Prediction of Bidirectional Shear Strength of Rectangular RC Columns Subjected to Multidirectional Earthquake Actions for Collapse Prevention

Article

Prediction of Bidirectional Shear Strength of Rectangular RC Columns Subjected to Multidirectional Earthquake Actions for Collapse Prevention

Yingbo Pang [1], **Iftikhar Azim** [2], **Momina Rauf** [3], **Muhammad Farjad Iqbal** [2,4], **Xinguang Ge** [2],
Muhammad Ashraf [4], **Muhammad Atiq Ur Rahman Tariq** [5,6] **and Anne W. M. Ng** [5,6,*]

[1] College of Civil Engineering and Architecture, Guangxi Vocational Normal University,
Nanning 530007, China; gxopalo@163.com
[2] State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering,
Shanghai Jiao Tong University, Shanghai 200240, China; iftikhar.azim@sjtu.edu.cn (I.A.);
farjad51@sjtu.edu.cn (M.F.I.); gxgzlr.2008@163.com (X.G.)
[3] Department of Civil Engineering, Military College of Engineering, NUST, Risalpur 23200, Pakistan;
mominababri@gmail.com
[4] Department of Civil Engineering, GIK Institute of Engineering Sciences and Technology, Topi,
Swabi 23460, Pakistan; matanoli@giki.edu.pk
[5] Institute for Sustainable Industries & Liveable Cities, Victoria University, Melbourne 8001, Australia;
atiq.tariq@yahoo.com
[6] College of Engineering, IT & Environment, Charles Darwin University, Darwin 0810, Australia
[*] Correspondence: anne.ng@cdu.edu.au

Abstract: The understanding of the effects of multidirectional loadings imposed on major load bearing elements such as reinforced concrete (RC) columns by seismic actions for collapse prevention is of utmost importance, and a few simplified models are available in the literature. In this study, the distinguishing features of two machine-learning (ML) methods, namely, multi expression programming (MEP) and adaptive neuro-fuzzy inference system (ANFIS) are exploited for the first time to develop eight novel prediction models (M1-to M4-MEP and M1-to M4-ANFIS) with different combinations of input parameters to predict the biaxial shear strength of RC columns (V). The performance of the developed models was assessed using various statistical indicators and by comparing them with the experimental values. Based on the statistical analysis of the developed models, M1-ANFIS and M1-MEP performed very well and exhibited the best overall efficiency of the studied ML methods. Simple mathematical formulations were also provided by the MEP algorithm for the prediction of V, using which the M1-MEP model was finalized based on its performance, accuracy, and generalization capability. A parametric analysis was also performed for the model to show that the mathematical formulation provided by MEP accurately represents the system under consideration and is imperative for prediction purposes. Based on its performance, the model can thus be recommended to update the current code provisions and engineering practices.

Keywords: evolutionary algorithms; predictive modelling; biaxial shear strength; RC columns

1. Introduction

Studying the structural behavior of RC columns subjected to biaxial seismic conditions is of great importance. Damage caused by bidirectional seismic loading is very dangerous in comparison to unidirectional seismic loading, as demonstrated by various investigations carried out on earthquakes and experimental studies for collapse prevention. Shear failure can occur in columns (especially in short columns) due to seismic loading, which can even trigger progressive collapse in structures [1]. Thus, prediction models are necessary to accurately predict the shear strength of RC columns and assist in the improvement of design methodologies and the evolution of codes.

Numerous experimental studies have been conducted to study the cyclic behavior of RC columns under biaxial cyclic loading with constant and variable axial loads. Recently, Shi et al. [2] investigated the responses of seven high-strength RC column specimens subjected to both reversed cyclic flexure and constant axial compressive loads with a new arrangement of high-grade transverse reinforcement. An improvement in the confined specimens' ductility along with a change in the failure mode of columns (i.e., from brittle to ductile) was reported as compared to unconfined specimens. Lee and Han [3] investigated lightly reinforced old columns retrofitted with deficient lap splices that were subjected to unidirectional and bidirectional loadings. The effects of the type of loading and the level of axial load were evaluated in detail. It was concluded that with an increase in the level of axial load, the bidirectional specimens received more damage. Pham and Li [4] carried out experimental and numerical studies on RC columns with lightly transverse reinforcement with stress on the effects of the directions of seismic loading on the failure mechanism of the specimens. They also concluded that the loading direction significantly affect the seismic failure mechanisms of RC columns. Rodrigues et al. [5] demonstrated that the nonlinear behavior and shear capacity of columns is greatly affected by the variation in axial load in combination with biaxial loading. Similarly, other researchers [6] also reported that the inelastic response of RC columns highly depends upon the axial load coupled with the horizontal cycle actions.

Some researchers proposed various modeling strategies to estimate the bidirectional shear strength of columns. In the inelastic behavior of RC elements in member-type modeling, a one-to-one correspondence between structural members and the elements exist. As for the simplest case of non-shear critical structural members subjected to uniaxial bending with no axial load, a 1-component lumped, or inelasticity form reached a level of maturity. Based on Classical/Bounding Surface Plasticity/Multi-Surface, comparatively better models have been developed to date for biaxial bending with a varying axial load in the literature. Nevertheless, a few unsolved numerical problems and the high number computations limit the applications of such methods. Mark [7] presented a revised truss model by considering a spatial distribution that permits the consideration of stresses in concrete, transverse, and longitudinal reinforcements with variations in the shear load inclination. A generalized formula for biaxial shear was also proposed. Galal and Ghobarah [8] presented a global element to model/simulate the biaxial shear and flexure behavior of columns subjected to varying axial loads based on the plasticity theory. The model accounted for stiffness degradation upon load reversals. The model was verified using experimental results. It was shown that it can be used to predict the behavior of axially loaded RC columns subjected to variable axial loads. Based on the analysis of the literature [9–11], it can be noted that there are difficulties in formulating simple and consistent models for the bidirectional shear strength of RC columns. Hence, simple and accurate models based on soft-computing techniques are required to predict the biaxial shear strength (V) of RC columns.

Recently, Murad [12] developed two models (with different input parameters) to predict the biaxial shear strength of RC columns based on a Gene Expression Programming (GEP) algorithm. The models were evaluated via a comparison with the experimental data using various statistical indicators. It was shown that the models can be used to predict the shear strength of RC columns; however, the GEP technique has certain limitations. A previous study concluded that GEP is highly dependent on the range of the datasets. The deviating datasets had to be excluded to avoid negative impacts on the final model efficacy [13]. Moreover, the evolutionary algorithm (EA) can consider a single expression while modelling. This limitation tends to make it suitable only for problems with a simple correlation between the input and output [14]. To cater to these issues, Multi Expression Programming (MEP), with the capability to encode multiple chromosomes in a single program, was implemented in the present study to model the biaxial shear strength of RC columns. MEP is advantageous to GEP, since it does not require the final form of the expression to be specified at the initial stage. This feature makes it suitable to model

complex engineering phenomena with high accuracy [15,16]. Moreover, it can eliminate errors, the decoding process is simple and practical expressions are developed as a final modelling result.

In this study, prediction models for estimating the biaxial shear strength of RC columns were developed by using the MEP and ANFIS algorithms with various combinations of input parameters, including the following: compressive strength of concrete, column cross-sectional area, longitudinal and shear reinforcement percentages, yield strength of rebars, column axial load, column height, width of column web, and depth of column. Based on the MEP algorithm, validated empirical equations were established and the suitability of ANFIS was also investigated for the first time. The performance of the models was evaluated using various statistical indicators. Coupled with parametric and in-depth statistical analyses, a model was finalized. The developed models were then validated with the experimental data for predicting the uniaxial shear strength of RC columns. The development of such reliable models will by-pass the need for complex experiments and can be beneficial in the pre-design stage. Further, the methodology presented in this study will pave the way for accurately predicting analogous complex engineering phenomena.

1.1. MEP

In this section, an overview of the MEP and ANFIS algorithms is presented. Based on the Darwinian principle of natural selection, an extension of the genetic algorithm (GA) was proposed by Koza [17]. The difference between the two techniques was the utilization of non-linear parse tree in genetic programming (GP) as opposed to fixed length strings in GA. However, the basic functioning of these genetic algorithm remained the same, i.e., reproduction, mutation, and crossover, with minor modifications in each process. The modelling process initiates by specifying fitness functions, process parameters and terminal conditions. A random population of parse tree was developed to ensure the models were diverse and to consider as many scenarios as possible [18]. The trees with poor fitness were excluded at the implementation stage. The best fit individuals were selected as parents to reproduce new individuals based on pre-defined criteria. The process developed is cyclic and continues until the best model is achieved. Although the GP-based models serve the primary purpose of empirically correlating different parameters, there are certain drawbacks associated with using this technique. The process is time consuming since only the tree cross-over operator is utilized, resulting in the development of a large population. The proposed expressions are often complex for practical applications since the GP works simultaneously as the genotype and phenotype. The complexity of output must be pre-specified irrespective of the nature of correlation among parameters. Moreover, a high fit model can be achieved as the number of genes increases up to a certain extent. Afterwards, it becomes impossible to increase the number of chromosomes without compromising the simplicity of the final expression.

To address the shortcomings of these GAs, MEP was proposed [14] in which the individuals can be represented as linear variable length entities enabling it to distinguish between genotype and phenotype [19]. MEP has the ability to search for a greater number of solutions due to its capability of encoding multiple solutions in a single chromosome. This feature makes the technique effective, particularly as the complexity of the problem is unknown. A common problem in genetic algorithms is their inability to manage mathematical exceptions. For instance, division by zero, invalid expressions, exceeding storage capacity. However, as the exception is generated by a gene, it mutates into a randomly chosen terminal symbol due to which no infertile individuals enter the next generation in MEP. This allows for a margin of change in the chromosome structure during evaluation [14]. The modelling process of MEP involves the random generation of a population of chromosomes, selection using the binary procedure, recombining, generation, mutation, and the replacement of worst-fit individuals. The process continues until the solution converges based on fitness criteria [20]. The process is summarized in Figure 1.
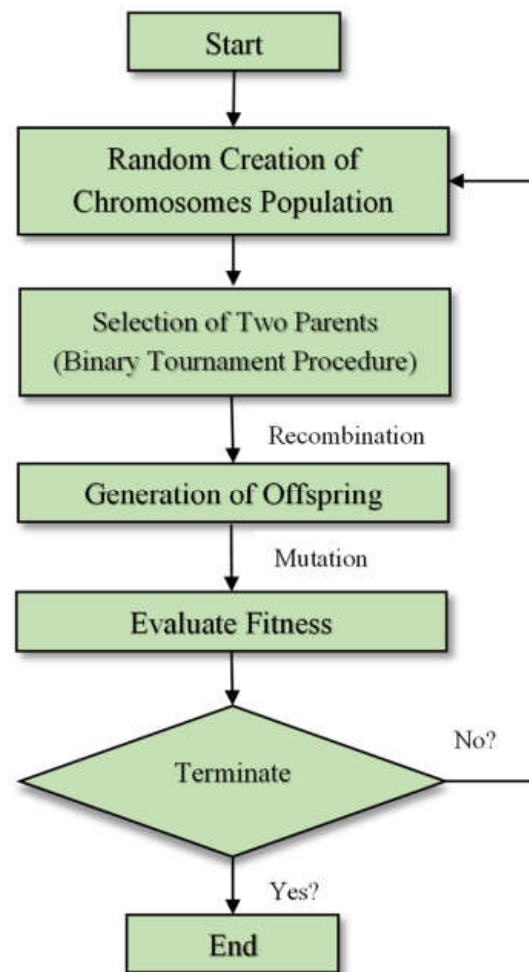
Figure 1. Schematic diagram of MEP algorithm.

It should be noted that MEP is decoded in a similar manner as the C and Pascal compilers, which translate an empirical relationship to the machine coding [21]. The output from MEP simulation can be defined as a linear string of instructions which are a combination of variables (terminal) or mathematical operators (functions). The number of genes per chromosomes governs the length of chromosomes, while the elements in the terminal and function set are encoded by the gene. The structure of the chromosome is such that the first symbol is the terminal symbol. A function gene has pointers towards function arguments, and function parameters have indices of values smaller than the position of the function in the chromosomes [20]. A better understanding of MEP can be achieved with the example described below using a set of mathematical functions F = {+, ×, ^} and a set of terminals T = {Z_1, Z_2, Z_3, Z_4}. The combination of chromosomes in the example is as follows [22]:

0: Z_1
1: Z_2
2: + 0,1
3: Z_3
4: × 2,3
5: Z_4
6: ^ 4,5

To decode the chromosome, it should be read from the top node to the bottom node. It can be seen from the above example that the genes 0, 1, 3 and 5 encode a simple expression

with a single terminal, i.e., Z_1, Z_2, Z_3 and Z_4, respectively [22]. The expressions for these genes are shown in Equation (1).

$$G_0 = Z_1; \; G_1 = Z_2; \; G_3 = Z_3; \; G_5 = Z_4 \tag{1}$$

Gene 2 (G_2) shows the operator + on the operands located at the position 0 and 1 on the chromosomes. Similarly, Gene 4 (G_4) and Gene 6 (G_6) represent the operator $\times$ and ^ on the operands located at positions 2, 3 and 4, 5, respectively [22]. The expressions encoded by these genes are shown in Equations (2)–(4).

$$G_2 = Z_1 + Z_2 \tag{2}$$

$$G_4 = (Z_1 + Z_2) \, Z_3 \tag{3}$$

$$G_6 = [(Z_1 + Z_2) \, Z_3]^{Z_4} \tag{4}$$

Each MEP chromosomes is encoded with several expressions specified by the length of the chromosomes.

Therefore, each MEP chromosome is viewed as a forest of tree rather than a single tree. Figure 2 shows the forest of expressions encoded by the above-mentioned MEP chromosome. It is pertinent to mention that each of these expressions can be a possible solution to the problem. Furthermore, the fitness of a particular expression in an MEP chromosome is defined as the fitness of best expression in that chromosome [14,20].
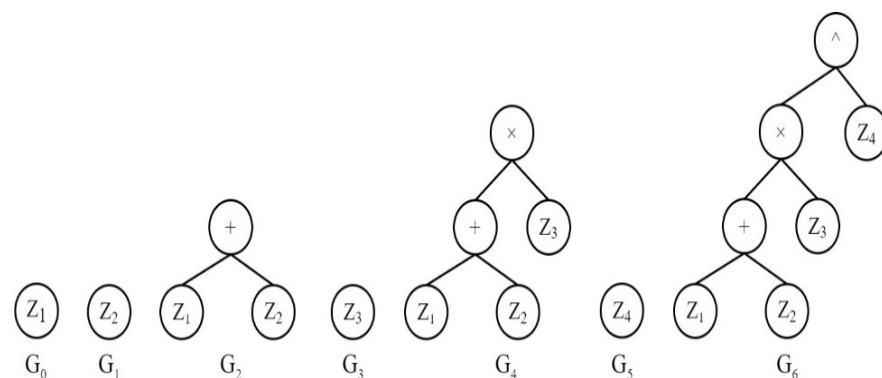

Figure 2. Representation of expressions encoded by MEP chromosomes as forest trees.

1.2. ANFIS

ANFIS is a widely used soft-computing technique that is based on the principles of the artificial neural network (ANN) and Sugeno-type rules of if–then, which can combine itself and human expertise. It blends the Fuzzy Inference System (FIS) and the learning ability of ANN. Identical to ANNs, ANFIS learns with training data by utilizing complex mathematical models. The results are projected into a FIS [23]. It is a group of multi-layer feedforward adaptive networks that estimates continuous functions accurately. This is because the algorithm is composed of five layers, wherein the nodes of every layer are connected to another layer by direct links. Thus, to produce the output for a single node, every node performs a specific function for the incoming signals.

The primary aim of the ANFIS is to ascertain the optimum values of the equivalent FIS parameters by employing a learning algorithm. Figure 3 depicts the architecture of ANFIS. There are five layers in ANFIS, namely, fuzzification, product, normalized, de-fuzzification, and the output layer.
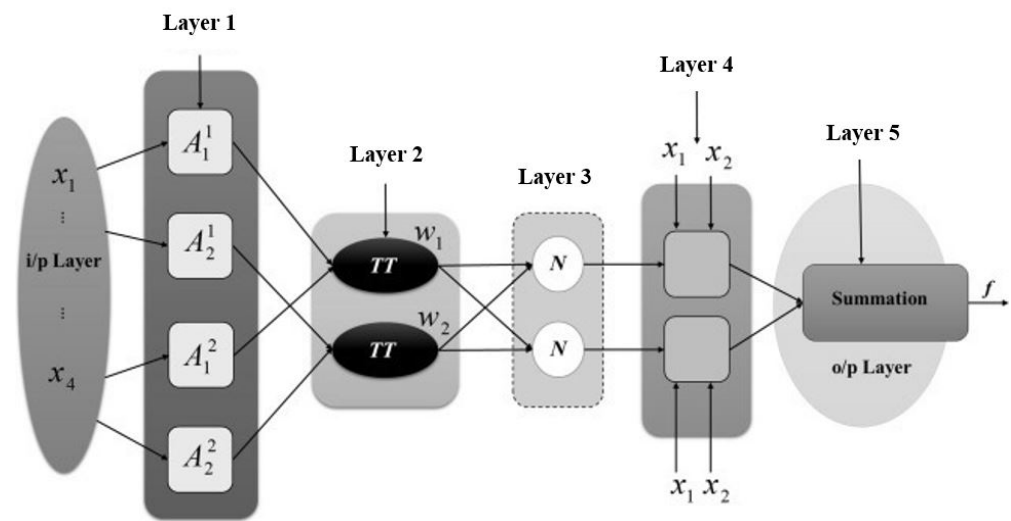
Figure 3. Sample architecture of an ANFIS network with two inputs and one output [24].

Layer 1 is an adaptive layer as the parameters involved need to be adjusted in the training process. In this layer, every node "*i*" represents an adaptive node with a node membership function as shown in Equations (5) and (6) below [24]:

$$O_i^1 = \mu_{Ai}(x_1), i = 1, 2, \ldots, n \tag{5}$$

$$O_i^1 = \mu_{Bi-2}(x_2), i = 3, 4, \ldots, n \tag{6}$$

Layer 2 calculates the firepower of the predetermined fuzzy rules via $\prod$ operator. It is a non-adaptive layer [24]. For the ANFIS network shown in Figure 3, we have:

$$O_i^2 = w_i = \mu_{Ai}(x_1) \times \mu_{Bi}(x_2), i = 1, 2, \ldots, n \tag{7}$$

Layer 3 is also a non-adoptive layer, which calculates the firepower of a rule from layer 2. Every node in this layer is a fixed node labelled as 'N' [24]. The outputs are the normalized firing power of the rules expressed, as given in Equation (8) below:

$$O_i^3 = \overline{w}_i = \frac{w_i}{\sum w_i}, i = 1, 2, \ldots, n \tag{8}$$

Layer 4 is an adaptive layer in which each node represents a consequent part of the fuzzy rule. The outputs are the products of the weights (normalized) into the node's function, which can be represented as given below [24]:

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i \{ p_k(x_1) + q_k(x_2) + r_k \}, i = 1, 2, \ldots, n \tag{9}$$

Layer 5 is a non-adaptive layer whose output is the final output of the ANFIS network. For the network shown in Figure 3 with one output, defuzzification is performed by the node ($\sum$), which is shown below [24]:

$$O_i^5 = f = \sum_{i=1}^{n} \overline{w}_i f_i = \sum_{i=1}^{n} \overline{w}_i \{ pk(x_1) + qk(x_2) + rk \} \tag{10}$$

Generally, GEP or ANN are utilized for machine-learning-based modelling in the realm of civil engineering [13,25–33]. The obvious advantages of MEP and ANFIS over other EA's and ANNs would result in the development of accurate models. They have been utilized in a very few studies to predict the mechanical properties of concrete, classify soils, constitutive models for the deformation modulus of soil, design models for columns, and the uplift capacity of suction caissons [16,34–37]. Due to their ability to create accurate models when the complexity of target expression is unknown (common phenomena in

the field of structural engineering) and the ability to search in wider space to predict the output, they can be used to model complex structural behavior. In this study, MEP and ANFIS were utilized to develop models for predicting the shear strength of biaxial-loaded columns. The efficacy of the developed models was assessed based on a rigorous statistical analysis. These studies also ensured that the models apprehended the underlying physical phenomena of the problem in hand. Moreover, a comparative study was performed based on the published experimental and modelling results to verify the universality of the proposed models.

2. Methodology

2.1. Fitting Parameters

Several trials were performed in this study and the final settings for each model were selected based on the overall high performance and lowest statistical errors of the models. The fitting parameters for the MEP- and ANFIS-based models based on the distribution of data and required accuracy are given below:

2.1.1. MEP Parameters

The settings of the four finalized models are summarized in Table 1. The number and size of the subpopulation are highly sensitive parameters that control the overall complexity and the accuracy of the models. A run with a larger value of these two parameters would require significant time to converge and would produce an accurate model. However, the issue of over-fitting may arise, and the model may not perform well on the un-seen data. The number of generations in the M1-MEP were selected by monitoring the fitness function. It was noted that no significant improvement was observed in the model beyond 1000 generations. Therefore, it was considered as an optimal value. However, M2-MEP, M3-MEP, and M4-MEP showed significant improvement in the correlation coefficient beyond 1000 generations. Therefore, 2000, 5000 and 3000 were selected as the optimal number of generations for these models, respectively. The mutation and crossover rate were kept as 0.01 and 0.9, respectively, to ensure the probability of offspring to undergo these operations while modelling [34]. Additionally, the code length was kept at 45, 35, 30, and 35 for M1-MEP, M2-MEP, M3-MEP, and M4-MEP, respectively. However, the final model was simplified by applying the basic rules of mathematics. Several combinations of these settings were implemented on the training data and the final settings for each model are summarized in Table 1.

Table 1. Parameter settings for MEP algorithm.

Parameters	Settings			
	M1-MEP	**M2-MEP**	**M3-MEP**	**M4-MEP**
Number of sub-population	30	30	30	30
Size of subpopulation	200	200	200	200
Code length	45	35	30	35
Crossover probability	0.9	0.9	0.9	0.9
Mathematical operators	$+, -, \times, \div, \sqrt{}$	$+, -, \times, \div, \sqrt{}$	$+, -, \times, \div, \sqrt{}$	$+, -, \times, \div, \sqrt{}$
Mutation probability	0.01	0.01	0.01	0.01
Tournament size	4	4	2	2
Operators	0.5	0.5	0.5	0.5
Mutation probability	0.5	0.5	0.5	0.5
Number of generations	1000	2000	5000	3000

2.1.2. ANFIS Parameters

The ANFIS modeling limits the response to one parameter (output), similarly to the MEP algorithm. Similar input parameters were considered for ANFIS modelling as well as for all four of the models. Moreover, a uniform distribution of the dataset was ensured in the training and testing tests. The ANFIS modelling begins with the creation of an FIS that

is developed based on the grid population considering the complexity and variable range of input parameters in the four models [38,39]. A hybrid optimization methodology (back propagation and root mean square error was evaluated) was employed in the modelling. It is important to mention that the structuring triangular membership function (trimf) was used to train the models to ensure the models were accurate. The error goal was kept as 0 for all the models, whereas the number of linear and non-linear parameters, and fuzzy rules were pre-set based on the modeling technique. For instance, the fuzzy rules were kept as 729, 243, 243, and 729 for M1-ANFIS, M2-ANFIS, M3-ANFIS, and M4-ANFIS, respectively. Table 2 below presents a summary of the parameter settings for the four ANFIS models.

Table 2. Parameter settings for ANFIS algorithm.

Parameters	Settings			
	M1-ANFIS	**M2-ANFIS**	**M3-ANFIS**	**M4-ANFIS**
Linear parameters	729	243	243	729
Non-linear parameters	54	45	45	54
Fuzzy rules	729	243	243	729
Nodes	1503	524	524	1503
Epochs	50	50	50	50
Error goal	0	0	0	0
Type of MF	Trimf	Trimf	Trimf	Trimf
Structure of fuzzy	Sugeno	Sugeno	Sugeno	Sugeno
Type of FIS	Grid Partition	Grid Partition	Grid Partition	Grid Partition
Optimization technique	Backpropagation and least square	Backpropagation and least square	Backpropagation and least square	Backpropagation and least square
Type of output function	Linear	Linear	Linear	Linear

2.2. Performance Evaluation of Models

Several statistical error measures were implemented to assess the accuracy and prediction capability of the developed models. An issue of concern in AI-based modelling is that a model may satisfy one error measure while it fails another. Moreover, to limit the overfitting of data, an objective function (OF) and a performance index (ρ) were utilized and minimized for the developed models to ensure that the models performed equally well on all three datasets, i.e., training, validation, and testing. These two parameters consider the number of datapoints in each three sets and the simultaneous impact of multiple statistical measure to check the validity of the models. The expressions and the criteria of these parameters are summarized in Table 3, where e_i, m_i, $\bar{e}_i$, and $\overline{m}_i$ are the ith experimental, predicted, mean experimental, and mean predicted values, respectively, while n is the total number of data points used for modelling. The subscripts n_L and n_T represent the number of learning (training and validation) and testing datapoints, respectively. The ρ_L and ρ_T denote the performance index of learning and testing sets, respectively.

Table 3. Statistical measure and evaluation criteria.

Parameter	Expression	Criteria
Correlation coefficient (R)	$\dfrac{\sum_{i=1}^{n}(e_i-\bar{e}_i)(m_i-\overline{m}_i)}{\sqrt{\sum_{i=1}^{n}(e_i-\bar{e}_i)^2 \ \sum_{i=1}^{n}(m_i-\overline{m}_i)^2}}$	>0.8 [40]
Mean absolute error (MAE)	$\dfrac{\sum_{i=1}^{n}|e_i-m_i|}{n}$	Minimum [41]
Root mean square error (RMSE)	$\sqrt{\dfrac{\sum_{i=1}^{n}(e_i-m_i)^2}{n}}$	Minimum
Relative root mean square error (RRMSE)	$\dfrac{1}{|\bar{e}|}\sqrt{\dfrac{\sum_{i=1}^{n}(e_i-m_i)^2}{n}}$	0–0.1 (Excellent) or 0.11–0.2 (Good) [42]
Performance index (ρ)	$\rho = \dfrac{\text{RRMSE}}{1+R}$	<0.2 [34]
Objective function (OF)	$\left(\dfrac{n_L-n_T}{n}\right)\rho_L + 2\left(\dfrac{n_T}{n}\right)\rho_T$	Close to zero [25]

2.3. Modelling

The most influential input parameters were identified based on the original sources where data were collected for the purpose of modeling. As discussed above, by utilizing the MEP and ANFIS algorithms, prediction models (four each) for V were developed in this study. The basic form of V for all the models is given as:

Model 1:
$$V = f\left(f_c', \ A_g, \ \rho_l, \ \rho_w, \ f_y, \ N\right) \tag{11}$$

Model 2:
$$V = f\left(f_c', \ A_g, \ \rho_w, \ N, \ H\right) \tag{12}$$

Model 3:
$$V = f\left(f_c', \ A_g, \ \rho_w, \ f_y, \ N\right) \tag{13}$$

Model 4:
$$V = f\left(f_c', \ b_w, \ d, \rho_w, \ f_y, \ N\right) \tag{14}$$

where,

f_c' = concrete compressive strength (MPa)
A_g = gross sectional area of column (mm^2)
ρ_l = longitudinal reinforcement percentage (%)
ρ_w = shear reinforcement percentage (%)
f_y = yield strength of longitudinal reinforcement (MPa)
N = axial load of column (kN)
b_w = width of column web (mm)
d = depth of column (mm)
H = column height (mm)

Table 4 shows the descriptive statistics and limits of the input parameters used in the four models, while Table 5 shows the R values between the input parameters to check the parameters' independency as various parameters may be dependent. Interdependency, also known as multi-collinearity, leads to obstructions in the explanation of a model. If the value of $R > 0.80$, then there exists a strong correlation between two of the input parameters. For a better model, the value of R (both negative and positive) should be less than 0.80. From Table 5, it can be noted that all the values of R were less than 0.80, showing no risk of interdependency in the models.

Commercial MEPX software was used for the application of the MEP algorithm, while MATLAB was used for ANFIS. Several initial runs were performed on the database for the optimization of the empirical models. It should be noted that the whole database was randomly divided into training, validation, and testing sets. Overall, 70% of the data were used for training while 30% of the data were used in the validation and testing sets.

Table 4. Descriptive statistics of the input parameters involved in modeling.

Parameter	f'_c (MPa)	N (kN)	b_w (mm)	d (mm)	A_g (mm^2)	ρ_l (%)	f_y (MPa)	ρ_w (%)	H (mm)
Mean	19.23	8.44×10^5	1.85	28.05	1.16×10^5	1.568	440.93	0.25	1493.86
Standard Error	0.694	8.02×10^4	0.070	0.693	6.77×10^3	0.115	10.82	0.03	79.38
Median	23	1.00×10^6	1.755	27	1.20×10^5	1.056	400.00	0.13	1700.00
Mode	23	1.00×10^6	2.06	38.2	9.00×10^4	2.547	400.00	0.11	1700.00
Standard Deviation	6.13	7.08×10^5	0.615	6.12	4.34×10^4	0.739	69.30	0.18	508.28
Sample Variance	37.60	5.02×10^{11}	0.378	37.51	1.88×10^9	0.546	4801.83	0.03	258,343.76
Kurtosis	−1.1036	1.30×10^1	−0.556	−0.572	5.57	−1.312	0.45	−0.58	−0.69
Skewness	−0.710	2.83	0.379	0.120	1.37	0.724	0.08	0.94	−0.62
Minimum	8.85	1.00×10^5	0.7	15.92	4.00×10^4	0.848	276.00	0.09	570.00
Maximum	25.7	4.29×10^6	3.35	38.2	2.92×10^5	2.777	575.60	0.63	2438.40

Table 5. Correlation coefficients (*R*) between the input parameters.

Parameter	f'_c (MPa)	N (kN)	b_w (mm)	d (mm)	A_g (mm^2)	ρ_l (%)	f_y (MPa)	ρ_w (%)	H (mm)
f'_c (MPa)	1.00	0.22	−0.32	0.21	−0.11	−0.30	−0.07	−0.20	0.03
N (kN)	0.22	1.00	0.10	0.17	0.14	0.24	0.29	−0.04	0.49
b_w (mm)	−0.32	0.10	1.00	0.16	0.76	0.05	0.51	−0.28	0.36
d (mm)	0.21	0.17	0.16	1.00	0.73	−0.26	−0.32	−0.52	0.49
A_g (mm^2)	−0.11	0.14	0.76	0.73	1.00	−0.11	0.11	−0.43	0.51
ρ_l (%)	−0.30	0.24	0.05	−0.26	−0.11	1.00	0.34	0.67	−0.22
f_y (MPa)	−0.07	0.29	0.51	−0.32	0.11	0.34	1.00	0.14	0.25
ρ_w (%)	−0.20	−0.04	−0.28	−0.52	−0.43	0.67	0.14	1.00	−0.47
H (mm)	0.03	0.49	0.36	0.49	0.51	−0.22	0.25	−0.47	1.00

In the case of ANFIS modelling, the selection of the number of epochs is a primary parameter during the modelling process. The higher the number of epochs, the longer the run time and the higher the complexity of the model and vice versa. Therefore, plenty of initial trial runs were initiated to optimize the epochs for the final models. For instance, the number was kept as 50 to ensure sufficient accuracy was achieved in the final optimized models. In addition to the built-in back propagation optimization, an external statistical error analysis was performed for each trial model. It is quite common for a model to perform well on one of the error measures and fail on other criteria. Therefore, multiple errors were calculated for each developed model and models with high accuracy for each criterion were selected as the final models.

In the case of MEP, the models were initially run after considering the basic mathematical operators (addition, multiplication, and subtraction). However, the range of the experimental values of *V* was large. Therefore, the models were unable to incorporate the impact of the deviating datasets. To counteract this issue, additional operators (such as square root and division) were incorporated. As discussed above, there are different parameters that need to be specified in MEPX. These include the number of sub-populations, sub-population size, code length, cross-over probability, and the error measure. For the developed models, a trial and error-based approach was adopted to optimize the results. For instance, the code length varied in the range of 30–50. The final lengths were selected as 45, 35, 30, and 35 for Model 1, 2, 3, and 4, respectively.

It is pertinent to mention that the number of sub-populations varied from 30 to 50. Based on the statistical checks, a population size of 30 was found to be optimal for providing mathematical formulations with a high accuracy and relation among the parameters. The number of generations varied from 100 to 10,000 depending on the complexity of the dataset for each combination. Mean absolute error and correlation co-efficient were selected as the measure of error. As a rule of thumb, the minimization of these error measures was considered as a first accuracy check for the developed models. It should be noted that for each combination of the parameters, data were imported into the software. The software ran for 5 h or until there was no change in the correlation for a period of one hour. Additionally, for the four finalized models, excel sheets are provided as Supplementary Materials with this manuscript, where the derived equations for the models were applied explicitly. For every developed model, mathematical formulations were developed and parametric analyses were conducted to compare the correlation of inputs and outputs with the actual physical phenomenon. The outputs for the prediction of *V* obtained in the form of C++ code were decoded and expressed in terms of an empirical formulation such as Equations (15)–(18).

M1-MEP:

$$V \ (kN) = 2\sqrt{\frac{\rho_l}{\rho_w} - \rho_w - \sqrt{\rho_w}} + \frac{\sqrt{A_g} - \sqrt{a} + f_y + f'_c}{N - \sqrt{a}} - b' + d'\sqrt{\rho_w} + \frac{\rho_l}{\rho_w} - e \quad (15)$$

where,

$$a = [(\rho_w + \rho_l)(\rho_w + \sqrt{\rho_w})]^2 (f_y + \rho_w)$$

$$b' = \frac{\sqrt{A_g} - a + f_y + f'_c}{\left(\frac{\rho_l}{\rho_w}\right)^2}$$

$$c = \frac{\sqrt{a} + N + f'_c}{\left(\sqrt{A_g} - f_y - f'_c\right) + \left[(\rho_w + \rho_l)(\rho_w + \sqrt{\rho_w})\right]^2}$$

$$d' = (\rho_w + \rho_l)\left[\frac{\rho_l}{\rho_w} - \rho_w - \sqrt{\rho_w}\right] + c^2 + \sqrt{A_g} - \sqrt{a}$$

$$e = \frac{\left(\sqrt{A_g} - f_y - f'_c\right) + \left[(\rho_w + \rho_l)(\rho_w + \sqrt{\rho_w})\right]^2 - f_y}{(\rho_w + \rho_l)^2(\sqrt{a} + N) - c^2 - \sqrt{A_g} + \sqrt{a}}$$

M2-MEP:

$$V\,(kN) = f + 2h - g + i \tag{16}$$

where,

$$f = \frac{2f_y - 2\rho_w - A_g + 2A_g\rho_w}{N - \sqrt{A_g - A_g\rho_w}} + 2f_y - 2\rho_w - A_g + 3A_g\rho_w + 2N - \sqrt{A_g - A_g\rho_w}$$

$$g = f'_c \frac{2f_y - 2\rho_w - A_g + 3A_g\rho_w}{A_g}$$

$$h = \frac{2N - \sqrt{A_g - A_g\rho_w}}{f'_c} - \frac{2f_y - 3\rho_w}{2N - \sqrt{A_g - A_g\rho_w}}$$

$$i = \frac{N(f_y - \rho_w) + A_g - A_g\rho_w}{2f_y - 2\rho_w - A_g + 2A_g\rho_w}$$

M3-MEP:

$$V\,(kN) = \frac{4N}{f'_c} - \frac{4N}{H} + \sqrt{A_g\rho_w + H + 6N} - j - k \tag{17}$$

where,

$$j = \frac{A_g\rho_w + H + 6N}{2H - 2N} + \frac{\rho_w(A_g\rho_w + H + 6N)}{H}$$

$$k = \frac{(H - N)(2H - 2N)^2}{A_g\rho_w(A_g\rho_w + H + 6N)}$$

M4-MEP:

$$V\,(kN) = \frac{f'_c}{l} + k \tag{18}$$

where,

$$l = \frac{2d - 2f'_c + N - f_y}{f_y - b}$$

$$m = \frac{8b + 4N}{f_y - b} + \sqrt{\rho_w}(d - f'_c)$$

$$k = m + 2\frac{d}{f_y - b} - \sqrt{f_y} - \frac{(f_y - b)(2d - 2f'_c + N - l)}{8b + 4N}$$

3. Results Analysis and Discussion

3.1. ANFIS Modeling Results

Figure 4a–d shows the results of the ANFIS modeling for the training, validation, and testing datasets while Table 6 shows the statistical performance of the models. It can

be noted that the performance of the models was generally satisfactory in every stage. Multiple statistical parameters were selected to measure the performance of the models, since some parameters such as R, MAE, and RMSE are sensitive to multiplication and division, allot higher weightage to low error values, and higher weightage to high error values, respectively [23].
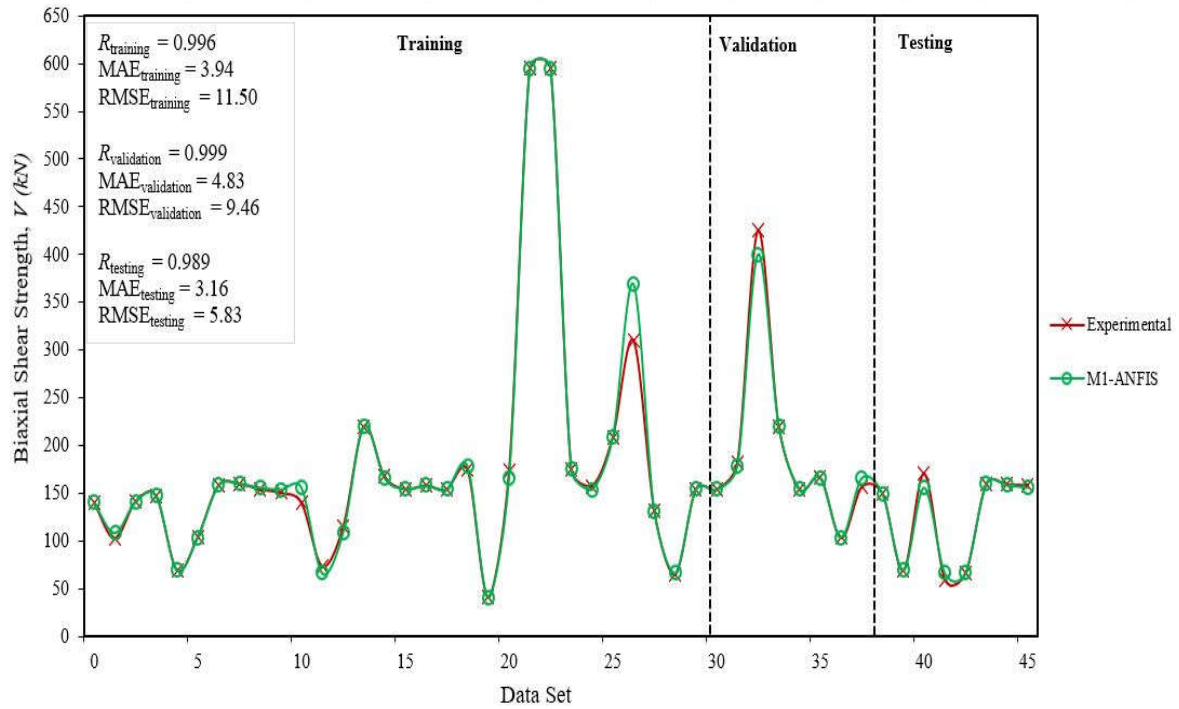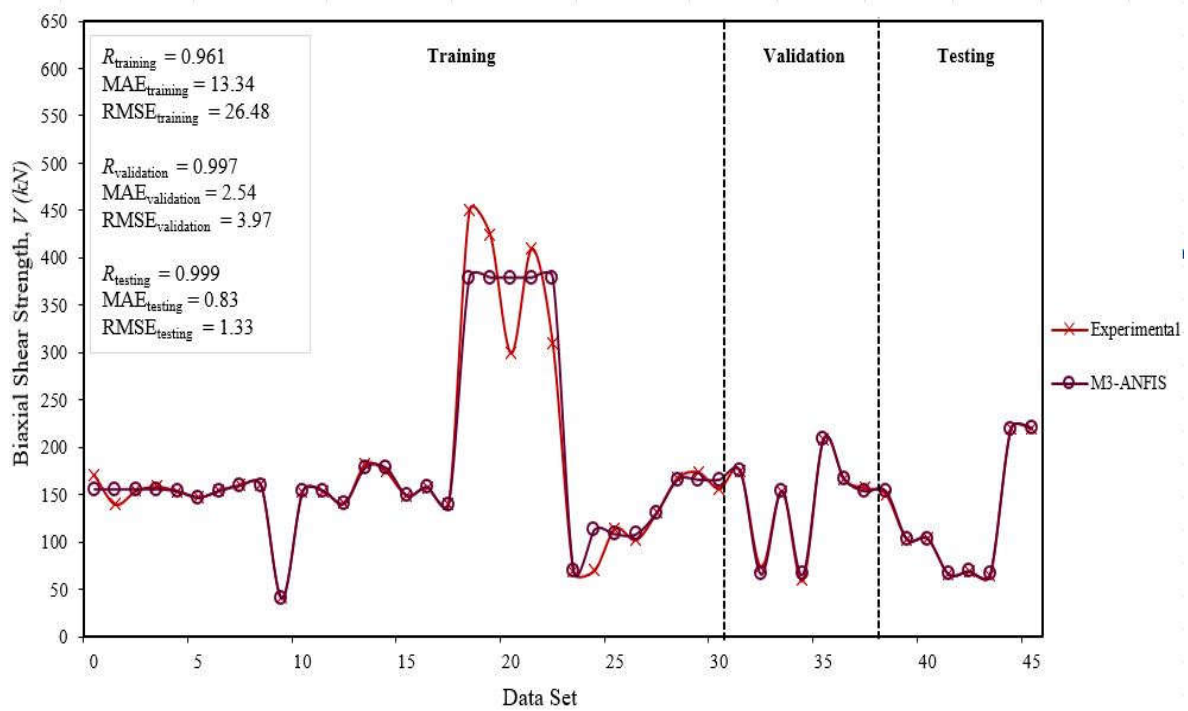

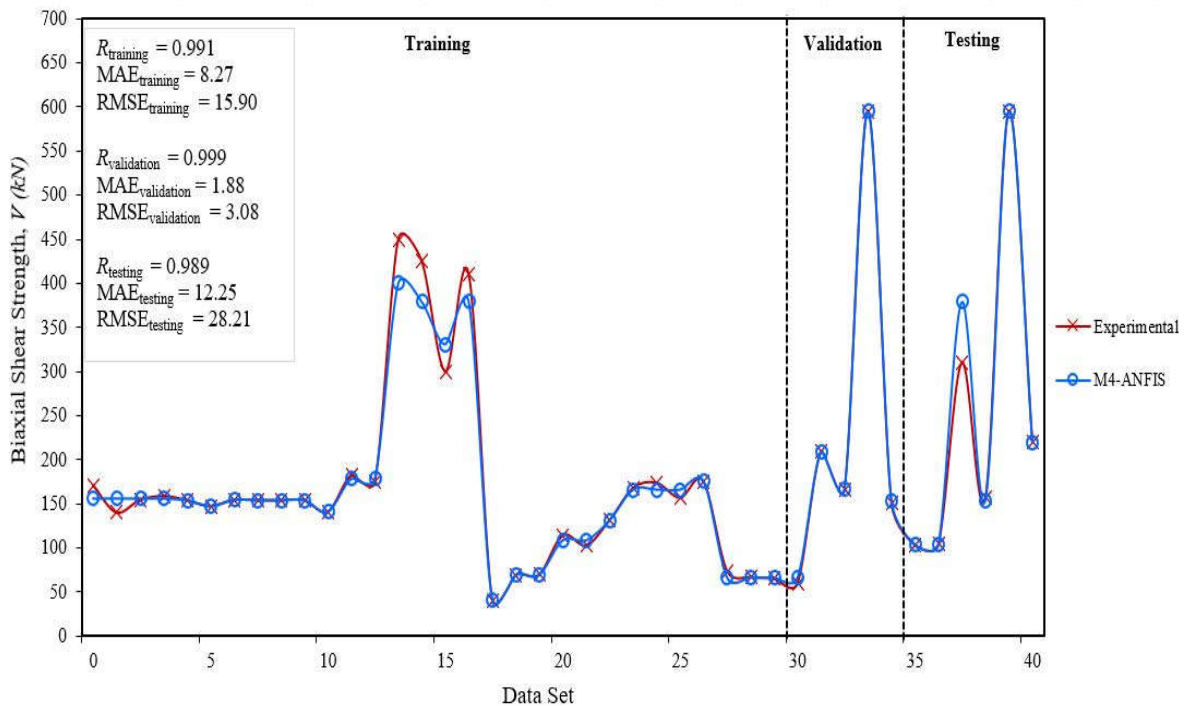(**a**)
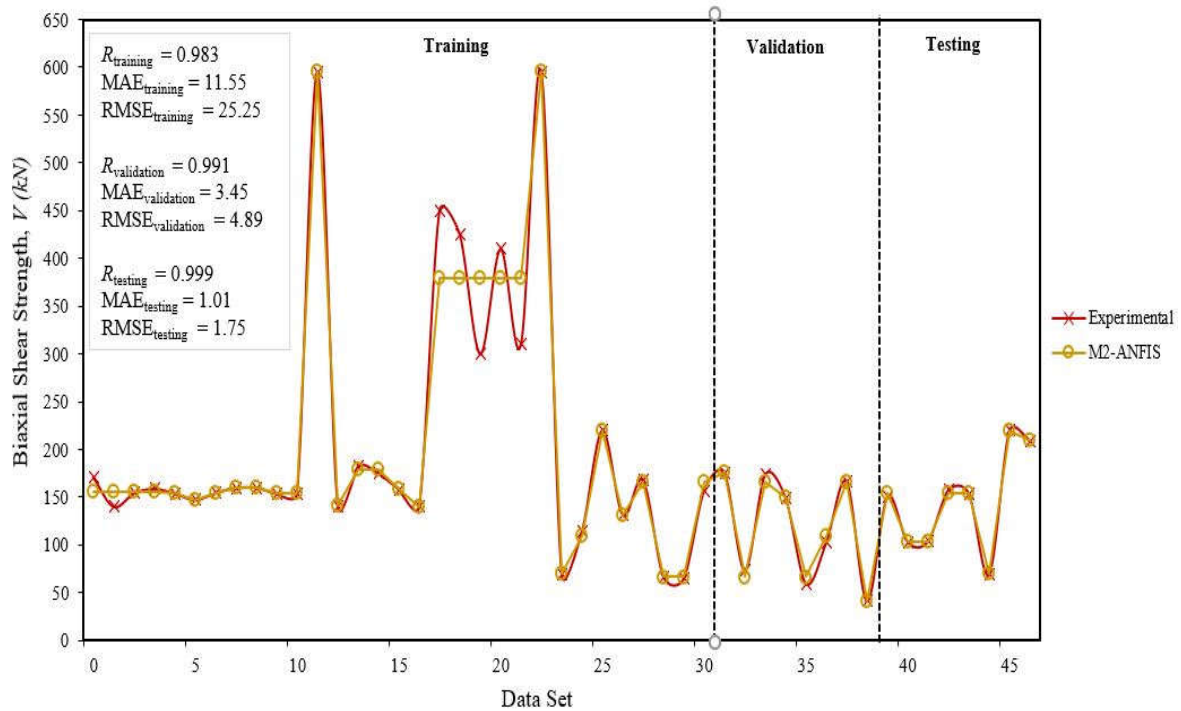
(**b**)

Figure 4. *Cont.*

(**c**)

(**d**)

Figure 4. Comparison between experimental and various ANFIS models results: (**a**) M1-ANFIS vs. experimental results, (**b**) M2-ANFIS vs. experimental results, (**c**) M3-ANFIS vs. experimental results, (**d**) M4-ANFIS vs. experimental results.

From Figure 4 and Table 6, it can be noted that the *R* values approached 1 for all the models, clearly indicating the high prediction ability of the ANFIS algorithm. However, as discussed above, *R* cannot be considered as a sole parameter with which to judge a model's performance. From the table, it can be noted that the performance of the M1-ANFIS model

stands out for all the three datasets, as it displays the highest generalization capability as compared to the remaining three models. The values of all the statistical parameters were close to each other in every dataset. The RRMSE values indicate that the M1-ANFIS model is in the 'excellent' range, as its values were in the range [0.05, 0.07] (vid. Table 3). Meanwhile, for the remaining three models, the RRMSE values were >0.10 for at least one of the sets, which falls into the 'good' range. The values of ρ were minimal and approached zero for all three datasets of all the models. Similarly, the values of OF were minimal, and in the range [0.015, 0.04] for all the models.

**Table 6.** Statistical parameters of the four ANFIS models.

| Model	Dataset	R	MAE	RMSE	RSE	RRMSE	ρ	OF
M1-ANFIS	Training	0.992	3.94	11.50	0.01	0.07	0.03	
	Validation	0.999	4.83	9.46	0.011	0.05	0.02	0.026
	Testing	0.994	3.16	5.83	0.016	0.05	0.02	
M2-ANFIS	Training	0.983	11.55	25.25	0.03	0.12	0.06	
	Validation	0.996	3.45	4.89	0.009	0.04	0.02	0.015
	Testing	0.999	1.01	1.75	0.001	0.01	0.01	
M3-ANFIS	Training	0.961	13.34	26.48	0.08	0.15	0.08	
	Validation	0.997	2.54	3.97	0.006	0.03	0.01	0.03
	Testing	0.999	0.83	1.33	0.0005	0.01	0.01	
M4-ANFIS	Training	0.991	8.27	15.90	0.03	0.09	0.05	
	Validation	0.999	1.88	3.08	0.0003	0.01	0.01	0.04
	Testing	0.989	12.25	28.21	0.027	0.11	0.06	

3.2. MEP Modeling Results

A comparison of the training, validation, and testing sets of the experimental data with the MEP models is provided in Figure 5. The datasets used in the ANFIS modeling were used for MEP modeling. The values of the statistical parameters are also shown for the three datasets of each model. From the figure and the analysis of the statistical parameters, we can note that the M1-MEP model exhibited the highest accuracy of all the models as the value of *R* was very close to 1, while the values of other parameters were low in comparison to other models. M3- and M4-MEP also exhibit excellent correlation for training, validation, and testing datasets; however, the values of other indicators such as MAE and RMSE were very high. Hence, a higher value of the correlation coefficient does not ensure that a model is good. A combination of different statistical indicators is recommended to check the performance of a model, as discussed above.

In comparison to ANFIS models, the values of *R* are very low for all the MEP models except M1-MEP, while the values of MAE and RMSE are very high. The M1-MEP model performed well as compared to the remaining three models. Hence, this model is discussed in detail below. Additionally, empirical equations and a parametric analysis were performed for this model.

Figure 6 presents a comparison of the experimental and predicted values of *V* for M1-MEP. It can be seen from the figure that the finalized model accurately incorporated the influence of all six parameters for the prediction of *V*. For all three datasets, a strong correlation was achieved as depicted by the slope of regression lines (1 for the ideal model), i.e., 0.9923, 0.9953, and 0.989 for training, validation, and testing, respectively.
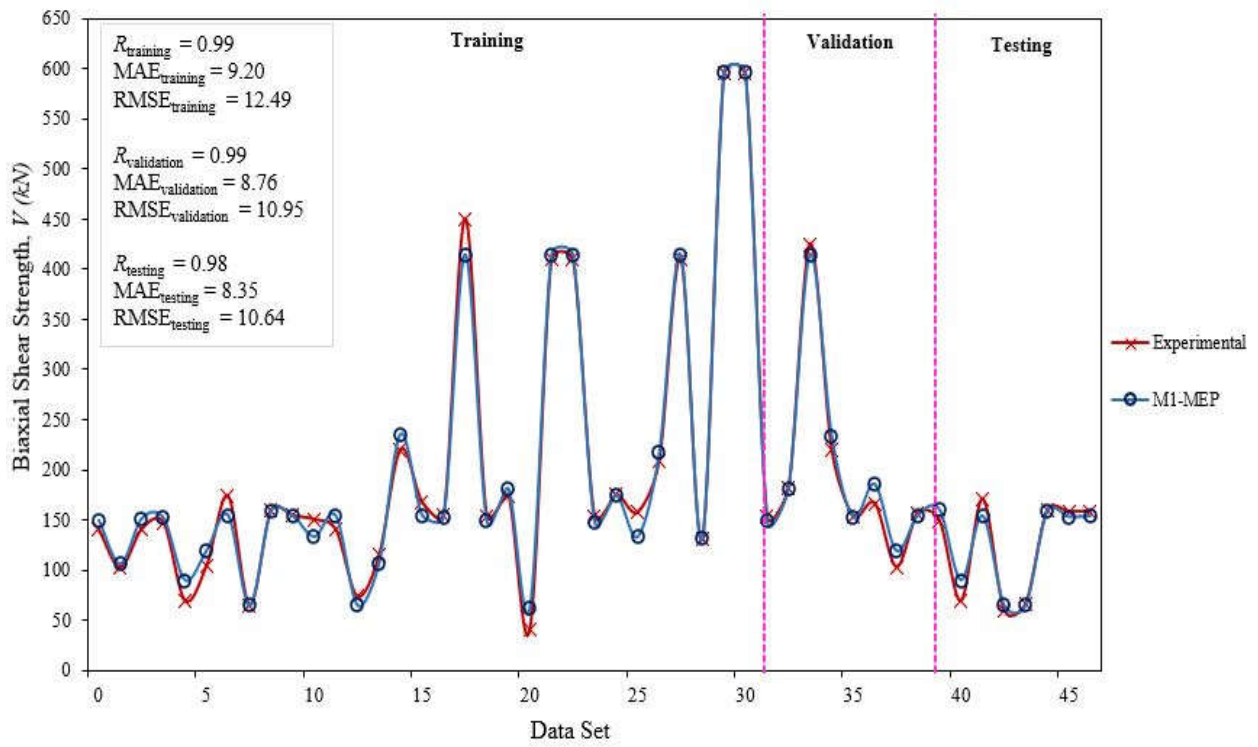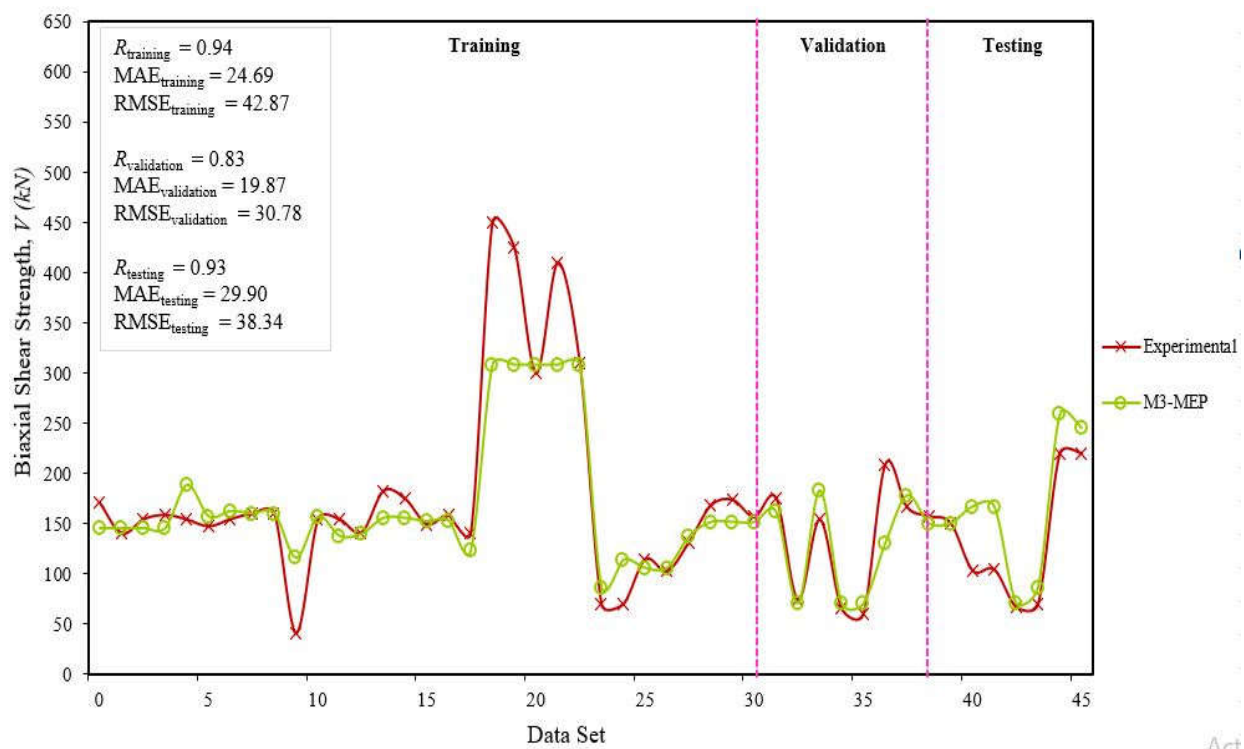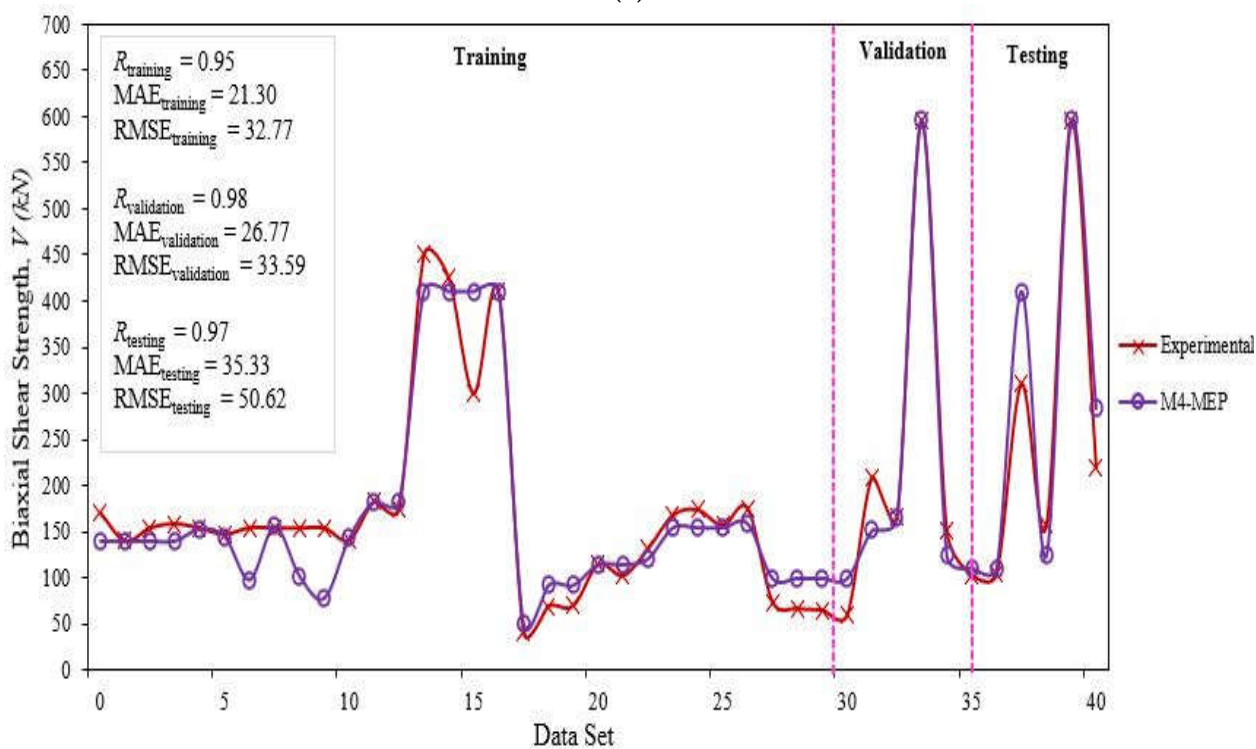
(**a**)
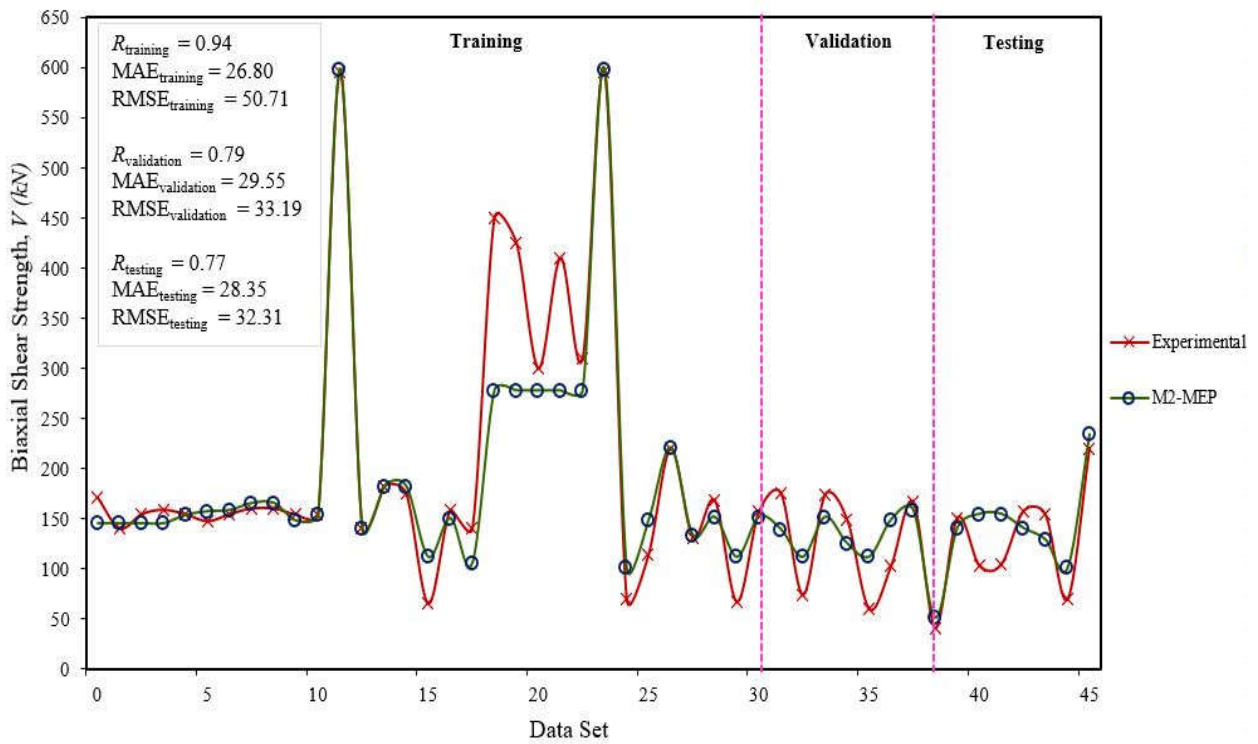
(**b**)

Figure 5. *Cont*.

(**c**)

(**d**)

Figure 5. Comparison between experimental and various MEP models results: (**a**) M1-MEP vs. experimental results, (**b**) M2-MEP vs. experimental results, (**c**) M3-MEP vs. experimental results, (**d**) M4-MEP vs. experimental results.
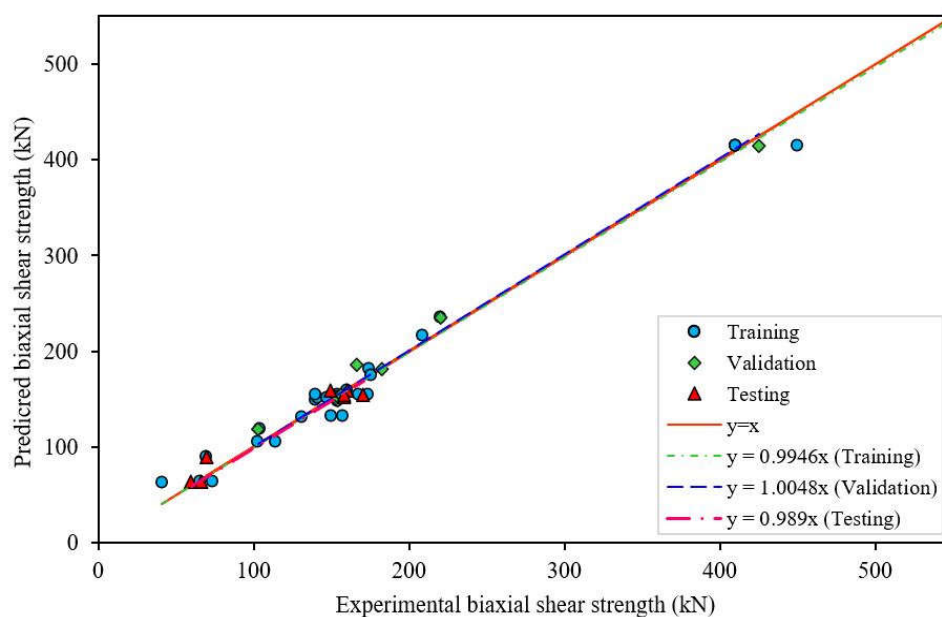
Figure 6. Experimental vs. predicted biaxial shear strength using M1-MEP.

As explained in Section 3, a rigorous statistical analysis was performed to assess the efficacy of the developed models. The results of the statistical analysis for the training, validation, and testing sets for M1-MEP are shown in Table 7. It can be inferred from the results that the model possesses the highest correlation among the predicted and the experimental values, as evident by values for *R* of 0.98, 0.97, and 0.96 for the training, validation, and testing sets, respectively. The values of MAE and RMSE were low and close to each other for all three sets, thereby indicating a high prediction and generalization capability. The MAE values were 9.20, 8.76, and 8.34 while the RMSE was recorded as 12.49, 10.95, and 10.64, respectively, for the three datasets. The RRMSE values were 0.06, 0.06, and 0.09, respectively, for the three sets. Based on these RRMSE values, the model can be categorized as being in the excellent range since the values were less than 0.10 (vid. Table 3). The values of *ρ* lie in the range of 0.03–0.04, indicating that the model satisfies the combined impact of multiple statistical checks (i.e., *R* and RRMSE). The value of OF was calculated as 0.059, which is very close to zero. This parameter validated the overall performance of the model and addressed the issue of overfitting of the data, which arises due to the superior performance on the training set and inferior performance on the testing or un-seen data. An issue of concern in the machine-learning-based modeling is that the model satisfies the statistical checks, however, the absolute errors may still be significantly large.

Table 7. Statistical indicators for training, validation and testing sets of M1-MEP.

Model	Dataset	*R*	MAE	RMSE	RSE	RRMSE	*ρ*	*OF*
Biaxial	Training	0.993	9.20	12.49	0.01	0.06	0.03	
shear	Validation	0.995	8.76	10.95	0.014	0.06	0.03	0.059
strength	Testing	0.977	8.35	10.64	0.054	0.09	0.04	

To study the absolute errors, the results of the M1-MEP are graphically represented in Figure 7. It can be seen from the figure that the mean error in the predicted values was approximately 5% with an absolute error for most of the dataset of less than 10 kN. Moreover, only two datapoints presented an error of greater than 10 kN which accounts for 4% of the total database used for modelling. The frequency of occurrence of maximum errors was significantly low. For instance, 90% of the datapoints had error of less than 6 kN, which is a testimony of the accuracy of the proposed formulation. The amount of data used for modelling is also an indicator of reliability and generalization of the model. It is

recommended in the literature that the number of datapoints should be greater than five times the number of input variables for the development of an effective model [43]. In this study, the number was 7.83, thereby meeting the stated criteria.
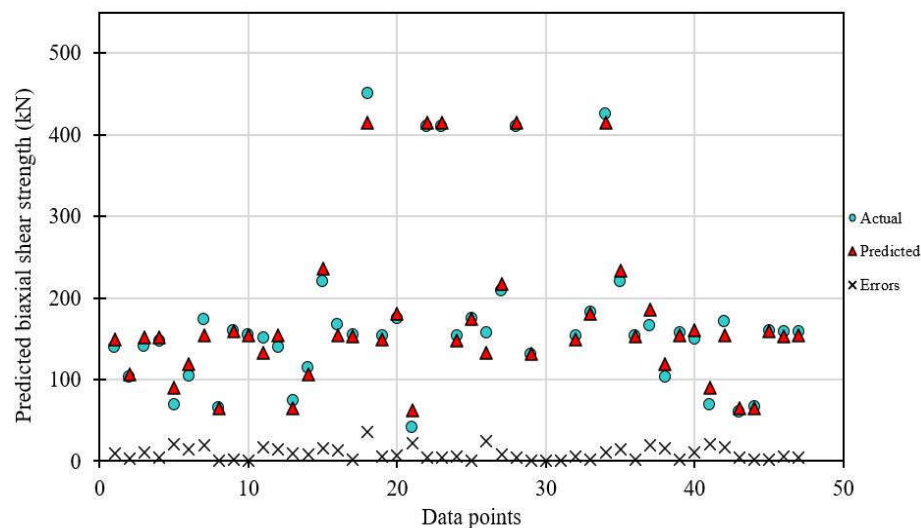

Figure 7. Graphical representation of absolute error between experimental and predicted values.

ACI-318 [44] proposed formulations for the prediction of unidirectional shear strength of columns. A comparative study was conducted by Murad et al. [12], wherein the results of the two GEP-based models for predicting the bi-directional shear strength predictions were compared with those predicted using ACI formulations. It was concluded that the *R* for the shear strength predicted by the ACI formulation was low, i.e., 0.592. The values of *R* were reported to be 0.932 and 0.854 for the two proposed GEP models. As shown above, M1-ANFIS and M1-MEP exhibited superior performance in this study. The values of *R* for the learning, validation, and testing datasets of M1-ANFIS and M1-MEP were 0.993, 0.995, 0.977, and 0.992, 0.999, 0.994, respectively, while the values of the other parameters such as MAE, RMSE, RRMSE, and ρ for the both the models were also lower as compared with the other models available in the literature. Therefore, it can be concluded that ANFIS and MEP are superior in performance compared to GEP and ACI-318. However, one of the drawbacks of the ANFIS algorithm is that since it belongs to the category of black-box models, its interpretation is not easy as compared to the MEP algorithm, which belongs to the grey-box category [45]. As for the MEP algorithm, the equations are lengthy as compared to GEP. For this purpose, excel sheets are provided, which can help the readers to apply the models with ease.

4. Parametric Analysis of MEP-Based Models

A parametric analysis was carried out to show that the proposed model represents the physics of the problem under consideration well and that it is not merely a combination of the independent variables. The procedure presented in [13,25] was adopted to carry out the analysis. The results of the analysis are shown in Figure 8. It must be noted that a parametric analysis was performed for all the models. Similar trends were observed for the common input variables among all the models. Hence, the trend of the six parameters of the best model, i.e., M1-MEP, is shown along with the remaining variables (*H*, b_w, and *d*) used in the other models in Figure 8. From the figure, it can be noted that *V* increased with an increase in all the input parameters for the four models. From Figure 8a,b, we can see that *V* increased linearly with an increase in f_c' and *N*, respectively. An increase in f_c' and *N* proved to be beneficial to the shear strength of the RC columns, and the observations are valid from an engineering viewpoint and conforms with the conclusions obtained by [4,46]. Similarly, *V* increased due to an increase in the percentage of both longitudinal

and transverse reinforcement in the specimens (vid. Figure 8d,e). This can be attributed to the fact that an increasing ρ_l allows the specimens to sustain greater damage in the core, while an increasing ρ_w confines the longitudinal reinforcement and prevents it from buckling [47,48]. Lastly, the biaxial shear strength increased due to increasing A_g and f_y, which is also correct from an engineering viewpoint. Hence, the results of the parametric analyses of the models show that the MEP algorithm accurately presents the system under consideration. The results of the parametric study of the common parameters in this study agree well with the results of Murad et al. [12].
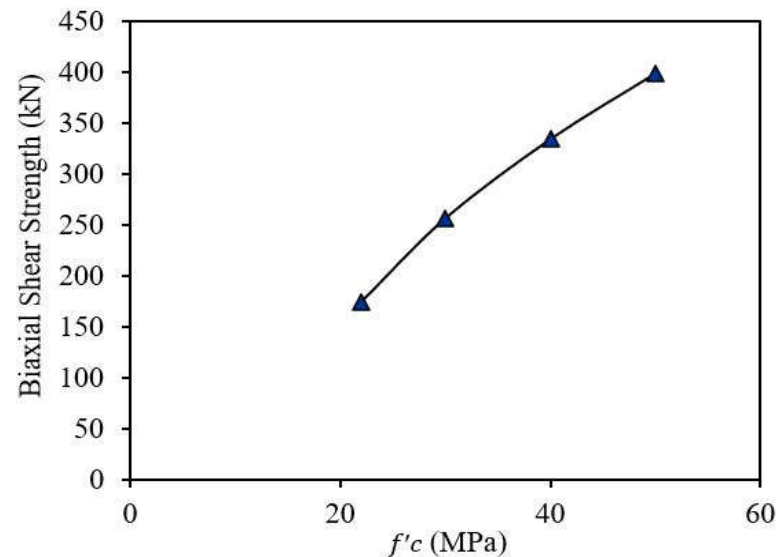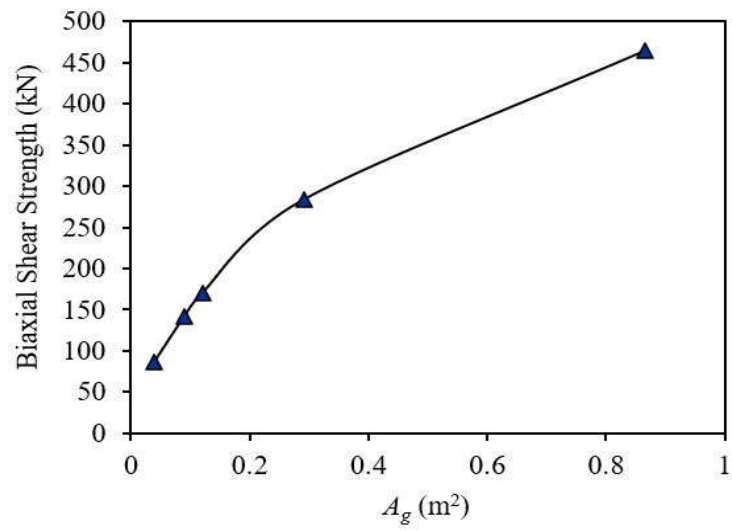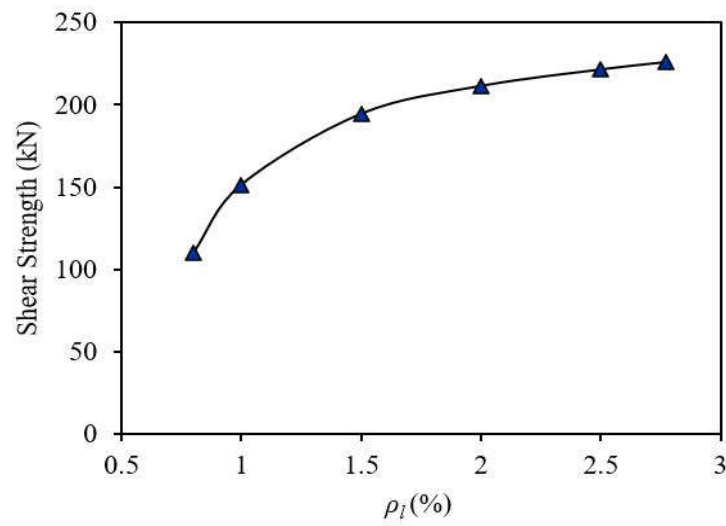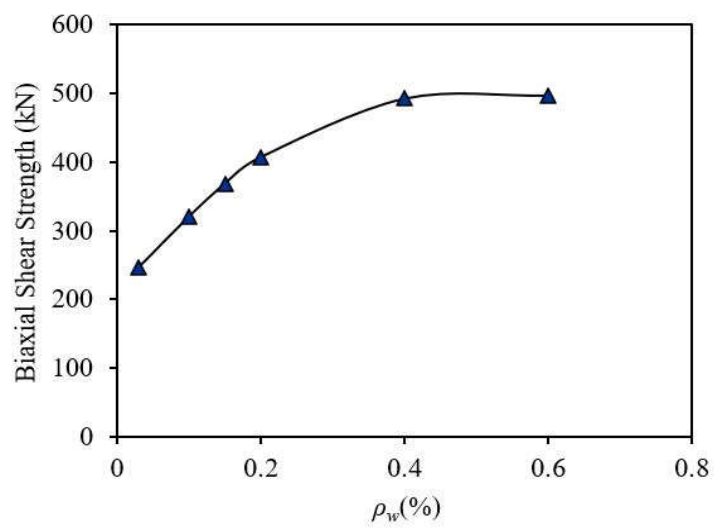

(**a**)
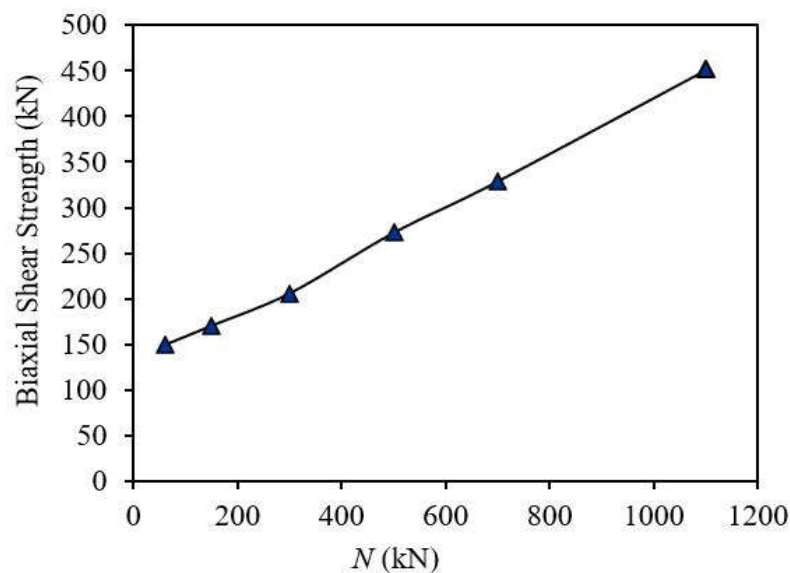
(**b**)

Figure 8. *Cont*.

(**c**)

(**d**)

(**e**)

Figure 8. *Cont*.
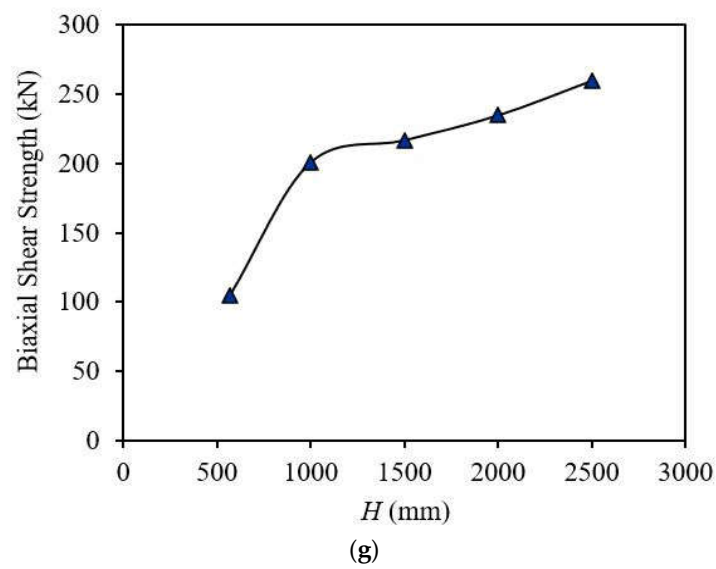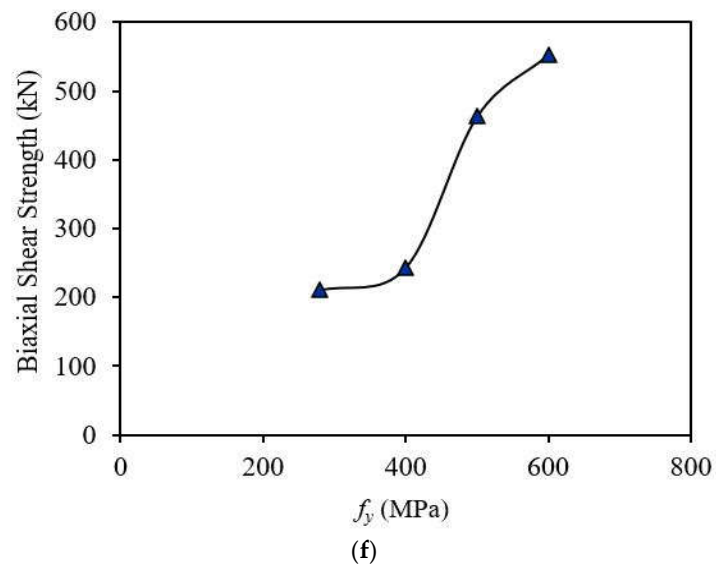
(**f**)

(**g**)

(**h**)

Figure 8. *Cont.*
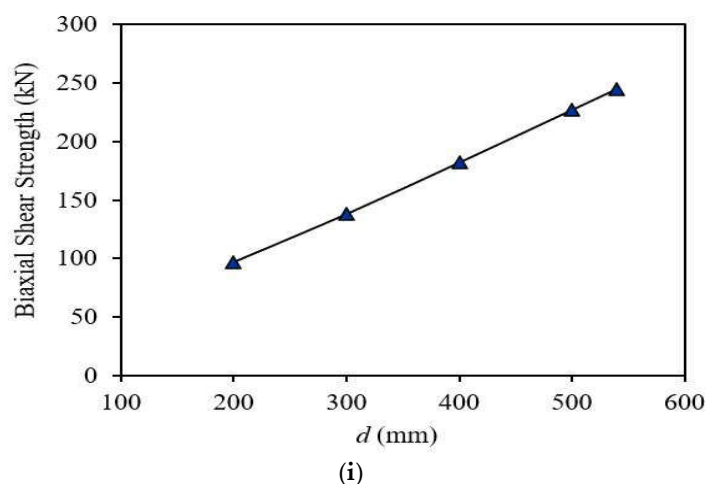
(i)
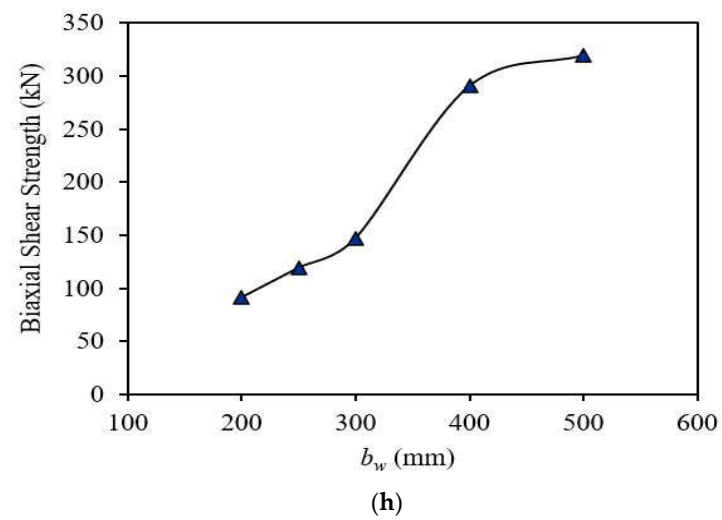
Figure 8. Results of parametric analysis. (**a**) Compressive Strength of Concrete, (**b**) Axial Load (**c**) Column Cross-sectional Area, (**d**) Longitudinal Reinforcement Percentage, (**e**) Shear Reinforcement Percentage, (**f**) Yield Strength Longitudinal Reinforcement, (**g**) Column Height, (**h**) Width of Column Web (mm), (**i**) Depth of Column (mm).

5. Conclusions

In this paper, we adopted new machine-learning methods to develop novel models to formulate the biaxial shear strength of RC columns. For this purpose, a total of eight prediction models using ANFIS and MEP algorithms were developed to model the biaxial shear strength with different combinations of input parameters and mathematical functions in the architecture of the models. The performance of the developed models was assessed by calculating various statistical measures such as R, MAE, RMSE, RSE, RRMSE, ρ, and OF. It is concluded that the M1-ANFIS and M1-MEP models had the highest statistical reliability and accuracy as they performed well in the three datasets compared to the other models since the R values were higher while the other statistical indicators were very low. All the models were also validated against independent experimental data by checking their performance on unseen data. This also demonstrated their robustness/generalization capability. Thus, from the study based on the two algorithms, it can be concluded that the biaxial shear strength of RC columns can be well predicted by using the following parameters as input: f'_c, A_g, ρ_l, ρ_w, f_y, N.

One of the main advantages of the MEP algorithm over the ANFIS algorithm is that it belongs to the grey-box category of models. The accuracy of the ANFIS-based models was higher; however, they cannot provide a mathematical structure of the developed model, which makes them inferior as compared to MEP. In this study, empirical equations were also derived for M1-MEP and other MEP-based models. A parametric analysis of the proposed equations was also performed for M1-MEP, which showed that the proposed model accurately considered the system under consideration. Additionally, the proposed empirical equations can ensure an effective and accurate foundation with which to improve the utilization of machine-learning techniques to predict the biaxial shear strength of RC columns. Based on a comparison of the developed models with the other models available in the literature, it was shown that the model possesses the highest accuracy, since it duly considers the effect of bidirectional lateral loading.

Author Contributions: Conceptualization—Y.P. and X.G.; Methodology—M.F.I., X.G., M.R. and M.A.; Software—Y.P., M.F.I. and I.A.; Validation—Y.P., M.A.U.R.T. and M.F.I.; Formal Analysis—Y.P., X.G. and M.A.U.R.T.; Investigation—Y.P., I.A. and M.F.I.; Resources—I.A., M.A., A.W.M.N. and M.A.U.R.T.; Data Curation—Y.P., I.A. and M.F.I.; Writing original draft preparation—Y.P., I.A., M.R.;

Visualization—M.F.I., M.R. and I.A.; Supervision—I.A. and A.W.M.N. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

RC	Reinforced Concrete
MEP	Multi Expression Programming
V	Biaxial Shear Strength
GEP	Gene Expression Programming
GA	Genetic Algorithm
GP	Genetic Programming
G_2	Gene 2
G_4	Gene 4
G_6	Gene 6
x_1 & x_2	Sample inputs in ANFIS
μ_{Ai} & μ_{Bi-2}	Weights obtained while connecting fuzzy membership functions
$\overline{w}_i$	Firing strength
f_i	Linear function
p_k, q_k & r_k	Linear function parameters for particular rule 'k'
OF	Objective Function
ρ	Performance Index
e_i	ith Experimental
m_i	ith Predicted
$\overline{e}_i$	ith Mean Experimental
$\overline{m}_i$	ith Mean Predicted
n_L	Number of learning (training and validation) data
n_T	Number of testing dataset
ρ_L	Performance index of learning dataset
ρ_T	Performance index of testing dataset
R	Correlation coefficient
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
f'_c	Concrete Compressive Strength
A_g	Gross Sectional Area of Column
ρ_l	Longitudinal Reinforcement Percentage
ρ_w	Shear Reinforcement Percentage
f_y	Yield Strength of Longitudinal Reinforcement
N	Axial Load Of Column
b_w	Width Of Column Web
d	Depth Of Column
H	Column Height

References

1. Azim, I.; Yang, J.; Bhatta, S.; Wang, F.; Liu, Q.-F. Factors influencing the progressive collapse resistance of RC frame structures. *J. Build. Eng.* **2020**, *27*, 100986. [CrossRef]
2. Shi, Q.; Ma, L.; Wang, Q.; Wang, B.; Yang, K. Seismic performance of square concrete columns reinforced with grade 600 MPa longitudinal and transverse reinforcement steel under high axial load. *Structures* **2021**, *32*, 1955–1970. [CrossRef]
3. Lee, C.S.; Han, S.W. Cyclic behaviour of lightly-reinforced concrete columns with short lap splices subjected to unidi-rectional and bidirectional loadings. *Eng. Struct.* **2019**, *189*, 373–384. [CrossRef]
4. Pham, T.P.; Li, B. Seismic Behavior of Reinforced Concrete Columns with Light Transverse Reinforcement under Different Lateral Loading Directions. *ACI Struct. J.* **2013**, *110*, 833. [CrossRef]
5. Rodrigues, H.; Furtado, A.; Arêde, A. Behavior of Rectangular Reinforced-Concrete Columns under Biaxial Cyclic Loading and Variable Axial Loads. *J. Struct. Eng.* **2016**, *142*, 04015085. [CrossRef]
6. Bonet, J.L.; Barros, M.H.F.M.; Romero, M.L. Comparative study of analytical and numerical algorithms for designing rein-forced concrete sections under biaxial bending. *Comput. Struct.* **2006**, *84*, 2184–2193. [CrossRef]
7. Mark, P. Shear-resistant design of biaxially loaded RC beams. *Mag. Concr. Res.* **2007**, *59*, 21–28. [CrossRef]
8. Galal, K.; Ghobarah, A. Flexural and shear hysteretic behaviour of reinforced concrete columns with variable axial load. *Eng. Struct.* **2003**, *25*, 1353–1367. [CrossRef]
9. Okamura, H.; Higai, T. Proposed design equation for shear strength of reinforced concrete beams without web rein-forcement. In Proceedings of the Japan Society of Civil Engineers; Japan Society of Civil Engineers: Tokyo, Japan, 1980; pp. 131–141.
10. Hansapinyo, C.; Chaisomphob, T.; Maekawa, K.; Pimanmas, A. Experimental Investigation on Rectangular Reinforced Concrete Beam subjected to Bi-axial Shear. In *Proceedings of the Eight Asia-Pacific on Structural Engineering and Construction*; Nanyang Technological University: Singapore, 2001.
11. Tinini, A.; Minelli, F.; Belletti, B.; Scolari, M. Biaxial shear in RC square beams: Experimental, numerical and analytical program. *Eng. Struct.* **2016**, *126*, 469–480. [CrossRef]
12. Murad, Y.Z. Predictive model for bidirectional shear strength of reinforced concrete columns subjected to biaxial cyclic loading. *Eng. Struct.* **2021**, *244*, 112781. [CrossRef]
13. Iqbal, M.F.; Liu, Q.-F.; Azim, I.; Zhu, X.; Yang, J.; Javed, M.F.; Rauf, M. Prediction of mechanical properties of green concrete incorporating waste foundry sand based on gene expression programming. *J. Hazard. Mater.* **2020**, *384*, 121322. [CrossRef] [PubMed]
14. Oltean, M.; Grosan, C. A comparison of several linear genetic programming techniques. *Complex Syst.* **2003**, *14*, 285–314.
15. Arabshahi, A.; Gharaei-Moghaddam, N.; Tavakkolizadeh, M. Development of applicable design models for concrete columns confined with aramid fiber reinforced polymer using Multi-Expression Programming. *Structures* **2020**, *23*, 225–244. [CrossRef]
16. Iqbal, M.F.; Javed, M.F.; Rauf, M.; Azim, I.; Ashraf, M.; Yang, J.; Liu, Q.-F. Sustainable utilization of foundry waste: Forecasting mechanical properties of foundry sand based concrete using multi-expression programming. *Sci. Total Environ.* **2021**, *780*, 146524. [CrossRef] [PubMed]
17. Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*; The MIT Press: Cambridge, MA, USA, 1992.
18. Jalal, F.E.; Xu, Y.; Iqbal, M.; Jamhiri, B.; Javed, M.F. Predicting the compaction characteristics of expansive soils using two genetic programming-based algorithms. *Transp. Geotech.* **2021**, *30*, 100608. [CrossRef]
19. Oltean, M.; Dumitrescu, D. Multi Expression Programming. 2002. Available online: https://www.researchgate.net/publication/2918165_Multi_Expression_Programming (accessed on 5 October 2021).
20. Alavi, A.H.; Gandomi, A.H.; Sahab, M.G.; Gandomi, M. Multi expression programming: A new approach to formulation of soil classification. *Eng. Comput.* **2010**, *26*, 111–118. [CrossRef]
21. Aho Alfred, V.; Ravi, S.; Ullman Jeffrey, D. *Compilers, Principles, Techniques*; Addison Wesley: Boston, MA, USA, 1986; Volume 7, p. 9.
22. Sharifi, S.; Abrishami, S.; Gandomi, A.H. Consolidation assessment using Multi Expression Programming. *Appl. Soft Comput.* **2020**, *86*, 105842. [CrossRef]
23. Khan, M.A.; Aslam, F.; Javed, M.F.; Alabduljabbar, H.; Deifalla, A.F. New prediction models for the compressive strength and dry-thermal conductivity of bio-composites using novel machine learning algorithms. *J. Clean. Prod.* **2022**, *350*, 131364. [CrossRef]
24. Sayyaadi, H. *Modeling, Assessment, and Optimization of Energy Systems*; Academic Press: Cambridge, MA, USA, 2020.
25. Azim, I.; Yang, J.; Javed, M.F.; Iqbal, M.F.; Mahmood, Z.; Wang, F.; Liu, Q.-F. Prediction model for compressive arch action capacity of RC frame structures under column removal scenario using gene expression programming. *Structures* **2020**, *25*, 212–228. [CrossRef]
26. Azim, I.; Yang, J.; Iqbal, M.F.; Mahmood, Z.; Javed, M.F.; Wang, F.; Liu, Q.-F. Prediction of Catenary Action Capacity of RC Beam-Column Substructures under a Missing Column Scenario Using Evolutionary Algorithm. *KSCE J. Civ. Eng.* **2021**, *25*, 891–905. [CrossRef]
27. Azim, I.; Yang, J.; Iqbal, M.F.; Javed, M.F.; Nazar, S.; Wang, F.; Liu, Q.-F. Semi-analytical model for compressive arch action capacity of RC frame structures. *Structures* **2020**, *27*, 1231–1245. [CrossRef]
28. Liu, Q.-F.; Iqbal, M.F.; Yang, J.; Lu, X.-Y.; Zhang, P.; Rauf, M. Prediction of chloride diffusivity in concrete using artificial neural network: Modelling and performance evaluation. *Constr. Build. Mater.* **2021**, *268*, 121082. [CrossRef]
29. Golafshani, E.M.; Behnood, A.; Arashpour, M. Predicting the compressive strength of normal and High-Performance Concretes using ANN and ANFIS hybridized with Grey Wolf Optimizer. *Constr. Build. Mater.* **2020**, *232*, 117266. [CrossRef]

30. Morcous, G.; Lounis, Z. Prediction of Onset of Corrosion in Concrete Bridge Decks Using Neural Networks and Case-Based Reasoning. *Comput. Aided Civ. Infrastruct. Eng.* **2005**, *20*, 108–117. [CrossRef]

31. Khan, M.I. Mix proportions for HPC incorporating multi-cementitious composites using artificial neural networks. *Constr. Build. Mater.* **2012**, *28*, 14–20. [CrossRef]

32. Bilim, C.; Atis, C.; Tanyildizi, H.; Karahan, O. Predicting the compressive strength of ground granulated blast furnace slag concrete using artificial neural network. *Adv. Eng. Softw.* **2009**, *40*, 334–340. [CrossRef]

33. Murad, Y.; Imam, R.; Abu Hajar, H.; Habeh, D.; Hammad, A.; Shawash, Z. Predictive compressive strength models for green concrete. *Int. J. Struct. Integr.* **2019**, *11*, 169–184. [CrossRef]

34. Gandomi, A.; Faramarzifar, A.; Rezaee, P.G.; Asghari, A.; Talatahari, S. New design equations for elastic modulus of concrete using multi expression programming. *J. Civ. Eng. Manag.* **2015**, *21*, 761–774. [CrossRef]

35. Gandomi, A.H.; Alavi, A.H.; Yun, G.J. Formulation of uplift capacity of suction caissons using multi expression programming. *KSCE J. Civ. Eng.* **2011**, *15*, 363–373. [CrossRef]

36. Baykasoğlu, A.; Güllü, H.; Canakci, H.; Özbakır, L. Prediction of compressive and tensile strength of limestone via genetic programming. *Expert Syst. Appl.* **2008**, *35*, 111–123. [CrossRef]

37. Alavi, A.H.; Mollahasani, A.; Gandomi, A.H.; Bazaz, J.B. Formulation of secant and reloading soil deformation moduli using multi expression programming. *Eng. Comput.* **2012**, *29*, 173–197. [CrossRef]

38. Jalal, F.E.; Xu, Y.; Iqbal, M.; Javed, M.F.; Jamhiri, B. Predictive modeling of swell-strength of expansive soils using artificial intelligence approaches: ANN, ANFIS and GEP. *J. Environ. Manag.* **2021**, *289*, 112420. [CrossRef] [PubMed]

39. Iqbal, M.; Onyelowe, K.C.; Jalal, F.E. Smart computing models of California bearing ratio, unconfined compressive strength, and resistance value of activated ash-modified soft clay soil with adaptive neuro-fuzzy inference system and ensemble random forest regression techniques. *Multiscale Multidiscip. Model. Exp. Des.* **2021**, *4*, 207–225. [CrossRef]

40. Gandomi, A.; Tabatabaei, S.M.; Moradian, M.H.; Radfar, A.; Alavi, A.H. A new prediction model for the load capacity of castellated steel beams. *J. Constr. Steel Res.* **2011**, *67*, 1096–1105. [CrossRef]

41. Iqbal, M.; Zhao, Q.; Zhang, D.; Jalal, F.E.; Jamal, A. Evaluation of tensile strength degradation of GFRP rebars in harsh alkaline conditions using non-linear genetic-based models. *Mater. Struct.* **2021**, *54*, 190. [CrossRef]

42. Despotovic, M.; Nedić, V.; Despotović, D.; Cvetanović, S. Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renew. Sustain. Energy Rev.* **2016**, *56*, 246–260. [CrossRef]

43. Frank, I.; Todeschini, R. *The Data Analysis Handbook*; Elsevier: Amsterdam, The Netherlands, 1994; Volume 14.

44. ACI 318-11. *Building Code Requirements for Structural Concrete and Commentary, PCA Notes on ACI 318-11*; with Design Applications; ACI International: Farmington Hills, MI, USA, 2011.

45. Zhang, Q.; Barri, K.; Jiao, P.; Salehi, H.; Alavi, A.H. Genetic programming in civil engineering: Advent, applications and future trends. *Artif. Intell. Rev.* **2021**, *54*, 1863–1885. [CrossRef]

46. Yazoghli-Marzouk, O.; Vulcano-Greullet, N.; Cantegrit, L.; Friteyre, L.; Jullien, A. Recycling foundry sand in road construction–field assessment. *Constr. Build. Mater.* **2014**, *61*, 69–78. [CrossRef]

47. Haitham, A.-T. Effect of Transverse Reinforcement on the Axial Compressive Strength of Reinforced Concrete Columns. *Al Qadisiyah J. Eng. Sci.* **2016**, *9*, 119–134.

48. Woods, C.; Matamoros, A.B. Effect of longitudinal reinforcement ratio on the failure mechanism of R/C columns most vulnerable to collapse. In Proceedings of the 9th US National and 10th Canadian Conference on Earthquake Engineering, Toronto, ON, Canada, 25–29 July 2010.