Hybrid ANFIS and Random Forest Algorithms Optimized by Prairie Dog Optimization for Predicting Punching Shear Strength in Reinforced Concrete Slab-Column Connections

Yan Li[1, *], Dong Liu[2] and Bin Li[2]
[1]School of Transportation Engineering, Wuhan Technical College of Communications, Wuhan 430065, China;
[2]Changjiang Institute of Survey, Planning, Design and Research Co., Ltd, Wuhan, 430010, China
E-mail: 15926383656@163.com
[*]Corresponding author

Keywords: shear reinforcement, slab-column, punching shear strength, prairie dog, feature importance

Received: August 31, 2025

This study presents a data-driven framework for predicting the punching shear resistance (Vn) associated with failure modes (FMs) in reinforced concrete (RC) slab-column connections with shear reinforcement. A curated database of 327 experimental tests was compiled, incorporating nine critical input parameters related to the punching shear mechanism. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets to construct, tune, and assess the predictive models. Two machine learning approaches—random forests (RF) and adaptive neuro-fuzzy inference system (ANFIS)—were optimized using the Prairie Dog Algorithm (PDA) to enhance hyperparameter selection. Model performance was evaluated using statistical indicators including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Results demonstrate that both RF-PDA and ANFIS-PDA achieved high predictive accuracy, with ANFIS-PDA marginally outperforming RF-PDA (R^2 = 0.9893, RMSE = 0.0666 on the test set) compared to RF-PDA (R^2 = 0.9753, RMSE = 0.1057). Comparative analysis against existing baseline models further confirmed the superiority of the proposed hybrid approaches. These findings highlight the potential of metaheuristic-optimized machine learning schemes as reliable tools for evaluating the punching shear resistance of RC slab-column connections.

Povzetek: Študija pokaže, da napredni modeli strojnega učenja omogočajo zelo zanesljivo napoved prebojne strižne nosilnosti armiranobetonskih plošča–steber spojev.

1 Introduction

Smooth slabs are commonly preferred in the design and building of Reinforced Concrete (RC) constructions because of their economic viability and optimal performance [1], [2]. These slabs are positioned immediately above the pillars, obviating the necessity for beams and facilitating a more direct transfer of bars from the slab to the pillars. The absence of beams offers several advantages, including reduced construction height, seamless integration of vertical shafts, improved layout flexibility, optimized improvement placement, accelerated construction timeline, and effective form design [3], [4]. The configuration of a smooth slab is predominantly influenced by the structural response of slabs and foundations to concentrated loads, leading to shear failure near the slab and pillar's junction, resulting from elevated shear powers. The profound incursion of forces causes an entire decrease of shear strength at the intersection between the slab and pillar, resulting in an abrupt and fragile rupture of this connection [5]. This defeat subsequently initiates a redistribution of bars onto adjacent structural elements, potentially resulting in a cascading chain of failures. Assessing the failure mechanism of structural components such as slabs and foundations subjected to shear from concentrated stresses in smooth slabs is a complex endeavor, since it encompasses numerous aspects, including excessive reinforcement bars and inadequate cementitious composite strength [6], insufficient slab width [7], lack of shear enhancement [8], tiny pillar tops, and substandard construction practices. These agents must be thoroughly examined to accurately gauge the smooth slabs' performance and integrity. Several experimental techniques have been utilized to formulate technical specifications [9]. The current experimental methods [10], [11], [12] have been derived from empirical data through forecasting. The efficacy of their function depends on the particular data foundation employed to compute the structural reaction of slabs and foundations to concentrated loads, resulting in shear failure. The experimental methodologies have demonstrated inconsistencies in their results, resulting in either the underestimate or overestimation of the structural response of slabs and foundations to concentrated loads, culminating in shear failure [13]. Numerous methodologies are available for evaluating an FM in structural components, including slabs and foundations, subjected to shear from concentrated loads, as well as the

efficacy of smooth slabs; yet, these methods are constrained to certain scenarios. Machine learning (ML) methodologies can be utilized to address the issues inherent in experimental approaches [14], [15], [16], [17], [18], [19], [20], [21]. Scholars have employed artificial neural networks (ANNs) in forecasting the loading capability of enhanced cement-based composite components [22], elucidating the flexible properties of ordinary also great-resistance cement-based composite [23], examining the constructional characteristics of slabs [13], estimating the maximum power of beams [24], anticipating the grooving function of asphalt combinations with steel slag aggregates [20], and projecting the conduct of shear fasteners in cement-based composite [25]. ANN architectures have shown impressive efficacy in evaluating the operational capabilities of diverse constructional elements. ANNs have been successfully employed to estimate the loading capability of constructional members, such as the resistance estimation of enhanced cement-based composite beams [26], also pillars [27], [28], [29]. Nonetheless, ANNs face difficulties with regional enhancement, which may obstruct comprehensive data collection and analysis, resulting in erroneous estimations. Improving the performance of an artificial neural network can be accomplished by employing heuristic methods. Heuristic processes have substantial advantages over artificial neural networks, including their ability to optimize network weights through refinement and their proficiency in circumventing local minima and managing multivariate challenges [30].

The cohesion of RC beams has been forecasted using a combined technique that mixed Adaptive Neuro Fuzzy Inference System (ANFIS) with a genetic procedure and particle congestion enhancement [31]. The integrated technique shown enhanced accuracy in assessing cohesiveness relative to individual procedures. A separate research study employed a sophisticated Bat ANN to predict the structural response of slabs and foundations to concentrated loads, resulting in shear failure, and the resistance of reinforced concrete smooth slabs devoid of shear improvements [32]. The study examined 30 unique configurations of the model to identify the most precise estimation scheme with reduced faults and the greatest R^2 quantities. Concha et al. [33] utilized a mixture of neural network and particle congestion enhancement to forecast the cohesion of the iron-enhanced agent RC deep beams [34]. The blended technique precisely estimated the resistance of an iron-reinforced cement-based composite deep beam, showcasing a robust correlation coefficient of 0.997. These composite forecasting approaches are reliable for estimating structural performance and have proven valuable in civil engineering applications. Sandeep and colleagues [35] employed ML tactics to forecast the cohesion of RC beams. The scholars utilized the atom search optimizer (ASO) procedure in conjunction with a neural network to make forecasts regarding the cohesion of beams. The outcomes were subsequently compared with the forecasted outcomes of different combined and independent techniques, such as ANN, Genetic Algorithm (GA) and Support Vector Machine (SVM).

A recent study underscores the extensive utilization of machine learning in engineering, encompassing crack identification in concrete via $GCN - GLCM$ models with an accuracy of 98.99%, as well as intrusion detection and economic forecasting [36], [37]. Machine learning-based surrogate models in structural analysis attained R^2 values of 0.996 with a limited number of sensors [38], [39]. These results validate the efficacy of AI in improving accuracy and efficiency while encouraging the implementation of hybrid metaheuristic-optimized models such as PDA-based RF and $ANFIS$ in this research.

The principal objective of this project is to create and validate a hybrid, data-driven framework for precisely forecasting the punched shear strength of reinforced concrete slab-column connections with shear reinforcement. Consequently, two machine learning algorithms, RF and $ANFIS$ are combined with a metaheuristic optimizer, the PDA, to improve hyperparameter tweaking and model efficacy. The present research indicates that metaheuristic optimization (via PDA) markedly enhanced the learning accuracy and resilience of RF and $ANFIS$ models, rendering them more dependable instruments for forecasting punch shear resistance. The effective depth of the slab, together with the specifications of reinforcement and concrete strength, are variables that influence punch shear resistance as input factors.

Table 1. provides a summary of previous studies related to the topic of this article.

Table 1: A comparative summary of related works

Study	Dataset Size	Method(s) Used	Domain	Optimization	R^2	RMSE
Ahmad et al. (2018) [22]	150	ANN	RC design prediction	✗	0.89	0.21
Faridmehr et al. (2022) [32]	180	Bat-ANN	RC slabs (no reinforcement)	✓ Bat Algorithm	0.92	0.19
Concha et al. (2023) [33]	250	ANN + PSO	Steel fiber deep beams	✓ PSO	0.94	0.17
Sandeep et al. (2023) [36]	300	ANN, SVM, PSO	RC beam shear strength	✓ ASO + others	0.95	0.16

Li et al. (2023) [35]	200	ANFIS-GA–PSO	Concrete beam shear	✓ GA-PSO	0.96	0.15
Karimipour et al. (2021) [27]	120	ANN	GFRP RC columns	✗	0.91	0.22
This Study	327	RF + PDA / ANFIS + PDA	RC slabs with shear reinforcement	✓ Prairie Dog Algorithm	0.9893 (ANFIS)	0.0666

Table. 1 presents a comparative overview of the most pertinent prior research in the field of shear strength estimation. Previous models have primarily concentrated on unreinforced slabs or specialized components like beams, with only a limited number integrating metaheuristic optimization. Significantly, none have incorporated the PDA algorithm with ANFIS or RF for forecasting punching shear strength in reinforced concrete slab-column connections. This study seeks to address this gap by creating PDA-optimized hybrid models that exhibit enhanced predictive performance and generalization.

1.1 Objective of the present study

Slabs supported directly on columns without beams are known as RC slab-column connections, and they demonstrate a straightforward construction method. To date, relatively few publications have been created to describe the punching shear resistance (V_n) linked to FMs in slab-column connections with shear reinforcement utilizing machine learning techniques. The present work presents a data-driven model that forecasts the V_n linked to FMs in slab-column connections that include shear reinforcement. To determine the V_n, several machine learning methods that drew inspiration from fuzzy and tree-based methods were created throughout the investigation. The tree-based Random Forests (RF) paradigm and the ANFIS paradigms were heavily considered during this exploration. Two well-established and reliable models were used for estimation: RF and ANFIS. Employing a computational database of 327 test results, nine input variables corresponding to the punching shear mechanism are discovered. Throughout the application of RF and $ANFIS$ techniques that the application of metaheuristic methodologies has refined, the project seeks to provide structural engineers with a more dependable tool for designing reinforced concrete structures that are safer and more effective. With a greater capacity to predict occurrences, choices about the evaluation and design of buildings may be made more effectively, minimizing errors and maximizing resource use during construction.

This document is organized as follows:

- Dataset description and pre-processing are presented in Part 2.
- Part 3 presents a description of the prairie dog algorithm (PDA).
- In Part 4, the prediction models (Hybrid Process and Base) are determined.
- The study's applicable indicators are presented in Part 5.
- Outcomes and explanations are presented in Part 6.
- Remarks are discussed in Section 7.

2 Dataset description and pre-process

The provided text presents a comprehensive analysis of the literature exploring the punching resistance and FMs in slab-column connections with various shear reinforcement configurations. The study utilized a dataset of 327 experimental data points from multiple sources to develop comparative models that offer deeper insights into the factors affecting the punching resistance of reinforced concrete slabs [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50]. The research emphasizes the importance of meticulous data selection and curation to guarantee the quality and representativeness of the database. The researchers gathered independent experimental samples of two-way RC slabs with shear reinforcement, establishing a reliable database for developing predictive schemes. Incorporating high-fidelity experimental data is crucial for creating data-driven schemes that accurately depict the complex behavior of RC structures under punching shear. By integrating insights from the literature review and experimental data, the researchers successfully developed comparative schemes highlighting the critical factors influencing punching resistance, contributing to the progress of civil engineering. The database of 327 laboratory tests was meticulously trained and validated to ensure the highest data quality and reliability. The researchers implemented a rigorous data partitioning strategy, allocating 70% (231 data rows) for training, 15% (48 data rows) for validation, and 15% (48 data rows) for testing, enabling a comprehensive assessment of the schemes' performance and generalization capabilities. The evaluation metrics, including minimum, maximum, standard deviation, kurtosis, range, mean, and median, were analyzed using the non-normalized data, offering key understandings into the statistical properties of the database. In addition to the 70/15/15 partitioning, we conducted a Kolmogorov–Smirnov (KS) test to statistically confirm that the training, validation, and testing subsets were drawn from the same underlying distribution. This ensured that the models were evaluated on representative and non-biased samples, thereby reducing the risk of overfitting to specific subsets.

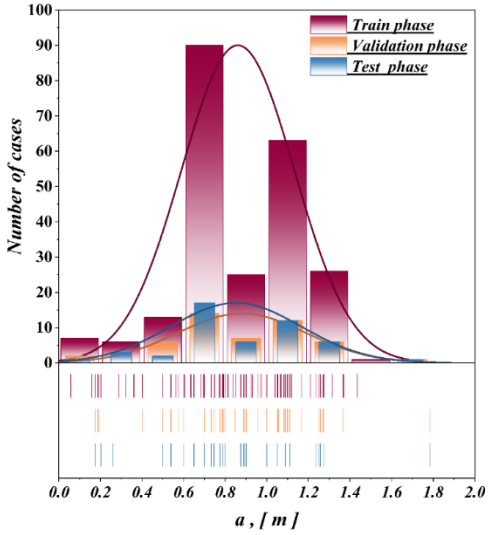
The parameters examined include the slab's adequate depth, radial distance from column face to bearing point, equivalent width of a column, ratio of flexural reinforcements, cross-sectional domain of shear reinforcement within column face, concrete compression

strength, yield strength of flexural reinforcement, yield strength of shear bar, failure mode of slabs, and punching resistance. Overall, the research presents a meticulous and comprehensive literature analysis, emphasizing the importance of high-quality experimental data and rigorous data analysis for developing accurate predictive schemes for RC structures and punching resistance. Table 2 showcases the statistical analysis of 9 inputs and one output to project punching resistance of the slab-column connections, such as min, max, range, average, skewness, kurtosis, and standard deviation. As can be seen, d, a, c, ρ_t, $A_{sw,d}$, f_{lc}, f_y, $f_{y,sw}$, FM and V_n are slab's adequate depth, radial distance from column face to the bearing point, equivalent width of a column, ratio of flexural reinforcements, cross-sectional domain of the shear reinforcement within the column face d, concrete compression strength, yield strength of the flexural reinforcement, yield strength of the shear bar, failure mode of the slabs (P_{max} = 0, P_{cs} = 1, and P_{out} = 2) and punching resistance, respectively.

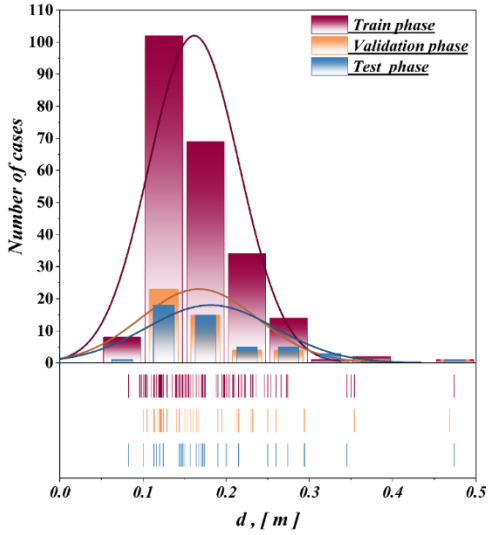
Table 2: Statistical descriptions of attributes introduced to models for estimating V_n as target

Phase	Index	Attributes									
		Input									Output
		a	d	c	ρ_t	$A_{sw,d}$	f_{lc}	f_y	$f_{y,sw}$	FM	V_n
		m	m	m	%	cm^2	MPa	MPa	MPa	$-$	MN
Train	$Min.$	0.0568	0.0825	0.0354	0.335	1	13.3	270	208	0	0.147
	$Max.$	1.435	0.474	0.52	3.02	53.4	92.4	917	1100	2	3.132
	$Range$	1.3782	0.3915	0.4846	2.685	52.4	79.1	647	892	2	2.985
	$Avg.$	0.861	0.162	0.237	1.209	12.859	34.413	521.608	481.974	1.048	0.907
	$St.D.$	0.271	0.054	0.078	0.477	9.897	13.069	108.353	145.129	0.627	0.549
	$Kurt.$	0.121	5.045	0.430	2.161	3.254	3.790	3.211	4.692	-0.451	1.108
	$Skew.$	-0.472	1.608	0.230	1.111	1.696	1.743	1.080	1.427	-0.035	1.181
Validate	$Min.$	0.175	0.1	0.106	0.335	1.56	13.9	270	280	0	0.259
	$Max.$	1.784	0.469	0.45	3.02	101.79	66.15	889	900	2	3.35
	$Range$	1.609	0.369	0.344	2.685	100.23	52.25	619	620	2	3.091
	$Avg.$	0.877	0.166	0.237	1.111	16.928	31.948	525.334	481.775	1.020	0.927
	$St.D.$	0.303	0.069	0.070	0.427	19.070	10.131	108.296	141.162	0.654	0.676
	$Kurt.$	0.835	7.154	0.459	7.133	10.759	2.215	1.572	-0.027	-0.605	3.436
	$Skew.$	0.103	2.285	0.228	1.645	2.956	1.279	0.595	0.627	-0.021	1.722
Test	$Min.$	0.175	0.0825	0.0895	0.335	1.46	14.6	278	278	0	0.144
	$Max.$	1.784	0.474	0.45	2.72	81.43	91.3	917	709	2	3.08
	$Range$	1.609	0.3915	0.3605	2.385	79.97	76.7	639	431	2	2.936
	$Avg.$	0.852	0.181	0.248	1.089	15.899	34.015	563.255	473.916	0.917	1.014
	$St.D.$	0.311	0.077	0.081	0.405	14.521	15.229	121.934	124.029	0.571	0.617
	$Kurt.$	0.956	3.580	-0.207	4.073	7.609	5.366	2.035	-1.110	0.133	1.494
	$Skew.$	0.175	0.0825	0.0895	0.335	1.46	14.6	278	278	0	0.144

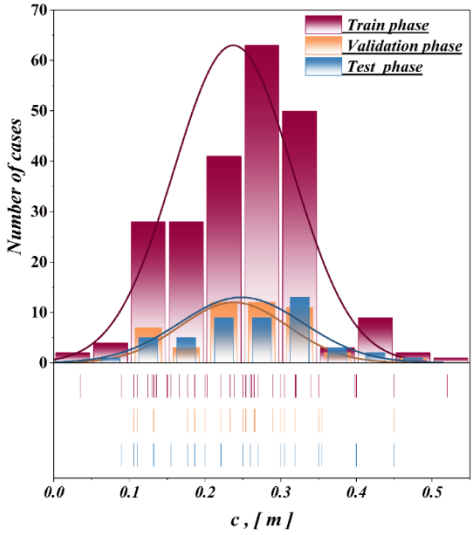
The provided images present a series of bar and normal distribution charts depicting the results of an exploration on punching shear strength of slab-column connections, as shown in Fig. 1. These visualizations appear to be generated from a database that was split into three phases: train, validation, and test. The charts display the distribution of various parameters, such as d, a, c, ρ_t, $A_{sw,d}$, f_{lc}, f_y, $f_{y,sw}$, FM, and V_n, across the different phases of the study. The color-coding and layout of the charts indicate a comprehensive database analysis to understand the relationships and patterns within the data. The detailed data visualizations presented in these charts offer valuable insights for developing estimation algorithms to project the punching shear strength of slab-column connections. The distributions of key parameters and their relationships can inform the choice of suitable simulation tactics and feature engineering strategies. The clear separation of the data into training, validation, and testing phases further suggests a structured approach to model development and evaluation, which is pivotal for ensuring the robustness and generalizability of the estimation algorithms. These visualizations can serve as a foundation for further data analysis, feature selection, and implementing advanced machine learning or statistical modeling tactics to project the punching shear strength of slab-column connections precisely.



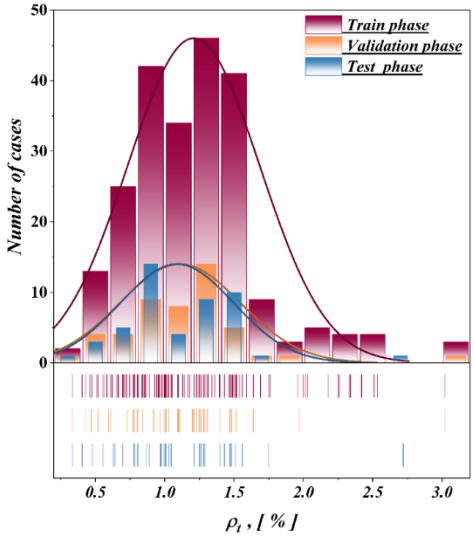
(I_1)



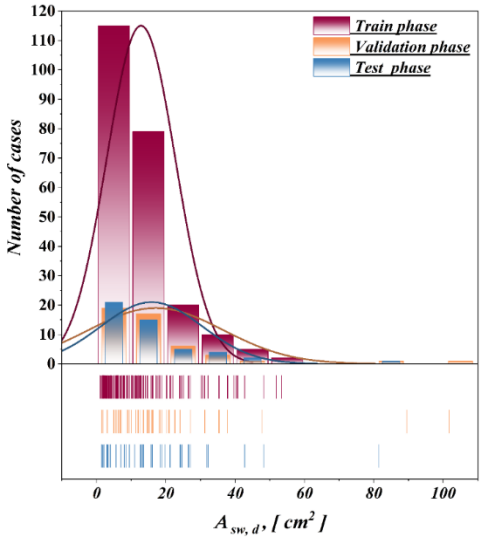
(I_2)



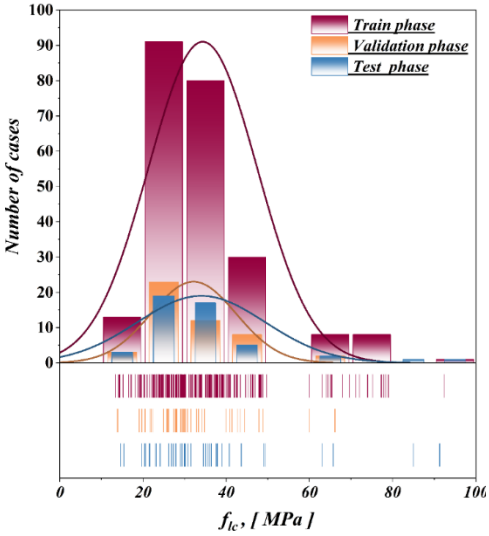
(I_3)



(I_4)



(I_5)



(I_6)

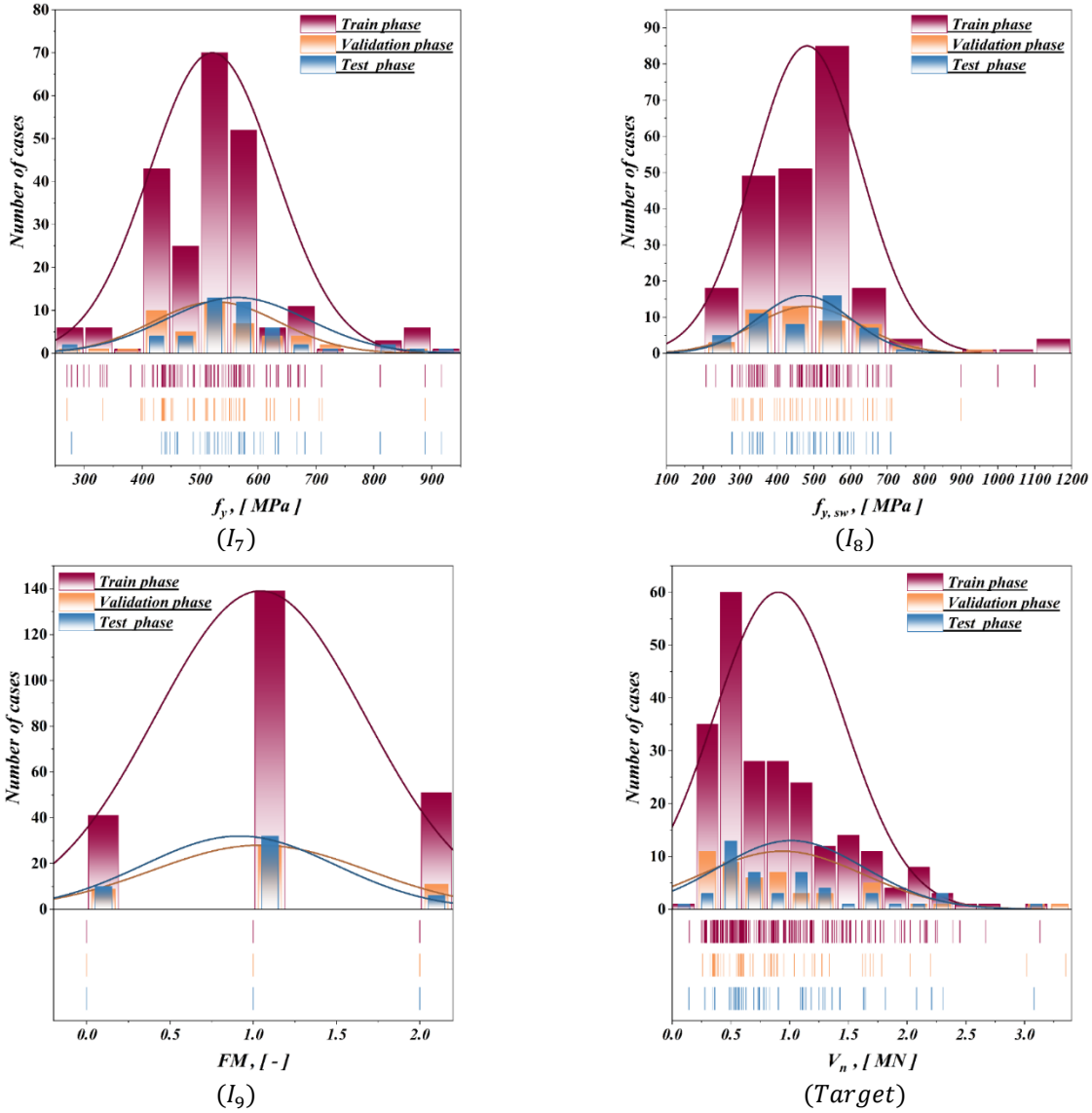


Figure 1: Normal distribution plots of attributes in three phases and V_n as target

Spearman correlation is a non-parametric measure of rank correlation that examines the monotonic link between two variables. Unlike the Pearson correlation, which gauges the linear link, Spearman correlation evaluates the strength and direction of the association between variables, drawing on their ranked positions rather than their true values. This makes Spearman correlation more robust to outliers and non-linear relationships. Regarding the provided figure, it appears to be a correlation matrix displaying the Spearman correlation coefficients between various variables. The matrix is color-coded, with the color and intensity indicating the strength and direction of the correlations. Typically, a correlation coefficient varies from -1 to 1, where -1 showcases a robust negative correlation, 0 showcases no correlation, and 1 showcases a robust positive correlation. The specific variables and meanings are unclear from the image alone, but the matrix visually showcases their links. The correlation matrix in Fig. 2 showcases low to strong correlations that could have significant implications for constructing a practical artificial intelligence (AI) model. One noteworthy finding is the strong positive correlation (0.85) between the variables V_n and d, indicating a close relationship between these parameters. This reveals that shifts in one variable tend to correspond with shifts in the other, which can be leveraged to enhance the model's predictive capabilities. Another significant correlation is the relatively low negative relationship (-0.35) between the variables d and FM, indicating an inverse relationship where a rise in one variable results in a drop in the other. Understanding this dynamic can help the AI model capture underlying patterns in the data, potentially leading to more accurate predictions and a more robust model. Additionally, the matrix reveals a moderate positive correlation (0.57) between the variables V_n and c, suggesting that the AI model can be utilized to enhance its understanding of the system and improve its overall performance. By carefully analyzing the Spearman correlation coefficient matrix, researchers can identify the most influential variables and their relationships, which can then be incorporated into the design and training of the AI model to reflect the multifaceted interactions between the metrics better and

ultimately boost the scheme's accuracy, interpretability, and overall effectiveness. The visual representation of the correlation matrix provides a clear and concise way to understand the relationships between the variables, making it a helpful tool for both researchers and *AI* model developers to build more effective schemes.

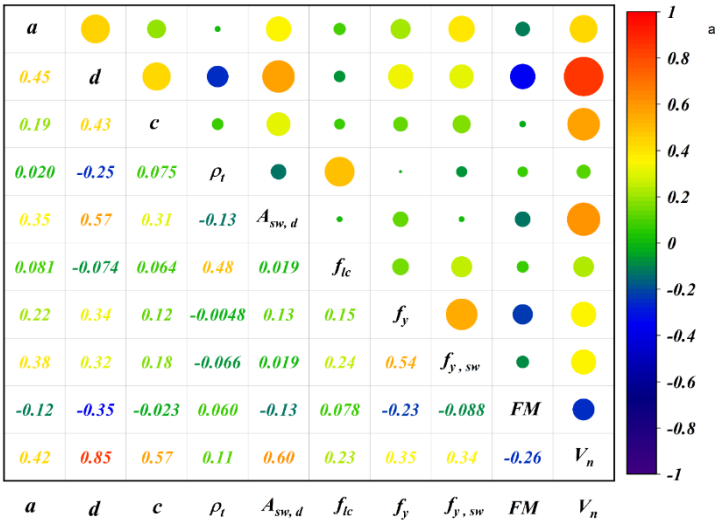


Figure 2: Correlation matrix values via Spearman analysis

3 Prairie dog algorithm (*PDA*)

Prairie Dog algorithm (PDA) is an innovative method that draws imagination from the natural conduct of grassland dogs (Fig. 3). This strategy is designed to efficiently ascertain the optimal resolution for a certain enhancement issue [51]. Prairie dogs display a behavioral pattern characterized by emerging from their burrows, traversing to different locations, and subsequently returning to their subterranean habitats. In the prairie dog's enhancement strategy, a collection of probable solutions is utilized to identify the optimal resolution for a particular issue. These prospective resolutions are perpetually refined and developed to ascertain the enhanced resolution.

The Prairie dog population comprises n individuals (PDs) grouped into m cliques. Each PD operates as a component of a collective within its designated groups. Consequently, a vector can represent the status of the ith PD within a specific clique. Formula (1) demonstrates the matrix representation of the locations of all groups (CTs) within the group.

$$CT = \begin{bmatrix} CT_{1,1} & CT_{1,2} & \dots & \dots & \dots & CT_{1,d} \\ CT_{2,1} & CT_{2,2} & \dots & \dots & \dots & CT_{2,d} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ CT_{m,1} & CT_{m,2} & \dots & \dots & \dots & CT_{m,d} \end{bmatrix} \quad (1)$$

The notation $CT\ i;\ j$ represents the jth size of the ith group inside the population. Formula (2) depicts the arrangement of all the prairie dogs inside a group.

$$PD = \begin{bmatrix} PD_{1,1} & PD_{1,2} & \dots & \dots & \dots & PD_{1,d} \\ PD_{2,1} & PD_{2,2} & \dots & \dots & \dots & PD_{2,d} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ PD_{m,1} & PD_{m,2} & \dots & \dots & \dots & PD_{m,d} \end{bmatrix} \quad (2)$$

P Di; j showcases the jth coordinate of the ith grassland dog in a group, with n also m being unique variables. The distribution of group and prairie dog situations is achieved through a uniform allocation, as illustrated in Formulas (1) & (2).

$$CT\ i;\ j = U(0,1) \times \left(UB_j - LB_j\right) + LB_j \quad (3)$$

$$P\ Di;\ j = U(0,1) \times \left(UB_j - LB_j\right) + LB_j \quad (4)$$

The notations UB_j & LB_j are used to signify the maximum & minimum limits of the jth size in an improvement dilemma. The highest limit UB_j is computed as $UB_j\ /m$, whereas the least limit LB_j is established as $LB_j\ /m$. The notation $U(0,1)$ signifies an accidental value produced with an even allocation ranging from 0 to 1.

The prairie dog improvement method modifies its approach by alternating between discovery & utilization, drawing on four specific situations. The whole value of repetitions is segmented into four divisions, with the first two focused on discovery and the final two on utilization. The investigation is additionally segmented into two approaches, ascertained by the criteria iter $<$ Maxiter/4 & Maxiter/4 $<$ iter $<$ Maxiter/2. The utilization is likewise split into two approaches, regulated by the criteria Maxiter/2 $<$ iter $<$ 3Maxiter/4 & 3Maxiter/4 $<$ iter $<$ Maxiter.

Discovery stage

Assessing the typicality of potential power resources is conducted, leading to the optimal choice for searching purposes. The typicality of the chosen power resources

influences the decision to construct new underground. Formula (5) symbolizes the adjustment of situations as part of the discovery stage in the searching method.

$$PD_i + 1, j + 1 = G\ BEST_j - eCBESTi, j \times p - CPDi, j \times Levy(n)Aiter < \frac{Maxiter}{4} \quad (5)$$

The next approach includes examining the caliber of earlier discovered power supplies and the excavation capability. Fresh tunnels are subsequently established according to this excavation ability, which declines as the count of repetitions uplifts, constraining the count of tunnels that can be established. Formula (6) illustrates the modification of places for the tunnel structure.

$$PDi + 1, j + 1 = GBestj, j \times rPD \times DS \times Levy(n)A\frac{Maxiter}{4} \leq iter \leq \frac{Maxiter}{2} \quad (6)$$

The present optimal resolution achieved on a universal scale is denoted as G Besti; j, whereas the influence of the excellent resolution at present is assessed using eC Besti; j, as illustrated in Formula (7). The power resource alarm, labeled as q also operating at a constant frequency of 0.1 kHz, serves as a crucial component in the ecosystem. The situation of a particular resolution is symbolized by rPD, while the combined impact of all prairie dogs within the community is expressed as CPDi; j, as outlined in Formula (8). The excavation capability of the group, called DS, depends on the caliber of the power supply and is stochastically established via Formula (9). The Levy allocation, denoted as Levy(n), is deployed to improve the thorough exploration of the issue domain with greater performance.

$$eCBest\ i, j, j = GBest\ i, j \times \Delta + \frac{PDi, j \times mean(PDn, m)}{GBest \times (UBj - LBj) + \Delta} \quad (7)$$

$$CPDi, j = \frac{GBest\ i, j - rPDi, j}{GBest\ i, j + \Delta} \quad (8)$$

$$DS = 1.5 \times r \times \left(1 - \frac{iter}{Maxiter}\right)^{2\frac{iter}{Maxiter}} \quad (9)$$

The factor ′r′ introduces unpredictability into the formula, thereby promoting discovery by alternating between -1 & 1 depending on the present repetition. When the repetition is odd, ′r′ takes the value of -1, whereas, for even repetitions, ′r′is equal to 1. On the other hand, ′D′ accounts for potential variations among the grassland dogs, despite the assumption in PDA's execution that all prairie dogs are identical. In this context, ′iter′ showcases the present repetition, while ′Maxiter′ signifies the highest permissible value of repetitions.

Utilization stage

The objective of the utilization strategies employed by PDO is to investigate the promising areas identified in the discovery phase extensively. Formulas (10) & (11) outline the two approaches used in this stage. As mentioned earlier, PDO switches among these two strategies depending on the situation Maxiter/2 ≤ iter < 3 Maxiter/4 & 3Maxiter/4 ≤ iter ≤ Maxiter, in the same order.

$$PDi + 1, j + 1 = GBesti, j - eCBest\ i, j \times \varepsilon - CPDi, j \times randA3\frac{Maxiter}{4} \leq iter < 3\frac{Maxiter}{4} \quad (10)$$

$$PDi + 1, j + 1 = GBesti, j - PE \times randA3\frac{Maxiter}{4} \leq iter < Maxiter \quad (11)$$

In this scenario, G Besti; j showcases the present ideal resolution discovered, also eCBesti; j implies the influence of the present best resolution obtained. According to Formula (10), ε showcases the goodness of the nourishment resource, whereas CPDi; j is the cumulative impact of all grassland dogs in the group, as explained in Formula (11). The hunter impact, explained in Formula (12), is denoted by PE, also rand is an accidentally generated value ranging from 0 to 1.

$$PE = 1.5\left(1 - \frac{iter}{Maxiter}\right)^{2\frac{iter}{Maxiter}} \quad (12)$$

The recent repetition is represented by ′iter′, while the highest value of repetitions permitted is represented as ′Maxiter′.

The pseudocode of the PDO algorithm is presented in Algorithm 1.

Algorithm 1. Pseudocode of the PDO algorithm

Setting up
Adjust the PDO variables: n,m,ρ, ε
Adjust Gbest and Cbest as φ
Setting up the applicant answers CT and PD
While $iter < Max_{iter}$ **do**
  For $(i = 1\ to\ m)$ **do**
    For $(i = 1\ to\ n)$ **do**
      Determine the fitness of PD
      Find the ideal answer so far (Cbest)
      Revise (Gbest)
      Revise DS and PE using formulas: $CPD_{i,j} = \frac{Gbest_{i,j} - rPD_{i,j}}{Gbest_{i,j} + \Delta}$ and $PD_{i+1,j+1} = Gbest_{i,j} \times PE \times rand\forall 3\frac{Max_{iter}}{4} \leq iter < Max_{iter}$
      Revise $CPD_{i,j}$ using formula: $eCbest_{i,j} = Gbest_{i,j} \times \Delta + \frac{PD_{i,j} \times mean(PD_{m,n})}{Gbest_{i,j} \times (UB_j - LB_j)} + \Delta$
      If $\left(iter < \frac{Max_{iter}}{4}\right)$ **then** (searching behaviour)

$$PD_{i+1,j+1} = Gbest_{i,j} - eCbest_{i,j} \times \rho - CPD_{i,j} \times levy(n)$$

Else if $\left(\frac{Max_{iter}}{4}\right) \leq iter < (\frac{Max_{iter}}{2})$ **then** (burrowing behaviour)

$$PD_{i+1,j+1} = Gbest_{i,j} - eCbest_{i,j} \times DS \times levy(n)$$

Else if $\left(\frac{Max_{iter}}{2} \leq iter < 3\frac{Max_{iter}}{4}\right)$ **then** (power supply alert)

$$PD_{i+1,j+1} = Gbest_{i,j} - eCbest_{i,j} \times \varepsilon - CPD_{i,j} \times rand$$

Else (antipredation alert)

$$PD_{i+1,j+1} = Gbest_{i,j} \times PE \times rand$$

End if

End for

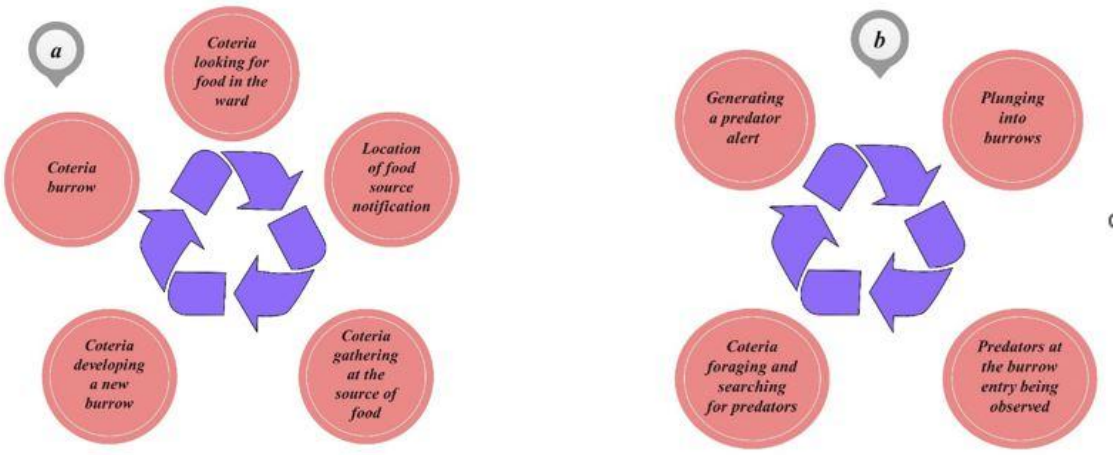
End for

Iter=iter+1

End while

Return ideal answer (Gbest)

End



(a) Exploration (b) Exploitation

Figure 3: Strategies of *PDA* during optimization

4 Prediction schemes (base and hybrid process)

4.1 *RF* analysis

RF is an ML tool that utilizes variance for quantitative learning and simulation. Breiman introduced it [52] RF is based on a flowchart-like structure used to make decisions or predictions, consisting of multiple independently constructed DTs. Consequently, the categorization outcomes of the input data are influenced by numerous random values within the forest. The description of a sub-decision tree in the RF method is outlined below:

$$f(x, \partial_k), k = 1,2, \ldots, n \quad (13)$$

x showcases the entrance vector, $f(x, \partial_k)$ showcases the fundamental categorizer of a sub-choice tree, and ∂_k comprises a collection of k uncorrelated specimens.

The RF method is comprised of four distinct stages:

A) Packing theory involves randomly selecting the main data collection to create several specimen collections. Initially, training specimens are randomly chosen using the packing method, and then the main data collection is subdivided to produce multiple sub-collections of training data. Training collections are created by randomly selecting small specimens from various classes and combining them each time. This process involves extracting data from large specimens within each class. Numerous training collections and methods can be obtained through this approach through multiple repetitions. As a result, the RF method can effectively address the issue of uneven data dispersion.

B) Every training sub-collection is utilized for categorizer training. If the specimen has M properties, M properties are chosen accidentally from all attributes while training the DT, and the tree is split based on the least Gini substandard rule. The least Gini substandard rule is defined as:

$$x_i = 1 - \sum_{i=1}^{m} p(i)^2 \quad (14)$$

The ratio of data specimens belonging to class-i in the training data collection is denoted by $p(i)$. Also, m is the number of classes in the same node.

C) The forecast outcomes have been acquired. The procedure above is repeated K times until K DTs are produced and also merged to create an RF, with a mix of techniques (like voting also weighted voting) utilized to produce categorization forecast outcomes.

D) The outcomes of the forecast are evaluated. The result of an accidental forest model is decided through either the maximum voting or averaging.

The accidental forest method offers robust accretion to tackle the issue of imbalanced specimens data.

The aim of creating an RF analysis scheme is to supply a strong and adaptable ML approach that can reliably and precisely perform tasks related to classification and regression. Improving feature important prediction precision and managing missing data, along with scalability and flexibility, are some of the key purposes of deploying RF. Either the trial-and-error or enhancement tactics may be deployed to adjust the RF framework hyperparameters, which are significantly affect their efficacy. The following steps are involved in adjusting an RF scheme's hyperparameters:

A. It was found that $n\_estimators$, max_{depth}, and $max\_features$ were the hyperparameters that could be altered. Each individual's range was set utilizing data from the RF library and research findings.
B. A 70/15/15 split separated the database into learning, validation, and assessment subgroups.
C. Subsequently, a tuning approach based on enhancement tactics was considered. The current endeavor sought to optimize hyperparameters via the combination of RF and a recently created optimization method, PDA.
D. Learning databases were then introduced to develop the first schemes. Various mixtures of hyperparameters will be used to train many schemes, and every model's efficacy will be assessed.
E. The ideal hyperparameters were identified by analyzing the search outcomes after it was finished, considering the $RMSE$ values as the objective function.
F. The final scheme's performance on an independent evaluation set was reviewed to identify its applicability to unobserved data.

4.2 ANFIS

Jang pioneered the implementation of the adjustment based on network fuzzy inference technology, representing a noteworthy computational intelligence framework that amalgamates the learning aptitudes of artificial neural networks (ANNs) with the logical reasoning abilities of fuzzy *log*ic [53]. ANFIS demonstrates superior investigation abilities and showcases a more effective method for tackling curvilinear complex issues with enhanced accuracy [54]. By employing input-result pairs and a sequence of IF − THEN fuzzy laws, ANFIS integrates the human − esque reasoning approach of systems of fuzzy inference (FIS). ANFIS is an enhanced version of FIS that addresses the limitation of adaptability to changing foreign surroundings. While FIS contains organized science with each fuzzy law defining the system's native conduct, ANFIS incorporates neural network (NN) learning tactics to enhance its compatibility. This integration of NN learning tactics allows ANFIS to dynamically adjust its conduct based on the changing foreign surroundings, making it a more robust and versatile system. Returned diffusion, which is the fundamental learning method of the network, aims to minimize the forecasting issue. ANFIS, on the other hand, mixes the learning abilities of an NN with the reasoning capabilities of fuzzy logic, as explained earlier. The ANFIS approach has shown significant effectiveness in numerous engineering fields, particularly when dealing with inconsistent or curvilinear data that traditional schemes struggle to handle due to complexity. Several modifications were implemented to enhance the efficiency of the ANFIS method and minimize fault margins.

The ANFIS architecture consists of multiple layers, as depicted in Fig. 4. This chart illustrates the ANFIS framework with two inputs and one result, comprising four membership subordinates and four laws. Based on the ANFIS framework depicted in Fig. 4, the layer arrangement of ANFIS is elaborated as follows. The initial layer is the fuzzy transformation layer, which utilizes membership subordinates to produce fuzzy sets from input quantities. The 2nd layer in the system is known as the law layer, where the discharge forces of the laws are computed based on the membership quantities obtained from the fuzzy transformation layer. Following this, the 3rd, normalizing layer determines the standardized discharge forces associated with each law. The standardized number is calculated as the ratio of the ith law discharge resistance to the total discharge resistance. The fuzzy transformation layer, the 4th layer, is where weighted quantities of laws are computed in each node. Following this, the 5th layer, the many layer, is responsible for summing the outcomes attained for each law in the fuzzy transformation layer to derive the real result of ANFIS.

Three crucial steps are involved in creating $ANFIS$ schemes: preparing the data, doing simulated learning, and adjusting the hyperparameters. Finding the hyperparameters with the greatest potential advantages may be accomplished by using the $ANFIS$ optimization approach to increase the simulation's efficacy.

The activities carried out in preparing the incoming database included handling missing values and adding details about coding classes. The data collection process produced many learning, validating, and evaluating subsets to assess the methodology's effectiveness. Once it was discovered that more modifications were needed, many hyperparameters, including the overall count of membership functions (MFs) and the count of fuzzy rules were determined to need modification. An effective objective measure for assessing productivity was the $RMSE$ metric. The shift may be upward or downward based on the hyperparameters under consideration. When setting the hyperparameters, a method called the PDA performed well. With the help of the whole learning

database and the optimal hyperparameters, we created $ANF-PDO$ model. The ultimate model's suitability for data never seen before was determined by evaluating its effectiveness on several independent assessments.

The PDO was chosen for hyperparameter tuning because to its benefits over GA and PSO. While successful in global search, GA sometimes has sluggish convergence rates and premature stalling in local optima in high-dimensional or nonlinear parameter spaces. PSO explores efficiently early on but converges and loses variety later on, making it difficult to fine-tune complicated model structures like $ANFIS$. In contrast, PDO balances exploration with exploitation like prairie dog foraging and digging. Alternating methods and stochastic variability help it avoid local minima and converge efficiently. PDO's dual capabilities allows it to develop $ANFIS$ and RF systems better than GA or PSO, narrowing error distributions and improving generalization. This research found that PDO not only expedited the search for ideal hyperparameters but also provided more stable and resilient solutions, making it a good alternative for the challenge at hand.

Table 3 presents the details of hyperparameter optimization for the developed models.

Table 3: Hyperparameter optimization details for developed models

Hybrid model	Parameters	Values
PDO	Rho	0.005
	$epsPD$	0.1
	$beta$	[1-2]
	Iterations	100
	Populations	50
$ANF-PDO$	Fuzzy rules	20
	MFs	35
	Epoch	40
	Type of MF	$trimf$
$RF-PDO$	$n\_estimators$	42
	max_depth	27
	$min\_samples\_split$	2
	$max\_features$	19

5 Indicators

To examine the efficiency of the RF, and $ANFIS$ schemes that were developed, a great number of factors were taken into consideration and computed. Every single one of the following measurements was included in these measurements: R^2, $RMSE$, RAE, Root relative square error ($RRSE$), MAE, Symmetric Mean Absolute Percentage Error ($SMAPE$), Mean Absolute Scaled Error ($MASE$), Mean Squared Logarithmic Error ($MSLE$), and Mean Relative Error (MRE).

$$R^2 = \left(\frac{\sum_{i=1}^{m}(N_i - \bar{N})(S_i - \bar{S})}{\sqrt{[\sum_{i=1}^{m}(N_i - N)^2][\sum_{i=1}^{m}(S_i - \bar{S})^2]}} \right) \quad (15)$$

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(S_i - N_i)^2} \quad (16)$$

$$RAE = \frac{\sum_{i=1}^{m}|N_i - S_i|}{\sum_{i=1}^{m}|N_i - \bar{N}|} \quad (17)$$

$$RRSE = \sqrt{\frac{\sum_{i=1}^{m}(N_i - S_i)^2}{\sum_{i=1}^{m}(N_i - \bar{N})^2}} \quad (18)$$

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|S_i - N_i| \quad (19)$$

$$SMAPE = \frac{1}{m}\sum_{i=1}^{m}\frac{|N_i - S_i|}{\frac{(|N_i| + |S_i|)}{2}} \times 100 \quad (20)$$

$$MASE = \frac{\frac{1}{m}\sum_{i=1}^{m}|N_i - S_i|}{\frac{1}{m-1}\sum_{i=2}^{m}|N_i - N_{i-1}|} \quad (21)$$

$$MSLE = \frac{1}{m}\sum_{i=1}^{m}\big(In(N_i + 1)\big) - In(S_i + 1)^2 \quad (22)$$

$$MRE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{N_i - S_i}{N_i}\right| \quad (23)$$

The introduced values are the overall count of observations (m), the value that was anticipated (S_i), and the average of the amount that was anticipated ($\bar{S}$). The average of the real V_n is displayed by $\bar{N}$, while the real V_n is displayed by N_i.

6 Results and justifications

It was possible to gauge the punching V_n by integrating the RF and $ANFIS$ methods with the PDA approach. These methods are also referred to as $RF-PDO$ and $ANF-PDO$, respectively. For the $RF-PDO$ and $ANF-PDO$ approaches, the observed and predicted V_n values are displayed in Fig. 4. These values were obtained throughout the experiment's validation, learning, and assessment phases. As an additional feature, it showcases the projected measured V_n ratio for each participant throughout the length of the inquiry. Table 4 showcases the findings from evaluating the constructions attained via training, validating, and evaluating the product development approach. Furthermore, the present research provided the variance percentage for every single scheme at every phase to boost the precision of the integrated schemes. XGB and $WOA-XGB$ [55] were considered to appraise the schemes' reliability and strength. The outcomes of the present inquiry on the produced schemes were also compared with the research results that were previously available to appraise the schemes' reliability and strength.

Based on the data, it is very likely both of which $RF-PDO$ and $ANF-PDO$ can accurately estimate the V_n. $ANF-PDO$ is a distinct technique which is thus more precise and trustworthy than $RF-PDO$, as evidenced by computed measurements for performance analysis. $RF-PDO$ can achieve lesser values (0.9731, 0.9755, and 0.9753) for the learn, validate, and assess phases, yet

$ANF-PDO$ can attain the greatest R^2 values (0.9869, 0.9938, and 0.9893).

A comparison of the $ANF-PDO$ model with another model led to the discovery of this outcome. The $ANF-PDO$ produced the lowest $RRSE$ index values for learning, validating, and evaluation; the corresponding values were 0.117, 0.0839, and 0.1079. Compared to the numbers 0.1769, 0.1806, and 0.1713 that $RF-PDO$ obtained in learning, validating, and evaluation, respectively, these figures displayed lower precision. The differences in percentages of the two schemes, which were created for these measures, are at least 31%; in some instances, the disparity is dropped by 68%, showcasing the potential of the $ANF-PDO$ to project competence and dependability. Both schemes are accurate and reliable, with $ANF-PDO$ being slightly better, as displayed by reasoning and evaluation metrics.

A full comparison with the current body study that considers the schemes XGB and $WOA-XGB$ is performed [55] to establish the reliability of the schemes. This makes it possible to determine whether or not the schemes are reliable. After carefully examining Table 4, it is evident that $ANF-PDO$ put out in this investigation yielded better results than those obtained by previous research that was part of the body of work being provided here. This conclusion was accomplished by using comparable measures, namely R^2, MAE, and $RMSE$, produced from the learning and evaluation data phase, respectively. Results that are produced by the superior model ($ANF-PDO$) are more dependable and resilient than those that are produced by XGB and $WOA-XGB$. This is because the $ANF-PDO$ model is more accurate. One may see this by observing the higher R^2 values as well as the decreased $RMSE$ and MAE values that are presented in [55]. For example, MAE reduction for XGB [55] during learning from 0.125 to 0.0451 and assessment from 0.149 to 0.0488. Furthermore, metrics calculated from the error based on $RMSE$ measurements dropped from 0.203 to 0.0643 during the learning phase and from 0.242 to 0.0666 during the evaluation phase. An in-depth comparison between $ANF-PDO$ and $WOA-XGB$ [55] could be formed by looking at the results in the learning, validating, and assessing datasets, where substantial gains were observed by boosting R^2 and reducing $RMSE$ and MAE.

The $ANFIS-PDO$ model outperformed the $RF-PDO$ framework in training, validation, and testing. Several variables explain this advantage. First, the $ANFIS$ design automatically blends neural network learning with fuzzy inference system rule-based reasoning to capture extremely nonlinear and complicated relationships between the nine input parameters and punching shear resistance. This hybrid structure allows for more flexibility in approximating nonlinear mappings, which tree-based approaches like RF struggle with. RF can handle noisy datasets and provide baseline predictions, but its piecewise structure restricts its capacity to capture intricate nonlinear connections compared to adaptive fuzzy rules in $ANFIS$.

Integrating the Prairie Dog Optimization (PDO) algorithm refined hyperparameters beyond trial-and-error procedures, improving performance. The optimization approach reduced overfitting and enhanced convergence. $ANFIS-PDO$ had shorter residual plot ranges and lower $RMSE$ values across all subgroups, indicating better error distributions. It seems that $ANFIS-PDO$ generalizes better to unknown data, lowering prediction variance and improving performance across slab-column configurations.

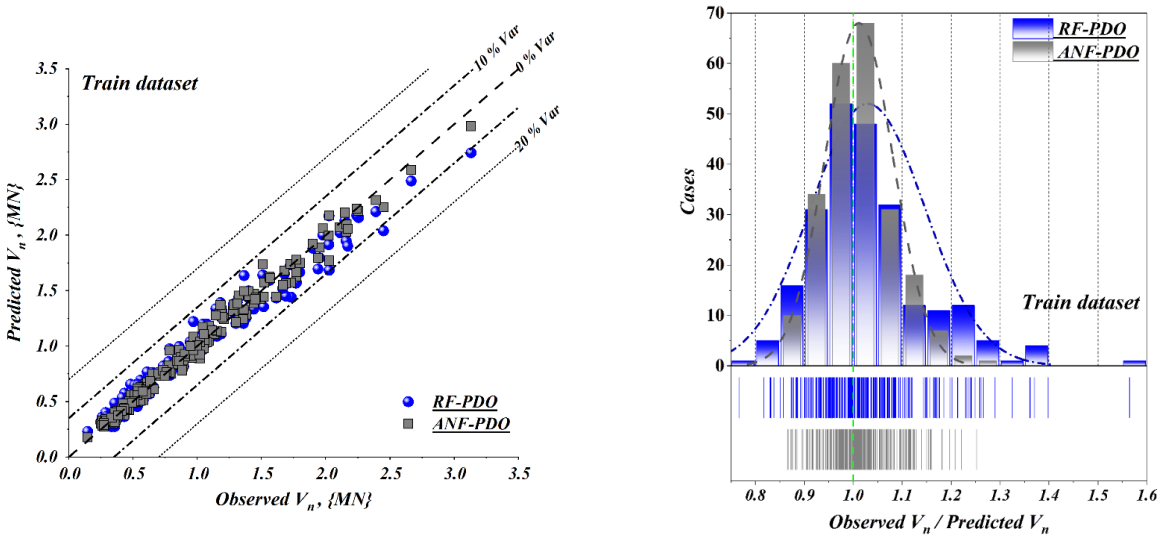
In comparison to past research, $RF-PDO$ and $ANFIS-PDO$ improved significantly. Early methods used empirical formulations or traditional machine learning models with little optimization, which underestimated or overestimated punching shear resistance. The suggested models not only improved R^2 values but also decreased error dispersion. Sensitivity research confirmed that slab depth, reinforcing characteristics, and concrete compressive strength strongly affect forecast accuracy. Because of its flexibility, $ANFIS-PDO$ could dynamically alter rule sets to changes in these factors, giving it a slight but constant advantage over $RF-PDO$.

Fig. 4 illustrates the ratio of measured V_n to projected V_n, which is monitored during learning, validation, and assessment. To assess the scheme's resilience, we provide a smaller distribution with a prominent peak and tightly restricted upper and lower boundaries below the central figure. The outcomes unequivocally showcase that $ANF-PDO$ consistently beat the $RF-PDO$ scheme in every step. This is shown by a more conspicuous summit and well-defined limits, especially when the ratio is one

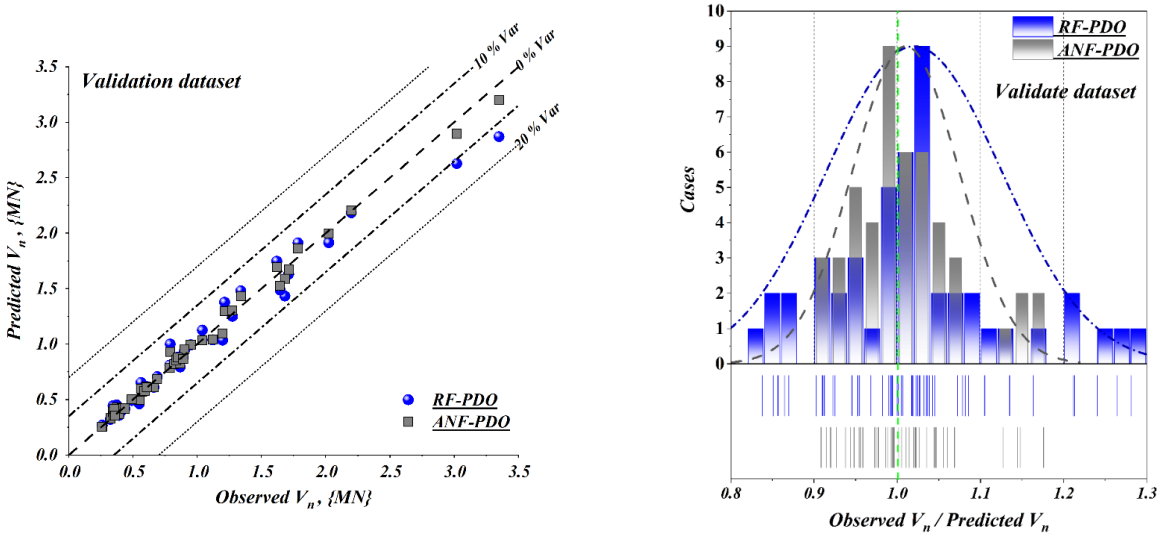
Table 4: The workability of the generated frameworks and comparison with the literature

Indicators	Sub-section	Created schemes			Publications		Eurocode 2	ACI 318-19
		$RF-PDO$	$ANF-PDO$	Differences (%)	XGB [55]	$WOA-XGB$ [55]		
R^2	Train	0.9731	0.9869	1.418	0.884	0.994	0.81	0.67
	Validation	0.9755	0.9938	1.876				
	Test	0.9753	0.9893	1.435	0.8682	0.9642		
$RMSE$	Train	0.0972	0.0643	-33.848	0.203	0.045	0.26	0.53
	Validation	0.1226	0.057	-53.507				
	Test	0.1057	0.0666	-36.991	0.242	0.125		
RAE	Train	0.0749	0.0499	-33.378				
	Validation	0.0812	0.0434	-46.552				

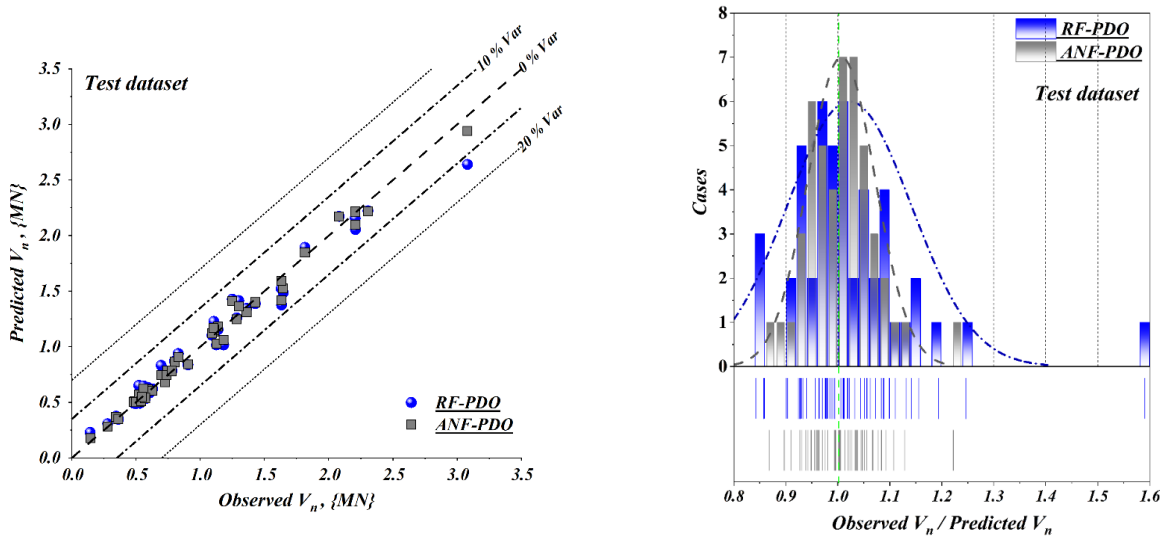
	Test	0.07	0.0481	-31.286				
RRSE	Train	0.1769	0.117	-33.861				
	Validation	0.1806	0.0839	-53.544				
	Test	0.1713	0.1079	-37.011				
MAE	Train	0.0678	0.0451	-33.481	0.125	0.033	0.17	0.41
	Validation	0.0761	0.0407	-46.518				
	Test	0.071	0.0488	-31.268	0.149	0.087		
SMAPE	Train	8.1302	5.2544	-35.372				
	Validation	7.5509	4.5664	-39.525				
	Test	7.337	4.9681	-32.287				
MASE	Train	0.1096	0.073	-33.394				
	Validation	0.1239	0.0662	-46.570				
	Test	0.129	0.0887	-31.240				
MSLE	Train	0.0019	8.49E-04	-55.316				
	Validation	0.0021	6.62E-04	-68.476				
	Test	0.0018	8.20E-04	-54.444				
MRE	Train	8.4911	5.3393	-37.119				
	Validation	7.7529	4.6555	-39.952				
	Test	7.7008	5.0154	-34.872				



(a)



(b)



(c)

Figure 4: The workability of generated schemes, a) Train phase, b) Validation phase, c) Test phase

A technique used in machine learning to identify how diverse input parameters and variables affect the output and effectiveness of the network is called sensitivity analysis. Several schemes were developed utilizing diverse input variables, and these schemes were included in the best-performing model ($ANF-PDO$) in the current study. $MASE$, $MSLE$, and MRE metrics were generated and compared utilizing $ANF-PDO$ to appraise the effects of diverse inputs (Table 5). The greater the impact of missing items on productivity, the greater the disparities for metrics. The outcomes underscore the significance of all attributes in projecting V_n, since most input factors hurt outcomes when compared to $ANF-PDO$. Most notably, when the d and ρ_t variables are taken out of the input group, there is a discernible gain in $MASE$, $MSLE$, and MRE metrics. The removal of the d throughout the training phase resulted in a rise in the MRE values from 5.3393 to 8.7804, 4.6555 to 9.8099 throughout the validating phase, and 5.0154 to 8.1239 throughout the evaluation phase. The outcomes imply that erasing every attribute parameter may lessen the schemes' dependability and comprehensiveness.

The $SHAP$ summary plot (Fig. 5) shows that the variable d has the strongest influence on the model's output, with both positive and negative impacts depending on its value. Other important features include ρ_t, $A_{sw,d}$, and c, which also display considerable variation in their contributions. Features such as f_c', f_y, a, and $f_{y,sw}$ have relatively smaller but still notable effects. Overall, the analysis highlights d as the dominant predictor, while the remaining variables provide secondary contributions to the model's decision-making.

Table 5: The sensitivity analysis of different scenarios on $ANF-PDO$

Index	Base model	Removed attribute								
	$ANF-PDO$	a	d	c	ρ_t	$A_{sw,d}$	f_{lc}	f_y	$f_{y,sw}$	FM
Train database										
$MASE$	0.073	0.0794	0.1172	0.0898	0.0928	0.0921	0.0852	0.0797	0.0808	0.0859
$MSLE$	8.49E-04	0.001	0.0023	0.0015	0.0015	0.0014	0.0011	9.6E-04	0.0011	0.0012
MRE	5.3393	5.8292	8.7804	6.5658	7.3135	6.4722	6.4381	5.9261	6.0575	6.1774
Validation database										
$MASE$	0.0662	0.0766	0.1354	0.08	0.0981	0.0835	0.0833	0.0723	0.0717	0.0729
$MSLE$	6.62E-04	9.7E-04	0.0043	0.0012	0.0016	8.9E-04	0.0011	7.5E-04	7.3E-04	7.5E-04
MRE	4.6555	5.5702	9.8099	6.0404	8.2436	5.3198	6.0478	4.9832	5.0006	5.0447
Test database										
$MASE$	0.0887	0.092	0.1346	0.1027	0.1244	0.0891	0.0913	0.1078	0.0969	0.0848
$MSLE$	8.20E-04	8.8E-04	0.0027	0.0011	0.0018	8.0E-04	9.0E-04	0.0018	9.6E-04	8.0E-04
MRE	5.0154	5.1404	8.1239	5.8218	7.7325	5.2199	5.0963	6.4097	5.6103	4.9549

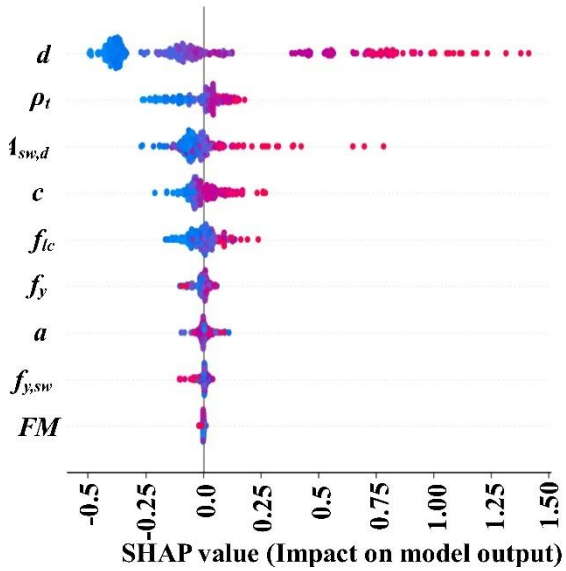


Figure 5: The $SHAP$ value to assess the impact of each model

7 Remarks

The purpose of this research is to recognize and examine the tree-based, fuzzy-based, and ML methods that have been demonstrated to be the most effective in lowering V_n of slab-column connections with shear reinforcement. This will be done to achieve the goal of reducing shear resistance. $ANFIS$ and Random forests analysis (RFA) are two pieces of software that were employed to achieve this objective. During this inquiry, the metaheuristic optimization techniques that were used included the Prairie dog algorithm, also known as PDA. RF and $ANFIS$ analyses were included in these methods to recognize the proper values for the parameters being considered for decision-making processes.

The Spearman correlation coefficient matrix reveals that the parameters have not shown considerable monotonic correlations, suggesting the variables are relatively independent and may not have strong non-linear relationships. This indicates the need for an AI model that can effectively capture complex, non-linear interactions to achieve accurate predictions, as the variables do not appear to have strong associations based on their ranked positions.

The data show that all qualities predict V_n, as most input parameters negatively affect outcomes compared to $ANF-PDO$. Removing the d and p_t parameters from the input group lead to significant improvements in $MASE$, $MSLE$, and MRE measures. The elimination of the d during training increased MRE values from 5.3393 to 8.7804, 4.6555 to 9.8099 when validating, and 5.0154 to 8.1239 during evaluating. The data suggest that erasing all attribute parameters may drop the scheme's reliability and comprehensiveness.

The results suggested that both $RF-PDO$ and $ANF-PDO$ can properly predict V_n. Efficiency study computations show that $ANF-PDO$ is more accurate and reliable than $RF-PDO$. Although $RF-PDO$ achieves lower values (0.9731, 0.9755, and 0.9753) throughout the learn, validate, and evaluate stages, $ANF-PDO$ achieves the highest R^2 values (0.9869, 0.9938, and 0.9893).

The $ANF-PDO$ has the minimum training, validation, and evaluation $RRSE$ index values of 0.117, 0.0839, and 0.1079. These measurements were more accurate than $RF-PDO$'s 0.1769, 0.1806, and 0.1713 in the training, validating, and appraisal stages. The $ANF-PDO$ can forecast competence and reliability since the two schemes for these measures vary by at least 31%, and in certain cases, 68%. Reasoning and assessment metrics show that $ANF-PDO$ is marginally more accurate and dependable.

This result was reached utilizing equivalent measurements (R^2, MAE, and $RMSE$) from the learn and assessment phases. XGB and $WOA-XGB$ adopted from the literature gives less reliable and robust results than $ANF-PDO$. Since the $ANF-PDO$ model is more precise. The increased R^2 values and lower $RMSE$ and MAE values demonstrate this. For XGB, MAE reduced from 0.125 to 0.0451 during learning and from 0.149 to 0.0488 during evaluation. $RMSE$-based error measures reduced from 0.203 to 0.0643 during learning and from 0.242 to 0.0666 during assessment.

References

[1] P. Wolfgang and H. Gerhard, "Composite Shear Head Systems for Improved Punching Shear Resistance of Flat Slabs," Apr. 26, 2012. doi: doi:10.1061/40826(186)22.

[2] R. S. Benemaran, M. Esmaeili-Falak, and M. S. Kordlar, "Improvement of recycled aggregate concrete using glass fiber and silica fume," *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 2023, doi: 10.1007/s41939-023-00313-2.

[3] P. Weerasinghe, K. Nguyen, P. Mendis, and M. Guerrieri, "Large-scale experiment on the behaviour of concrete flat slabs subjected to standard fire," *Journal of Building Engineering*, vol. 30, p. 101255, 2020, doi: https://doi.org/10.1016/j.jobe.2020.101255.

[4] S. Amiri and S. Behzad Talaeitaba, "Punching shear strengthening of flat slabs with EBROG and EBRIG – FRP strips," *Structures*, vol. 26, pp. 139–155, 2020, doi: https://doi.org/10.1016/j.istruc.2020.04.017.

[5] C.-C. Chen and C.-Y. Li, "Punching shear strength of reinforced concrete slabs strengthened with glass fiber-reinforced polymer laminates," *ACI Struct J*, vol. 102, no. 4, p. 535, 2005.

[6] R. Z. Alrousan and B. R. Alnemrawi, "The influence of concrete compressive strength on the punching shear capacity of reinforced concrete flat slabs under different opening configurations and loading conditions," *Structures*, vol. 44, pp. 101–119, 2022, doi: https://doi.org/10.1016/j.istruc.2022.07.091.

[7] G. Birkle, "Flat slabs: the influence of the slab thickness and the stud layout," 2004.

[8] R. T. S. Mabrouk, A. Bakr, and H. Abdalla, "Effect of flexural and shear reinforcement on the punching behavior of reinforced concrete flat slabs," *Alexandria Engineering Journal*, vol. 56, no. 4, pp. 591–599, 2017, doi: https://doi.org/10.1016/j.aej.2017.05.019.

[9] J. Einpaul, C. E. Ospina, M. Fernández Ruiz, and A. Muttoni, "Punching shear capacity of continuous slabs," *ACI Struct J*, pp. 861–872, 2016.

[10] P. Code, "Eurocode 2: design of concrete structures-part 1–1: general rules and rules for buildings," *British Standard Institution, London*, vol. 668, pp. 659–668, 2005.

[11] H. Marzouk, E. Rizk, and R. Tiller, "Design of shear reinforcement for thick plates using a strut-and-tie model," *Canadian Journal of Civil Engineering*, vol. 37, no. 2, pp. 181–194, Feb. 2010, doi: 10.1139/L09-120.

[12] A. S. A. Jabbar, M. A. Alam, and K. N. Mustapha, "A new equation for predicting punching shear strength of R/C flat plates," in *Proceedings National Graduate Conference*, 2012, pp. 8–10.

[13] I. Alkroosh and H. Ammash, "Soft computing for modeling punching shear of reinforced concrete flat slabs," *Ain Shams Engineering Journal*, vol. 6, no. 2, pp. 439–448, 2015, doi: https://doi.org/10.1016/j.asej.2014.12.001.

[14] Y. Dawei, Z. Bing, G. Bingbing, G. Xibo, and B. Razzaghzadeh, "Predicting the CPT-based pile set-up parameters using HHO-RF and PSO-RF hybrid models," *Structural Engineering and Mechanics*, vol. 86, no. 5, pp. 673–686, 2023.

[15] D. Li, X. Zhang, Q. Kang, and E. Tavakkol, "Estimation of unconfined compressive strength of marine clay modified with recycled tiles using hybridized extreme gradient boosting method," *Constr Build Mater*, vol. 393, p. 131992, 2023, doi: https://doi.org/10.1016/j.conbuildmat.2023.131992.

[16] M. Esmaeili-Falak and R. Sarkhani Benemaran, "Ensemble Extreme Gradient Boosting Based models to predict the Bearing Capacity of Micropile Group," *Applied Ocean Research*, 2024.

[17] R. Liang and B. Bayrami, "Estimation of frost durability of recycled aggregate concrete by hybridized Random Forests algorithms," *Steel and Composite Structures*, vol. 49, no. 1, pp. 91–107, 2023, doi: https://doi.org/10.12989/scs.2023.49.1.091.

[18] K. Zhang, Y. Zhang, and B. Razzaghzadeh, "Application of the optimal fuzzy-based system on bearing capacity of concrete pile," *Steel and Composite Structures*, vol. 51, no. 1, p. 25, 2024.

[19] M. Esmaeili-Falak and R. Sarkhani Benemaran, "Application of optimization-based regression analysis for evaluation of frost durability of recycled aggregate concrete," *Structural Concrete*, vol. 25, no. 1, pp. 716–737, Jan. 2024, doi: https://doi.org/10.1002/suco.202300566.

[20] N. K. Oghli and M. Esmaeili-Falak, "Predicting Liquefaction Triggering Potential Using Metaheuristic GMDH Approaches," *Geotechnical and Geological Engineering*, vol. 43, no. 6, p. 308, 2025, doi: 10.1007/s10706-025-03275-z.

[21] M. Esmaeili-Falak and A. Letafat, "Prediction of Unconfined Compressive Strength in Pozzolanic Geopolymer-Stabilized Granular Materials Using Tree-Based Models," *Transportation Infrastructure Geotechnology*, vol. 12, no. 7, p. 247, 2025, doi: 10.1007/s40515-025-00701-w.

[22] A. Ahmad, G. Kotsovou, D. M. Cotsovos, and N. D. Lagaros, "Assessing the accuracy of RC design code predictions through the use of artificial neural networks," *International Journal of Advanced Structural Engineering*, vol. 10, no. 4, pp. 349–365, 2018, doi: 10.1007/s40091-018-0202-4.

[23] F. Demir, "Prediction of elastic modulus of normal and high strength concrete by artificial neural networks," *Constr Build Mater*, vol. 22, no. 7, pp. 1428–1435, 2008, doi: https://doi.org/10.1016/j.conbuildmat.2007.04.004.

[24] R. Perera, M. Barchín, A. Arteaga, and A. De Diego, "Prediction of the ultimate strength of reinforced concrete beams FRP-strengthened in shear using neural networks," *Compos B Eng*, vol. 41, no. 4, pp. 287–298, 2010, doi: https://doi.org/10.1016/j.compositesb.2010.03.003.

[25] M. Shariati *et al.*, "Application of a Hybrid Artificial Neural Network-Particle Swarm Optimization (ANN-PSO) Model in Behavior Prediction of Channel Shear Connectors Embedded in Normal and High-Strength Concrete," 2019. doi: 10.3390/app9245534.

[26] M. Shariati *et al.*, "Identification of the most influencing parameters on the properties of corroded concrete beams using an Adaptive Neuro-Fuzzy Inference System (ANFIS)," *Steel Compos Struct*, vol. 34, no. 1, p. 155, 2020.

[27] A. Karimipour, J. Mohebbi Najm Abad, and N. Fasihihour, "Predicting the load-carrying capacity of GFRP-reinforced concrete columns using ANN and evolutionary strategy," *Compos Struct*, vol. 275, p. 114470, 2021, doi: https://doi.org/10.1016/j.compstruct.2021.114470.

[28] J. H. Haido, "Prediction of the shear strength of RC beam-column joints using new ANN formulations," *Structures*, vol. 38, pp. 1191–1209, 2022, doi: https://doi.org/10.1016/j.istruc.2022.02.046.

[29] L. Chen, P. Fakharian, D. Rezazadeh Eidgahee, M. Haji, A. Mohammad Alizadeh Arab, and Y. Nouri, "Axial compressive strength predictive models for recycled aggregate concrete filled

circular steel tube columns using ANN, GEP, and MLR," *Journal of Building Engineering*, vol. 77, p. 107439, 2023, doi: https://doi.org/10.1016/j.jobe.2023.107439.

[30] X.-S. Yang, "Review of meta-heuristics and generalised evolutionary walk algorithm," *International Journal of Bio-Inspired Computation*, vol. 3, no. 2, pp. 77–84, Jan. 2011, doi: 10.1504/IJBIC.2011.039907.

[31] J. Li *et al.*, "Predicting the shear strength of concrete beam through ANFIS-GA–PSO hybrid modeling," *Advances in Engineering Software*, vol. 181, p. 103475, 2023, doi: https://doi.org/10.1016/j.advengsoft.2023.103475.

[32] I. Faridmehr, M. L. Nehdi, and M. Hajmohammadian Baghban, "Novel informational bat-ANN model for predicting punching shear of RC flat slabs without shear reinforcement," *Eng Struct*, vol. 256, p. 114030, 2022, doi: https://doi.org/10.1016/j.engstruct.2022.114030.

[33] N. Concha, J. R. Aratan, E. M. Derigay, J. M. Martin, and R. E. Taneo, "A hybrid neuro-swarm model for shear strength of steel fiber reinforced concrete deep beams," *Journal of Building Engineering*, vol. 76, p. 107340, 2023, doi: https://doi.org/10.1016/j.jobe.2023.107340.

[34] M. Imran Waris, V. Plevris, J. Mir, N. Chairman, and A. Ahmad, "An alternative approach for measuring the mechanical properties of hybrid concrete through image processing and machine learning," *Constr Build Mater*, vol. 328, p. 126899, 2022, doi: https://doi.org/10.1016/j.conbuildmat.2022.126899.

[35] M. S. Sandeep, K. Tiprak, S. Kaewunruen, P. Pheinsusom, and W. Pansuk, "Shear strength prediction of reinforced concrete beams using machine learning," *Structures*, vol. 47, pp. 1196–1211, 2023, doi: https://doi.org/10.1016/j.istruc.2022.11.140.

[36] R. Pourhanasa and A. Monadipour, "Concrete crack detection via graph representation learning and texture analysis," *Innovative Infrastructure Solutions*, vol. 10, no. 8, pp. 1–12, 2025.

[37] E. Rivandi and R. Jamili Oskouie, "A Novel Approach for Developing Intrusion Detection Systems in Mobile Social Networks," *Available at SSRN 5174811*, 2024.

[38] M. B. Bagherabad, E. Rivandi, and M. J. Mehr, "Machine Learning for Analyzing Effects of Various Factors on Business Economic," *Authorea Preprints*, 2025.

[39] A. Hashemi, J. Beheshti, and M. Mohammadi, "Physics-Based, AI-Driven Surrogate Modeling for Structural Displacement Prediction in Mechanical Systems with Limited Sensor Data," *IEEE Access*, 2025.

[40] H. Lima, R. Palhares, G. Sales de Melo, and M. Oliveira, "Experimental analysis of punching shear in flat slabs with variation in the anchorage of shear reinforcement," *Structural Concrete*, vol. 22, no. 2, pp. 1165–1182, Apr. 2021, doi: https://doi.org/10.1002/suco.202000158.

[41] P. M. Lewiński and P. P. Więch, "Finite element model and test results for punching shear failure of RC slabs," *Archives of Civil and Mechanical Engineering*, vol. 20, no. 2, p. 36, 2020, doi: 10.1007/s43452-020-00037-x.

[42] S. Kinnunen, *Punching of Structural Concrete Slabs: Technical Report*, vol. 12. fib Fédération internationale du béton, 2001.

[43] Y. Jin, W. J. Yi, and L. Hu, "Experimental study of performance of reinforced concrete slab-column connection with punching shear keys," *Ind. Constr*, vol. 47, pp. 60–65, 2017.

[44] J.-I. Jang and S.-M. Kang, "Punching Shear Behavior of Shear Reinforced Slab–Column Connection with Varying Flexural Reinforcement," *Int J Concr Struct Mater*, vol. 13, no. 1, p. 29, 2019, doi: 10.1186/s40069-019-0341-4.

[45] V. Hugo Dalosto de Oliveira, H. Jorge Nery de Lima, and G. Sales Melo, "Punching shear resistance of flat slabs with different types of stirrup anchorages such as shear reinforcement," *Eng Struct*, vol. 253, p. 113671, 2022, doi: https://doi.org/10.1016/j.engstruct.2021.113671.

[46] M. P. Ferreira, G. S. Melo, P. E. Regan, and R. L. Vollum, "Punching of reinforced concrete flat slabs with double-headed shear reinforcement," *ACI Struct J*, vol. 111, no. 2, p. 363, 2014.

[47] T.-S. Eom, S.-M. Kang, T.-W. Choi, and H.-G. Park, "Punching shear tests of slabs with high-strength continuous hoop reinforcement," *ACI Struct J*, vol. 115, no. 5, p. 1295, 2018.

[48] T. X. Dam, J. K. Wight, and G. J. Parra-Montesinos, "Behavior of Monotonically Loaded Slab-Column Connections Reinforced with Shear Studs.," *ACI Struct J*, vol. 114, no. 1, 2017.

[49] R. Cantone, M. Fernández Ruiz, J. Bujnak, and A. Muttoni, "Enhancing punching strength and deformation capacity of flat slabs," *ACI Struct J*, vol. 116, no. 5, pp. 261–274, 2019.

[50] M. Bartolac, D. Damjanović, and I. Duvnjak, "Punching strength of flat slabs with and without shear reinforcement," *Građevinar*, vol. 67, no. 08., pp. 771–786, 2015.

[51] A. E. Ezugwu, J. O. Agushaka, L. Abualigah, S. Mirjalili, and A. H. Gandomi, "Prairie Dog Optimization Algorithm," *Neural Comput Appl*, vol. 34, no. 22, pp. 20017–20065, 2022, doi: 10.1007/s00521-022-07530-9.

[52] C. Rao, M. Liu, M. Goh, and J. Wen, "2-stage modified random forest model for credit risk assessment of P2P network lending to 'Three Rurals' borrowers," *Appl Soft Comput*, vol. 95, p. 106570, 2020, doi: https://doi.org/10.1016/j.asoc.2020.106570.

[53] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Trans Syst Man*

Cybern, vol. 23, no. 3, pp. 665–685, 1993, doi: 10.1109/21.256541.

[54] M. Esmaeili-Falak, H. Katebi, M. Vadiati, and J. Adamowski, "Predicting triaxial compressive strength and Young's modulus of frozen sand using artificial intelligence methods," *Journal of Cold Regions Engineering*, vol. 33, no. 3, p. 4019007, 2019, doi: https://doi.org/10.1061/(ASCE)CR.1943-5495.0000188.

[55] H. Yan, N. Xie, and D. Shen, "Hybrid Machine Learning Algorithms for Prediction of Failure Modes and Punching Resistance in Slab-Column Connections with Shear Reinforcement," *Buildings*, vol. 14, no. 5, p. 1247, 2024.