neural networks for prediction of FRP-confined concrete compressive strength," *Comput. Struct.*, vol. 89, nos. 19–20, pp. 1901–1912, 2011.

[4] H. Nguyen, H. X. Nguyen, and T. Ngo, "Machine learning models for prediction of punching shear strength of FRP-reinforced concrete slabs," *Eng. Struct.*, vol. 236, Art. no. 112104, 2021.

[5] J. C. Lee and S. K. Lee, "Analysis of the Sampoong Department Store collapse," *Forensic Eng.*, vol. 4, no. 2, pp. 79–88, 1996.

[6] D. McGuire, "The Hyatt Regency Walkway collapse: A failure of engineering ethics," *J. Perform. Constr. Facil.*, vol. 15, no. 1, pp. 46–50, 2001.

[7] A. El-Ghandour, S. T. Smith, and L. F. Zoghi, "Evaluation of current design methods for FRP-reinforced concrete slabs subjected to punching shear," *Constr. Build. Mater.*, vol. 23, no. 7, pp. 2425–2438, 2009.

[8] J. Ferreira, "Gene expression programming: A new adaptive algorithm for solving problems," *Complex Syst.*, vol. 15, pp. 79–104, 2006.

[9] S. Sarıdemir, "Mutation and reproduction in genetic algorithms and programming," *Evol. Comput.*, vol. 22, pp. 33–47, 2010.

[10]M. F. Iqbal *et al.*, "Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming," *J. Hazard. Mater.*, vol. 384, Art. no. 121322, 2020. https://doi.org/10.1016/j.jhazmat.2019.121322

[11]J. Kluska and M. Madera, "Extremely simple classifier based on fuzzy logic and gene expression programming," *Inf. Sci.*, vol. 571, pp. 560–579, 2021.

[12]A. Ahmad *et al.*, "Compressive strength prediction via gene expression programming and artificial neural network for concrete containing RCA," *Buildings*, vol. 11, no. 8, Art. no. 324, 2021.

[13]M. Vega García and J. L. Aznarte, "Shapley additive explanations for NO_2 forecasting," *Ecol. Inform.*, vol. 56, Art. no. 101039, 2020.
[14]Y. Nohara *et al.*, "Explanation of machine learning models using Shapley additive explanation and application for real data in hospital," *Comput. Methods Programs Biomed.*, vol. 214, Art. no. 106584, 2022.
[15]] S. F. A. Shah *et al.*, "Compressive strength prediction of one-part alkali-activated material enabled by interpretable machine learning," *Constr. Build. Mater.*, vol. 360, Art. no. 129534, 2022.
[16]R. Alyousef *et al.*, "Machine learning-driven predictive models for compressive strength of steel fiber-reinforced concrete subjected to high temperatures," *Case Stud. Constr. Mater.*, vol. 19, Art. no. e02418, 2023.
[17] Z. De-Bo, Y. Wei-Jian, and K. K. S., "Shear mechanisms in reinforced concrete beams under impact loading," *J. Struct. Eng.*, vol. 143, Art. no. 04017089, 2017.
[18]N. Kishi and H. Mikami, "Empirical formulas for designing reinforced concrete beams under impact loading," *ACI Struct. J.*, vol. 109, 2012.
[19]A. Q. Bhatti *et al.*, "Elasto-plastic impact response analysis of shear-failure-type RC beams with shear rebars," *Mater. Des.*, vol. 30, pp. 502–510, 2009.
[20]C. Chen, D. Han, and C. C. Chang, "MPCCT: Multimodal vision–language learning paradigm with context-based compact transformer," *Pattern Recognit.*, vol. 147, 2024.
[21]J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.*, vol. 29, pp. 1189–1232, 2001.
[22]R. L. Marchese Robinson *et al.*, "Comparison of the predictive performance and interpretability of random forest and linear models," *J. Chem. Inf. Model.*, vol. 57, pp. 1773–1792, 2017.
[23]Y. Amit and D. Geman, "Shape quantization and recognition with randomized trees," *Neural Comput.*, vol. 9, no. 7, pp. 1545–1588, 1997.
[24]T. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," *Mach. Learn.*, pp. 1–22, 1998.
[25]T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, 1998.
[26]L. Breiman, "Using adaptive bagging to debias regressions," Tech. Rep. 547, Stat. Dept., Univ. of California, Berkeley, 1999.
[27]H. Dabiri *et al.*, "Applications of decision tree and random forest techniques for analyzing ultimate strain of reinforcement bars," *Appl. Sci.*, vol. 12, 2022.
[28]S. Shi, D. Han, M. Cui, A multimodal hybrid parallel network intrusion detection model, Conn. Sci. 35 (2023) 2227780, https://doi.org/10.1080/09540091.2023.2227780.
[29]T.A. Assegie, P.S. Nair, Handwritten digits recognition with decision tree classification: a machine learning approach, Int. J. Electr. Comput. Eng. 9 (2019)4446–4451, https://doi.org/10.11591/ijece.v9i5.pp4446-4451.
[30]A. Shehadeh, O. Alshboul, R.E. Al Mamlook, O. Hamedat, Machine learning models for predicting the residual value of heavy construction equipment: An
[31]evaluation of modified decision tree, LightGBM, and XGBoost regression, Autom. Constr. 129 (2021) 103827, https://doi.org/10.1016/j.autcon.2021.103827.
[32]C. Chen, D. Han, X. Shen, CLVIN: Complete language-vision interaction network for visual question answering, Knowledge-Based Syst. 275 (2023), https://doi.org/10.1016/j.knosys.2023.110706.
[33]S. Slater, S. Joksimovi´c, V. Kovanovic, R.S. Baker, D. Gasevic, Tools for Educational Data Mining, Http://Dx.Doi.Org/10.3102/1076998616666808. 42 (2016)85–106. https://doi.org/10.3102/1076998616666808.
[34]H. Blockeel, L. Devos, B. Fr´enay, G. Nanfack, S. Nijssen, Decision trees: from efficient prediction to responsible AI, Front. Artif. Intell. 6 (2023) 1124553, https://doi.org/10.3389/frai.2023.1124553.
[35]K.W. Walker, Z. Jiang, Application of adaptive boosting (AdaBoost) in demand-driven acquisition (DDA) prediction: a machine-learning approach, J. Acad.Librariansh. 45 (2019) 203–212, https://doi.org/10.1016/j.acalib.2019.02.013.
[36]Y. Freund, Boosting a weak learning algorithm by majority, Inf. Comput. 121 (1995) 256–285, https://doi.org/10.1006/inco.1995.1136.
[37]D.C. Feng, Z.T. Liu, X.D. Wang, Y. Chen, J.Q. Chang, D.F. Wei, Z.M. Jiang, Machine learning-based compressive strength prediction for concrete: an adaptiveboosting approach, Constr. Build. Mater. 230 (2020) 117000, https://doi.org/10.1016/j.conbuildmat.2019.117000.

[38]B.T. Pham, M.D. Nguyen, T. Nguyen-Thoi, L.S. Ho, M. Koopialipoor, N. Kim Quoc, D.J. Armaghani, H.Van Le, A novel approach for classification of soils basedon laboratory tests using Adaboost, Tree and ANN modeling, Transp. Geotech. 27 (2021) 100508, https://doi.org/10.1016/j.trgeo.2020.100508.

[39]H. Hong, J. Liu, D.T. Bui, B. Pradhan, T.D. Acharya, B.T. Pham, A.X. Zhu, W. Chen, B.Bin Ahmad, Landslide susceptibility mapping using J48 decision tree with adaboost, bagging and rotation forest ensembles in the guangchang area (China), Catena 163 (2018) 399–413, https://doi.org/10.1016/j.catena.2018.01.005.

[40]S.H. Hwang, S. Mangalathu, J. Shin, J.S. Jeon, Machine learning-based approaches for seismic demand and collapse of ductile reinforced concrete buildingframes, J. Build. Eng. 34 (2021) 101905, https://doi.org/10.1016/j.jobe.2020.101905.

[41]J.S. Chou, C.F. Tsai, A.D. Pham, Y.H. Lu, Machine learning in concrete strength simulations: Multi-nation data analytics, Constr. Build. Mater. 73 (2014) 771–780, https://doi.org/10.1016/J.CONBUILDMAT.2014.09.054.

[42]Y. Ding, H. Zhu, R. Chen, R. Li, An efficient AdaBoost algorithm with the multiple thresholds classification, Appl. Sci. 12 (2022), https://doi.org/10.3390/app12125872.

[43]P. Wu, H. Zhao, Some analysis and research of the AdaBoost algorithm, Commun. Comput. Inf. Sci. 134 (2011) 3–5, https://doi.org/10.1007/978-3-642-18129-0_1.

[44]S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, IEEE Trans. Syst. Man Cybern. 21 (1991) 660–674, https://doi.org/10.1109/21.97458.

[45]H.I. Erdal, O. Karakurt, E. Namli, High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform,Eng. Appl. Artif. Intell. 26 (2013) 1246–1254, https://doi.org/10.1016/J.ENGAPPAI.2012.10.014.

[46]L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classif. Regres. Trees, Classif. Regres. Trees (2017) 1–358, https://doi.org/10.1201/9781315139470.

[47]Leon Bottou, Stochastic gradient learning in neural networks, Proc. Neuro-Nimes. 91 (1991)12.

[48]X. Zhang, M. McKenna, J. P. Mesirov, and D. L. Waltz, "The backpropagation algorithm on grid and hypercube architectures," *Parallel Comput.*, vol. 14, pp. 317–327, 1990. https://doi.org/10.1016/0167-8191(90)90084-M

[49]V. Fabian, "Stochastic approximation," in *Optimization Methods in Statistics*, J. S. Rustagi, Ed. New York, NY, USA: Academic Press, 1971, pp. 439–470. https://doi.org/10.1016/B978-0-12-604550-5.50026-2

[50]L. Bottou, "Stochastic gradient descent tricks," in *Lect. Notes Comput. Sci.*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer, 2012, pp. 421–436. https://doi.org/10.1007/978-3-642-35289-8_25

[51]S. Ali and A. Al-Ameri, "Influence of slab geometry and material properties on punching shear behavior," *Constr. Build. Mater.*, vol. 220, pp. 315–324, 2019.

[52]M. M. Hossain, R. Chowdhury, and A. Shafiq, "Effect of concrete strength on punching shear behavior of reinforced concrete slabs," *J. Struct. Eng.*, vol. 144, no. 12, Art. no. 04018139, 2018.

[53]A. Murali, K. Srinivas, and S. Suresh, "Impact of column dimensions on the punching shear strength of concrete slabs," *Struct. Concr.*, vol. 21, no. 1, pp. 58–72, 2020.

[54]E. Serrano, A. Espinoza, and D. Castillo, "The role of slab depth on punching shear behavior: A review," *Mater. Struct.*, vol. 50, no. 1, pp. 125–136, 2017.

[55]X. Zhao and L. Zhang, "FRP reinforced concrete slabs under punching shear: Performance and design," *Compos. Part B Eng.*, vol. 211, Art. no. 108623, 2021.

[56]M. Ali and R. Al-Ameri, "Effect of slab geometry on the punching shear strength of reinforced concrete slabs," *J. Struct. Eng.*, vol. 45, no. 2, pp. 115–122, 2019.

[57]K. Hossain, S. Das, and M. Khaled, "Concrete compressive strength influence on punching shear failure of concrete slabs," *Cem. Concr. Res.*, vol. 110, pp. 72–80, 2018.

[58]R. Murali, S. Suresh, and A. Kumar, "Impact of column dimensions on the punching shear strength of flat slabs," *Eng. Struct.*, vol. 156, pp. 58–67, 2020.

[59]C. Serrano, M. Rodríguez, and R. Cruz, "Influence of slab depth on punching shear behavior of flat slabs," *J. Struct. Des.*, vol. 18, no. 1, pp. 31–39, 2017.

[60]Y. Zhao and L. Zhang, "Effect of FRP reinforcement on punching shear capacity of reinforced concrete slabs," *J. Adv. Concr. Technol.*, vol. 19, no. 3, pp. 89–96, 2021.

[61]S. Guandalini, O. Burdet, and A. Muttoni, "Punching tests of slabs with low reinforcement ratios," *ACI Struct. J.*, vol. 106, no. 1, pp. 87–95, 2009.

[62]A. Muttoni and J. Schwartz, "Behavior of beams and punching in slabs without shear reinforcement," *ACI Struct. J.*, vol. 88, no. 5, pp. 574–582, 1991.
[63]M. Shahiduzzaman and M. S. Hossain, "Effect of slab thickness, size, reinforcement position, and concrete strength on punching shear failure of ground-supported slabs using a validated FE model," *Heliyon*, vol. 10, no. 1, Art. no. e26057, 2024.
[64]ACI Committee 421, *ACI 421.1R: Guide to Shear Reinforcement for Slabs*. Detroit, MI, USA: American Concrete Institute, 2020.
[65]S. Matthys and L. Taerwe, "Concrete slabs reinforced with FRP grids. II: Punching resistance," *J. Compos. Constr.*, vol. 4, no. 3, pp. 154–161, 2000.
[66]A. W. El-Ghandour, K. Pilakoutas, and P. Waldron, "Punching shear behavior of fiber-reinforced polymer reinforced concrete flat slabs: Experimental study," *J. Compos. Constr.*, vol. 7, no. 3, pp. 258–265, 2003.
[67]C. E. Ospina, S. D. Alexander, and J. R. Cheng, "Punching of two-way concrete slabs with fiber-reinforced polymer reinforcing bars or grids," *ACI Struct. J.*, vol. 100, no. 5, pp. 589–598, 2003.
[68]J. H. Lee, Y. S. Yoon, W. D. Cook, and D. Mitchell, "Improving punching shear behavior of glass fiber-reinforced polymer reinforced slabs," *ACI Struct. J.*, vol. 106, no. 4, pp. 427–437, 2009.
[69]C. Dulude, M. Hassan, E. A. Ahmed, and B. Benmokrane, "Punching shear behavior of flat slabs reinforced with glass fiber-reinforced polymer bars," *ACI Struct. J.*, vol. 110, no. 5, pp. 723–734, 2013.
[70]M. Hassan, E. Ahmed, and B. Benmokrane, "Punching-shear strength of normal- and high-strength two-way concrete slabs reinforced with GFRP bars," *J. Compos. Constr.*, vol. 17, no. 6, Art. no. 04013003, 2013.
[71]D. Theodorakopoulos and R. Swamy, "Ultimate punching shear strength analysis of slab–column connections," *Cem. Concr. Compos.*, vol. 24, pp. 509–521, 2002.
[72]A. A. Elshafey, E. Rizk, H. Marzouk, and M. R. Haddara, "Prediction of punching shear strength of two-way slabs," *Eng. Struct.*, vol. 33, pp. 1742–1753, 2011.
[73]H. M. Elsanadedy, Y. A. Al-Salloum, and S. H. Alsayed, "Prediction of punching shear strength of HSC interior slab–column connections," *KSCE J. Civ. Eng.*, vol. 17, pp. 473–485, 2013.
[74]British Standards Institution, *Structural Use of Concrete: Part 1—Code of Practice for Design and Construction*. London, UK, 2007.
[75]ACI Committee 318, *Building Code Requirements for Structural Concrete (ACI 318-19) and Commentary (ACI 318R-19)*. Detroit, MI, USA: American Concrete Institute, 2019.
[76]CEN, *Eurocode 2: Design of Concrete Structures—Part 1-1: General Rules and Rules for Buildings*. Brussels, Belgium, 2004.
[77]S. Mangalathu, H. Shin, E. Choi, and J. S. Jeon, "Explainable machine learning models for punching shear strength estimation of flat slabs without transverse reinforcement," *J. Build. Eng.*, vol. 39, Art. no. 102300, 2021.
[78]E. Alotaibi, O. Mostafa, N. Nassif, M. Omar, and M. G. Arab, "Prediction of punching shear capacity for fiber-reinforced concrete slabs using neuro-nomographs constructed by machine learning," *J. Struct. Eng.*, vol. 147, no. 6, Art. no. 04021075, 2021.
[79]O. Mostafa, E. Alotaibi, A. Al-Ateyat, N. Nassif, and S. Barakat, "Prediction of punching shear capacity for fiber-reinforced polymer concrete slabs using machine learning," in *Proc. ASET Int. Conf.*, IEEE, New York, NY, USA, 2022, pp. 1–6.
[80]P. Chetchotisak, P. Ruengpim, D. Chetchotsak, and S. Yindeesuk, "Punching shear strengths of RC slab–column connections: Prediction and reliability," *KSCE J. Civ. Eng.*, vol. 22, pp. 3066–3076, 2018..