Civil Engineering Infrastructures Journal xx,xx

DOI: 10.22059/ceij.2025.390540.2253

Algorithms in Machine Learning for Predicting the Pull-Out Energy of Twin-Twisted Fibers within Cementitious Composites

Hemmatian, A.[1], Jalali, M.[1*] and Naderpour, H.[2,3]

[1]Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

[2] Faculty of Civil Engineering, Semnan University, Semnan, Iran

[3] Department of Civil Engineering, Toronto Metropolitan University, Ontario, Canada

Received: 14/02/2025

Revised: 05/04/2025

Accepted: 21/05/2025

Accepted / Not Edited

ABSTRACT

This research examines the pull-out characteristics of twisted twin fibers within concrete employing advanced soft computing methods. The study highlights the necessity for precise predictive models in fiber-reinforced concrete scenarios, considering the intricate interactions between fibers and their surrounding matrix. Artificial Neural Networks (ANN) and Gene Expression Programming (GEP), were created to forecast the pull-out energy needed for fiber extraction. A detailed dataset comprising 228 experimental samples was used, and various models were trained, including 51 ANN designs and 10 GEP configurations. For the first time, a mathematical formula was established using GEP to estimate pull-out energy, showcasing high accuracy with minimal error margins. The ANN model, especially the one utilizing a log-sigmoid activation function, achieved the highest correlation coefficient (0.995), surpassing the GEP model, which also demonstrated a robust correlation (0.98). Sensitivity analysis indicated that compressive strength had the most substantial effect on pull-out energy, accounting for 18.5% of the observed variance. The results offer a new and precise method for predicting fiber pull-out energy, improving the comprehension of fiber-matrix interactions in cement-based materials. Future investigations should aim to broaden the dataset and examine additional fiber shapes to enhance predictive accuracy.

Keywords: Fiber reinforce concrete; Pull-out energy; Twin-twisted fibers; Artificial neural network; Gene expression programming.

*** Corresponding author:** Meysam Jalali, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran. mjalali@shahroodut.ac.ir

Accepted / Not Edited

1. Introduction

Concrete is a widely used building material consisting of cement (typically Portland cement), sand, crushed stone or gravel (aggregates), chemical admixtures, and water (Banchhor, 2018). Due to its extensive use globally, large quantities of various types of concrete are typically produced. As a result, many researchers have focused on investigating its engineering properties. Conventional concrete is recognized as a brittle material, failing without warning. This brittleness results in remarkably low flexural strength. However, the inclusion of steel fibers in concrete transforms its brittle nature into a more ductile one . Fiber-reinforced concrete (FRC) is composed of cement, mortar, and uniformly dispersed fibers that are randomly oriented and discontinuous (Anita Jessie and Santhi, 2019). The interaction between steel fibers and the concrete matrix plays a crucial role in determining the performance of fiber-reinforced concrete. Consequently, the single fiber pull-out test is commonly employed to study the behavior of fiber-reinforced composites and evaluate fiber-matrix interactions. The force-slip curves derived from the single pull-out test help in determining key parameters such as maximum pull-out force, maximum slip, pull-out energy, bond strength, breaking force, and more. Recently, there has been an increasing focus on experimental research in this domain, highlighting the need for utilizing soft computing methods. In recent years, a growing number of studies have been dedicated to FRC experiments, emphasizing the importance of employing soft computing techniques. Therefore, this numerical investigation was conducted to predict the pull-out energy of twin-twist fibers from cement-based matrices.

The extraction energy of steel fibers in cement-based composites plays a crucial role in assessing the mechanical properties and interfacial interactions of these materials. This energy is affected by various elements, such as the bond between the fiber and matrix, the type of fiber used, the strength of the matrix, and the inclusion of cement replacement materials (Le and Kim, 2019). In steel fiber-reinforced composites, the strength of the matrix plays a pivotal role in determining the pullout performance of the embedded fibers. In particular, an increase in matrix strength is typically associated with an enhancement in both the pullout load and the energy required for pullout (Jamee et al., 2022). The morphology and geometric configuration of steel fibers are pivotal in influencing the energy required for their extraction. Empirical studies have demonstrated that fibers characterized by hooked extremities exhibit superior pullout resistance in comparison to their straight counterparts, as the hooks facilitate an enhanced mechanical interlock with the cement matrix. The optimization of fiber morphology has the potential to markedly improve the residual tensile strength and pullout resistance,

consequently augmenting the overall energy required for extraction (Wang et al., 2024). Sengul (2018) investigated the use of waste steel fibers recovered from scrap tires in Slurry Infiltrated Fiber Concrete (SIFCON) and demonstrated their effectiveness in enhancing mechanical properties. The study found that these hybrid fibers significantly improved flexural strength, residual strength, and toughness, confirming their potential for sustainable construction applications. Abdallah et al. (2016) proposed an analytical model to predict how the geometry of hooked-end steel fibers influences pull-out behavior, with validation against experimental results from single fiber pull-out tests. Lee and Kighuta (2017) conducted experiments on the pull-out behavior of fibers, comparing straight and hooked-end fibers to investigate the impact of the twin-twist effect on pull-out resistance. Their findings demonstrated that twin-twisted fibers exhibited significantly higher maximum and residual pullout loads compared to straight and hooked-end fibers. Extensive research has been conducted by Yoo et al. (2020) on the behavior of single-fiber pullout at different embedment lengths. In order to achieve this, experiments were carried out using two types of steel fibers in ultra-high-performance concrete (UHPC) at three embedment lengths and two inclination angles. At an embedment length of 15 mm, it was observed that the half-hooked fiber experienced premature rupture, resulting in lower pullout energy and equivalent bond strength compared to the straight fiber. Soft computing methods have been widely employed in numerous studies to predict the behavior and mechanical characteristics of concrete. Dantas et al. (2013) utilized ANNs to predict the compressive strength of concrete containing construction and demolition waste. They employed a dataset of 1178 data points, demonstrating ANN's promising application for forecasting compressive strength over different curing periods. Adaptive neurofuzzy inference system (ANFIS) are investigated for Earthquake magnitude prediction by (Mirrashid, 2014). The seismic data from 1950 to 2013 in the area located at 2°E longitude and 4°N latitude in Iran was analyzed. Three different algorithms, namely grid partition (GP), subtractive clustering (SC), and fuzzy C-means (FCM), were employed to create models based on ANFIS structure. It was found that the ANFIS-FCM model demonstrated a remarkable accuracy in predicting earthquake magnitudes. Cascardi et al. (2017) developed an ANN-based model to predict the compressive strength of FRP-confined concrete circular columns, outperforming existing design models. Bu et al. (2021) used ANNs to predict the compressive strength of recycled aggregate concrete. Khan et al. (2021) employed gene expression programming (GEP) to predict the compressive strength of geopolymer concrete produced with fly ash, demonstrating superior predictive capabilities over traditional methods. Ahmad et al. (2021) applied supervised machine learning, GEP, and ANN techniques to predict the compressive strength of concrete using recycled coarse aggregates, with the GEP model showing superior prediction accuracy. Al-hashem et al. (2022) utilized ANN and GEP models to forecast the compressive strength of concrete containing fly ash and rice husk ash. The ANN model outperformed the GEP model. Huang et al. (2024) employed Evolutionary Algorithms to forecast the strength of Geopolymer Concrete. A total of seven unique input variables were utilized for the modeling objective. In this study, the prediction models based on GEP demonstrated superior performance, accuracy, and generalization ability when compared to those based on MEP. The models based on GEP showed greater correlation coefficients (R) for forecasting the compressive and split tensile strengths, yielding values of 0.89 and 0.87, respectively. Hemmatian et al. (2023) conducted a study utilizing soft computing techniques to examine the pull-out behavior of straight, hooked, and spiral steel fibers in fiber-reinforced cementitious composites. They developed and trained Artificial Neural Network (ANN) models using an extensive dataset comprising 382 experimental results. The results revealed that the geometric configuration of the fibers, along with their tensile strength, emerged as the most significant factors among the eight input variables influencing the outputs. Taffese et al. (2024) utilized machine learning techniques to forecast the aging factor of concrete. Their research presents seven distinct algorithms aimed at predicting these aging factors. Significantly, the LightGBM algorithm demonstrated superior performance in Scenario III, achieving an impressive mean

Accepted / Not Edited

absolute error (MAE) of 0.110, a mean squared error (MSE) of 0.018, a root mean squared error (RMSE) of 0.133, and a coefficient of determination (R2) of 0.818. Shahrokhishahraki et al. (2024) employed machine learning techniques, including Artificial Neural Networks and regression algorithms, to predict cement content for concrete targeting full compressive strength at 90 days rather than the conventional 28 days. Their study achieved prediction accuracies of 94% and 90% for machine learning and deep learning methods, respectively, with the Elastic Net algorithm demonstrating superior performance. A case study on a mid-sized reinforced concrete structure showed that optimizing for a 90-day strength target reduced cement usage and carbon emissions by approximately 10%, highlighting significant sustainability benefits. Hoang and Tran (2024) compared the performance of prominent machine learning models, demonstrating their potential to reduce the time and effort required for laboratory testing. Utilizing 11 historical datasets and conducting 20 independent data sampling runs, their study evaluated model generalization. Their findings highlighted that gradient boosting machines, particularly the extreme gradient boosting model, achieved superior performance across five datasets.

Rahman Sobuz et al. (2025) investigated graphene nano-engineered hybrid fiber reinforced concrete (GNFRC) with 0–0.06% graphene and 0–0.5% hybrid fibers, achieving up to 22%, 36%, and 14% increases in compressive strength, tensile strength, and elastic modulus, respectively, at 90 days. Microstructural analysis revealed improved hydration products as the primary cause. Machine learning models, particularly XGBoost (R^2 = 0.994), accurately predicted compressive strength using 438 mix proportions, with cement and curing age identified as key influencers via SHAP and PDP analyses. These findings highlight GNFRC's potential for durable, high-performance concrete composites. Soft computing techniques have found extensive application across various fields of engineering, demonstrating their versatility and effectiveness in solving complex problems (Chandra Nayak et al., 2024; Ghorbani et al., 2024; Ismael Jaf et al., 2024; Mastan et al., 2023; Nasir et al., 2020; Nguyen et al., 2023; Shahrokhishahraki et al., 2024).This study presents a comprehensive investigation into the pull-out behavior of twisted twin steel fibers from concrete matrices. Soft computing models, particularly ANN and GEP, were developed to predict the energy required for fiber extraction. A total of 51 ANN-based models and 10 GEP models were created, and for the first time, a mathematical correlation was proposed to estimate pull-out energy, showing low error rates and high accuracy. The models were compared to experimental data, indicating significant precision in predictions. This research represents a new phase in the study of this emerging fiber type.

This document is organized as follows: Section 2 outlines the method for assessing bond strength and pull-out energy in cement composites. Section 3 introduces ANNs for predicting pull-out energy, detailing their structure and training. Section 4 presents GEP as a modeling technique using evolutionary algorithms. Section 5 describes the dataset for training and testing, including fiber type and matrix composition. Section 6 discusses proposed ANN models for pull-out energy prediction. Section 7 compares ANN models for accuracy and efficiency. Section 8 introduces GEP models for pull-out energy prediction. Section 9 presents GEP results, evaluating prediction accuracy. Section 10 compares ANN and GEP predictions with experimental data. Finally, Section 11 summarizes the findings, highlighting the effectiveness of ANN and GEP models.

2. Materials and methods

2.1 Fiber pull-out tests

A detailed series of experiments has been undertaken to explore the different anchoring mechanisms exhibited by fibers when integrated into a cement matrix. These investigations utilize specially crafted specimens that feature a matrix body with a deliberate discontinuity extending across its cross-section. This gap is effectively

Accepted / Not Edited

bridged by one or more fibers, which are crucial for preserving the structural integrity and stability of the composite material. The testing procedure involves securely anchoring one end of the specimen to ensure the reliability of the results. Following this setup, a controlled load is progressively applied to the specimen, leading to a separation between the matrix components and the embedded fibers. In Figure 1, various kinds of fibers are shown.

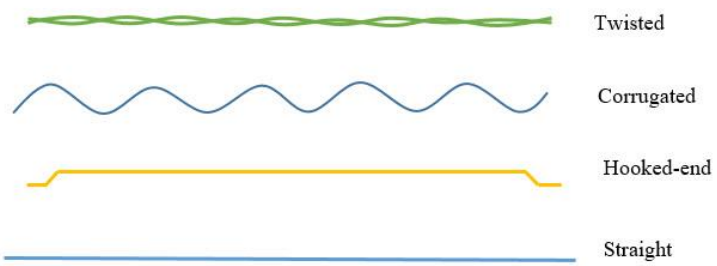


Fig. 1. Different types of fibers

Throughout the experimental process, it is crucial to precisely record both the magnitude of the applied load and the degree of separation between the fibers and the cement matrix. The degree of separation refers to the amount of debonding or separation occurring between the fibers and the matrix during the pullout process. This parameter is measured by monitoring the displacement of the fiber relative to the matrix using displacement transducers (LVDT) alongside the applied load. This comprehensive documentation is essential for gaining a deeper understanding of the performance characteristics and behavior of the fibers within the cement matrix under various stress conditions. By analyzing these parameters, researchers can gain valuable insights into the effectiveness of the anchoring mechanisms and the overall durability of the fiber-reinforced cementitious materials. Figure 2 illustrates the structure of a fiber pull-out test featuring a dog-bone sample.

Accepted / Not Euited

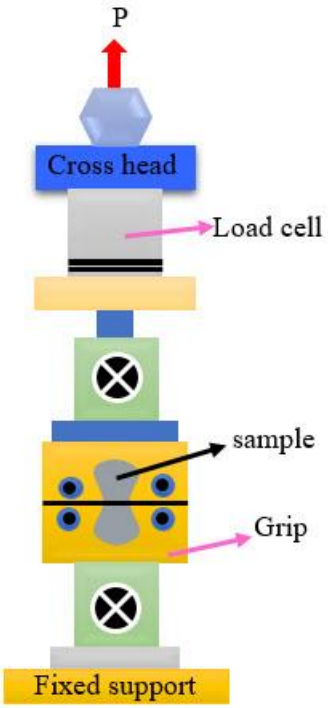


Fig. 2. Fiber pull-out test setup with a dog-bone sample

2.2 Artificial neural networks

The modeling of neural networks is based on the biological processes of the human brain (Khademi et al., 2016). An Artificial Neural Network (ANN) is a machine learning algorithm that aims to replicate the intricate workings of the human nervous system. Its primary objective is to process experimental data through various tasks, including classification, clustering, regression, and prediction (Sobhani et al., 2010). The multi-layer feed-forward perception network is a typical type of artificial neural network (ANN). Figure 3 shows an example of a multilayer perceptron network with two hidden layers.Its topological structure comprises an input layer, one or more hidden layers, and an output layer. Within these layers, a significant number of neurons are distributed (Naderpour et al., 2019; Qu et al., 2018).

Accepted / Not Edited

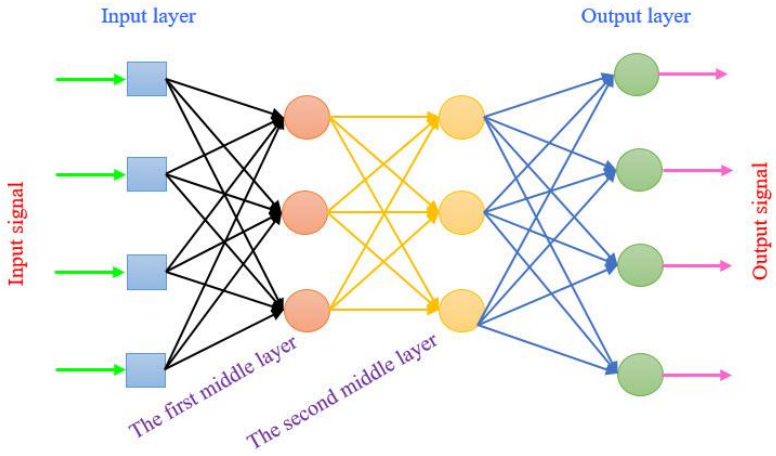


Fig. 3. A multilayer perceptron with two hidden layers

Within the network, every layer's neurons are interconnected with the subsequent layer, while no interconnection exists between neurons within the same layer. The transmission from the input to the output layer is achieved by modifying the weights and thresholds across the layers. If the problem is learnable, a consistent set of weights will be established, leading to a favorable outcome. The backpropagation neural network is a fundamental and crucial neural network that employs nonlinear training methods for multi-layer feed-forward networks. It utilizes a supervised learning approach to minimize the discrepancies between the calculated and experimental values by iteratively adjusting the weights in small increments (Vidivelli and Jayaranjini, 2016). Similar successful applications of Artificial Neural Networks can be found in recent studies. For example, Al-Zwainy et al. (2021) effectively predicted the residual strength of self-consolidating concrete (SCC) exposed to elevated temperatures using a backpropagation neural network. In another study, (Al-Zwainy, F. M., Al-khazrajy, M. G., Hussein, N. M., Mohamed, S., Sarhan, M. M., Al-Musawi, T. J., and Hayder, 2024) applied ANN models to forecast earned value indicators during the early phases of bridge projects, demonstrating the model's strong predictive capabilities across various construction-related parameters.

An artificial neural network (ANN) operates by using a set of examples from a training database as input, a learning algorithm to adjust the weights, and an activation function to produce an output. Changing the connection weight between neurons will result in an alteration of the network's output relative to its input. The process of fine-tuning the connection weights by exposing the network repeatedly to known input-output data is referred to as training. The error back-propagation learning method is recognized as the most popular and effective training technique.

2.3 Gene expression programming

In summary, genetic algorithms serve as a powerful tool for tackling intricate problems by simulating evolutionary mechanisms. Through the iterative process of selection and reproduction, these algorithms refine potential solutions over successive generations. Ultimately, they provide a practical approach to problem-solving in situations where traditional methods may fall short, making them invaluable in various fields of research and application (Scott M. Thede, 2014).

Accepted / Not Edited

The Gene expression programming (GEP) is recognized as an exceptionally effective instrument within the domain of evolutionary computation. It presents a unique hybrid mechanism that possesses the remarkable ability to adjust and cater to an extensive array of applications. This dual foundation of GEP not only enhances the overall functionality of the algorithm but also significantly expands its range of applicability across numerous fields that demand advanced and sophisticated problem-solving techniques. As a result, GEP proves to be an invaluable resource for tackling complex challenges in various disciplines (Ferreira, 2001). In Genetic Expression Programming , every individual gene is composed of two distinct categories of symbols: one category includes fixed-length variables, while the other consists of constants that serve as terminal sets. Additionally, there are arithmetic operations that function as the function sets. A fundamental characteristic that defines GEP is the generation of chromosomes, which possess the remarkable ability to represent any parse tree. This is accomplished through the use of the Karva language, which facilitates the reading and expression of the information that is encoded within these chromosomes. Following this, the chromosomes undergo a transformation process, resulting in the formation of branched structures known as expression trees(ET) (Gholampour et al., 2017).The proposed GEP algorithm is depicted through its workflow diagram in Figure 4.

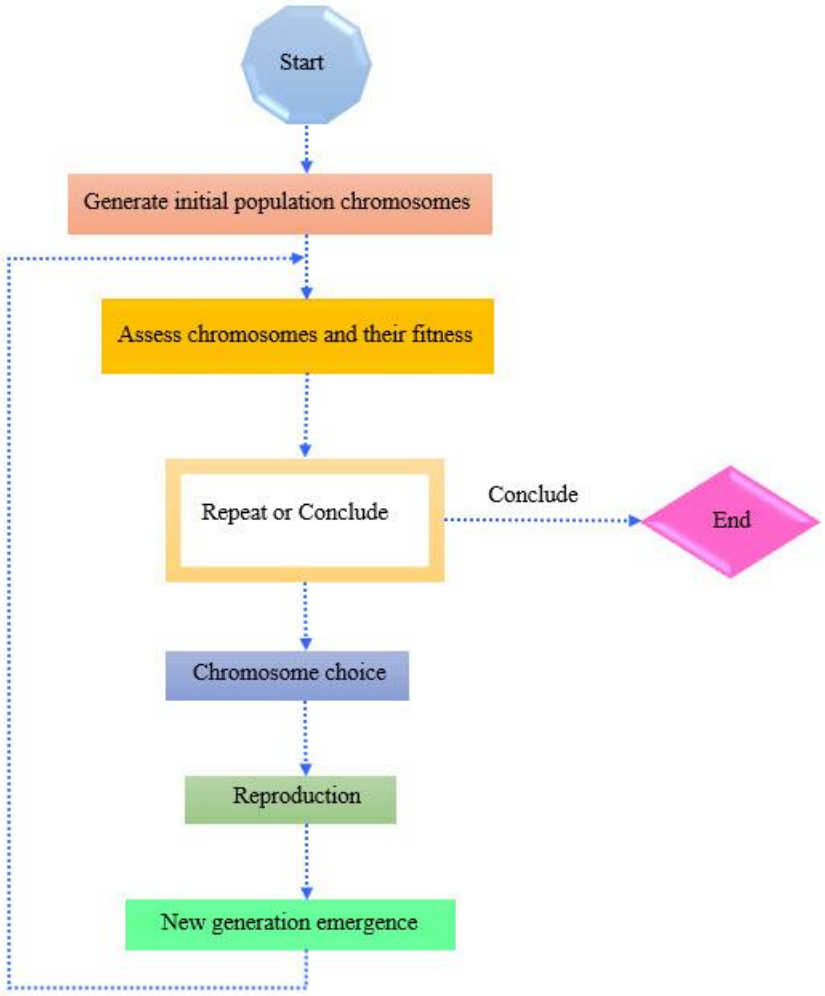


Fig. 4. The proposed GEP algorithm

Accepted

The cycle continues with the newly created individuals over multiple generations until a satisfactory solution is achieved. Throughout this iterative process, genetic diversity within the population is maintained by implementing various genetic operations, including mutation, rotation, and crossover, on the selected individuals. These operations are crucial for introducing variability and enhancing the potential for discovering optimal solutions.

2.4 Experimental dataset

In this paper, a total of 228 experimental datasets were gathered and utilized for training and testing the models. The fiber pull-out test specimens of this data set were cylindrical. The input data are: fiber diameter, number of fiber twists per 10 mm length, embedment length in the matrix, fiber length-to-diameter ratio, water-to-cement ratio, and compressive strength of concrete, while the target is the pull-out energy
A detailed presentation of the 228 experimental test results, including all relevant parameters, is provided in Appendix A1**.**

A summary of the selected database is shown in the Table 1. The input parameters for distribution in relation to the output for each dataset are illustrated in Figure 5.

Table 1. The statistical attributes associated with experimental findings

Statistical indicators	**d_f (mm)**	**n_{10}**	**L_e (mm)**	**l_f/d_f**	**W/C**	**f_c` (Mpa)**	**W_p (N.mm)**
Min	0.41	2	10	69.81	0.22	47.7	919.25
Max	0.57	4	20	97.56	0.42	109.82	6323.58
Mean	0.48	2.95	14.54	84.70	0.30	76.43	2773.10
Standard deviation	0.066	0.81	4.02	11.85	0.09	25.74	1225.31
Coefficient of variation	0.137	0.274	0.277	0.140	0.303	0.337	0.442

Prior to initiating the training phase, the chosen dataset, encompassing both the input variables and the corresponding target outcomes, undergoes a normalization or scaling process that is guided by the interrelationships present within the data. This critical procedure seeks to minimize the variations among the values of the examined variables, especially those that function on disparate scales, thus establishing a more consistent foundation for analysis.
In the course of this research study, we employ Eq. (1) as a method to normalize the data we are working with. This particular relationship allows us to scale all of the data points so that they fall within a specified range, specifically between the values of 0.1 and 0.9.

Accepted / Not Edited

$$X_n = 0.8\frac{X_r - X_{min}}{X_{max} - X_{min}} + 0.1 \quad (1)$$

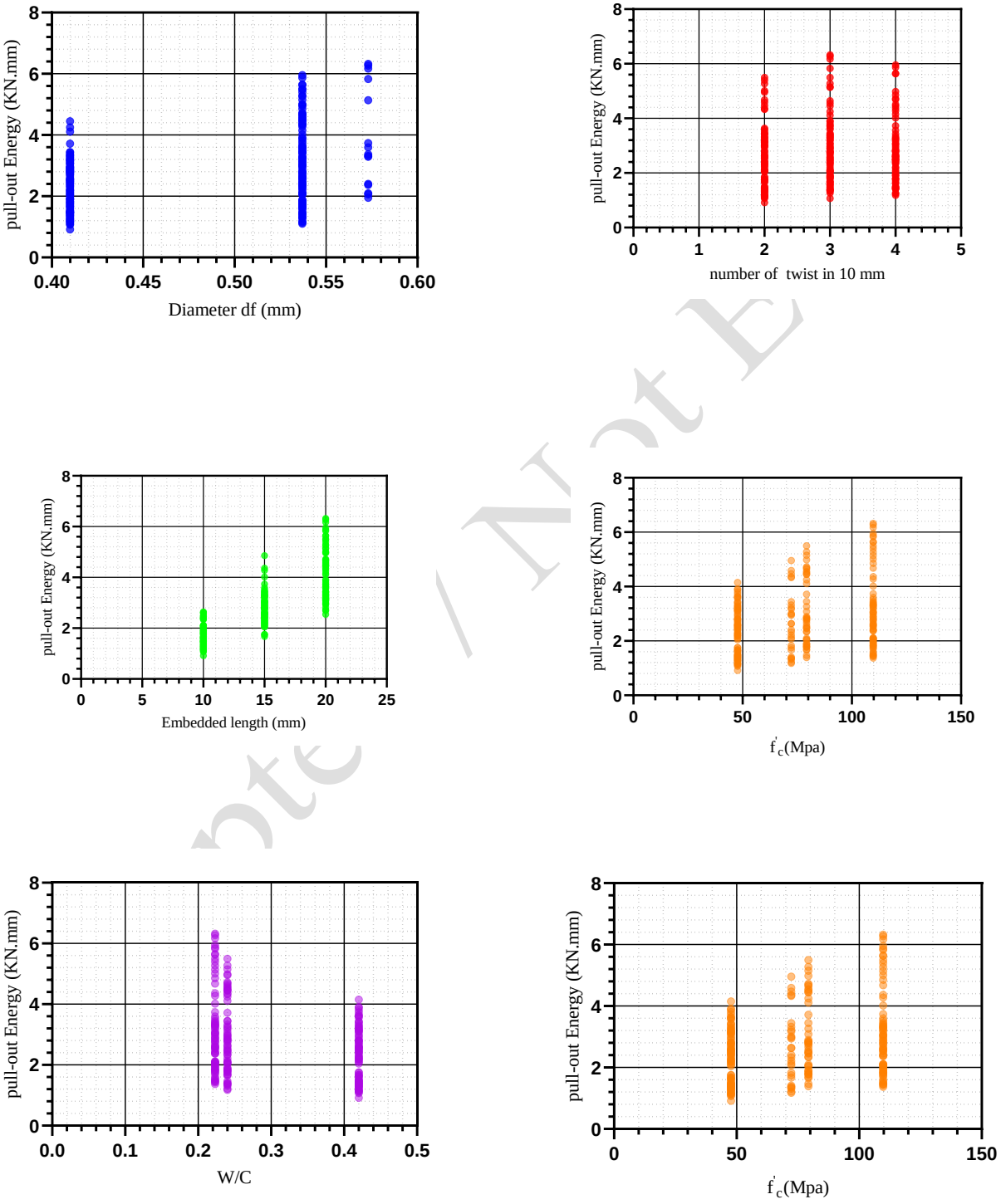


Fig. 5. The allocation of the output to the input parameters

The investigation undertaken in this study involved the examination of fibers through a method that entailed the twisting of two straight fibers using a direct current (DC) motor, which was calibrated to function at a speed of 10 revolutions per minute (rpm). A custom-designed steel disk, measuring 3 inches in diameter and half a centimeter in thickness, was employed to facilitate this twisting process. This disk featured two precisely positioned holes that enabled a secure and direct attachment of the fibers to the motor.Subsequently, the ends of the fibers were affixed to a dynamometer with a maximum capacity of 150 Newtons, which was strategically placed 600 millimeters away from the motor. This arrangement allowed for accurate measurement and analysis of the forces exerted during the twisting process, thereby contributing valuable data to the overall research objectives(Ataee et al., 2024). Figure 6 illustrates an overall perspective of the development of this fiber.

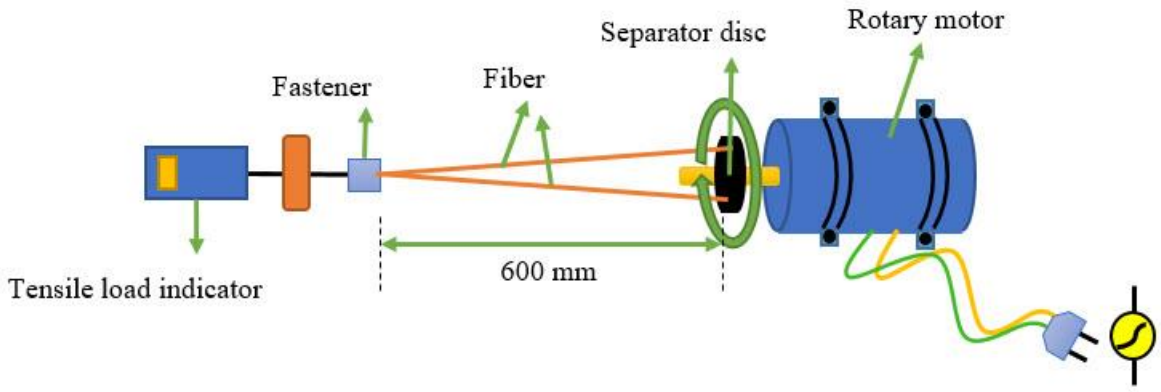


Fig. 6. The perspective of the development of twin-twisted fiber (Ataee et al., 2024)

The straight fibers that are utilized for the production of this particular kind of specialized fibers are constructed from steel, featuring diameters measuring 0.38 mm and 0.29 mm, along with an elasticity modulus of 200 GPa. Following the specified Eq. (2), twisted steel fibers have been manufactured, possessing diameters of 0.537 mm and 0.41 mm.

$$D=\sqrt{2*d_{sf}^{2}} \tag{2}$$

These twisted fibers have been specifically prepared for the purpose of creating samples that are designed for pull-out testing, as illustrated in Figure 7.

Accepted - Edited

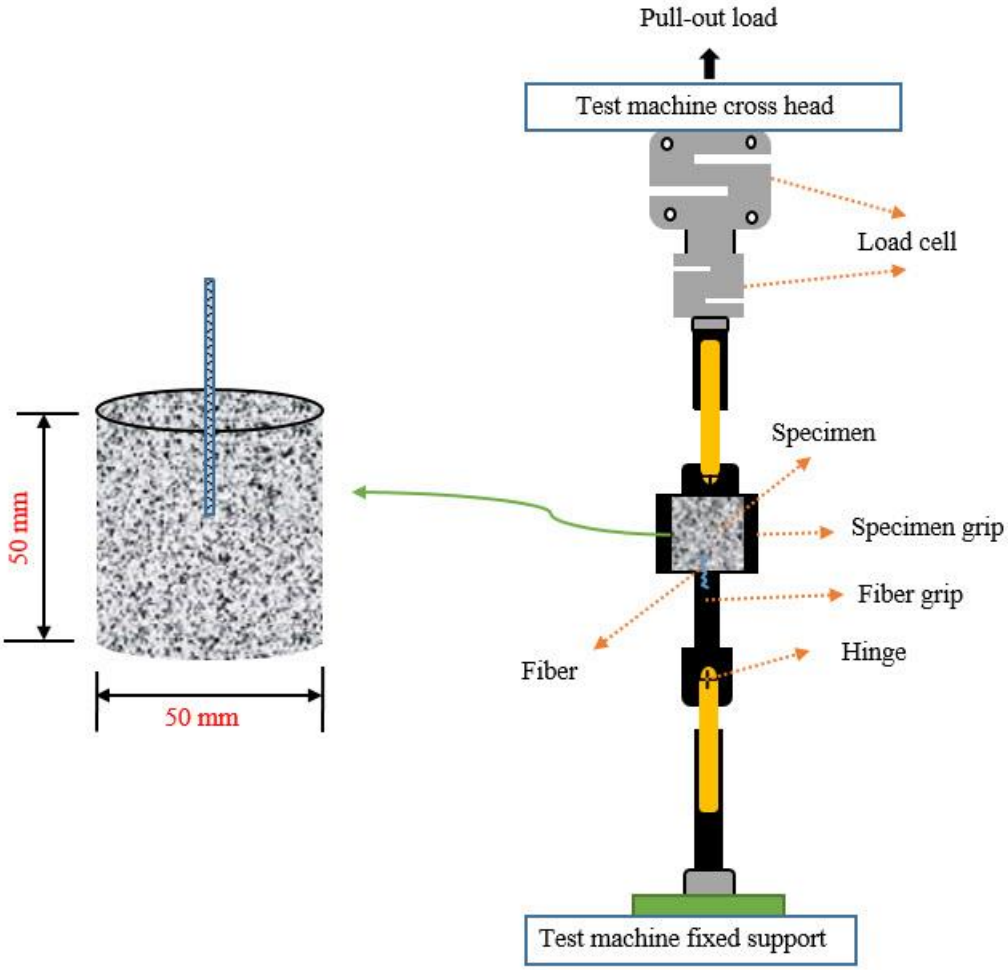


Fig. 7. Modeling the specimen and conducting the fiber pull-out test (Ataee et al., 2024)

3. Proposed ANN Models

The architecture of the suggested neural network is depicted in Figure 8. In this specific section, a single hidden layer was employed. The number of neurons in the hidden layer ranged from 4 to 20. The configuration utilized in this study was named **ANN6-n-1**, where the first digit represents the number of input nodes, 'n' represents the number of hidden neurons, and the third digit represents the number of output nodes.

Accepted

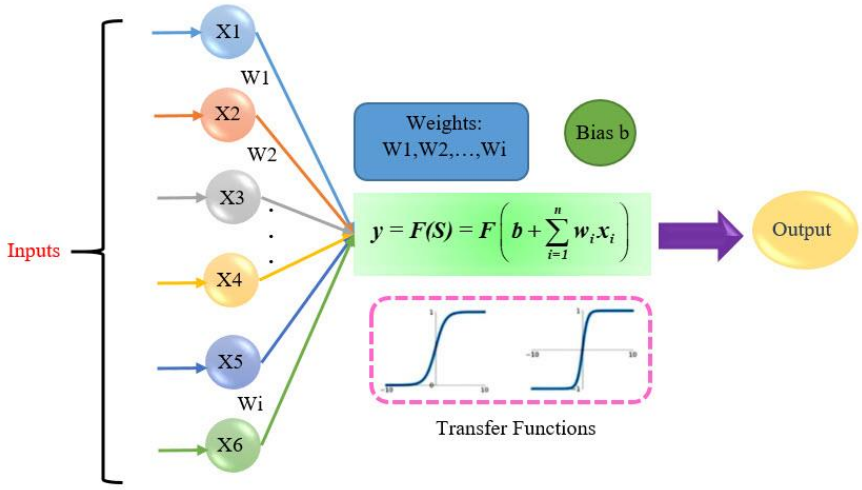


Fig. 8. Modeling the specimen and conducting the fiber pull-out test

This section presents an analysis of three distinct models of artificial neural networks: (1) the neural network utilizing the Tansig translate function, (2) the neural network employing the Pierlin translate function, and (3) the neural network based on the Logsig translate function. For each model, a total of 17 individual networks have been developed by varying the number of neurons within the range of 4 to 20.

The critical parameters for assessing the pull-out energy of twin fibers encompass several factors, including the diameter of the fibers (d_f), the number of twists per 10 mm of fiber length (n_{10}), the length of fiber embedment within the matrix (L_e), the ratio of fiber length to diameter (d_f/l_f), the water-to-cement ratio (W/C), and the compressive strength of the concrete (f'_c). These parameters are essential for understanding the mechanical interactions and performance characteristics of the fiber-reinforced composite materials.

3.1 Method 1: neural network (ANN6-18-Tansig)

In the current research endeavor, the Levenberg–Marquardt algorithm is employed with the back-propagation neural network methodology to develop an Artificial Neural Network (ANN). This particular strategy is grounded in empirical curves and regression equations. Like other statistical computing techniques, the dataset undergoes normalization as indicated by Eq. (1). Among the diverse array of networks constructed during this study, the one that featured a hidden layer comprising eighteen nodes exhibited the most promising results, as illustrated in Figure 2. The transfer functions that were utilized for this specific ANN were of the tangent sigmoid variety, commonly referred to as Tansig. As a result of these characteristics, this network has been designated as ANN6-18-Tansig.

Accepted / Not Edited

The findings derived from the neural network that has been presented indicate that this particular network demonstrates a commendable level of performance when it comes to predicting various values. This effectiveness is clearly illustrated in the regression graph depicted in Figure 9. Additionally, the process of learning that the network undergoes is also represented in Figure 10, providing further insight into its operational capabilities. Furthermore, Figure 11 illustrates both the performance metrics and the current training status of the neural network that has been proposed in this study.

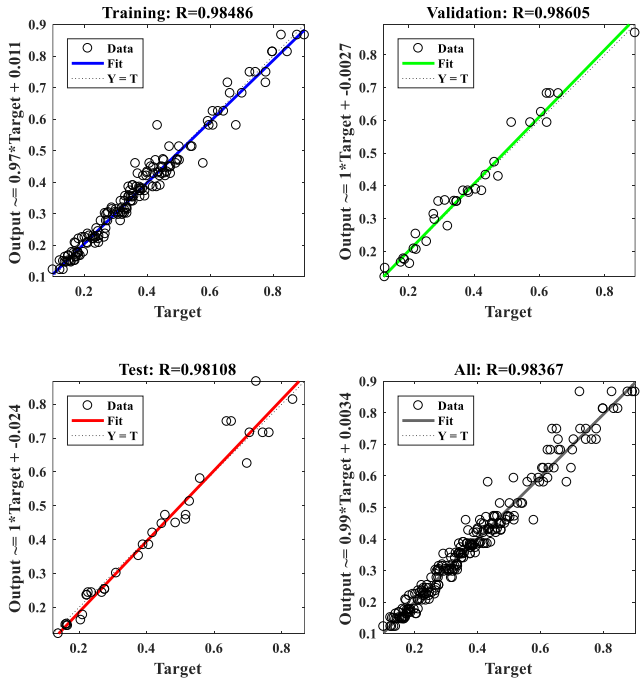


Fig. 9. Simulated regressions of training, validation, and test datasets by ANN6-18-Tansig

Accepted

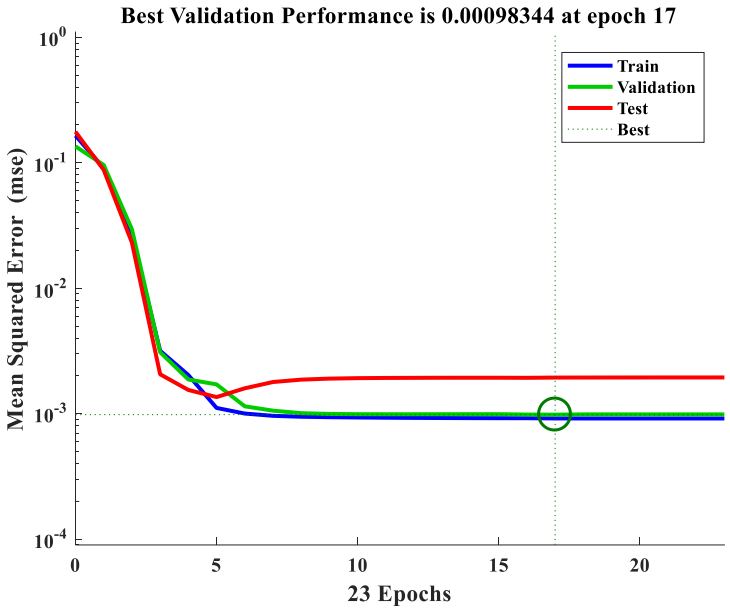


Fig. 10. The effectiveness of network functionality ANN6-18-Tansig

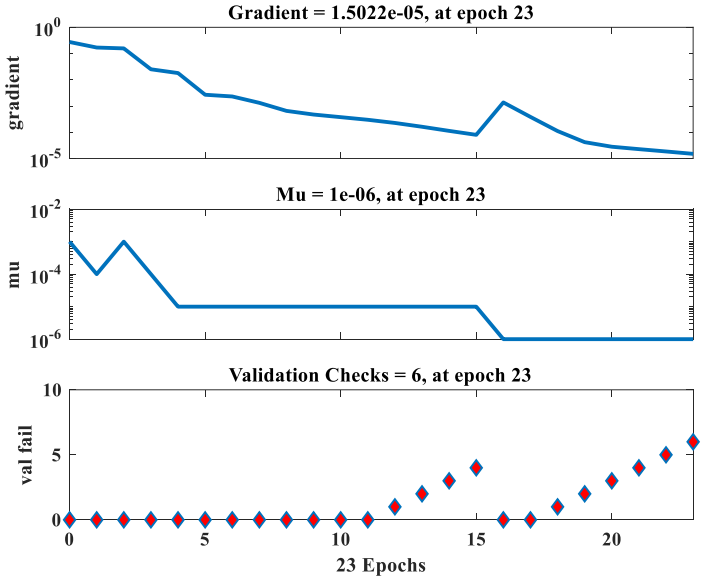


Fig. 11. The current status of training in artificial neural networks (ANN6-18-Tansig)

3.2 Method 2: neural network (ANN6-11-Purelin)

In the process of constructing the neural network, Eq. (1) plays a pivotal role as the normalization standard that effectively guides the entire development process. Through a comprehensive series of iterative evaluations and assessments, the optimal configuration for the network was determined to consist of eleven nodes, which have

been designated as ANN6-11-Purelin. A particularly crucial aspect of this network is its exclusive reliance on Purelin functions, which are utilized consistently throughout the entire structure of the network. For the sake of clarity and enhanced comprehension, it is essential to emphasize that the output layer of ANN6-11-Purelin also employs the Logsig function, maintaining uniformity across the network's architecture. The overall layout of this network is illustrated in Figure 12.

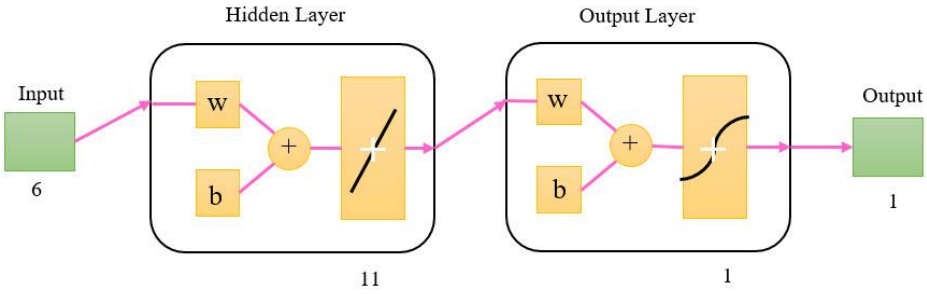


Fig. 12. The overall layout of the ANN6-11-Purelin

Specifically, Matrix 1 (Eq. (3)) represents the weight matrix of the hidden layer, while Matrix 2 (Eq. (4)) corresponds to the weight matrix of the output layer. Matrix 3 (Eq. (5)) contains the bias matrix of the hidden layer, and Matrix 4 (Eq. (6)) represents the bias matrix of the output layer. These matrices offer a summary of the ANN setup and have been mentioned in Eqs. (3 - ۶).

$$Weight_{hidden} = \begin{bmatrix} 0.47378 & -0.27746 & 0.57666 & -0.64487 & -0.41903 & -0.74381 \\ -0.37444 & 0.49426 & 0.82967 & 0.041564 & 0.29652 & 0.068794 \\ 0.028713 & 0.28088 & 0.20571 & 0.21736 & 0.081931 & -0.069864 \\ -0.68953 & 0.39578 & -0.58343 & -0.69405 & -0.056643 & 0.64412 \\ -0.13099 & -0.74409 & -0.78273 & 0.59404 & -0.24965 & -1.1373 \\ -0.063696 & -0.70258 & 0.0078301 & -0.269 & 0.026365 & -0.29166 \\ -0.58612 & -0.44521 & 0.47815 & 0.25025 & 0.96948 & -0.10699 \\ -0.72706 & -0.038511 & -0.092904 & 0.3227 & 0.54538 & 0.027011 \\ -0.11829 & 0.29782 & -0.81357 & 0.34957 & 0.25611 & 0.56681 \\ -0.86633 & -0.12957 & -0.4283 & 0.13349 & -0.073218 & -0.39637 \\ 0.07824 & 0.59668 & 0.57149 & 0.52306 & 0.094873 & 0.65693 \end{bmatrix} \quad (۳)$$

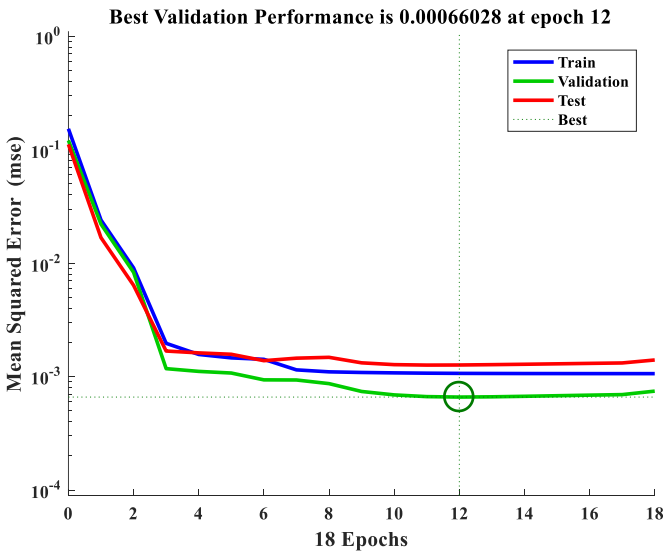
$$Weight_{output} \begin{bmatrix} -0.34571 & 0.25547 & -0.071154 & -0.25237 & -0.41979 & 0.37458 & -0.27023 & -0.10958 & -0.41364 & 0.27437 & -0.012786 \end{bmatrix} \quad (4)$$

Accepted / Not Edited

$$Bias_{hidden} = \begin{bmatrix} -0.072879 \\ -0.9484 \\ -0.64451 \\ 0.43254 \\ 0.055748 \\ -0.69976 \\ -0.34815 \\ 0.19414 \\ -0.57134 \\ 0.47325 \\ -0.48615 \end{bmatrix} \quad (۵) \qquad Bias_{output} = \begin{bmatrix} -0.22975 \end{bmatrix} \quad (۶)$$

3.3 Method 3: neural network (ANN6-16-Logsig)

In this section, a well-suited neural network architecture was implemented, which consists of a single hidden layer along with an output layer. For the functioning of these two distinct layers, we made use of specific transfer functions: the log sigmoid function was applied to the hidden layer, while the Tan sigmoid function was utilized for the output layer. This configuration was chosen to effectively process the data and produce the desired results. The illustrations presented in Figures 13 and 14 provide a detailed overview of the performance metrics and the training conditions associated with the proposed neural network model. Additionally**,** Figure 15 showcases the regression plots that correspond to the training, validation, testing, and comprehensive dataset. These plots are derived from the results obtained using the selected neural network configuration, which consists of six nodes, and they effectively represent the actual values extracted from the database.



Accepted / Not Edited

Fig. 13. The effectiveness of network functionality ANN6-16-Logsig

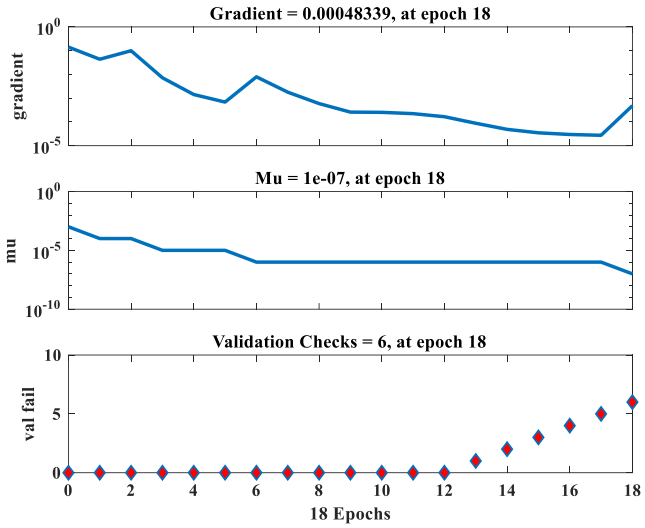


Fig. 14. The current status of training in artificial neural networks (ANN6-16-Logsig)

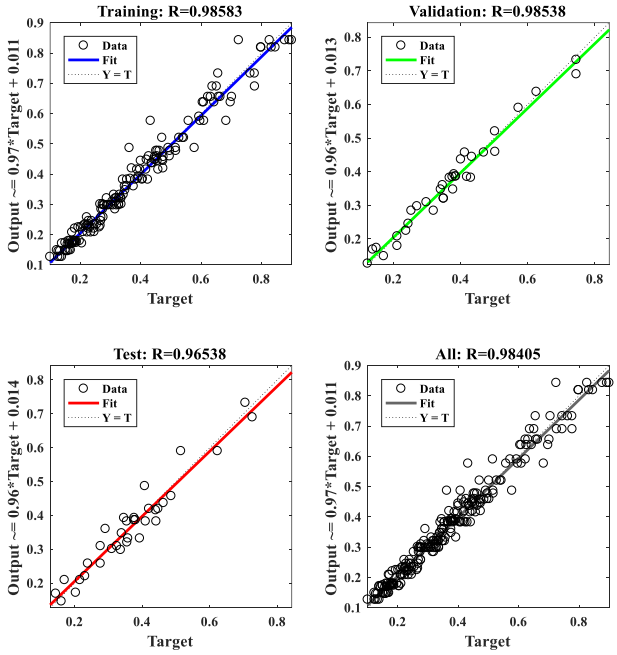


Fig. 15. Simulated regressions of training, validation, and test datasets by ANN6-16-Logsig

4. Comparison of models ANNs

In this particular section, a comprehensive comparison is made between the results obtained from various artificial neural network models, utilizing the Eqs. (7-9). The findings from this analysis are systematically organized and displayed in Table 4 for clarity and ease of understanding.

$$MSE=\frac{1}{n}\sum_{i=1}^{n}\left(V_{calc}-V_{test}\right)^{2} \tag{7}$$

$$RMSE=\sqrt{\frac{\sum_{i=1}^{n}\left(V_{calc}-V_{test}\right)^{2}}{n}} \tag{8}$$

$$R^{2}=1-\frac{\sum_{i=1}^{n}\left(V_{calc}-V_{test}\right)^{2}}{\sum_{i=1}^{n}\left(V_{calc}\right)^{2}} \tag{9}$$

The results highlighted above indicate that, among the different models developed through the artificial neural network approach, the ANN6-16-Logsig model demonstrates a commendable level of accuracy. This model shows great potential and can effectively serve as a reliable estimation tool for determining the pull-out energy associated with twin fibers.

5. Proposed GEP Models

The dataset utilized in this particular research comprises a total of 228 distinct laboratory samples. Within the scope of this study, the various input parameters that have been taken into account include several critical factors such as the diameter of the fibers, the number of twists that the fibers undergo within a length of 10 mm, the length of the fibers that are embedded within the matrix, the ratio of fiber length to fiber diameter, the ratio of water to cement, the compressive strength of the materials, and the output parameter that is being measured, which is the pull-out energy. For the purposes of this research, 80% of the entire dataset is allocated for the training phase, while the remaining 20% is reserved specifically for the validation of the model. The specific range of the input and output variables is detailed and can be found in Table 2.

This research utilized a set of 10 unique gene expression programming models to carry out its analysis. A thorough description of the parameters that were modified for each of these models is presented in Table

Accepted / Not Edited

2, which offers an extensive summary of the adjustments implemented.The use of multiple models allows for a more nuanced understanding of the data, as each model may capture different aspects of the underlying biological processes. The detailed overview in Table 2 serves as a valuable resource for readers seeking to comprehend the specific alterations made to each model, thereby enhancing the transparency and reproducibility of the study's findings.

Table 2. Parameters settings for the GEN models

Parameter	Number of models									
	Gen1	Gen2	Gen3	Gen4	Gen5	Gen6	Gen7	Gen8	Gen9	Gen10
Chromosomes	40	30	35	45	50	40	30	35	45	50
Genes	3	3	3	3	3	4	4	4	4	4
Head size	7	7	7	7	7	10	10	10	10	10
Linking function	+	+	+	+	+	*	*	*	*	*
mutation	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
One-point recombination	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Two-point recombination	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Gene recombination	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
IS transposition	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
RIS transposition	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Gene transposition	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.013

An analysis of the proposed models reveals that GEN.9 exhibits the most favorable correlation coefficient alongside the minimal error rate. This indicates that GEN.9 provides the most reliable framework for predicting pull-out energy.

6. Modeling results with GEP:

Figure 16 illustrates the equations that have been derived from the GEP.9 model, which serves as a proposed method for forecasting the maximum pull-out energy. This model is recognized as the most precise gene expression coding model and is represented in the form of an expression tree. Among the ten models that were proposed, this particular model was chosen based on its performance in error estimation functions.

Accepted / Not Edited

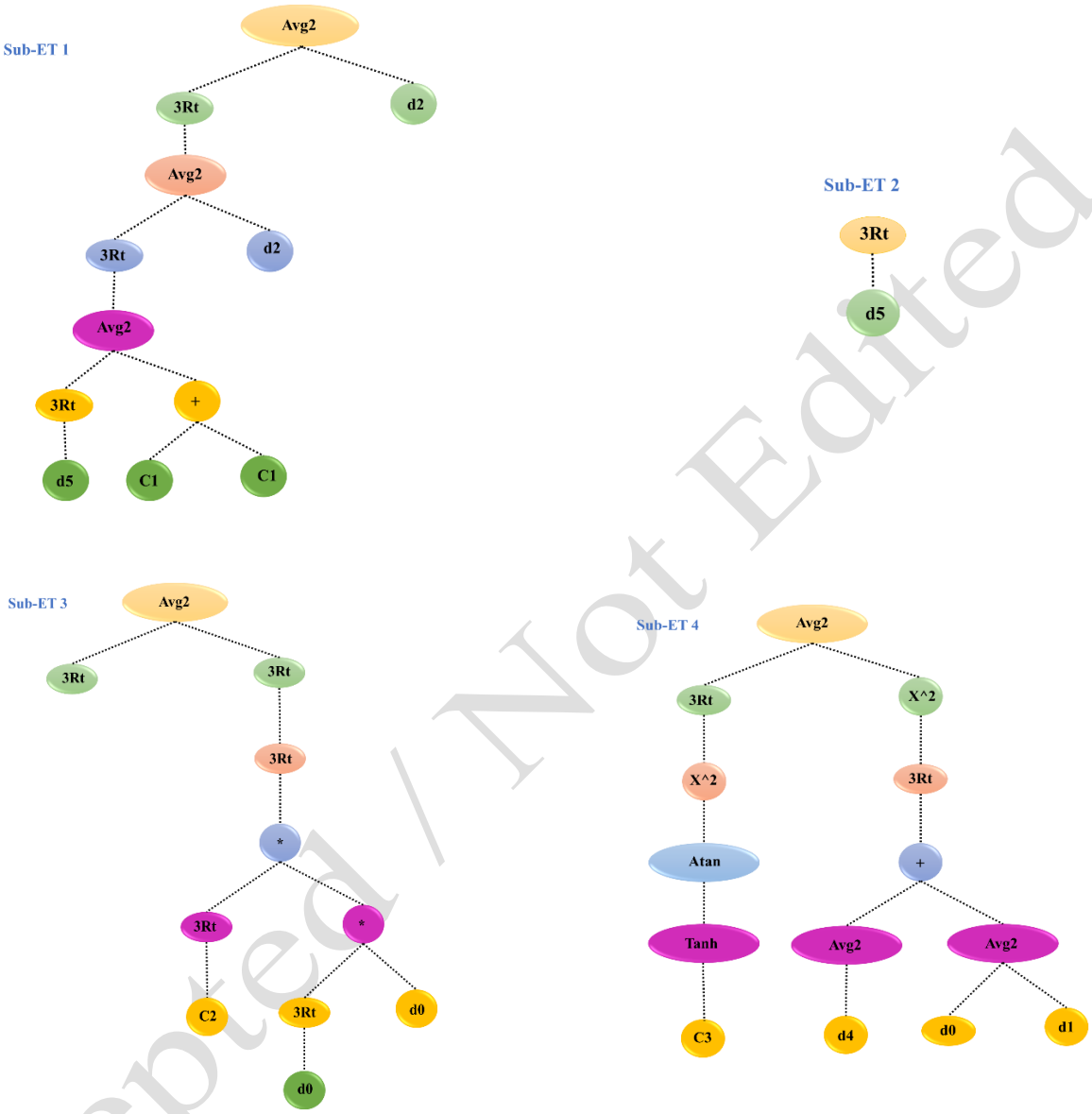


Fig. 16. ET corresponding to fourth formulation (GEP.9)

As illustrated in Figure 16, this particular model integrates a comprehensive total of ten distinct constant values. For those interested in a more thorough examination, a detailed presentation and breakdown of these constant values can be located in Table 3, where each value is clearly outlined and explained.

Accepted / Not Edited

Table 3. constant values for GEP.9

Gene No.	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9
Gene1	-6.5428	4.6568	9.7827	-1.1093	6.1436	6.4673	-8.7719	6.4482	-7.1259	-7.1575
Gene2	4.7361	-1.6873	-3.6246	6.3060	-7.3332	9.6093	7.2563	9.2364	7.5066	3.1431
Gene3	-8.9995	-2.8763	4.2967	-4.4267	5.7048	-7.8838	-5.7084	-1.9461	9.4994	8.1597
Gene4	-1.9541	-5.3538	0.4824	8.9806	-3.2932	-9.8547	1.1380	0.7498	-5.3971	-6.3975

6.1 Determining the relationship based on GEP

The equations that have been derived from the Genetic Expression Programming (GEP) model are illustrated in the form of an expression tree, which can be found in Figure 16. The fixed values associated with this model are detailed in Table 3. This particular model comprises a total of 400 constants, which are denoted by the letters ranging from C0 to C9, respectively. Following the extraction of the proposed relationships for each gene, it becomes possible to derive the extraction energy from Eq. (14).

Based on the equation that has been derived, it is possible to observe the correlation diagram that illustrates the relationship between the results obtained from the laboratory and the predictions made by this particular model, as depicted in Figure 18. The data presented in Figure 18 reveals a substantial and positive relationship between the predictions made by this model and the laboratory findings, highlighting a high degree of alignment between the two datasets**.**

Based on the relationships obtained in each gene**,** Eqs. (10 -13) are presented.

$$ET_1 = \left(Avg\left(L_e, \left(Avg\left(L_e, \left(Avg\left(\left(f_c' \right)^{1/3}, \left(2C_1 \right) \right)^{1/3} \right) \right) \right) \right) \right) \quad (10)$$

Accepted / Not Edited

$$ET2 = \left(f_c' \right)^{1/3} \tag{11}$$

$$ET3 = Avg\left(L_e^{1/3}, \left(\left(C_2^{1/3} * \left(d_f^{1/3} * \left(d_f^{1/3} * d_f \right) \right)^{1/3} \right)^{1/3} \right) \right) \tag{12}$$

$$ET4 = Avg\left(\left(Atan\left(Tanh\left(C_3 \right) \right)^2 \right)^{1/3}, \left(\left(Tanh\left(W/C \right) + Avg\left(d_f, n_{10} \right)^{1/3} \right)^2 \right) \right) \tag{13}$$

Once the suggested relationships associated with each gene within this particular model have been thoroughly extracted and analyzed, it becomes possible to determine the maximum energy that can be obtained from the pull-out of the fibers, which can specifically be derived from Eq. (14).

$$W = ET_1 * ET_2 * ET_3 * ET_4 \tag{14}$$

7. Results and Discussions

7.1 Artificial neural network results

The results of the pull-out energy for twin fibers within cement-based matrices, as forecasted by three artificial neural network (ANN) models, are presented in Figure 17. The Figure demonstrates that the ANN predictions align closely with the experimental data.

Accepted / Not Edited

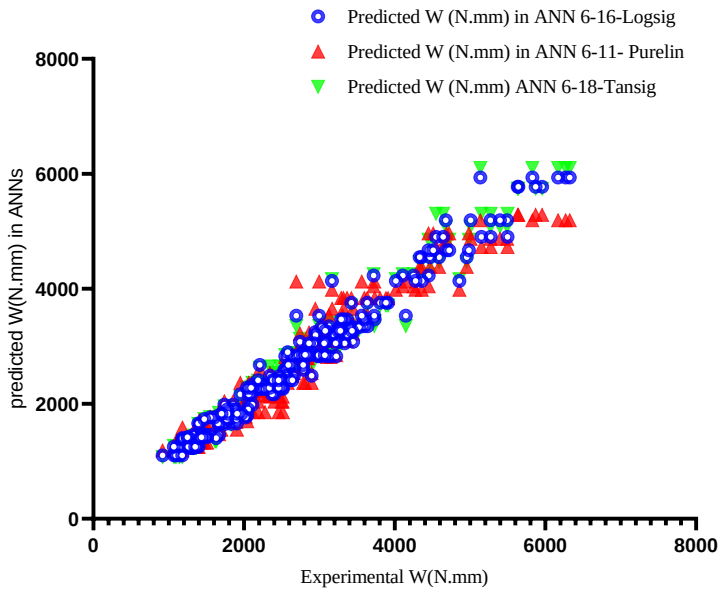


Fig. 17. A comparison of anticipated and observed parameters for ANN models

Additionally, a series of statistical analyses were executed to assess the accuracy of the predictions, employing metrics such as R^2, mean squared error (MSE), and root mean squared error (RMSE), with the findings summarized in Table 4.

Table 4. Statistical analyses to evaluate prediction accuracy (ANNs)

Index	ANN6-18-Tansig	ANN6-11-Purelin	ANN6-16-Logsig
R^2	0.994	0.988	0.995
RMSE (kN.m)	0.222	0.319	0.21
MSE (kN.m)	0.049	0.102	0.047

Table 4 presents the correlation coefficients for the actual versus estimated data in the ANN6-18-Tansig and ANN6-16-Logsig models, which are 0.994 and 0.995, respectively. These values indicate a robust correlation between the datasets. The Mean Squared Error (MSE) for the pull-out energy estimation utilizing the proposed Artificial Neural Network (ANN) model is noted to be 0.047 kN·m for the ANN6-16-Logsig model. Furthermore, the Root Mean Squared Error (RMSE) values are recorded as 0.222, 0.319, and 0.21 for the ANN6-18-Tansig, ANN6-11-Purelin, and ANN6-16-Logsig models, respectively. The relatively low MSE and RMSE values associated with the developed ANN model indicate its high accuracy and effectiveness in outcome prediction.

Accepted / Not cited

7.2 Gene expression programming results

The GEP model presented in Eq. (14) is proposed as a means to estimate the pull-out energy of twin fibers. The important point is that the above equation includes all the important parameters that affect the pull-out energy, including concrete compressive strength, fiber diameter, water-cement ratio, fiber overlap length, and number of fiber twists per 10 mm length.

Based on the information provided in Eq. (14), one can observe the correlation diagram that illustrates the relationship between the results obtained from the laboratory experiments and the predictions generated by this particular model. This diagram, which is crucial for understanding the performance of the model during both the training and validation phases, is presented in Figure 18.

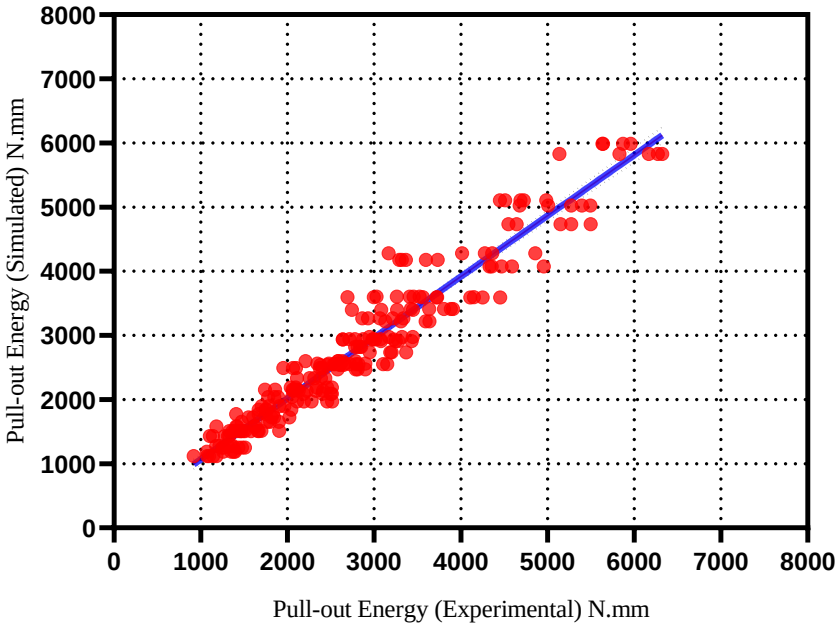


Fig. 18. ET corresponding to fourth formulation (GEP.9)

From the observation made in Figure 18, it is evident that there exists a remarkably strong correlation between the prediction results generated by this model and the data obtained from laboratory experiments. The specific values pertaining to the correlation coefficient, along with the corresponding error values, are detailed in Table 5 for further reference and analysis.

Table 5. Statistical analyses to evaluate prediction accuracy (GEPs)

Number of Modells	Training data				Validation data			
	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2
GEN.1	0.0585	0.0034	0.0470	0.908	0.0657	0.0043	0.0490	0.925
GEN.2	0.0660	0.0043	0.055	0.902	0.0691	0.0047	0.0593	0.868
GEN.3	0.0626	0.0039	0.0486	0.871	0.0813	0.0066	0.0633	0.866
GEN.4	0.0591	0.0034	0.0429	0.898	0.0538	0.0029	0.417	0.881
GEN.5	0.0719	0.0051	0.0573	0.861	0.0777	0.0060	0.0555	0.770
GEN.6	0.0537	0.0028	0.0427	0.912	0.0610	0.0037	0.0497	0.904
GEN.7	0.0543	0.0029	0.0426	0.929	0.069	0.0048	0.0530	0.903
GEN.8	0.0589	0.0034	0.0449	0.900	0.0544	0.0029	0.0457	0.902
GEN.9	0.0454	0.0020	0.0344	0.937	0.0541	0.0029	0.0389	0.918
GEN.10	0.0515	0.0026	0.0379	0.939	0.0705	0.0049	0.0500	0.894

Acording Table 5 The values of R^2 and MSE and RMSE for the GEP model in training and validation data were found to be 0.937,0.002,0.0454,0.918,0.0029,and 0.0541, for the GEN.9 respectively . The accuracy of this model and the correlation between the results estimated by this model and the actual results were acceptable, and the corresponding formula can be used in practice. The GEP model was to estimate the target parameter with an MAE of 0.002.

7.3 Sensitivity analysis

Sensitivity analysis serves as a valuable technique employed to assess the significance of various input parameters in relation to a particular target or outcome. In order to thoroughly investigate the relative importance of the six distinct parameters that play a role in determining pull-out energy, a comprehensive sensitivity analysis was conducted. This analysis utilized the weights that were derived from the Artificial Neural Network model, following the methodology put forth by Milne, as demonstrated in Eq. (15), (Milne, 1995).

Accepted / Not Edited

$$IIF = \frac{\sum_{j=1}^{n_{hidden}} \frac{Wji}{\sum_{i=1}^{n_{inputs}} |Wji|} . Woj}{\sum_{K=1}^{n_{inputs}} \left(\sum_{j=1}^{n_{hidden}} \left| \frac{wjk}{\sum_{i=1}^{n_{inputs}} |Wji|} . Woj \right| \right)} \quad (15)$$

This analytical framework provides a more profound understanding of how changes in the input parameters can influence the desired results. By implementing the structured approach established by Milne, the research effectively underscores the contributions of each individual parameter to the overall pull-out energy. This, in turn, aids in facilitating well-informed decision-making processes within the context of the analysis being conducted.

Figure 19 presents a comprehensive depiction of the impact of various parameters on pull-out energy, highlighting that, of all the factors analyzed, the compressive strength of concrete emerges as the most critical determinant affecting the total pull-out energy.

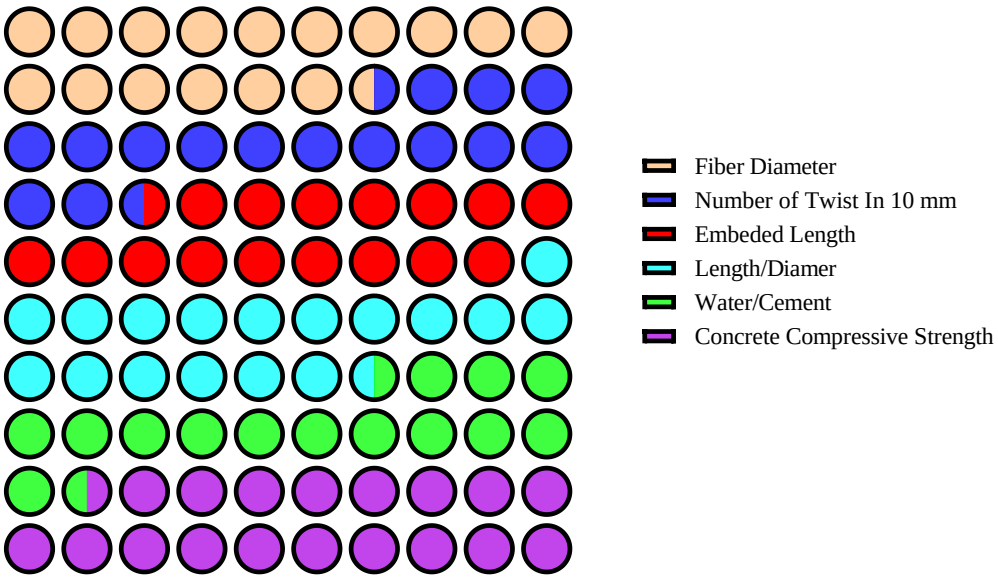


Fig. 19. Significance of input factors (d_f %16.4,n_{10} %16, Le %16.5, l_f/d_f %17.6, W/C %15 ,f'_c %18.5)

Accepted / Not Edited

8. Conclusion

In this study, two advanced machine learning approaches the Gene Expression Programming (GEP) and Artificial Neural Networks (ANN) were utilized to predict the pull-out energy of twin fibers embedded in cement-based matrices. The key findings of the research are as follows:

- The ANN model demonstrated outstanding performance with a high correlation coefficient of 0.995, indicating a very strong agreement between the predicted and experimental results. This confirms the potential of ANN for accurately predicting the pull-out energy in cement-based composites.
- The GEP model also showed satisfactory accuracy, achieving a correlation coefficient of 0.98. Despite a slightly lower correlation compared to the ANN model, the GEP approach proved to be effective and can be applied successfully in practical engineering applications.
- A comparison of the models revealed that while the GEP model yielded accurate predictions, its correlation coefficient was about 1.5% lower than that of the ANN model. This suggests that ANN offers slightly superior predictive capabilities for the given dataset.
- Sensitivity analysis identified that the compressive strength of the concrete had the most significant influence on the pull-out energy, highlighting its critical role in determining the bond strength between fibers and matrix.

While both models demonstrated high predictive accuracy, the reliability of predictions is closely tied to the volume of data available. One limitation encountered in this study was the relatively small dataset, which restricted the model's ability to fully capture the complex relationships between the input parameters and the pull-out energy. Therefore, future research should focus on expanding the dataset to enhance the robustness and generalizability of these models.The results of this study offer valuable insights into the predictive capabilities of ANN and GEP for fiber pull-out energy in cement-based matrices, contributing to more efficient material design in civil engineering. However, the practical implementation of these models requires larger datasets to improve prediction reliability, emphasizing the need for comprehensive experimental databases in future studies.

Notation			
d_f	Fiber diameter	V_{calc}	The value calculated from the prediction
d_{sf}	The diameter of steel fibers in the construction of twin fibers	V_{test}	The value calculated from the test
D	Equivalent diameter of twin fibers	W/C	Water to cement ratio
f'_c	Concrete compressive strength	W_p	Pull-out energy
Le	Fiber embedded length	x_{min}	The minimum value of the data x
l_f	Fiber length	x_{max}	The maximum value of the data x
l_f/d_f	Fiber aspect ratio (Fiber length to diameter ratio)	x_n	The normal value of the data x
n_{10}	Number of fiber twists per 10 mm length	x_r	The real value of the data x

Author Contributions

Abolfaz Hemmatian: Data curation, Investigation, Software, Validation, Writing – original draft. Meysam Jalali: Conceptualization, Project administration, Supervision, Formal analysis, Methodology, Writing – review and editing. Hossain Naderpour: Validation, Writing – review and editing.

Accepted / Not Edited

Acknowledgments

The authors express their gratitude for the technical assistance received from Shahrood University of Technology.

Conflicts of Interests

We have no conflicts of interest to disclose.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Funding: This research received no external funding.

References

Abdallah, S., Fan, M., & Rees, D. W. A. (2016). Analysis and modelling of mechanical anchorage of 4D/5D hooked end steel fibres. *Materials and Design*, *112*, 539–552. https://doi.org/10.1016/j.matdes.2016.09.107

Ahmad, A., Chaiyasarn, K., Farooq, F., Ahmad, W., Suparp, S., & Aslam, F. (2021). Compressive strength prediction via gene expression programming (Gep) and artificial neural network (ann) for concrete containing rca. *Buildings*, *11*(8), 1–18. https://doi.org/10.3390/buildings11080324

Al-hashem, M. N., Amin, M. N., Raheel, M., Khan, K., Alkadhim, H. A., Imran, M., Ullah, S., & Iqbal, M. (2022). Predicting the Compressive Strength of Concrete Containing Fly Ash and Rice Husk Ash Using ANN and GEP Models. *Materials*. https://doi.org/10.3390/ma15217713

Al-Zwainy, F. M., Al-khazrajy, M. G., Hussein, N. M., Mohamed, S., Sarhan, M. M., Al-Musawi, T. J., & Hayder, G. (2024). Utilizing Artificial Neural Networks for Predictive KPI Analysis in Bridge Projects. *Journal of Computational Analysis and Applications*, *33*(7).

Al-Zwainy, F. M. S., Salih, S. A., & Aldikheeli, M. R. (2021). Prediction of residual strength of sustainable selfconsolidating concrete exposed to elevated temperature using artificial intelligent technique.

Accepted / Not Edited

International Journal of Applied Science and Engineering, *18*(2), 1–15. https://doi.org/10.6703/IJASE.202106_18(2).012

Anita Jessie, J., & Santhi, A. S. (2019). Effect of temperature on compressive strength of steel fibre reinforced concrete. *Journal of Applied Science and Engineering*, *22*(2), 233–238. https://doi.org/10.6180/jase.201906_22(2).0004

Ataee, S., Jalali, M., & Nehdi, M. L. (2024). Pull-out behavior of twin-twisted steel fibers from various strength cement-based matrices. *Construction and Building Materials*, *445*. https://doi.org/10.1016/j.conbuildmat.2024.137855

Banchhor, S. (2018). Role of mineral admixtures and chemical admixtures in concrete. *International Journal of Advance Research*. www.IJARIIT.com

Bu, L., Du, G., & Hou, Q. (2021). Prediction of the compressive strength of recycled aggregate concrete based on artificial neural network. *Materials*, *14*(14). https://doi.org/10.3390/ma14143921

Cascardi, A., Micelli, F., & Aiello, M. A. (2017). An Artificial Neural Networks model for the prediction of the compressive strength of FRP-confined concrete circular columns. *Engineering Structures*, *140*, 199–208. https://doi.org/10.1016/j.engstruct.2017.02.047

Chandra Nayak, S., Dehuri, S., & Cho, S. B. (2024). An evolutionary functional link artificial neural network for assessment of compressive strength of concrete structures. *Ain Shams Engineering Journal*, *15*(3). https://doi.org/10.1016/j.asej.2023.102462

Dantas, A. T. A., Batista Leite, M., & De Jesus Nagahama, K. (2013). Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. *Construction and Building Materials*, *38*, 717–722. https://doi.org/10.1016/j.conbuildmat.2012.09.026

Ferreira, C. (2001). *Gene Expression Programming: a New Adaptive Algorithm for Solving Problems*. http://arxiv.org/abs/cs/0102027

Gholampour, A., Gandomi, A. H., & Ozbakkaloglu, T. (2017). New formulations for mechanical properties of

Accepted / Not Edited

recycled aggregate concrete using gene expression programming. *Construction and Building Materials*, *130*, 122–145. https://doi.org/10.1016/j.conbuildmat.2016.10.114

Ghorbani, A., Maleki, H., Naderpour, H., & Khatami, S. M. H. (2024). Advancing Compressive Strength Prediction in Self-Compacting Concrete via Soft Computing: A Robust Modeling Approach. *Journal of Soft Computing in Civil Engineering*, *8*(1), 126–140. https://doi.org/10.22115/SCCE.2023.396669.1646

Hemmatian, A., Jalali, M., Naderpour, H., & Nehdi, M. L. (2023). Machine learning prediction of fiber pull-out and bond-slip in fiber-reinforced cementitious composites. *Journal of Building Engineering*, *63*. https://doi.org/10.1016/j.jobe.2022.105474

Hoang, N. D., & Tran, D. V. (2024). Machine Learning-Based Estimation of Concrete Compressive Strength: A Multi-Model and Multi-Dataset Study. *Civil Engineering Infrastructures Journal*, *57*(2), 247–265. https://doi.org/10.22059/CEIJ.2023.354679.1910

Huang, B., Bahrami, A., Javed, M. F., Azim, I., & Iqbal, M. A. (2024). Evolutionary Algorithms for Strength Prediction of Geopolymer Concrete. *Buildings*, *14*(5). https://doi.org/10.3390/buildings14051347

Ismael Jaf, D. K., Abdalla, A., Mohammed, A. S., Abdulrahman, P. I., Rawaz Kurda, & Mohammed, A. A. (2024). Hybrid nonlinear regression model versus MARS, MEP, and ANN to evaluate the effect of the size and content of waste tire rubber on the compressive strength of concrete. *Heliyon*, *10*(4). https://doi.org/10.1016/j.heliyon.2024.e25997

Jamee, M., Alaee, F. J., & Zeighami, E. (2022). Effect of matrix strength on pull-out behaviour of hooked steel fibre-reinforced cementitious composites. *Canadian Journal of Civil Engineering*, *49*(11), 1703–1713. https://doi.org/10.1139/cjce-2021-0421

Khademi, F., Jamal, S. M., Deshpande, N., & Londhe, S. (2016). Predicting strength of recycled aggregate concrete using Artificial Neural Network, Adaptive Neuro-Fuzzy Inference System and Multiple Linear Regression. *International Journal of Sustainable Built Environment*, *5*(2), 355–369. https://doi.org/10.1016/j.ijsbe.2016.09.003

Accepted / Not Edited

Khan, M. A., Zafar, A., Akbar, A., Javed, M. F., & Mosavi, A. (2021). Application of gene expression programming (GEP) for the prediction of compressive strength of geopolymer concrete. *Materials*, *14*(5), 1–23. https://doi.org/10.3390/ma14051106

Le, H. V., & Kim, D. J. (2019). Effects of Matrix Strength, Fiber Type, and Fiber Content on the Electrical Resistivity of Steel-Fiber-Reinforced Cement Composites During Fiber Pullout. *KSCE J. Civ. Environ. Eng. Res*, *39*(6), 675–689. https://doi:10.12652/KSCE.2019.39.6.0675

Lee, J. H., & Kighuta, K. (2017). Twin-twist effect of fibers on the pullout resistance in cementitious materials. *Construction and Building Materials*, *146*, 555–562. https://doi.org/10.1016/j.conbuildmat.2017.04.147

Mastan, S., Anandh, S., & Sindhu Nachiar, S. (2023). Numerical method and validation using ANN of composite steel–concrete beam for optimized geometry and emplacement of web opening. *Asian Journal of Civil Engineering*. https://doi.org/10.1007/s42107-023-00860-6

Milne, L. (1995). Feature Selection Using Neural Networks with Contribution Measures. *Australian Conference on Artificial Intelligence*, *November*, 1–8. https://doi.org/https://doi.org/10.26190/unsworks/378 License:

Mirrashid, M. (2014). Earthquake magnitude prediction by adaptive neurofuzzy inference system (ANFIS) based on fuzzy C-means algorithm. *Natural Hazards*, *74*(3), 1577–1593. https://doi.org/10.1007/s11069-014-1264-7

Naderpour, H., Nagai, K., Fakharian, P., & Haji, M. (2019). Innovative models for prediction of compressive strength of FRP-confined circular reinforced concrete columns using soft computing methods. *Composite Structures*, *215*, 69–84. https://doi.org/10.1016/j.compstruct.2019.02.048

Nasir, M., Gazder, U., Maslehuddin, M., Baghabra Al-Amoudi, O. S., & Syed, I. A. (2020). Prediction of Properties of Concrete Cured Under Hot Weather Using Multivariate Regression and ANN Models. *Arabian Journal for Science and Engineering*, *45*(5), 4111–4123. https://doi.org/10.1007/s13369-020-

Accepted / Not Edited

04403-y

Nguyen, T. H., Nguyen, X. B., Nguyen, V. H., Nguyen, T. H. T., & Nguyen, D. D. (2023). Shear strength prediction of concrete beams reinforced with FRP bars using novel hybrid BR-ANN model. *Asian Journal of Civil Engineering*. https://doi.org/10.1007/s42107-023-00876-y

Qu, D., Cai, X., & Chang, W. (2018). Evaluating the effects of steel fibers on mechanical properties of ultra-high performance concrete using artificial neural networks. *Applied Sciences (Switzerland)*, *8*(7). https://doi.org/10.3390/app8071120

Rahman Sobuz, M. H., Aayaz, R., Rahman, S. A., Ahmed Shaikh, F. U., Islam Kabbo, M. K., & Hayet Khan, M. M. (2025). Assessment of hybrid fiber reinforced graphene nano-engineered concrete composites: From experimental testing to explainable machine learning modeling. *Journal of Materials Research and Technology*, *36*, 1409–1430. https://doi.org/10.1016/j.jmrt.2025.03.003

Scott M. Thede. (2014). An introduction to genetic algorithms. *Journal of Computing Sciences in Colleges*.

Sengul, O. (2018). Mechanical properties of slurry infiltrated fiber concrete produced with waste steel fibers. *Construction and Building Materials*, *186*, 1082–1091. https://doi.org/10.1016/j.conbuildmat.2018.08.042

Shahrokhishahraki, M., Malekpour, M., Mirvalad, S., & Faraone, G. (2024). Machine learning predictions for optimal cement content in sustainable concrete constructions. *Journal of Building Engineering*, *82*. https://doi.org/10.1016/j.jobe.2023.108160

Sobhani, J., Najimi, M., Pourkhorshidi, A. R., & Parhizkar, T. (2010). Prediction of the compressive strength of no-slump concrete: A comparative study of regression, neural network and ANFIS models. *Construction and Building Materials*, *24*(5), 709–718. https://doi.org/10.1016/j.conbuildmat.2009.10.037

Taffese, W. Z., Wally, G. B., Magalhães, F. C., & Espinosa-Leal, L. (2024). Concrete aging factor prediction using machine learning. *Materials Today Communications*, *40*. https://doi.org/10.1016/j.mtcomm.2024.109527

Accepted / Not Edited

Vidivelli, B., & Jayaranjini, A. (2016). Prediction of compressive strength of high performance concrete containing industrial by products using artificial neural networks. *International Journal of Civil Engineering and Technology, 7*(2), 302–314.

Wang, X., Xu, B., Luan, K., Mu, R., & Chen, J. (2024). Optimization of the Shape of Hooked-End Steel Fiber Based on Pulling Out and Reinforcing Cementitious Composites. *Materials, 17*(1). https://doi.org/10.3390/ma17010047

Yoo, D. Y., Je, J., Choi, H. J., & Sukontasukkul, P. (2020). Influence of embedment length on the pullout behavior of steel fibers from ultra-high-performance concrete. *Materials Letters, 276*. https://doi.org/10.1016/j.matlet.2020.128233

Accepted / Not Edited

Appendix A1

Table A1. Detailed experimental results for 228 test cases used in the study.

Number	Diameter df (mm)	number of twist in 10 mm	Embedded llength (mm)	lf/df	W/C	fc (Mpa)	pull-out Energy (n.mm)	Ref.
1	0.54	2.00	20.00	74.49	0.42	47.70	3639.23	(Ataee et al., 2024)
2	0.54	2.00	20.00	74.49	0.42	47.70	3308.22	
3	0.54	2.00	20.00	74.49	0.42	47.70	3592.21	
4	0.54	2.00	20.00	74.49	0.42	47.70	3127.33	
5	0.41	2.00	20.00	97.56	0.42	47.70	2722.44	
6	0.41	2.00	20.00	97.56	0.42	47.70	3104.89	
7	0.41	2.00	20.00	97.56	0.42	47.70	2819.00	
8	0.41	2.00	20.00	97.56	0.42	47.70	3150.30	
9	0.41	2.00	20.00	97.56	0.42	47.70	2546.40	
10	0.54	2.00	15.00	74.49	0.42	47.70	2111.50	
11	0.54	2.00	15.00	74.49	0.42	47.70	2437.00	
12	0.54	2.00	15.00	74.49	0.42	47.70	2310.54	
13	0.54	2.00	15.00	74.49	0.42	47.70	2255.05	
14	0.54	2.00	15.00	74.49	0.42	47.70	2386.21	
15	0.41	2.00	15.00	97.56	0.42	47.70	1747.10	
16	0.41	2.00	15.00	97.56	0.42	47.70	2044.37	
17	0.41	2.00	15.00	97.56	0.42	47.70	1668.00	
18	0.41	2.00	15.00	97.56	0.42	47.70	1759.00	
19	0.54	2.00	10.00	74.49	0.42	47.70	1329.34	
20	0.54	2.00	10.00	74.49	0.42	47.70	1312.48	
21	0.54	2.00	10.00	74.49	0.42	47.70	1278.21	
22	0.54	2.00	10.00	74.49	0.42	47.70	1106.37	
23	0.54	2.00	10.00	74.49	0.42	47.70	1141.47	
24	0.41	2.00	10.00	97.56	0.42	47.70	1072.13	
25	0.41	2.00	10.00	97.56	0.42	47.70	1143.73	
26	0.41	2.00	10.00	97.56	0.42	47.70	919.25	
27	0.41	2.00	10.00	97.56	0.42	47.70	1096.76	
28	0.41	2.00	10.00	97.56	0.42	47.70	1181.06	
29	0.54	2.00	20.00	74.49	0.24	72.29	4356.10	
30	0.54	2.00	20.00	74.49	0.24	72.29	4330.68	

Accepted / Not Edited

31	0.54	2.00	20.00	74.49	0.24	72.29	4955.37	(Ataee et al., 2024)
32	0.54	2.00	20.00	74.49	0.24	72.29	4588.36	
33	0.54	2.00	20.00	74.49	0.24	72.29	4470.21	
34	0.41	2.00	20.00	97.56	0.24	72.29	3170.76	
35	0.41	2.00	20.00	97.56	0.24	72.29	2945.75	
36	0.41	2.00	20.00	97.56	0.24	72.29	3321.14	
37	0.41	2.00	20.00	97.56	0.24	72.29	3446.14	
38	0.54	2.00	15.00	74.49	0.24	72.29	3013.11	
39	0.54	2.00	15.00	74.49	0.24	72.29	3225.47	
40	0.54	2.00	15.00	74.49	0.24	72.29	2638.13	
41	0.54	2.00	15.00	74.49	0.24	72.29	2979.34	
42	0.54	2.00	15.00	74.49	0.24	72.29	2638.01	
43	0.41	2.00	15.00	97.56	0.24	72.29	2409.95	
44	0.41	2.00	15.00	97.56	0.24	72.29	2131.13	
45	0.41	2.00	15.00	97.56	0.24	72.29	2328.00	
46	0.41	2.00	15.00	97.56	0.24	72.29	2067.00	
47	0.41	2.00	15.00	97.56	0.24	72.29	2219.89	
48	0.54	2.00	10.00	74.49	0.24	72.29	1831.86	
49	0.54	2.00	10.00	74.49	0.24	72.29	1667.76	
50	0.54	2.00	10.00	74.49	0.24	72.29	1751.21	
51	0.54	2.00	10.00	74.49	0.24	72.29	1409.62	
52	0.41	2.00	10.00	97.56	0.24	72.29	1216.17	
53	0.41	2.00	10.00	97.56	0.24	72.29	1311.24	
54	0.41	2.00	10.00	97.56	0.24	72.29	1368.77	
55	0.41	2.00	10.00	97.56	0.24	72.29	1177.61	
56	0.41	2.00	10.00	97.56	0.24	72.29	1349.44	
57	0.54	2.00	20.00	74.49	0.22	109.82	5493.50	
58	0.54	2.00	20.00	74.49	0.22	109.82	5272.37	
59	0.54	2.00	20.00	74.49	0.22	109.82	4676.90	
60	0.54	2.00	20.00	74.49	0.22	109.82	5007.55	
61	0.54	2.00	20.00	74.49	0.22	109.82	5396.90	
62	0.54	2.00	15.00	74.49	0.22	109.82	3453.12	
63	0.54	2.00	15.00	74.49	0.22	109.82	3263.08	
64	0.54	2.00	15.00	74.49	0.22	109.82	3410.91	
65	0.54	2.00	15.00	74.49	0.22	109.82	3525.21	
66	0.54	2.00	15.00	74.49	0.22	109.82	3028.03	
67	0.41	2.00	15.00	97.56	0.22	109.82	2791.75	
68	0.41	2.00	15.00	97.56	0.22	109.82	2534.79	
69	0.41	2.00	15.00	97.56	0.22	109.82	2596.61	
70	0.41	2.00	15.00	97.56	0.22	109.82	2543.98	
71	0.54	2.00	10.00	74.49	0.22	109.82	2061.43	
72	0.54	2.00	10.00	74.49	0.22	109.82	1857.61	

Accepted / Not Edited

73	0.54	2.00	10.00	74.49	0.22	109.82	1740.34	(Ataee et al., 2024)
74	0.54	2.00	10.00	74.49	0.22	109.82	2111.26	
75	0.54	2.00	10.00	74.49	0.22	109.82	2109.40	
76	0.41	2.00	10.00	97.56	0.22	109.82	1463.79	
77	0.41	2.00	10.00	97.56	0.22	109.82	1511.22	
78	0.41	2.00	10.00	97.56	0.22	109.82	1367.81	
79	0.41	2.00	10.00	97.56	0.22	109.82	1460.48	
80	0.41	2.00	10.00	97.56	0.22	109.82	1426.65	
81	0.54	3.00	20.00	74.49	0.42	47.70	3911.38	
82	0.54	3.00	20.00	74.49	0.42	47.70	3634.30	
83	0.54	3.00	20.00	74.49	0.42	47.70	3805.19	
84	0.54	3.00	20.00	74.49	0.42	47.70	3885.00	
85	0.54	3.00	20.00	74.49	0.42	47.70	3423.90	
86	0.41	3.00	20.00	97.56	0.42	47.70	3186.88	
87	0.41	3.00	20.00	97.56	0.42	47.70	3370.15	
88	0.41	3.00	20.00	97.56	0.42	47.70	2949.20	
89	0.41	3.00	20.00	97.56	0.42	47.70	3200.10	
90	0.54	3.00	15.00	74.49	0.42	47.70	2897.80	
91	0.54	3.00	15.00	74.49	0.42	47.70	2811.00	
92	0.54	3.00	15.00	74.49	0.42	47.70	2785.80	
93	0.54	3.00	15.00	74.49	0.42	47.70	2580.09	
94	0.41	3.00	15.00	97.56	0.42	47.70	2189.40	
95	0.41	3.00	15.00	97.56	0.42	47.70	2277.40	
96	0.41	3.00	15.00	97.56	0.42	47.70	2458.50	
97	0.41	3.00	15.00	97.56	0.42	47.70	2514.70	
98	0.41	3.00	15.00	97.56	0.42	47.70	2110.50	
99	0.54	3.00	10.00	74.49	0.42	47.70	1579.44	
100	0.54	3.00	10.00	74.49	0.42	47.70	1669.36	
101	0.54	3.00	10.00	74.49	0.42	47.70	1459.10	
102	0.54	3.00	10.00	74.49	0.42	47.70	1330.36	
103	0.54	3.00	10.00	74.49	0.42	47.70	1487.86	
104	0.41	3.00	10.00	97.56	0.42	47.70	1260.00	
105	0.41	3.00	10.00	97.56	0.42	47.70	1397.17	
106	0.41	3.00	10.00	97.56	0.42	47.70	1379.49	
107	0.41	3.00	10.00	97.56	0.42	47.70	1069.20	
108	0.41	3.00	10.00	97.56	0.42	47.70	1351.90	
109	0.54	3.00	20.00	74.49	0.24	79.29	4548.13	
110	0.54	3.00	20.00	74.49	0.24	79.29	5498.00	
111	0.54	3.00	20.00	74.49	0.24	79.29	5150.09	
112	0.54	3.00	20.00	74.49	0.24	79.29	4643.61	
113	0.54	3.00	20.00	74.49	0.24	79.29	5274.30	

Accepted / Not Edited

114	0.41	3.00	20.00	97.56	0.24	79.29	4108.46	(Ataee et al., 2024)
115	0.41	3.00	20.00	97.56	0.24	79.29	4452.49	
116	0.41	3.00	20.00	97.56	0.24	79.29	3719.77	
117	0.41	3.00	20.00	97.56	0.24	79.29	4252.30	
118	0.54	3.00	15.00	74.49	0.24	79.29	3447.87	
119	0.54	3.00	15.00	74.49	0.24	79.29	3080.55	
120	0.54	3.00	15.00	74.49	0.24	79.29	2743.76	
121	0.54	3.00	15.00	74.49	0.24	79.29	3264.10	
122	0.41	3.00	15.00	97.56	0.24	79.29	2896.90	
123	0.41	3.00	15.00	97.56	0.24	79.29	2661.46	
124	0.41	3.00	15.00	97.56	0.24	79.29	2345.36	
125	0.41	3.00	15.00	97.56	0.24	79.29	2561.27	
126	0.41	3.00	15.00	97.56	0.24	79.29	2476.00	
127	0.54	3.00	10.00	74.49	0.24	79.29	2055.00	
128	0.54	3.00	10.00	74.49	0.24	79.29	1847.35	
129	0.54	3.00	10.00	74.49	0.24	79.29	1892.10	
130	0.54	3.00	10.00	74.49	0.24	79.29	1780.10	
131	0.41	3.00	10.00	97.56	0.24	79.29	1391.55	
132	0.41	3.00	10.00	97.56	0.24	79.29	1699.98	
133	0.41	3.00	10.00	97.56	0.24	79.29	1655.81	
134	0.41	3.00	10.00	97.56	0.24	79.29	1906.00	
135	0.57	3.00	20.00	69.81	0.22	109.82	6270.67	
136	0.57	3.00	20.00	69.81	0.22	109.82	5827.68	
137	0.57	3.00	20.00	69.81	0.22	109.82	6323.58	
138	0.57	3.00	20.00	69.81	0.22	109.82	6167.96	
139	0.57	3.00	20.00	69.81	0.22	109.82	5137.81	
140	0.57	3.00	15.00	69.81	0.22	109.82	3597.30	
141	0.57	3.00	15.00	69.81	0.22	109.82	3366.83	
142	0.57	3.00	15.00	69.81	0.22	109.82	3319.53	
143	0.57	3.00	15.00	69.81	0.22	109.82	3290.21	
144	0.57	3.00	15.00	69.81	0.22	109.82	3733.56	
145	0.41	3.00	15.00	97.56	0.22	109.82	2713.69	
146	0.41	3.00	15.00	97.56	0.22	109.82	2873.73	
147	0.41	3.00	15.00	97.56	0.22	109.82	2990.60	
148	0.41	3.00	15.00	97.56	0.22	109.82	3071.42	
149	0.41	3.00	15.00	97.56	0.22	109.82	2779.16	
150	0.57	3.00	10.00	69.81	0.22	109.82	2401.13	
151	0.57	3.00	10.00	69.81	0.22	109.82	1951.21	
152	0.57	3.00	10.00	69.81	0.22	109.82	2099.40	
153	0.57	3.00	10.00	69.81	0.22	109.82	2063.53	
154	0.57	3.00	10.00	69.81	0.22	109.82	2373.30	
155	0.41	3.00	10.00	97.56	0.22	109.82	2022.40	

Accepted / Not Edited

156	0.41	3.00	10.00	97.56	0.22	109.82	1604.40	(Ataee et al., 2024)
157	0.41	3.00	10.00	97.56	0.22	109.82	1550.00	
158	0.41	3.00	10.00	97.56	0.22	109.82	1816.30	
159	0.41	3.00	10.00	97.56	0.22	109.82	1846.00	
160	0.54	4.00	20.00	74.49	0.42	47.70	3727.65	
161	0.54	4.00	20.00	74.49	0.42	47.70	3563.43	
162	0.54	4.00	20.00	74.49	0.42	47.70	4148.95	
163	0.54	4.00	20.00	74.49	0.42	47.70	2692.52	
164	0.54	4.00	20.00	74.49	0.42	47.70	2997.06	
165	0.41	4.00	20.00	97.56	0.42	47.70	3249.02	
166	0.41	4.00	20.00	97.56	0.42	47.70	3432.85	
167	0.41	4.00	20.00	97.56	0.42	47.70	3279.87	
168	0.41	4.00	20.00	97.56	0.42	47.70	3078.00	
169	0.54	4.00	15.00	74.49	0.42	47.70	2209.88	
170	0.54	4.00	15.00	74.49	0.42	47.70	2632.46	
171	0.54	4.00	15.00	74.49	0.42	47.70	2587.98	
172	0.54	4.00	15.00	74.49	0.42	47.70	2785.41	
173	0.54	4.00	15.00	74.49	0.42	47.70	2582.70	
174	0.41	4.00	15.00	97.56	0.42	47.70	2498.40	
175	0.41	4.00	15.00	97.56	0.42	47.70	2512.63	
176	0.41	4.00	15.00	97.56	0.42	47.70	2185.30	
177	0.41	4.00	15.00	97.56	0.42	47.70	2403.65	
178	0.54	4.00	10.00	74.49	0.42	47.70	1408.20	
179	0.54	4.00	10.00	74.49	0.42	47.70	1421.60	
180	0.54	4.00	10.00	74.49	0.42	47.70	1182.30	
181	0.54	4.00	10.00	74.49	0.42	47.70	1630.70	
182	0.54	4.00	10.00	74.49	0.42	47.70	1615.85	
183	0.41	4.00	10.00	97.56	0.42	47.70	1509.24	
184	0.41	4.00	10.00	97.56	0.42	47.70	1258.68	
185	0.41	4.00	10.00	97.56	0.42	47.70	1476.48	
186	0.41	4.00	10.00	97.56	0.42	47.70	1435.11	
187	0.41	4.00	10.00	97.56	0.42	47.70	1249.16	
188	0.54	4.00	20.00	74.49	0.24	79.29	4449.80	
189	0.54	4.00	20.00	74.49	0.24	79.29	4687.62	
190	0.54	4.00	20.00	74.49	0.24	79.29	4511.19	
191	0.54	4.00	20.00	74.49	0.24	79.29	4983.81	
192	0.54	4.00	20.00	74.49	0.24	79.29	4724.97	
193	0.41	4.00	15.00	97.56	0.24	79.29	2849.30	
194	0.41	4.00	15.00	97.56	0.24	79.29	2810.30	
195	0.41	4.00	15.00	97.56	0.24	79.29	2838.80	
196	0.41	4.00	15.00	97.56	0.24	79.29	2781.80	
197	0.54	4.00	10.00	74.49	0.24	79.29	2512.07	
198	0.54	4.00	10.00	74.49	0.24	79.29	2338.72	
199	0.54	4.00	10.00	74.49	0.24	79.29	2459.20	

Accepted / Not Edited

200	0.54	4.00	10.00	74.49	0.24	79.29	2038.70	(Ataee et al., 2024)
201	0.54	4.00	10.00	74.49	0.24	79.29	2094.34	
202	0.41	4.00	10.00	97.56	0.24	79.29	1472.96	
203	0.41	4.00	10.00	97.56	0.24	79.29	1800.36	
204	0.41	4.00	10.00	97.56	0.24	79.29	1900.76	
205	0.41	4.00	10.00	97.56	0.24	79.29	1770.04	
206	0.54	4.00	20.00	74.49	0.22	109.82	5634.20	
207	0.54	4.00	20.00	74.49	0.22	109.82	5640.26	
208	0.54	4.00	20.00	74.49	0.22	109.82	5957.49	
209	0.54	4.00	20.00	74.49	0.22	109.82	5870.50	
210	0.54	4.00	15.00	74.49	0.22	109.82	4360.80	
211	0.54	4.00	15.00	74.49	0.22	109.82	4857.50	
212	0.54	4.00	15.00	74.49	0.22	109.82	4015.50	
213	0.54	4.00	15.00	74.49	0.22	109.82	4276.13	
214	0.54	4.00	15.00	74.49	0.22	109.82	3166.40	
215	0.41	4.00	15.00	97.56	0.22	109.82	2931.14	
216	0.41	4.00	15.00	97.56	0.22	109.82	3337.38	
217	0.41	4.00	15.00	97.56	0.22	109.82	3064.11	
218	0.41	4.00	15.00	97.56	0.22	109.82	3215.70	
219	0.41	4.00	15.00	97.56	0.22	109.82	2862.26	
220	0.54	4.00	10.00	74.49	0.22	109.82	2604.80	
221	0.54	4.00	10.00	74.49	0.22	109.82	2643.70	
222	0.54	4.00	10.00	74.49	0.22	109.82	2378.40	
223	0.54	4.00	10.00	74.49	0.22	109.82	2600.60	
224	0.54	4.00	10.00	74.49	0.22	109.82	2459.30	
225	0.41	4.00	10.00	97.56	0.22	109.82	1784.31	
226	0.41	4.00	10.00	97.56	0.22	109.82	1964.83	
227	0.41	4.00	10.00	97.56	0.22	109.82	1908.78	
228	0.41	4.00	10.00	97.56	0.22	109.82	1702.46	

Accepted / Not Edited