WILEY | Hindawi

Research Article

Hybridized Deep Learning Model for Perfobond Rib Shear Strength Connector Prediction

Jamal Abdulrazzaq Khalaf ⬥,[1] **Abeer A. Majeed** ⬥,[2] **Mohammed Suleman Aldlemy** ⬥,[3] **Zainab Hasan Ali** ⬥,[4] **Ahmed W. Al Zand** ⬥,[5] **S. Adarsh,**[6] **Aissa Bouaissi** ⬥,[7,8] **Mohammed Majeed Hameed** ⬥,[9] **and Zaher Mundher Yaseen** ⬥[10]

[1]*Civil Engineering Department, Collage of Engineering, University of Anbar, Ramadi, Iraq*
[2]*Reconstruction and Projects Department, University of Baghdad, Baghdad, Iraq*
[3]*Department of Mechanical Engineering, Collage of Mechanical Engineering Technology, Benghazi, Libya*
[4]*College of Engineering, Civil Engineering Department, University of Diyala, Baquba, Iraq*
[5]*Department of Civil Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), UKM Bangi 43600, Selangor, Malaysia*
[6]*Department of Civil Engineering, TKM College of Engineering Kollam, Kollam, India*
[7]*School of Engineering, University of Plymouth, Plymouth PL4 8AA, UK*
[8]*UNA Developments Ltd., Airport Business Center, Plymouth Devon PL6 7PP, UK*
[9]*Department of Civil Engineering, Al-Maaref University College, Ramadi, Iraq*
[10]*Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam*

Correspondence should be addressed to Zaher Mundher Yaseen; zaheryaseen88@gmail.com

Accurate and reliable prediction of Perfobond Rib Shear Strength Connector (PRSC) is considered as a major issue in the structural engineering sector. Besides, selecting the most significant variables that have a major influence on PRSC in every important step for attaining economic and more accurate predictive models, this study investigates the capacity of deep learning neural network (DLNN) for shear strength prediction of PRSC. The proposed DLNN model is validated against support vector regression (SVR), artificial neural network (ANN), and M5 tree model. In the second scenario, a comparable AI model hybridized with genetic algorithm (GA) as a robust bioinspired optimization approach for optimizing the related predictors for the PRSC is proposed. Hybridizing AI models with GA as a selector tool is an attempt to acquire the best accuracy of predictions with the fewest possible related parameters. In accordance with quantitative analysis, it can be observed that the GA-DLNN models required only 7 input parameters and yielded the best prediction accuracy with highest correlation coefficient ($R = 0.96$) and lowest value root mean square error (RMSE = 0.03936 KN). However, the other comparable models such as GA-M5Tree, GA-ANN, and GA-SVR required 10 input parameters to obtain a relatively acceptable level of accuracy. Employing GA as a feature parameter selection technique improves the precision of almost all hybrid models by optimally removing redundant variables which decrease the efficiency of the model.

1. Introduction

Steel-concrete composite/hybrid systems have found wide application in several engineering works, due to the recent advancements in structural engineering. In this regard, the shear connector serves as an important component that ensures the development of composite actions by facilitating the shear transfer between the concrete slab and the steel profile [1–3]. At the site, conventional shear connectors (i.e., Nelson stud) are beneficial owing to their high level of automation; meanwhile, they are prone to certain problems, especially, in structures that are subjected to stress [3–5]. When compared with other connectors, Nelson stud somehow exhibits low resistance which can lead to the

design of girders with partial interaction. Considering this fact, many research studies have focused on how to improve the shear connector for hybrid composite systems [6, 7], with the first work dated back to the 1980s. The development of Perfobond, another form of a connector with higher resistance, was reported by Leonhardt, Andra, and Partners in 1987, when working on the third bridge that crossed the Caroni River in Venezuela [3, 6, 8]. A study by Vianna et al. [9] compared the economic costs of steel girders manufacturing using different types of connectors. From the outcome of the study, it was observed that Perfobond connectors are more cost-efficient to be used in the steel-concrete composite. Later, Vianna et al. reported another study on Perfobond and T-Perfobond rib shear connectors in terms of their ductility, resistance, and collapse modes [3]. The results showed that PRSC is both structurally efficient and economical in terms of shear transfer in hybrid and composite structures. Other research studies focused on the numerical and parametric evaluation of PRSC on 40 pushout samples [2, 10]. The study of a simple perforated plate PRSC contains several holes and transverse rebars; the sample also exhibited varying concrete compressive strengths. The study reported two major findings as it involved finite element (FE) method and regression analysis during the prediction of the PRSC shear capacity [11, 12]. Another experimental study on the structural response of PRSC was performed by [4]. The study reported an increase in the resistance of PRSC with increases in the number of holes, and based on this outcome, it was submitted that the resistance and ductility of PRSC can be increased by passing the reinforcement bars through the holes while reducing the upward displacement. Rodrigues and Laím focused on the influence of the holes number, rib holes, and transverse reinforcement, as well as the doubled PRSC at both ambient and high temperatures [6]. From the outcomes, higher temperatures significantly impacted PRSC in terms of its load-carrying capacity, especially the doubled PRSC. The study further showed that transverse reinforcement bars, when present in rib holes, cause a reduction in the capacity of PRSC to carry a load, especially at high temperatures. A parametric study on PRSC shear strength based on the FE method has been reported [13]. The developed model in this study was verified using experimental pushout tests, and from the results, a mathematical model was developed for the estimation of the shear capacity of PRSC. The shear behavior of PRSC has been investigated by [14], under both cyclic and static loadings. Based on the static tests, the results showed the shear capacity of pure concrete-based specimens to be about 65% of that of specimens with both concrete end-bearing zone and concrete dowels. Specimens with transverse rebars in holes exhibited shear capacity of about two times that of specimens with transverse rebars. From the cyclic tests, samples without transverse rebars showed a significant decrease in residual shear capacity compared to that of static shear capacity. Hence, specimens with transverse rebars exhibited residual shear capacity that was almost similar to their static shear capacity.

A parametric study has been reported on the circular-hole and long-hole PRSC [15]. From the outcome of the study, a relationship was established between the failure mode of both long-hole and circular-hole PRSC and the concrete failure. It was also reported that the increase in both height and diameter increases the shear stiffness of PRSC. Steel-concrete decks with PRSC have been investigated for dynamic characteristics in [16, 17]. The study considered PRSC with both normal-weight high-strength concrete and lightweight high-strength concrete. The considered characteristics include the natural frequencies, frequency response functions, and damping ratio; these were evaluated using a nondestructive approach. The experimental outcomes of natural frequencies were also compared using the FE model. From the results, the first mode with a damping ratio of almost 0.5% was found to be the most effective mode for both concrete types. Relying on these studies, it is evident that several factors influence the structural behavior of PRSC. Such factors include end-bearing force, rib spacing, rib arrangement, and concrete compressive strength, as well as the yield and area strength of the transverse rebars [18]. Having identified these factors that govern the behavior of PRSC, it becomes necessary that the resistance of PRSC should be predicted for proper implementation by the decision makers [19, 20]. The quantification of the PRSC resistance based on analytical methodologies has been introduced; however, there are certain limitations of such methods [21].

According to several studies published in the literature, their behavior of PRSC is affected by various contributing factors, including the area and the yield strength of transverse rears, the end-bearing force, the concrete compressive strength, the rib spacing, and the rib arrangement. In addition, several analytical and empirical models were developed to predict the resistance of PRSC; however, it provided undesirable predictions with an increase in calculated errors as well.

The advancements in technology have made computer-aid methods some of the optimistic alternatives for modelling several structural engineering-related problems, and the most famous among them is the artificial intelligence models which are easy, applicable, convenient, and strong predictive models [22–25]. AI models are beneficial as they can solve nonlinear, stochastic, and nonstationary problems that may not be addressed when the classical regression models are used [26–29]. Several AI models have been developed to determine the actual relationship between the predictors and perfobond rib shear strength. For instance, Köroğlu et al. [30] investigated the genetic programming (GP) model for ultimate shear capacity prediction in composite beams with profiled steel sheeting. The study compared the model's accuracy in an ultimate shear capacity prediction of the composite beams to that of the proposed GP model based on the employed test data. From the results, it was submitted that the new GP performed more accurate ultimate shear capacity prediction of the composite beams compared to the building codes. Another study by Ali [31] focused on the prediction of the shear strength of channel shear connectors in a composite beam that consists of concrete and steel sections based on adaptive neuro-fuzzy inference system (ANFIS) and linear regression (LR) which are nonlinear and linear modelling tools, respectively. From the results, ANFIS performed more accurate and precise predictions than LR. Although, there have been several

explorations of AI models for modelling shear strength-related structure and material problems [32, 33]. It can be observed for the existing approaches that the capacity of the used model is mainly influenced by the structure of the used approach and the selection of input parameters. Based on that, introducing a novel approach has capabilities to discover the complex relationship between predictors and target which is very important to increase the prediction precession [34]. Moreover, incorporating that approach with a novel algorithm to select the most important input parameters and accurately predict the Perfobond Rib Shear Strength Connector with high accuracy is very significant for structural engineering.

The research scope of computer aid is still limited, and the exploration of newly developed AI models is still ongoing research motivation. The DLNN model is a newly explored AI version that demonstrates a reliable machine learning model for solving nonlinear regression problems [35–39] and yet to be developed for the PRSC shear strength prediction. The current research is conducted based on the implementation of the deep learning model for predicting PRSC shear strength modelling. Prior to the step, to predict the PRSC shear strength, genetic algorithm is used with the DNN model to select the most important input parameters and then introduce these factors to the adopted approach. The proposed DLNN model is validated against several well-established machine learning models including M5Tree, ANN, and SVR. The investigation is extended with the integration of the genetic algorithm as a robust nature-inspired optimization algorithm for input parameters' selection. The obtained results are assessed and discussed comprehensively.

2. Experimental Dataset Description

To evaluate the shear strength capacity of a Perfobond connector (see Figure 1), data included 90 records related to the shear connector of steel-concrete structures. These records were collected from eight databases published in literature studies. These studies comprised data was collected from [1–5, 10, 40]. The input variables included concrete compressive strength (f_c), area of concrete dowels (A_D), rib holes number (n), area of cross reinforcement bars and yield stress of reinforcement bars ($A_{tr,r}f_{y,r}$), area of cross reinforcement bars and the tensile strength of cross rebar ($A_{tr}f_y$), area of the connector at the end-bearing zone (Ab), the ratio between the thickness of the concrete slab to connector height (b/h_{sc}), connector height (h_{sc}), the contact area between the connector and concrete (A_F), and coefficient of end-bearing force (α). These parameters were entered to the hybrid model to predict shear of Perfobond connector Q_{exp}. The description statistics of the experimental dataset are shown in Table 1. The data were separated into groups: training dataset 85% and testing dataset 15%.

3. Methodology Overview

3.1. Artificial Neural Network (ANN).
In the last few decades, artificial networks, such as neural networks, social networks,
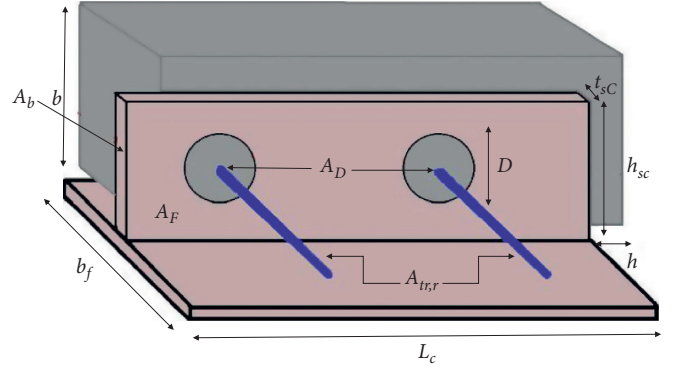

FIGURE 1: The systematic framework of the Perfobond rib shear strength connector.

and other algorithms, have established. The main merit of these technologies is their abilities to predict data and deal with complex systems. Neural network is a mathematical model applied to make the decision in the right way by mimicking the neurons in the biological brain [41]. The construction of a neural network is depending on the connection of several layers called neurons. Feedforward neural network with backpropagation learning algorithm is widely used by researchers. Backpropagation is a common algorithm in neural network applications due to its ability in training the network based on the supervised learning method [42]. In this algorithm, the predicted value is compared with the original variable to compute the error between them. The algorithm modified the weights in the neural network to decrease the error value to a small amount. The structure of the algorithm is explained as follows:

$$N - H_1 - H_2 - \cdots - H_{\text{NHL}} - M, \tag{1}$$

where N represents the input neuron, H is the number of hidden layers in the neural network, and M is the output variable. The construction of a feedforward neural network contains at least an input variable, output, and hidden layer. In this network, the information is transferred in one direction from input to hidden and output layer without loops in the network. The number of input variables and predicted labels are corresponded to the number of neurons in the input and output layer, respectively. Neurons in hidden layers are used for nonlinear transformation of the input variables. Hidden layers in the neural network are calculated by

$$v_i = \left(1 + \exp\left(-1 \times \sum_{i=1}^{1} x_i w_{ij}\right)\right)^{-1}, \tag{2}$$

where v_i refers to the hidden layer, x_i is the input variable, and w_{ij} represents the weight between the layers. The value of the output layer can be computed as below:

$$y = \left(1 + \exp\left(-1 \times \sum_{j=1}^{1} v_j w_{ij}\right)\right)^{-1}. \tag{3}$$

To design the network, a number of nodes and hidden layers are required. Various studies stated that one or two

TABLE 1: Descriptive statistics of experimental data.

	fc	AD	n	$A_{tr,r}f_{y,r}$	$A_{tr}f_y$	$Ab = t_{SC}h_{SC}$	b/h_{SC}	h_{SC}	A_F	α	q_{exp}
Training phase QUOTE hsc											
Max	52.6	9817.48	5	281486.7	876251.51	1950	3.6	150	112993.4	1	774.2
Min	20.91	0	0	0	0	137.5	1.2	25	8831.77	0	13.14
Mean	30.60	3601.63	2.2	33955.61	417272.072	1325.95	1.65	107.16	69384	0.85	387.15
SD	8.84	3100.35	1.37	74153.85	328887.75	566.76	0.84	38.06	31925.45	0.35	181.96
CV	0.28	0.86	0.62	2.18	0.79	0.42	0.51	0.36	0.46	0.42	0.47
Testing phase											
Max	52.6	9817.48	5	281486.7	876251.51	1950	3.6	150	112993.4	1	774.2
Min	26.6	0	0	0	302378.29	612	1.2	102	45500	1	282
Mean	30.07	5628.68	2.86	60056.78	799735.0807	1068.4	1.67	115.06	74868.07	1	505.94
SD	8.84	3646.44	1.85	103972.06	195079.3831	454.23	0.76	17.41	18781.76	0	174.36
CV	0.29	0.65	0.64	1.731	0.243930006	0.43	0.45	0.15	0.25	0	0.34

layers are enough to achieve better prediction performance [43, 44]. The best performance of the training process is based on the good selection of inputs' network. During this process, the relations between inputs and outputs are designed by the neural network. In every iteration phase, the modification of weights and biases is done by decreasing the error measure between actual and predicted outcomes. The error between actual and predicted values is presented below:

$$\text{Error} = 0.5(d - y)^2, \tag{4}$$

where d represents the real value and y is the estimated value obtained from the algorithm. In this study, one hidden layer was used with sigmoid activation function due to its validity in the regression process. Figure 2 shows the structure of the neural network.

3.2. Deep Learning Neural Networks (DLNN).

Recently, the study of neural network application is based on the concept of deep learning technique. The structure of deep neural network is an extension of classical neural network with the addition of extra hidden layer(s) to the network. Deep learning was introduced by Hinton et al. through proposing the layerwise greedy-learning method. By this method, the neural network is pretrained by unsupervised learning technique before the training process layer by layer. Deep learning technique is popular due to two reasons: (i) developing of huge technical data, which can solve the problem of overfitting and (ii) assigning of nonrandom value to the neural network before the unsupervised learning process [45]. Thus, better performance can be reached after the training phase. There are various types of deep learning tools, and in this study, the backpropagation neural network was used.

This approach is used in many types of application, the same multihidden layers' approach, and trained via backpropagation with gradient descent algorithm. The network contains input, output, and large numbers of neurons and hidden layers. The algorithm is worked based on the connection between the first layer and hidden layers which leads to yielding of a new variable that is transferred to the output layer. Then, the output layer predicts the result of the process. The specific thing in deep learning is the nonlinear
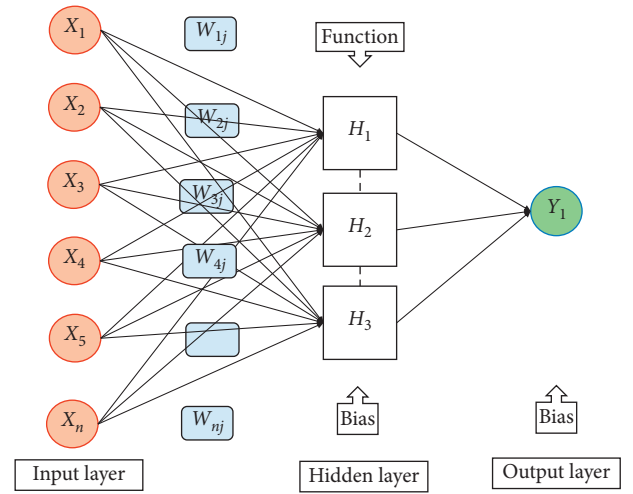

FIGURE 2: The schematic structure of the applied ANN predictive model.

relations between the multiple layers in the network that gave them the ability to deal with different nonlinear functions. This deep network can recognize complex patterns used in a complicated process. Figure 3 describes the general structure of a deep neural network with input, output, and multiple hidden layers. The mathematical process can be discussed as follows:

$$M^I = f\left(w^I \varphi^{I-1} + b^I\right), \quad \text{for } 0 < I < L, \tag{5}$$

where f represents the activation function, w is the weight matrix and, b is the bias. The input variable is denoted by 0 layer and L represents the output layer. In this research, hyperbolic tangent function is utilized as an activation function due to its ability to obtain better performance in the study problem.

3.3. M5 Rule Model.

M5 rule algorithm was developed by Holmes et al. [35] to forecast the numeric and nominal data. Building of the M5 rule is based on M5 tree by using the trees to build the model. The popular technique of M5 Model Tree (MT) works with classification cum regression principle. MT propounded by Quinlan [46] divides the complete domain
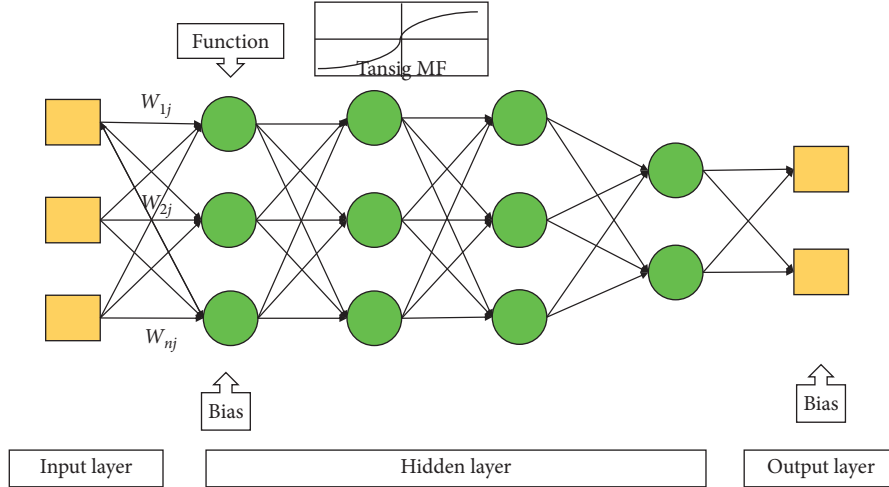
FIGURE 3: The systematic structure of the developed DLNN model.

into many subdomains, and multiple linear regression models are developed for each of them. In this case, nonlinear input-output relationships are approximated by a number of linear models.

Rule generation depended on the partial and regression tree (PART) model developed by [47]. The work of the algorithm depends on iterating model constructing and choosing the rule which has a good result at each iteration. In the training phase, the M5 model is applied, and then, the best leaf is chosen as a rule. The process continues until all instances are tested and utilized by the rules. The main merit in this approach is that the algorithm builds full trees and develops a small amount of dataset at the testing phase [48, 49]. In the first stage, in the development of the MT model, a decision tree is developed following a division criterion. Based on the criterion used for dividing the domain, a number of variants of model trees are available, and the one which follows standard deviation reduction (SDR) as the criterion is known as an M5 Model Tree [42, 43]. The SDR quantifies the reduction in error at each node while testing of attributes, and its computation can be made as follows:

$$\sigma^R = \sigma(N) - \sum \frac{|T_i|}{|N|} \sigma(T_i), \qquad (6)$$

where σ^R = standard deviation reduction, N is the number of training samples, T_i is the training samples of ith subdomain, and = $\sigma(N)$ and $\sigma(T_i)$ are the SDs of total samples and ith subdomain sample. The resulting model for the subdomain can be represented as $O = a_0 + a_1 x_1 + a_2 x_2 + \cdots$, where O is the output, $a_1, a_2 \ldots$ are the coefficients of linear regression, and $x_1, x_2 \ldots$ are the inputs. The procedure of computation is illustrated in Figure 4, which shows the division to the number of subdomains followed by development of different models considering x_1 and x_2 as inputs.

The partitioning process should be continued till the variation in the class values of all the instances that are associated with a node becomes negligible. Then, the models are refined by the "pruning" and "smoothing" processes, which may help to alleviate the "overfitting" and abrupt

changes between individual subclasses [50]. The complete theoretical description of M5 model trees is available in literature [43, 51]. This method does not demand any control parameter settings, while on the contrary, its application results in more user friendly linear models [45, 48].

3.4. Support Vector Regression (SVR).

Support vector regression has been developed by [52] as an algorithm based on using a hyperplane to separate a dataset and calculate the distance from the hyperplane and the nearest variable. In recent year, SVR algorithm has been intensively used by many researchers for solving different engineering problems and shown better prediction than other machine learning algorithms [53, 54]. SVR estimates the error between the input and output variable in the regression process by computing the distance from SVR margin. The mathematical expression of the SVR model is shown as follows:

$$M = \{x_1, y_1), x_2, y_2), \ldots, x_n y_n), \qquad (7)$$

where M denotes the dataset training and x and y represent the input and output variables. The SVR function that applied to the training dataset is

$$f(x) = w \cdot \varphi(x) + b, \qquad (8)$$

where w represents the weight vector and $\varphi(x)$ refers to the high dimension of input space, whereas b represents the scalar vector. In the regression problem, error deviation can be calculated by the following optimization algorithm:

$$\min \varphi(w, \xi) = 0.5w^2 + C\left(\sum_{i=1}^{n} \xi_i\right) \varphi(x_i) + b \qquad (9)$$

$$\text{subjected to} \quad M_i\left(w^M x_i + b\right) \geq 1 - \xi_i, \ \xi \geq 0,$$

where ξ represents the slag variable and C is equal to the error between regularization and empirical error. To optimize the SVR model, Lagrange multipliers and sequential minimal optimization were used. Figure 5 illustrates SVR regression with ε-insensitive loss function. In this study, the predicted
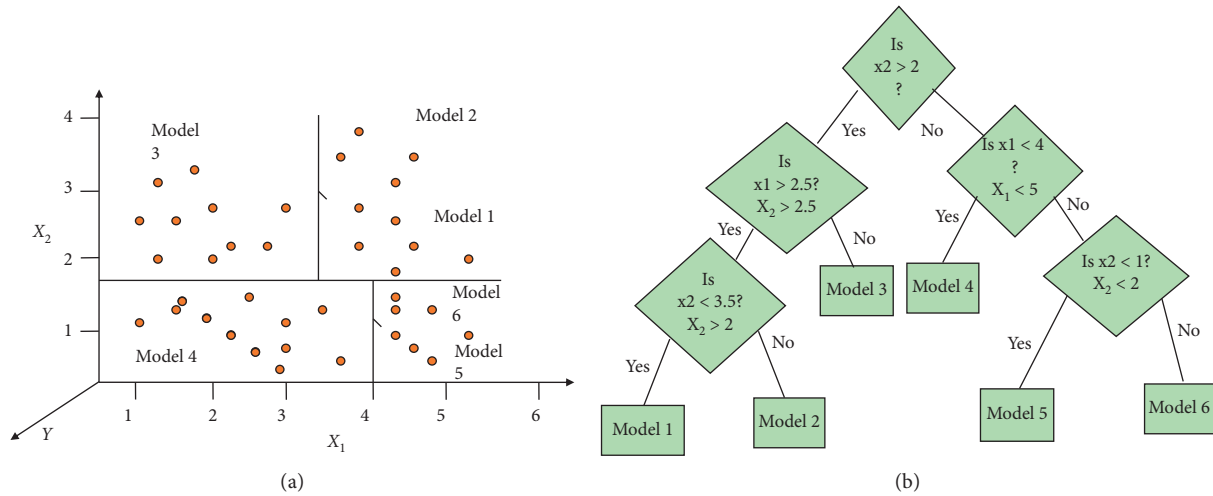
(a)

(b)

FIGURE 4: The systematic structure of M5 Tree modelling approach.
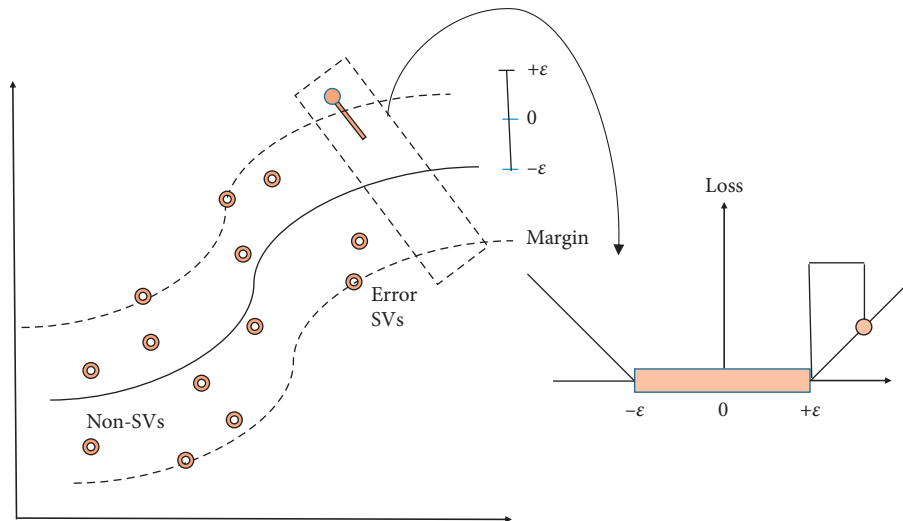

FIGURE 5: A basic structure of the SVR model and calibrating data points with ε-insensitive loss function.

problem is characterized by a nonlinear relationship between input and output variables, in which the nonlinear mapping of SVR can be suitable to calculate the data correlation. In the SVR model, 4 kernel functions can be used for nonlinear mapping during the training phase. These kernels include linear, sigmoid, polynomial, and radial basis functions [53]. Radial basis function is applied due to its efficiency and ability to deal with complex regression problems [55].

3.5. Hybridized Genetic Algorithm (GA) with AI Models. Genetic algorithm is an evolutionary algorithm used to optimize solutions in complicated systems by finding on biological selections [56]. GA has been employed in various research areas such as pattern recognition, image processing, and control system [57]. In several research studies in engineering and science applications, GA demonstrated a reliable method in feature selection than other selection tools [58, 59]. The efficiency of GA can be discussed by its ability to explore the search space and concentration on the global optimization which led to a better investigation, utilizing the search space. The main idea included the application of natural selection, such as the creation of chromosomes, crossover, and mutation in solving complex processes. These processes are employed to reduce the features which are transferred to binary string [60, 61].

There are three main phases at natural selection (see Figure 6). Firstly, use crossover to produce offspring, then mutation may occur to the generated individuals and, finally, the fittest individual is selected. The first step in genetic programming is the population initialization. In this step, many individuals are generated randomly. Then, the fittest individual is chosen to produce the offspring. This phenomenon can be applied for search space. We produce many solutions to the study problem and the best solution is selected from them. In this study, a genetic algorithm is employed to choose the highly correlated
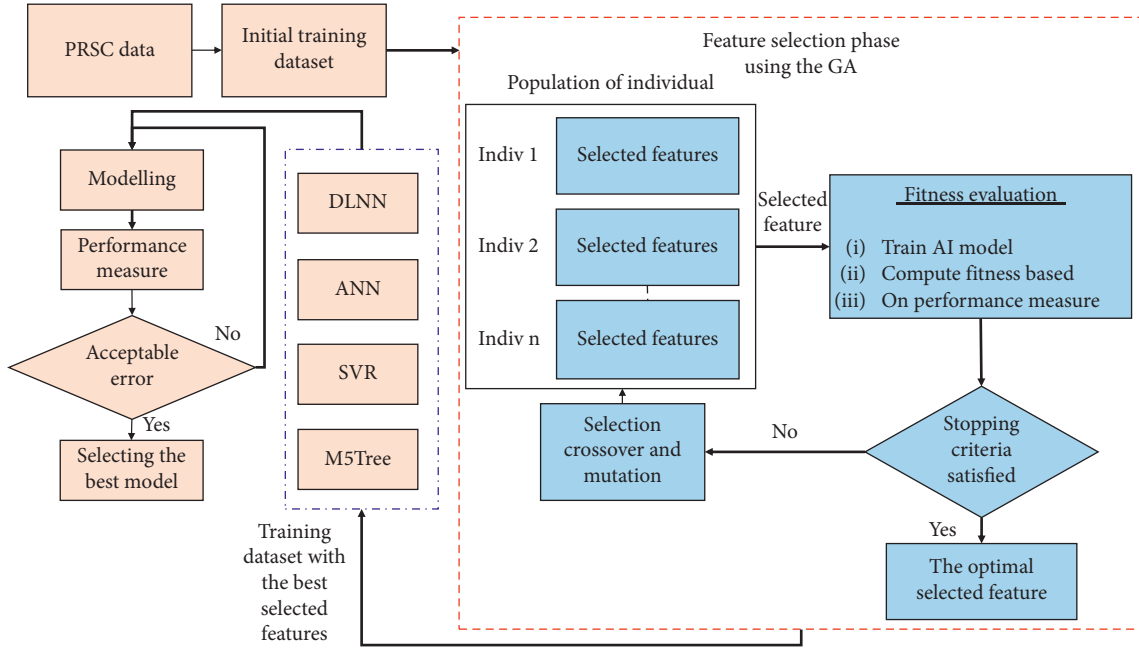
FIGURE 6: Process of selecting best input variables of each AI modelling using GA algorithm.

variables and the process is begun from two variables. Then, the models ANN, DLNN, SVR, and M5 rule are applied [62, 63].

3.6. Performance Metrics. In order to select the best predictive modelling approaches, five statistical indicators are used in this study which are root mean square error (RMSE), mean absolute (MAE), mean absolute percentage error (MAPE), relative root mean squared error (RRMSE), relative error (RE), mean relative error (MRE), Nash–Sutcliffe efficiency (NSE), and BIAS. The mathematical expressions of these measures can be seen below [64–66]:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)^2},$$

$$MAE = \frac{1}{N} \sum_{i=1}^{n} \left|\left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)\right|,$$

$$MAPE\% = \frac{1}{N} \sum_{i=1}^{n} \left|\frac{\left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)}{Q_{i \text{ exp}}}\right| * 100,$$

$$RRMSE = \frac{\sqrt{(1/N) \sum_{i=1}^{n} \left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)^2}}{\sum_{I=1}^{N} Q_{i \text{ exp}}},$$

$$BIAS = \frac{1}{N} \sum_{i=1}^{N} \left|\left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)\right| * 100,$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n} \left(Q_{i \text{ exp}} - Q_{i\text{pre}}\right)^2}{\sum_{i=1}^{n} \left(Q_{i \text{ exp}} - \overline{Q_{\text{exp}}}\right)^2}\right],$$

$$RE\% = 100 * \frac{Q_{i \text{ exp}} - Q_{i\text{pre}}}{Q_{i \text{ exp}}},$$

$$MRE\% = \frac{100}{N} \sum_{i=1}^{N} \left|\frac{Q_{i \text{ exp}} - Q_{i\text{pre}}}{Q_{i \text{ exp}}}\right|,$$

(10)

where $Q_{i \text{ exp}}$ and $Q_{i\text{pre}}$ are the observed and predicted ith values of PRSC capacity, $\overline{Q_{\text{exp}}}$ represents the average values of PRSC, and N is the total number of experimental samples. In this regard, eight statistical measures have been used to assess the performances of the adopted models. The mentioned measures are commonly used to assess the efficiency of AI models, including statistical parameter (i.e., RMSE and MAE) used for evaluating the forecasted error between actual and predictive values. Besides, the other measures such as NSE are employed to compute the degree of correlation between the predictive and actual points.

4. Results and Discussion

This section describes the results obtained from standard AI models according to all available data. The other scenario of this part of the study shows the incorporation of genetic algorithm as a tool used for selecting the most significant input parameters for adopted AI models. In both scenarios, the performance of each model is assessed based on quantitative assessment, using different statistical criteria, and visualized assessments, using different plots and figures.

4.1. First Scenario: Applying the Standard Models. The motivation of this current study is to accurately predict the shear strength of PRSC using different AI modelling approaches, including DLNN, SVR, ANN, and M5 Tree. In this

scenario, all mentioned predictive models have been established based on ten predictors (f_c, $A_D n$, $A_{tr,r} f_{y,r}$, $A_{tr} f_y$, Ab, b/h_{sc}, connector height h_{sc}, A_F, and α). The collected experimental samples were divided into two sets, during the stage of developed standalone AI models. The majority of samples (85%) were used for modelling construction, whereas the rest was used for validation purposes. In order, to evaluate the performance accuracy of each modelling technique separately, ten statistical metrics were used, including correlation measures and error measures. Simulated results obtained by four predictive models for both training and testing stages were illustrated in Table 2. It can be clearly seen that all models during the training stage yielded unpromising accuracy except DLNN models which provided the best accuracy of predictions and produced the highest values of NSE (0.957) and the lowest values of RMSE (0.047 KN), MAE (0.033 KN), and RMSRE (0.914). However, the testing phase is the most important stage in the evaluation of the accuracy of the predictive models. According to Table 2, the superiority of DLNN over other AI models can be easily observed during the testing phase. Moreover, DLNN models generated the highest accuracy of predicted shear strength values with the shortest magnitudes of RMSE (0.045 KN), MAE (0.020 KN), RRMSE (0.092), and the highest values of NSE (0.888).

To evaluate the performance of each developed model during the testing phase in a more rigorous way, several graphical visualizations were established including scatter plots, relative error plots, and Taylor diagram. The scatter plot is considered a very important figure in the evaluation of the variance between the predicted and the actual shear strength values. Based on Figure 7, DLNN modes presented less scatter and recorded a higher value of correlation coefficient ($R = 0.96$) than the other comparable models (R of 0.95, 0.95, and 0.94, respectively, for M5Tree, ANN, and SVR). Besides, among all AI models, the DLNN modelling approaches produced fewer relative error percentages (see Figure 8). Figure 8 clearly indicates that except for one sample (sample 12), the relative error of predictions by DLNN is ±20% indicating a success rate of 92%, while by other AI methods, multiple samples surpass the relative error limits of even 40%.

For better visual comparison, the Taylor diagram was established because it can summarize different statistical measures (correlation coefficient and standard deviation) in one figure thereby, facilitating the process of selecting the best model accuracy.

Taylor diagrams are polar plots that present the similarity between observed and predicted data based on the correlation coefficient and standard deviation in a 2D plane. From Figure 9, it is evident that the point corresponding to DLNN predictions is the closest to the point corresponding to the observed dataset, indicating the best performance by the standalone DNN model, i.e., DLNN generated more accurate predicted values and closer to the actual ones. Based on the mentioned result, DLNN models showed better generalization capabilities in comparison with the other AI models during training and testing phases. Conversely, SVR modelling approach exhibited the lowest level of prediction accuracy in comparison with all AI predictive models.

4.2. Second Scenario: Applying the Hybrid Models. This section of the paper investigates the capability of using GA as a bioinspired algorithm for assisting the four AI models in selecting the best combination of input parameters, which importantly affects the PRSC. As the AI models can efficiently learn from the behavior of the datasets, it is very essential to minimize the model complexity. Thus, any improvement in the model performance with the use of minimal input parameters can be considered as economical in the modelling process. With this objective, the potential of GA was used for developing hybrid models with different AI methods in this study. Accordingly, eight different combination models were developed by hybridizing different AI methods with GA (GA-ANN, GA-SVR, GA-DLNN, and GA-M5Tree. The model combinations for different hybrid methods were designated as M1 to M8 in this paper, whereas the number of input parameters varied from 2 to 9 as presented in Tables 3–6.

The performance of prediction abilities over the training and the testing phases for the hybrid models are summarized in Tables 7–10. The most remarkable note can be observed that the GA improved the performance of the most predictive models in comparison with pure AI models, which have been carried out in the first scenario of this paper. For instance, the hybrid (GA-M5Tree-M8) recorded good accuracy of predictions compared with standard (M5Tree) model and the statistical measures, such as RMSE and MAE reduced by 8.55% and 3.77%, respectively, during the testing phase. The robustness of GA in properly selecting the optimal input parameters can be clearly seen when GA-M5Tree-M6 (with seven parameters) model generated slightly higher predicting accuracy than standard M5Tree model (with ten input variables). For more comparative analysis, GA-ANN-M6 (with 7 input variables) performed better than the best standalone DLNN and GA-M5Tree-M6, respectively. With respect to SVR-GA models, they were slightly improved in comparison with standard SVR. Moreover, all these GA-SVR models (8 models) showed the lowest accuracy and none of them could outperform the standard DLNN. On top of that, GA-SVR models scored the highest values of relative error (ranging from 20.25 to 25.44%) in comparison with other modelling approaches. Although all GA-SVR models performed the lowest accuracy of performances in comparison with other hybrid models in this scenario, they also showed lower prediction accuracy than standard SVR models which have been carried out in the first scenario. Therefore, it can be concluded that there was no specific advantage in the hybridization of SVR with GA for this problem and dataset. On the contrary, the hybridization of GA with DLNN provided more excellent predicted results than all standalone models which performed in the first scenario.

Among the different hybrid DLNN models, the hybrid GA-DLNN- M6 model performed very well (with NSE of 0.914, RMSE of 0.039 KN, MAPE of 0.052, and MAE of 0.021 KN). Furthermore, it was noted that very fewer bias indicators were generated in GA-DLNN-M6, GA-DLNN-M7, and GA-DLNN-M8 (−0.0004, -0.0004, and -0.01) than standalone DLNN model (-0.011). It was evident that GA-based hybrid models can improve the performance when

TABLE 2: The performance prediction skills for the standalone predictive models over the training and testing phase.

| AI model	MAPE	RMSE (KN)	MAE (KN)	RRMSE (KN)	MRE	BIAS	NSE
Training phase							
DLNN	0.173	0.047	0.033	0.914	0.163	−0.026	0.957
M5Tree	0.636	0.089	0.064	4.176	0.528	0.000	0.848
ANN	0.283	0.090	0.065	1.312	0.178	0.000	0.845
SVR	1.235	0.108	0.086	8.377	1.120	0.000	0.776
Testing phase							
DLNN	0.047	0.045	0.020	0.093	0.027	−0.011	0.888
M5Tree	0.062	0.047	0.023	0.129	0.041	−0.005	0.880
ANN	0.064	0.047	0.023	0.130	−0.009	−0.003	0.878
SVR	0.096	0.052	0.023	0.252	0.096	−0.007	0.853
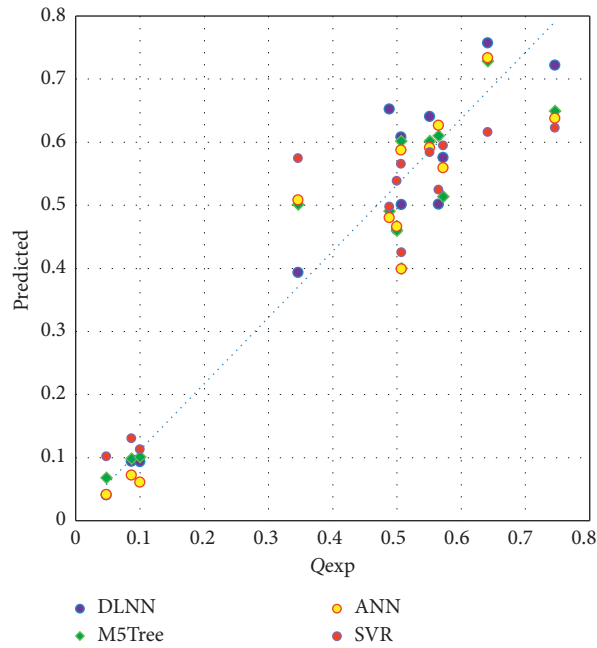
FIGURE 7: The scatter plots graphical presentation over the testing modelling phase: classical AI models.
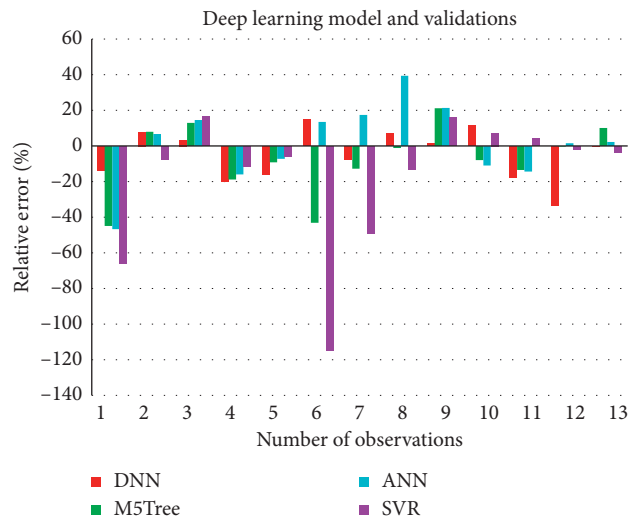

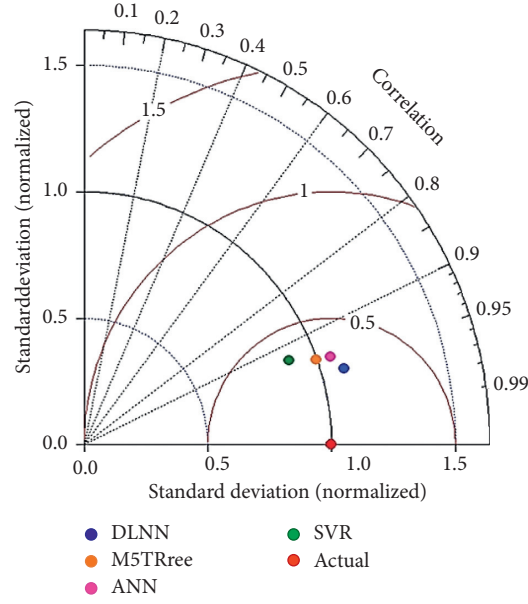FIGURE 8: Relative error distribution for each AI approach: testing phase.

FIGURE 9: Taylor diagram graphical presentation over the testing modelling phases: classical AI models.

TABLE 3: The abstraction of the influential input parameters using GA-DLNN for building the prediction process for the shear strength.

Inputs	Models
2	M_1: $Q_{exp} = f(A_D, \alpha)$
3	M_2: $Q_{exp} = f(A_D, A_F, \alpha)$
4	M_3: $Q_{exp} = f(A_D, n, A_{tr,r} f_{y,r}, \mathbf{h}_{SC})$
5	M_4: $Q_{exp} = f(A_D, n, A_{tr,r} f_{y,r}, \mathbf{h}_{SC}, \alpha)$
6	M_5: $Q_{exp} = f(A_D, n, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, A_F)$
7	M_6: $Q_{exp} = f(f_c, n, A_{tr,r} f_{y,r}, A_{tr} f_y, b/h_{SC}, h_{SC}, \alpha)$
8	M_7: $Q_{exp} = f(f_c, A_D, A_{tr,r} f_{y,r}, A_{tr} f_y, b/h_{SC}, h_{SC}, A_F, \alpha)$
9	M_8: $Q_{exp} = f(f_c, n, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, b/h_{SC}, h_{SC}, A_F, \alpha)$

TABLE 4: The abstraction of the influential input parameters using GA-M5tree for building the prediction process for the shear strength.

Inputs	Combinations
2	M_1: $Q_{exp} = f(A_D, \alpha)$
3	M_2: $Q_{exp} = f(A_D, A_F, \alpha)$
4	M_3: $Q_{exp} = f(f_C, n, A_{tr} f_y, A_F)$
5	M_4: $Q_{exp} = f(A_D, n, A_{tr,r} f_{y,r}, \mathbf{h}_{SC}, \alpha)$
6	M_5: $Q_{exp} = f(A_D, n, A_{tr,r} f_{y,r}, b/\mathbf{h}_{SC}, A_F, \alpha)$
7	M_6: $Q_{exp} = f(f_c, A_D, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, \mathbf{b}/h_{SC}, h_{SC})$
8	M_7: $Q_{exp} = f(f_c, n, A_{tr,r} f_{y,r}, A_{tr} f_y, b/h_{SC}, h_{SC}, A_F, \alpha)$
9	M_8: $Q_{exp} = f(f_c, n, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, b/h_{SC}, h_{SC}, A_F, \alpha)$

TABLE 5: The abstraction of the influential input parameters using GA-ANN for building the prediction process for the shear strength.

Inputs	Combinations
2	M_1: $Q_{exp} = f(A_D, \mathbf{Ab} = t_{sc}h_{sc})$
3	M_2: $Q_{exp} = f(A_D, A_{tr,r}, f_{y,r}, \alpha)$
4	M_3: $Q_{exp} = f(n, A_{tr,r} f_{y,r}, \mathbf{Ab} = t_{sc}h_{sc}, h_{sc})$
5	M_4: $Q_{exp} = f(f_c, A_D, n, A_{tr,r} f_{y,r}, \mathbf{h}_{sc})$
6	M_5: $Q_{exp} = f(f_C, n, A_{tr,r} f_{y,r}, A_{tr} f_y, \mathbf{Ab} = t_{SC}h_{SC}, h_{SC})$
7	M_6: $Q_{exp} = f(f_C, n, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, A_F, \alpha)$
8	M_7: $Q_{exp} = f(f_c, A_D, n, A_{tr,r} f_{y,r}, A_{tr} f_y, b/h_{SC}, A_F, \alpha)$
9	M_8: $Q_{exp} = f(f_c, A_D, n, A_{tr,r} f_{y,r}, A_{tr} f_y, Ab = t_{SC}\mathbf{h}_{SC}, b/h_{SC}, h_{SC}, \alpha)$

Table 6: The abstraction of the influential input parameters using GA-SVR for building the prediction process for the shear strength.

Inputs	Combinations
2	$M_1: Q_{exp} = f(A_D, \alpha)$
3	$M_2: Q_{exp} = f(f_c, A_D, A_F)$
4	$M_3: Q_{exp} = f(A_{tr,r}f_{y,r}, A_{tr}f_y, b/h_{SC}, h_{sc})$
5	$M_4: Q_{exp} = f(f_c, A_D, n, A_{tr}f_y, Ab = t_{SC}\mathbf{h}_{SC}, \alpha)$
6	$M_5: Q_{exp} = f(f_c, A_D, A_{tr,r}f_{y,r}, A_{tr}f_y, A_F, \alpha)$
7	$M_6: Q_{exp} = f(f_c, n, A_{tr,r}f_{y,r}, A_{tr}f_y, Ab = t_{SC}\mathbf{h}_{SC}, b/h_{SC}, \alpha)$
8	$M_7: Q_{exp} = f(f_c, A_D, A_{tr,r}f_{y,r}, A_{tr}f_y, Ab = t_{SC}\mathbf{h}_{SC}, \mathbf{h}_{SC}, A_F, \alpha)$
9	$M_8: Q_{exp} = f(f_c, n, A_{tr,r}f_{y,r}, A_{tr}f_y, Ab = t_{SC}\mathbf{h}_{SC}, b/h_{SC}, h_{SC}, A_F, \alpha)$

Table 7: The performance prediction skills for the hybrid GA-DLNN predictive model over the training and testing phases.

	MAPE	RMSE (kN)	MAE (kN)	RRMSE (kN)	MRE	BIAS	NSE
Training phase							
GA-DLNN-M1	0.675	0.123	0.094	4.025	0.506	−0.006	0.710
GA-DLNN-M2	1.099	0.122	0.097	7.470	0.955	−0.001	0.714
GA-DLNN-M3	0.823	0.115	0.089	5.415	0.693	−0.002	0.746
GA-DLNN-M4	0.414	0.110	0.086	2.100	0.275	0.005	0.766
GA-DLNN-M5	0.862	0.102	0.074	6.117	0.734	0.002	0.800
GA-DLNN-M6	0.519	0.086	0.064	3.388	0.407	0.001	0.859
GA-DLNN-M7	0.501	0.086	0.064	3.079	0.374	−0.001	0.859
GA-DLNN-M8	0.239	0.046	0.030	1.462	0.228	−0.022	0.959
Testing phase							
GA-DLNN-M1	0.117	0.084	0.039	0.256	−0.018	−0.004	0.611
GA-DLNN-M2	0.097	0.078	0.034	0.222	0.037	−0.005	0.663
GA-DLNN-M3	0.081	0.071	0.032	0.187	0.051	−0.006	0.718
GA-DLNN-M4	0.059	0.064	0.027	0.160	0.028	−0.005	0.776
GA-DLNN-M5	0.054	0.051	0.024	0.121	0.021	−0.003	0.858
GA-DLNN-M6	0.052	0.039	0.021	0.087	−0.013	0.000	0.914
GA-DLNN-M7	0.044	0.042	0.020	0.090	0.010	0.000	0.903
GA-DLNN-M8	0.042	0.040	0.019	0.082	0.011	−0.007	0.912

Table 8: The performance prediction skills for the hybrid GA-M5Tree predictive model over the training and testing phases.

	MAPE	RMSE (kN)	MAE (kN)	RRMSE (kN)	MRE	BIAS	NSE
Training phase							
GA-M5Tree-M1	0.674	0.121	0.095	4.023	0.522	0.000	0.718
GA-M5Tree-M2	0.319	0.123	0.095	0.978	0.153	0.006	0.706
GA-M5Tree-M3	0.409	0.111	0.086	1.350	0.275	0.000	0.762
GA-M5Tree-M4	1.365	0.123	0.096	9.442	1.221	0.000	0.710
GA-M5Tree-M5	0.465	0.115	0.090	2.353	0.323	0.000	0.743
GA-M5Tree-M6	0.636	0.089	0.064	4.176	0.528	0.000	0.848
GA-M5Tree-M7	1.178	0.110	0.087	8.547	1.018	−0.002	0.765
GA-M5Tree-M8	0.769	0.094	0.069	5.262	0.637	−0.001	0.830
Testing phase							
GA-M5Tree-M1	0.097	0.079	0.035	0.227	−0.006	−0.002	0.653
GA-M5Tree-M2	0.114	0.085	0.039	0.246	−0.031	0.003	0.605
GA-M5Tree-M3	0.107	0.080	0.041	0.208	0.004	−0.003	0.650
GA-M5Tree-M4	0.134	0.083	0.039	0.306	0.104	−0.006	0.616
GA-M5Tree-M5	0.072	0.073	0.030	0.188	0.044	−0.006	0.705
GA-M5Tree-M6	0.062	0.047	0.023	0.129	0.041	−0.005	0.880
GA-M5Tree-M7	0.096	0.063	0.028	0.236	−0.027	−0.005	0.781
GA-M5Tree-M8	0.071	0.043	0.024	0.129	−0.005	−0.004	0.900

hybridized with DLNN, NN, and M5 Model Tree with a fewer number of input parameters for this dataset. Moreover, the GA-DLNN-M6 was considered very simply when compared to the other models because fewer input parameters are needed, and it can achieve a significant improvement compared to the standard DLNN, in which the RMSE and the RMSRE reduced by 12.62% and 6.06%, respectively, whereas the NSE was increased by 2.98%. The superiority of this model (GA-DLNN) did not only appear in comparison to simple models, but also appeared when compared to hybrid models

TABLE 9: The performance prediction skills for the hybrid GA-ANN predictive model over the training and testing phases.

	MAPE	RMSE (kN)	MAE (kN)	RRMSE (kN)	MRE	BIAS	NSE
Training phase							
GA-ANN-M1	1.895	0.127	0.098	13.468	1.733	0.000	0.687
GA-ANN-M2	0.622	0.117	0.096	3.617	0.469	0.000	0.734
GA-ANN-M3	0.529	0.124	0.098	2.162	0.378	0.000	0.705
GA-ANN-M4	1.218	0.112	0.085	8.436	1.089	0.000	0.758
GA-ANN-M5	0.311	0.095	0.069	1.374	−0.105	0.000	0.826
GA-ANN-M6	0.701	0.093	0.070	4.675	−0.518	0.000	0.833
GA-ANN-M7	0.204	0.099	0.079	0.454	0.078	0.000	0.811
GA-ANN-M8	0.209	0.091	0.066	0.690	0.102	0.000	0.842
Testing phase							
GA-ANN-M1	0.167	0.094	0.045	0.400	0.137	−0.007	0.514
GA-ANN-M2	0.077	0.073	0.030	0.190	0.016	−0.003	0.706
GA-ANN-M3	0.106	0.076	0.033	0.247	0.088	−0.008	0.685
GA-ANN-M4	0.104	0.074	0.034	0.241	0.066	−0.005	0.696
GA-ANN-M5	0.066	0.051	0.026	0.125	0.001	−0.005	0.857
GA-ANN-M6	0.083	0.044	0.024	0.181	−0.035	−0.003	0.893
GA-ANN-M7	0.072	0.049	0.024	0.149	−0.018	−0.002	0.867
GA-NN-M8	0.065	0.048	0.023	0.133	−0.008	−0.004	0.874

TABLE 10: The performance prediction skills for the hybrid GA-SVR predictive model over the training and testing phases.

	MAPE	RMSE (kN)	MAE (kN)	RRMSE (kN)	MRE	BIAS	NSE
Training phase							
GA-SVR-M1	4.123	0.154	0.132	30.498	3.949	−0.002	0.544
GA-SVR-M2	5.964	0.169	0.135	44.736	5.805	−0.016	0.447
GA-SVR-M3	3.415	0.159	0.134	25.019	3.225	0.010	0.513
GA-SVR-M4	2.398	0.122	0.102	17.574	2.250	0.008	0.713
GA-SVR-M5	2.621	0.122	0.103	19.240	2.487	−0.003	0.713
GA-SVR-M6	1.689	0.121	0.100	11.560	1.551	0.005	0.715
GA-SVR-M7	1.691	0.111	0.087	12.229	1.575	0.000	0.760
GA-SVR-M8	1.483	0.122	0.099	9.990	1.349	0.003	0.712
Testing phase							
GA-SVR-M1	0.312	0.086	0.041	0.972	0.379	−0.018	0.589
GA-SVR-M2	0.465	0.112	0.049	1.525	0.593	−0.028	0.306
GA-SVR-M3	0.269	0.071	0.038	0.822	0.320	−0.014	0.719
GA-SVR-M4	0.157	0.054	0.026	0.471	0.176	−0.008	0.841
GA-SVR-M5	0.185	0.057	0.029	0.558	0.217	−0.011	0.822
GA-SVR-M6	0.148	0.050	0.026	0.413	0.164	−0.008	0.859
GA-SVR-M7	0.115	0.051	0.025	0.310	0.118	−0.007	0.854
GA-SVR-M8	0.124	0.050	0.024	0.340	0.138	−0.009	0.863

(GA-SVR, GA-ANN, and GA-M5Tree). Generally, the best input parameters improved the models comprising f_c, n, $A_{tr,r} f_{y,r}$, $A_{tr} f_y$, b/h_{sc}, h_{sc}, and α. These input combinations are considered the most efficient parameters which significantly affect the PRSC.

For visualization assessment, scatter plots for each hybrid model were shown in Figures 10–13. These figures are very important in evaluating the performance of each predictive model. Besides, the best model should be established based on fewer input variables as well as generating predicted values with less diversion from the actual ones. It can be seen from these figures that GA-DLNN-M6 produced the highest accuracy performance with R of 0.96 with respect to other hybrid models, they generated lower accuracy of performances, and, in most cases, they required many input parameters to gain slight improvements.

Figure 14 portrays more concrete and convincing statistical relationship between the forecasted and the actual shear strength, using the Taylor diagram. A visual comparison of the four plots shows that the points corresponding to the high-end hybrid models M6–M8 resulted in points closer to the point corresponding to the actual data. Among these models, M6 (7 input model) lies much closer except for the SVR-based hybrid model. Also, it is clearly evident that the GA-DLNN-M6 hybrid model recorded the closer predicted values to the actual ones. This also supported the selection of the adopted model (GA-DLNN-M6), which has been considered in this study, and its assessment was very consistent with other quantitative and visualized assessments which performed previously. Furthermore, it can be noted that the GA-DLNN-M6 model presented fewer input parameters with the best
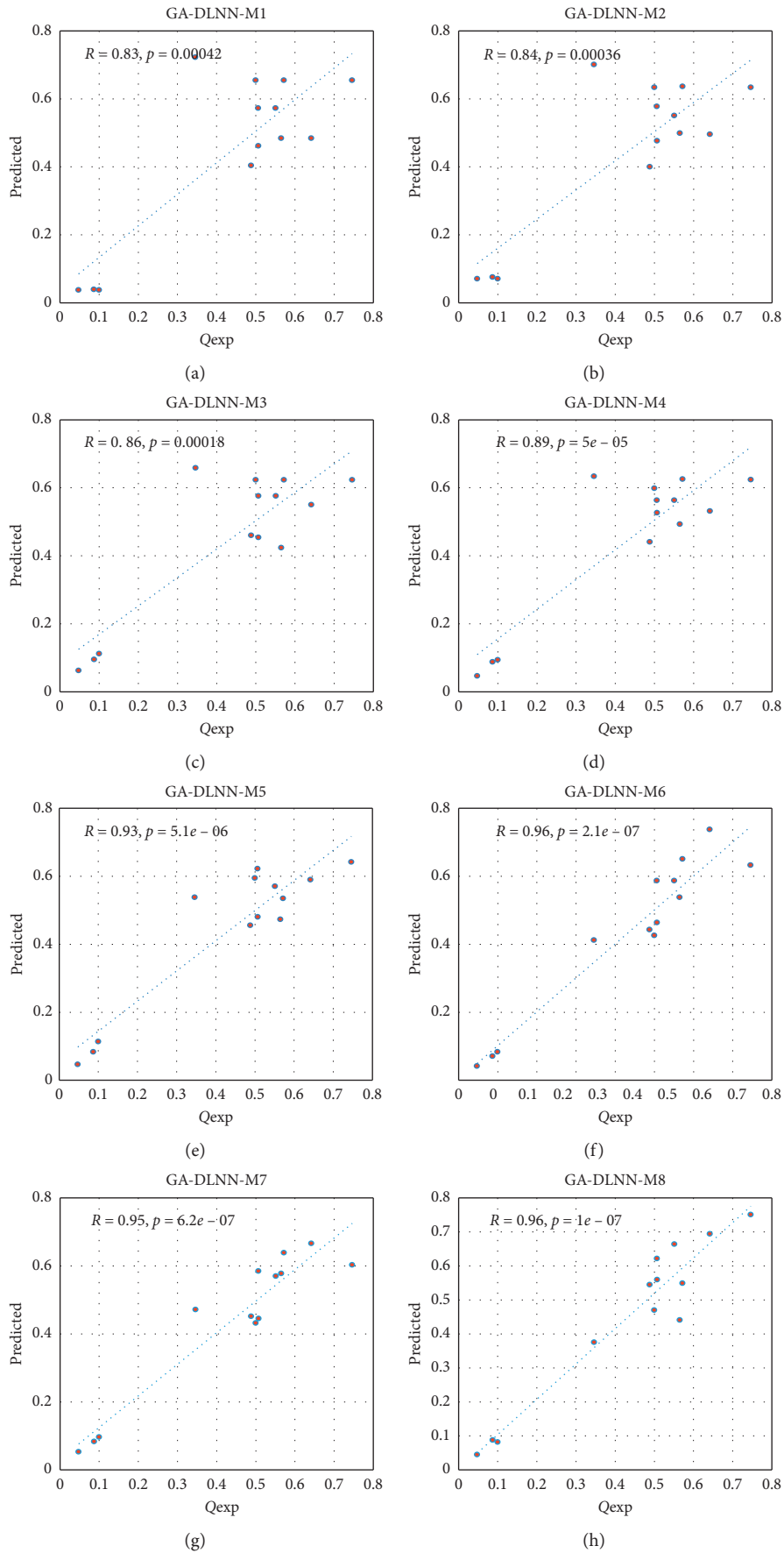
FIGURE 10: The scatter plots graphical presentation over the testing modelling phase: hybrid GA-DLNN models.
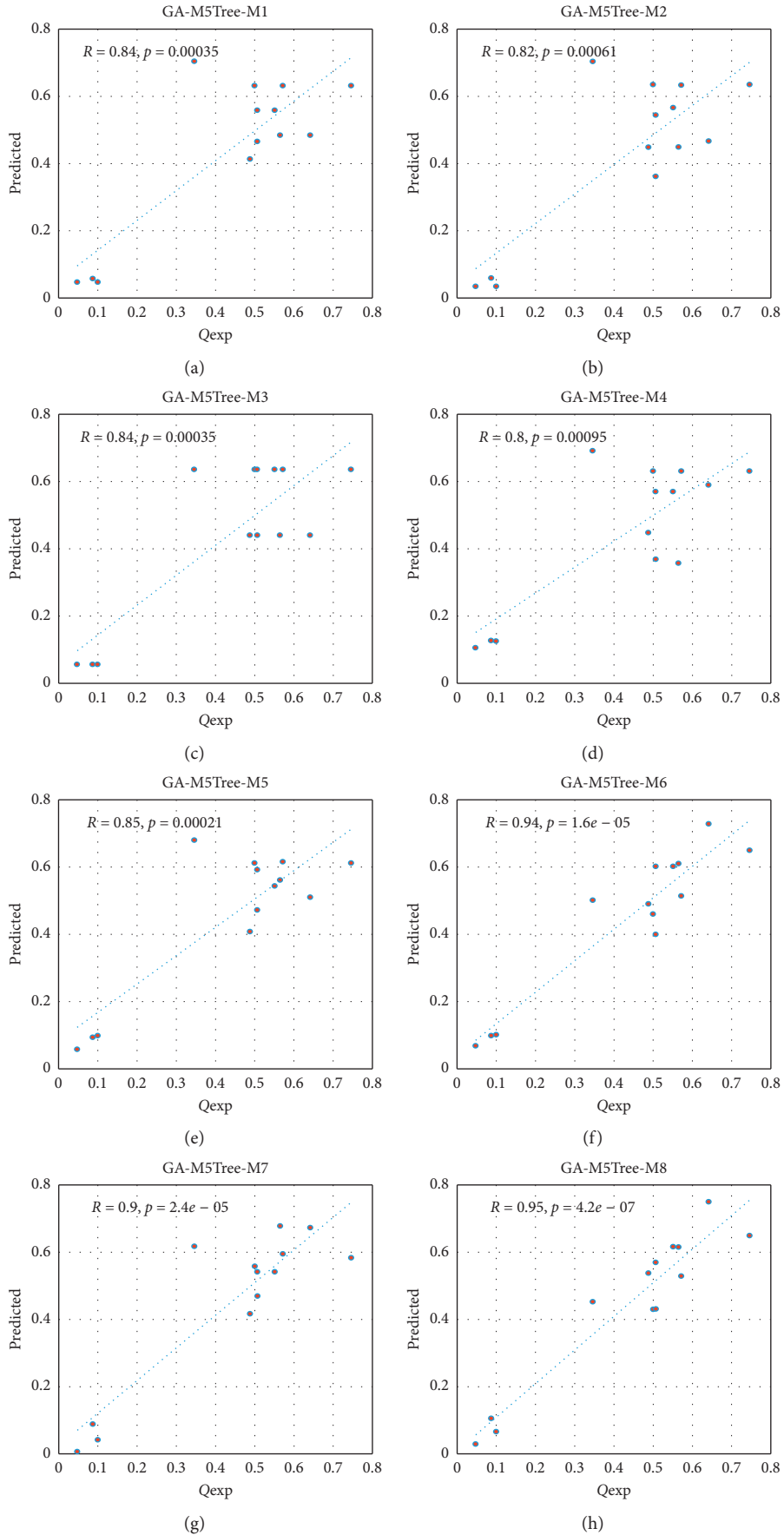
GA-M5Tree-M1
$R = 0.84, p = 0.00035$

GA-M5Tree-M2
$R = 0.82, p = 0.00061$

(a)   (b)

GA-M5Tree-M3
$R = 0.84, p = 0.00035$

GA-M5Tree-M4
$R = 0.8, p = 0.00095$

(c)   (d)

GA-M5Tree-M5
$R = 0.85, p = 0.00021$

GA-M5Tree-M6
$R = 0.94, p = 1.6e - 05$

(e)   (f)

GA-M5Tree-M7
$R = 0.9, p = 2.4e - 05$

GA-M5Tree-M8
$R = 0.95, p = 4.2e - 07$

(g)   (h)

FIGURE 11: The scatter plots graphical presentation over the testing modelling phase: hybrid GA-M5Tree models.
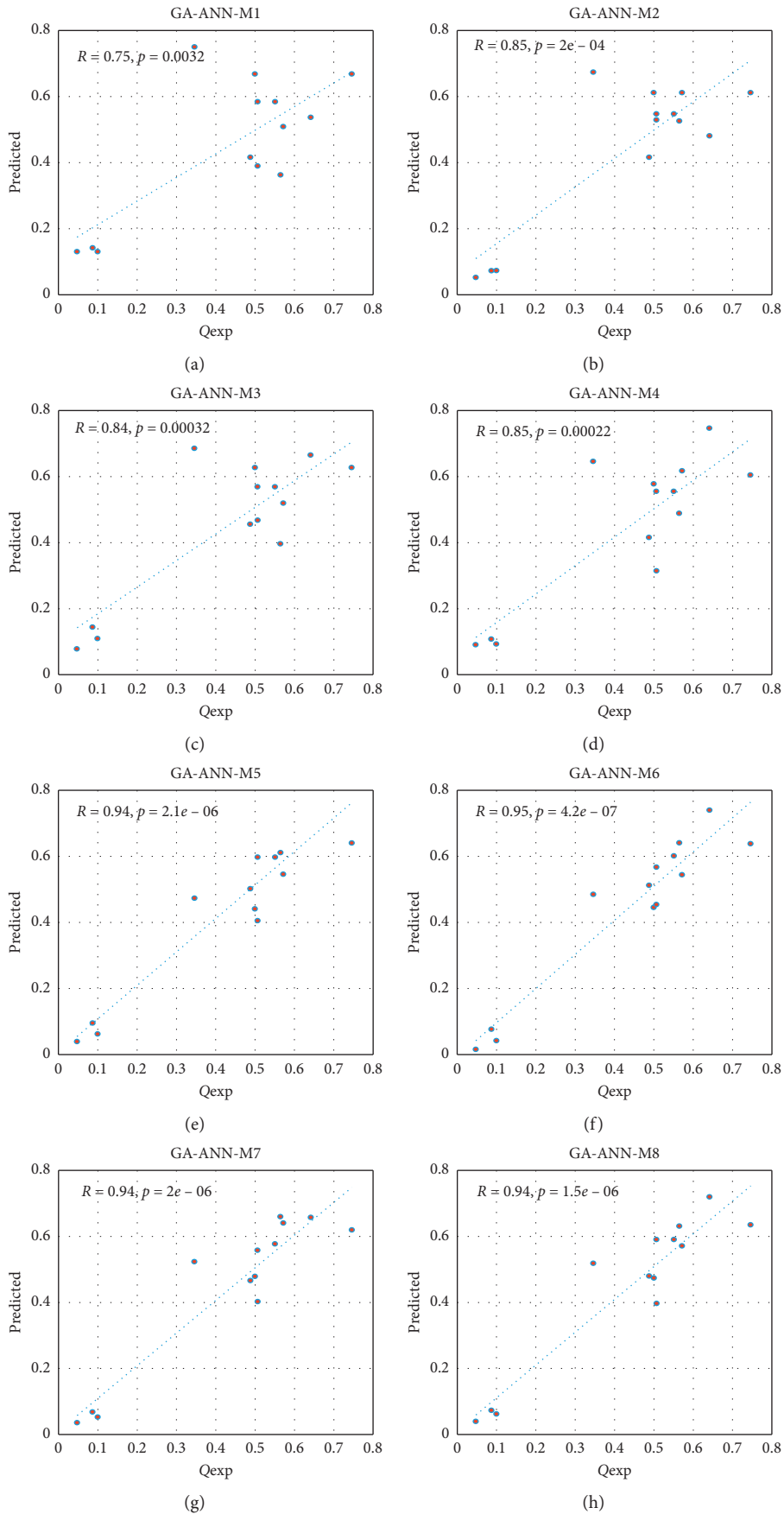
FIGURE 12: The scatter plots graphical presentation over the testing modelling phase: hybrid GA-ANN models.
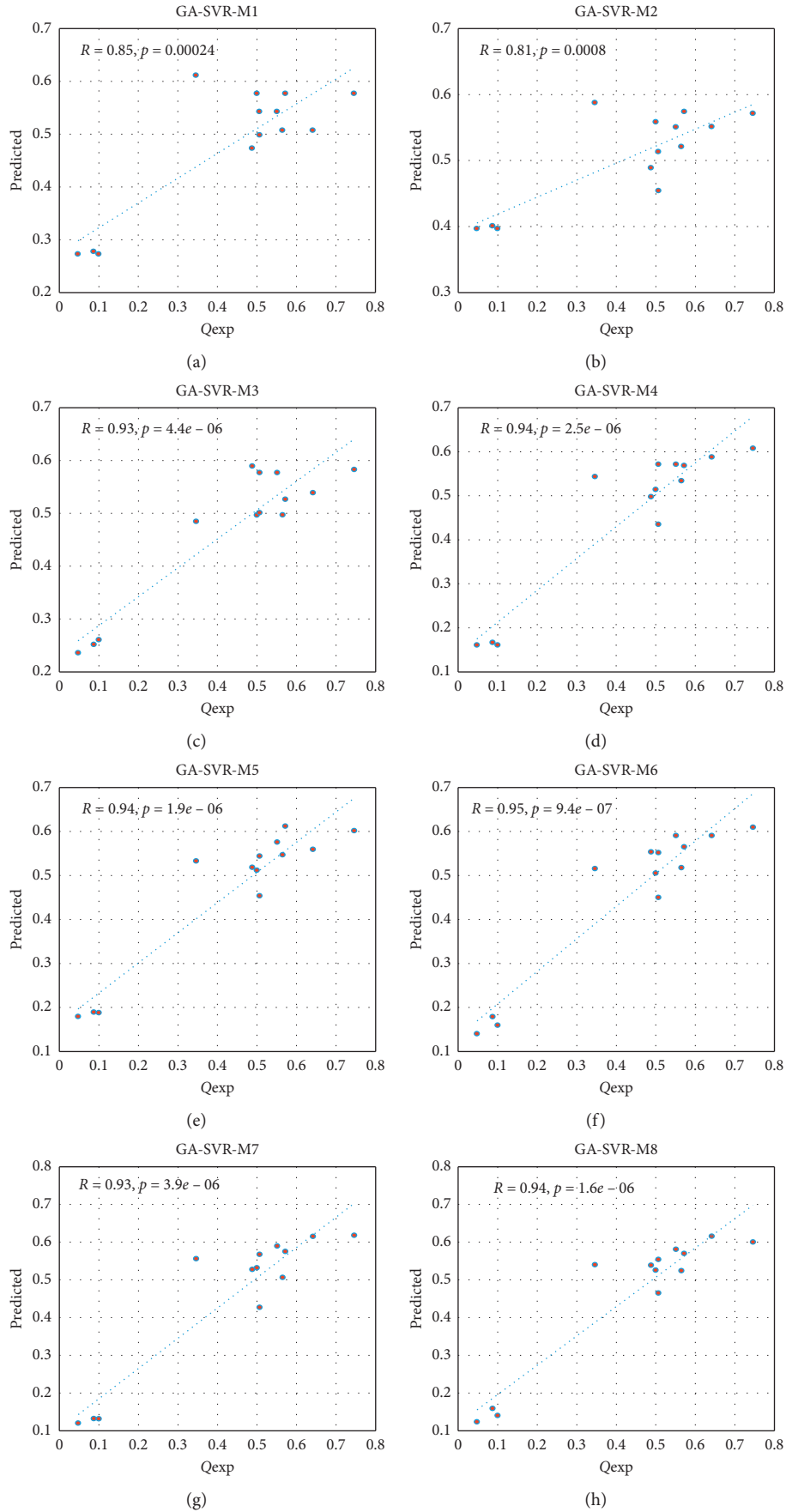
Figure 13: The scatter plots graphical presentation over the testing modelling phase: hybrid GA-SVR models.
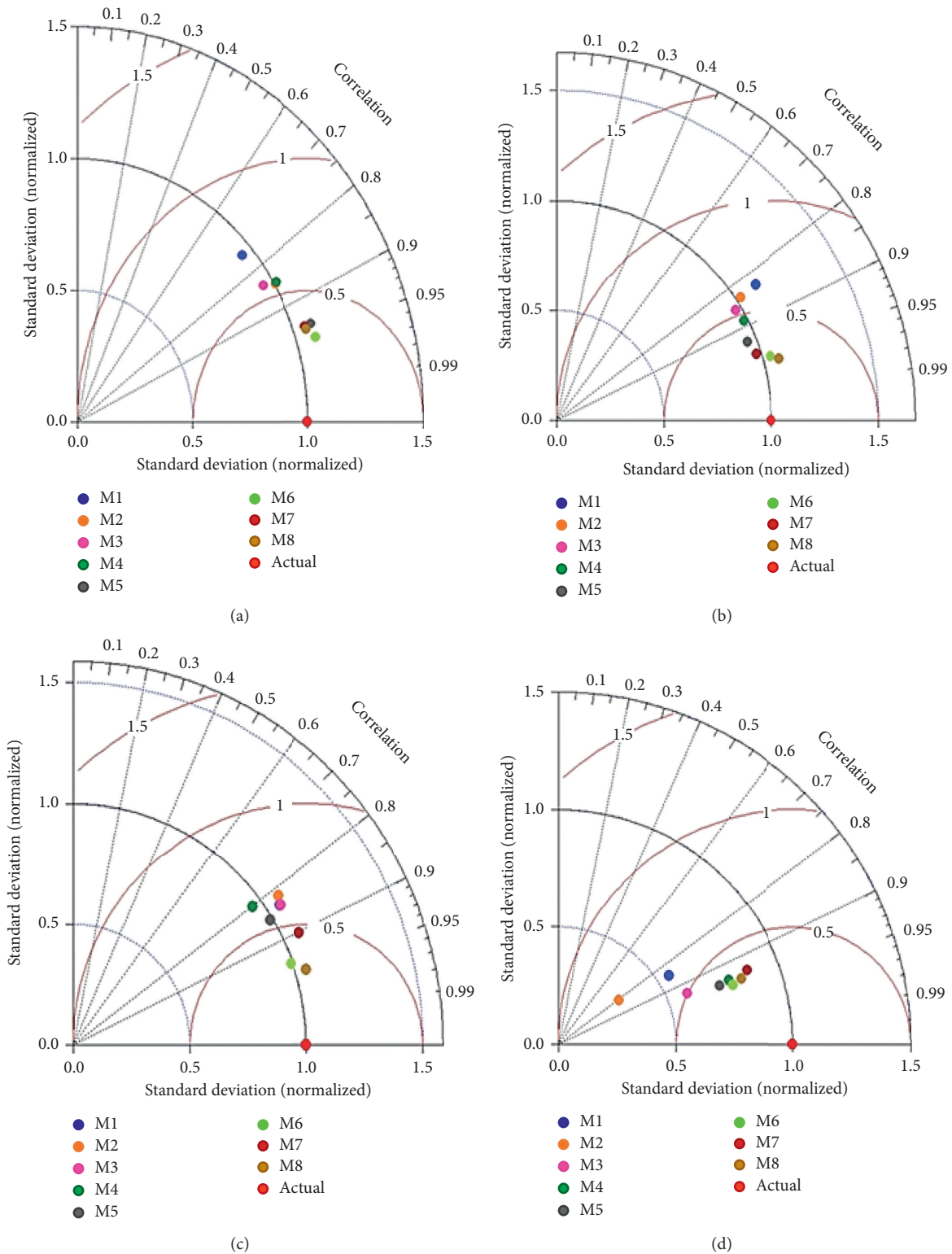
Figure 14: Taylor diagram graphical presentations over the testing phase by different hybrid models. (a) GA-ANN models. (b) GA-DLNN models. (c) GA-M5Tree models. (d) GA-SVR models.

accuracy in comparison with other comparable AI models and yielded more accurate predictions of PRSC values based on all quantitative and visualized assessments. The adopted technique for the selection of the best and most

suitable input parameters for suggested approaches has in general a significant influence on the predictive models' performances by removing the redundant information and hence producing sufficient and clean data to the
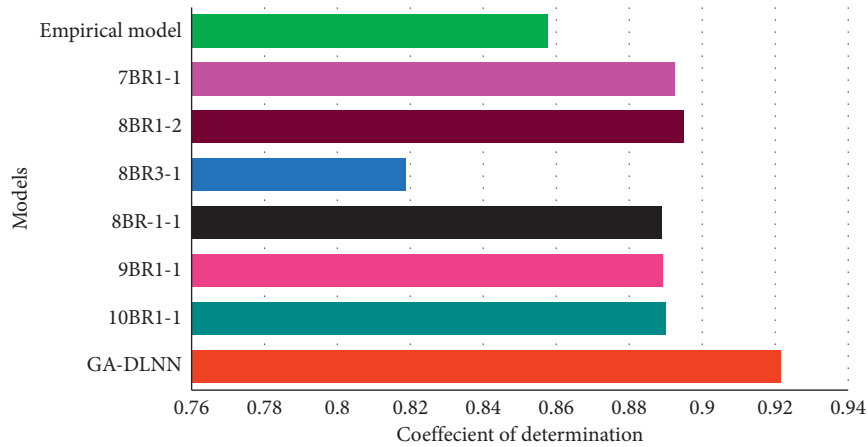
FIGURE 15: Validating the proposed GA-DLNN model against several models developed in the previous studies.

predictive models [67]. Finally, the integration of GA with a deep learning model was found to produce the best model in terms of minimizing the forecasted errors according to assessments carried out using different statistical measures.

In this study, a novel hybrid modelling framework for the Perfobond rib connectors based on several AI methods has been presented. This study performed rigorous sensitivity by considering eight different combination models, in which the input parameter is optimally selected by the effective utilization of GA and the following four AI methods (DLNN, M5Tree, ANN, and SVR) as prediction tools. The use of GA was very successful in selecting the optimal number and combination of the predictor dataset, which considerably reduced the model complexity. Increasing the number of input parameters alone will not help in improving the predictive power of AI models; instead a recognition of appropriate predictor dataset is very important. The DLNN displayed excellent generalization capabilities in understanding the nonlinear relationships between the candidate variables and PRSC. Hybridizing DLNN with GA successful in identifying the 7-input model (M6) was found to be the best for shear strength predictions, i.e., it was found to be successful in identifying the best model with the least number of input parameters with excellent prediction skill for the shear strength predictions of PRSCs. A multitude of statistical performance evaluation measures and graphical representations confirmed the robustness of the DLNN-GA hybrid model for prediction of shear strength of Perfobond rib connectors. This could solve many complexities and problems in structural engineering.

4.3. Validating the Proposed Model against Several Models Conducted in the Previous Research Studies. It is an important aspect to validate the reliability and accuracy of the suggested GA-DLNN model in predicting the PRSC capacity against the recognized researches over the literature studies. Herein, the results obtained by the GA-DLNN model over the testing phase are validated against some predictive models which were described in the literature.

Allahyari et al. [16] developed several models based on ANN approaches trained by Bayesian Regularization (BR) backpropagation algorithm to predict the capacity of PRSC. The main challenge was to probably select the best input combinations; therefore, they adopted classical and statistical methods. The adopted approaches may lead to select redundant information and decrease the efficiency of the predictive models. Subsequently, this study revealed that the best models were GA-DLNN based on 7 input parameters; therefore, we compared the proposed model (GA-DLNN) with several models established in [16] depending on 7 to 10 parameters (i.e., 7-BR1 and 8-BR1). The comparison assessment, as shown in Figure 15, reveals that the proposed model in this study outperformed the comparable models. Moreover, Oguejiofort and Hosaint [2] utilized empirical models to predict the shear strength capacity and yielded good accuracy of prediction with R^2 of 0.8577, as shown in Figure 15. In accordance with these comparative analyses, it can be observed that the adopted models of this study shows better performance prediction and yielded a higher value of R^2 than the comparable models.

5. Conclusions

Accurate prediction of Perfobond Rib Shear Strength is very important in structural engineering sectors. In this investigation, four AI approaches comprising ANN, SVR, DLNN, and M5Tree were developed using 90 experimental samples collected from previous studies. Ten input parameters were used in this current study including concrete compressive strength (f_c), area of concrete dowels (A_D), rib holes number (n), area of cross reinforcement bars and yield stress of reinforcement bars ($A_{tr,r}f_{y,r}$), area of cross reinforcement bars and the tensile strength of cross rebar ($A_{tr}f_y$), area of the connector at the end-bearing zone (Ab), the ratio between the thickness of the concrete slab to connector height (b/h_{sc}), connector height (h_{sc}), the contact area between the connector and the concrete (A_F), and coefficient of end-bearing force (α). The simulated result revealed that the DLNN model achieved a high prediction accuracy and outperformed the comparable

models based on ten input variables. Moreover, the SVR models were the worst predicted models during the training and the testing phases. In order to reduce the number of input parameters, we hybridized the AI models with a bioinspired natural optimization approach called a genetic algorithm (GA) for properly selecting the best input parameters for each AI model separately. Optimal selection of input combinations can effectively reduce the complexity of the model, thereby, obtaining better generalization capabilities, decreasing the computational cost, and increasing the quality of predictions, as well as removing the redundant information. The obtained results showed that the hybridization of AI models with GA importantly improved the prediction accuracy for all predicted models (GA-ANN, GA-M5Tree, and GA-DLNN) except the GA-SVR model. Moreover, the GA-DLNN produced a higher accuracy of performance than other hybrid modelling approaches. The obtained result revealed that the GA-DLNN models required only 7 input parameters to generate the best result accuracy in comparing other hybrid models and classical AI models (DLNN, ANN, SVR, and M5Tree) which were developed based on ten input parameters. Additionally, the outcomes of this study illustrated that removing three input parameters (area of the connector at the end-bearing zone, connector height, and concrete slab thickness) efficiently improved the prediction accuracy of GA-DLNN. The remarkable observation in this study is that is possible to accurately predict the Perfobond rib shear strength with fewer input parameters. This study found that the proper selection of input parameters has a great influence on the performances of AI models. Accordingly, the recommendations for future studies are to intensify the use of GA as a feature selection and combine that algorithm with different AI models to address the most difficult issues related to structural and material engineering.

Abbreviations

PRSC: Perfobond rib shear strength connector
DLNN: Deep learning neural network
SVR: Support vector regression
ANN: Artificial neural network
GA: Genetic algorithm
FE: Finite element method
GP: Genetic programming
ANFIS: Adaptive neuro-fuzzy inference system
LR: Linear regression
(f_c): Concrete compressive strength
(A_D): Area of concrete dowels
(n): Rib holes number
$(A_{tr,r}f_{y,r})$: Area of cross reinforcement bars and yield stress of reinforcement bars
$(A_{tr}f_y)$: Area of cross reinforcement bars and the tensile strength of cross rebar
(Ab): Area of the connector at the end-bearing zone
(b/h_{sc}): The ratio between the thickness of the concrete slab to connector height
(h_{sc}): Connector height
(A_F): The contact area between the connector and concrete
(α): Coefficient of end-bearing force
(MT): Model tree
PART: Partial and regression tree
SDR: Standard deviation reduction
RMSE: Root mean square error
MAE: Mean absolute
MAPE: Mean absolute percentage error
RRMSE: Relative root mean squared error
RE: Relative error
MRE: Mean relative error
NSE: Nash–Sutcliffe efficiency.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] S.-H. Kim and J.-H. Choi, "Experimental study on shear connection in unfilled composite steel grid bridge deck," *Journal of Constructional Steel Research*, vol. 66, no. 11, pp. 1339–1344, 2010.

[2] E. C. Oguejiofor and M. U. Hosain, "Numerical analysis of push-out specimens with perfobond rib connectors," *Computers & Structures*, vol. 62, no. 4, pp. 617–624, 1997.

[3] J. D. C. Vianna, L. F. Costa-Neves, P. C. G. Da S. Vellasco, and S. A. L. De Andrade, "Experimental assessment of Perfobond and T-Perfobond shear connectors' structural response," *Journal of Constructional Steel Research*, vol. 65, no. 2, pp. 408–421, 2009.

[4] J. P. S. Cândido-Martins, L. F. Costa-Neves, and P. C. G. Da Vellasco, "Experimental evaluation of the structural response of Perfobond shear connectors," *Engineering Structures*, vol. 32, no. 8, p. 1976, 1985.

[5] L. F. Costa-Neves, J. P. Figueiredo, P. C. G. D. S. Vellasco, and J. D. C. Vianna, "Perforated shear connectors on composite girders under monotonic loading: an experimental approach," *Engineering Structures*, vol. 56, pp. 721–737, 2013.

[6] J. P. C. Rodrigues and L. Laím, "Behaviour of perfobond shear connectors at high temperatures," *Engineering Structures*, vol. 33, no. 10, pp. 2744–2753, 2011.

[7] A. M. Ibrahim, W. D. Salman, and F. M. Bahlol, "Flexural behavior of concrete composite beams with new steel tube section and different shear connectors," *Tribunal de Justiça do Espírito Santo*, vol. 26, no. 1, pp. 51–61, 2019.

[8] W. Zellner, "Recent designs of composite bridges and a new type of shear connectors," *Composite Construction in Steel and Concrete*, vol. 34, pp. 52–240, 1988.

[9] J. C. Vianna, L. F. Costa Neves, P. Vellasco, and S. A. L. Andrade, "Estudo Comparativo de Conectores de Corte para Estruturas Mistas de Aço e Betão," *Construção Magazine*, vol. 23, pp. 23–30, 2008.

[10] E. C. Oguejiofor and M. U. Hosain, "A parametric study of perfobond rib shear connectors," *Canadian Journal of Civil Engineering*, vol. 21, no. 4, pp. 614–625, 1994.

[11] M. M. Jumaa'h, B. T. Kamil, and O. S. Baghabra, "Mechanical and structural properties of a lightweight concrete with different types of recycling coarse aggregate," *Tikrit Journal of Engineering Sciences*, vol. 26, no. 1, pp. 33–40, 2019.

[12] A. A. Shubbar, Z. S. Al-khafaji, M. S. Nasr, and M. W. Falah, "Using non-destructive tests for evaluating flyover footbridge: case study," *Knowledge-Based Engineering and Sciences*, vol. 1, no. 1, pp. 23–39, 2020.

[13] S. Y. K. Al-Darzi, A. R. Chen, and Y. Q. Liu, "Finite element simulation and parametric studies of perfobond rib connector," *American Journal of Applied Sciences*, vol. 4, no. 3, pp. 122–127, 2007.

[14] J.-H. Ahn, S.-H. Kim, and Y.-J. Jeong, "Shear behaviour of perfobond rib shear connector under static and cyclic loadings," *Magazine of Concrete Research*, vol. 60, no. 5, pp. 347–357, 2008.

[15] S. Zheng, Y. Liu, T. Yoda, and W. Lin, "Parametric study on shear capacity of circular-hole and long-hole perfobond shear connector," *Journal of Constructional Steel Research*, vol. 117, pp. 64–80, 2016.

[16] H. Allahyari, I. M. Nikbin, S. Rahimi, and A. Heidarpour, "A new approach to determine strength of Perfobond rib shear connector in steel-concrete composite structures by employing neural network," *Engineering Structures*, vol. 157, pp. 235–249, 2018.

[17] N. S. Radhi, K. F. Morad, M. H. Hafiz, A. A. Atiyah, and A. Bouaissi, "Optimization of nickel content on some properties of (niti) shape memory alloy," *Knowledge-Based Engineering and Sciences*, vol. 1, no. 1, pp. 40–47, 2020.

[18] Q. Zhang, S. Pei, Z. Cheng, Y. Bao, and Q. Li, "Theoretical and experimental studies of the internal force transfer mechanism of perfobond rib shear connector group," *Journal of Bridge Engineering*, vol. 22, no. 2, 2017.

[19] S.-H. Kim, S. Park, K.-S. Kim, and C.-Y. Jung, "Generalized formulation for shear resistance on Y-type perfobond rib shear connectors," *Journal of Constructional Steel Research*, vol. 128, pp. 245–260, 2017.

[20] M. Abdulrahman and S. Mahmood, "Strength of reinforced reactive powder concrete hollow beams," *Tikrit Journal of Engineering Sciences*, vol. 26, no. 2, pp. 15–22, 2019.

[21] S.-H. Kim, K.-S. Kim, D.-H. Lee, J.-S. Park, and O. Han, "Analysis of the shear behavior of stubby Y-type perfobond rib shear connectors for a composite frame structure," *Materials*, vol. 10, no. 11, p. 1340, 2017.

[22] H. Q. Nguyen, H.-B. Ly, V. Q. Tran, T.-A. Nguyen, T.-T. Le, and B. T. Pham, "Optimization of artificial intelligence system by evolutionary algorithm for prediction of axial capacity of rectangular concrete filled steel tubes under compression," *Materials*, vol. 13, no. 5, p. 1205, 2020.

[23] E. Miranda and J. Suñé, "Memristors for neuromorphic circuits and artificial intelligence applications," *Materials*, vol. 13, no. 4, p. 938, 2020.

[24] B. Savkovic, P. Kovac, B. Dudic et al., "Comparative characterestics of ductile iron and austempered ductile iron modeled by neural network," *Materials*, vol. 12, no. 18, p. 2864, 2019.

[25] A. A. Al-Musawi, "Determination of shear strength of steel fiber RC beams: application of data-intelligence models," *Frontiers of Structural and Civil Engineering*, vol. 11, 2019.

[26] B. Keshtegar, M. Bagheri, and Z. M. Yaseen, "Shear strength of steel fiber-unconfined reinforced concrete beam simulation: application of novel intelligent model," *Composite Structures*, vol. 212, 2019.

[27] A. A. Al-Musawi, A. A. H. Alwanas, S. Q. Salih, Z. H. Ali, M. T. Tran, and Z. M. Yaseen, "Shear strength of SFRCB without stirrups simulation: implementation of hybrid artificial intelligence model," *Engineering with Computers*, vol. 36, no. 1, 2020.

[28] Z. M. Yaseen, R. C. Deo, A. Hilal et al., "Predicting compressive strength of lightweight foamed concrete using extreme learning machine model," *Advances in Engineering Software*, vol. 115, pp. 112–125, 2018.

[29] Z. M. Yaseen, M. T. Tran, S. Kim, T. Bakhshpoori, and R. C. Deo, "Shear strength prediction of steel fiber reinforced concrete beam using hybrid intelligence models: a new approach," *Engineering Structures*, vol. 177, pp. 244–255, 2018.

[30] M. A. Köroğlu, A. Köken, M. H. Arslan, and A. Çevik, "Genetic programming based modeling of shear capacity of composite beams with profiled steel sheeting," *Adv Steel Constr*, vol. 7, no. 2, pp. 157–172, 2011.

[31] T. Ali, *Applications of the ANFIS and LR Models in the Prediction of Shear Connection in Composite beams/Ali Toghroli*, University of Malaya, Kuala Lumpur, Malaysia, 2015.

[32] M.-Y. Cheng and M.-T. Cao, "Evolutionary multivariate adaptive regression splines for estimating shear strength in reinforced-concrete deep beams," *Engineering Applications of Artificial Intelligence*, vol. 28, pp. 86–96, 2014.

[33] V. Le and L. Caracoglia, "A neural network surrogate model for the performance assessment of a vertical structure subjected to non-stationary, tornadic wind loads," *Computers & Structures*, vol. 231, Article ID 106208, 2020.

[34] A. Ashrafian, F. Shokri, M. J. Taheri Amiri, Z. M. Yaseen, and M. Rezaie-Balf, "Compressive strength of Foamed Cellular Lightweight Concrete simulation: new development of hybrid artificial intelligence model," *Construction and Building Materials*, vol. 230, 2020.

[35] C. Wang, D. Shi, and S. Li, "A study on establishing a microstructure-related hardness model with precipitate segmentation using deep learning method materials," 2020.

[36] C. Jaremenko, N. Ravikumar, E. Affronti, M. Merklein, and A. Maier, "Determination of forming limits in sheet metal forming using deep learning materials," 2019.

[37] Z. M. Yaseen, H. A. Afan, and M.-T. Tran, "Beam-column joint shear prediction using hybridized deep learning neural network with genetic algorithm," *In IOP Conference Series: Earth and Environmental Science*, vol. 143, p. 1, 2018.

[38] A. A. H. Alwanas, A. A. Al-Musawi, S. Q. Salih, H. Tao, M. Ali, and Z. M. Yaseen, "Load-carrying capacity and mode failure simulation of beam-column joint connection: application of self-tuning machine learning model," *Engineering Structures*, vol. 194, 2019.

[39] C. A. Jeyasehar and K. Sumangala, "Nondestructive evaluation of prestressed concrete beams using an artificial neural network (ANN) approach," *Structural Health Monitoring*, vol. 5, no. 4, pp. 313–323, 2006.

[40] J.-H. Ahn, C.-G. Lee, J.-H. Won, and S.-H. Kim, "Shear resistance of the perfobond-rib shear connector depending on concrete strength and rib arrangement," *Journal of Constructional Steel Research*, vol. 66, no. 10, pp. 1295–1307, 2010.

[41] A. De Fenza, A. Sorrentino, and P. Vitiello, "Application of artificial neural networks and probability ellipse methods for damage detection using lamb waves," *Composite Structures*, vol. 133, pp. 390–403, 2015.

[42] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of*

the *National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[43] R. Lippmann, "An introduction to computing with neural nets," *IEEE Assp Magazine*, vol. 4, no. 2, pp. 4–22, 1987.

[44] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Internal Representations by Error Propagation*, California Univ San Diego La Jolla Inst for Cognitive Science, San Diego, CL, USA, 1985.

[45] M. Fu, T. Fan, Z. Ding, S. Q. Salih, N. Al-ansari, and Z. M. Yaseen, "Deep learning data-intelligence model based on adjusted forecasting window scale," *Application in Daily Streamflow Simulation*, vol. 8, 2020.

[46] J. R. Quinlan, "Learning with continuous classes," *Machine Learning*, vol. 92, pp. 343–348, 1992.

[47] E. Frank and I. H. Witten, "Generating accurate rule sets without global optimization," 1998.

[48] G. Holmes, M. Hall, and E. Prank, "Generating rule sets from model trees," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*Springer, Berlin, Germany, 1999.

[49] Y. Ayaz, A. F. Kocamaz, and M. B. Karakoç, "Modeling of compressive strength and UPV of high-volume mineral-admixtured concrete using rule-based M5 rule and tree model M5P classifiers," *Construction and Building Materials*, vol. 21, 2015.

[50] V. Jothiprakash and A. S. Kote, "Effect of pruning and smoothing while using M5 model tree technique for reservoir inflow prediction," *Journal of Hydrologic Engineering*, vol. 16, no. 7, pp. 563–574, 2011.

[51] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[52] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 2000.

[53] S. Raghavendra and P. C. Deka, "Support vector machine applications in the field of hydrology: a review," *Applied Soft Computing*, vol. 19, pp. 372–386, 2014.

[54] G. Zhang, "Reinforced concrete deep beam shear strength capacity modelling using an integrative bio-inspired algorithm with an artificial intelligence model," *Engineering with Computers*, vol. 28, 2020.

[55] M. K. AlOmar, M. M. Hameed, N. Al-Ansari, and M. A. AlSaadi, "Data-driven model for the prediction of total dissolved gas: robust artificial intelligence approach," *Advances in Civil Engineering*, vol. 2020, Article ID 6618842, 2020.

[56] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, MIT press, Cambridge, MA, USA, 1992.

[57] A. T. Azar and S. Vaidyanathan, *Computational Intelligence Applications in Modeling and Control*, Springer, Berlin, Germany, 2015.

[58] P. L. Braga, A. L. I. Oliveira, and S. R. L. Meira, "A GA-based feature selection and parameters optimization for support vector regression applied to software effort estimation," *Proceedings of the 2008 ACM Symposium on Applied Computing*, vol. 8, pp. 1788–1792, 2008.

[59] G. Stein, B. Chen, A. S. Wu, and K. A. Hua, "Decision tree classifier for network intrusion detection with GA-based feature selection," *Proceedings of the 43rd Annual Southeast Regional Conference on - ACM-SE*, vol. 43, 2005.

[60] E. Bruderer and J. V. Singh, "Organizational evolution, learning, and selection: a genetic-algorithm-based model," *Academy of Management Journal*, vol. 39, no. 5, pp. 1322–1349, 1996.

[61] J. Yang and V. Honavar, "Feature subset selection using a genetic algorithm," in *Feature Extraction, Construction and Selection*, pp. 117–136, Springer, Berlin, Germany, 1998.

[62] R. K. Yin and G. B. Moore, "The use of advanced technologies in special education," *Journal of Learning Disabilities*, vol. 20, no. 1, pp. 60–63, 1987.

[63] M. S. Podder and C. B. Majumder, "The use of artificial neural network for modelling of phycoremediation of toxic elements As(III) and As(V) from wastewater using Botryococcus braunii," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 155, pp. 130–145, 2016.

[64] M. M. Hameed and M. K. AlOmar, "Prediction of compressive strength of high-performance concrete: hybrid artificial intelligence technique bt - applied computing to support industry," *Innovation and Technology*, vol. 35, pp. 323–335, 2020.

[65] B. Badarloo, A. Kari, and F. Jafari, "Experimental and numerical study to determine the relationship between tensile strength and compressive strength of concrete," *Civil Engineering Journal*, vol. 4, no. 11, pp. 2787–2800, 2018.

[66] H. Tao, A. A. Ewees, A. O. Al-Sulttani et al., "Global solar radiation prediction over North Dakota using air temperature: development of novel hybrid intelligence model," *Energy Reports*, vol. 7, pp. 136–157, 2021.

[67] M. K. Alomar, M. M. Hameed, and M. A. AlSaadi, "Multi hours ahead prediction of surface ozone gas concentration: robust artificial intelligence approach," *Atmospheric Pollution Research*, vol. 11, no. 9, 2020.