Article

Predicting the Lateral Load Carrying Capacity of Reinforced Concrete Rectangular Columns: Gene Expression Programming

Raheel Asghar [1], Muhammad Faisal Javed [1,*], Raid Alrowais [2], Alamgir Khalil [3], Abdeliazim Mustafa Mohamed [4,5], Abdullah Mohamed [6] and Nikolai Ivanovich Vatin [7]

1 Department of Civil Engineering, Abbottabad Campus, COMSATS University Islamabad, Abbottabad 22060, Pakistan; raheelasghar@cuiatd.edu.pk
2 Department of Civil Engineering, Jouf University, Sakaka, Al-Jouf 72388, Saudi Arabia; rnalrowais@ju.edu.sa
3 Department of Civil Engineering, University of Engineering and Technology, Peshawar 25120, Pakistan; alamgirkhalil@uetpeshawar.edu.pk
4 Department of Civil Engineering, College of Engineering, Prince Sattam Bin Abdulaziz University, Alkharj 16273, Saudi Arabia; a.bilal@psau.edu.sa
5 Building and Construction Technology Department, Bayan College of Science and Technology, Khartoum 210, Sudan
6 Research Centre, Future University in Egypt, New Cairo 11835, Egypt; mohamed.a@fue.edu.eg
7 Peter the Great St. Petersburg Polytechnic University, 195291 St. Petersburg, Russia; vatin@mail.ru
* Correspondence: arbabfaisal@cuiatd.edu.pk

Abstract: This research presents a novel approach of artificial intelligence (AI) based gene expression programming (GEP) for predicting the lateral load carrying capacity of RC rectangular columns when subjected to earthquake loading. To achieve the desired research objective, an experimental database assembled by the Pacific Earthquake Engineering Research (PEER) center consisting of 250 cyclic tested samples of RC rectangular columns was employed. Seven input variables of these column samples were utilized to develop the coveted analytical models against the established capacity outputs. The selection of these input variables was based on the linear regression and cosine amplitude method. Based on the GEP modelling results, two analytical models were proposed for computing the flexural and shear capacity of RC rectangular columns. The performance of both these models was evaluated based on the four key fitness indicators, i.e., coefficient of determination (R^2), root mean squared error ($RMSE$), mean absolute error (MAE), and root relative squared error ($RRSE$). From the performance evaluation results of these models, R^2, $RMSE$, MAE, and $RRSE$ were found to be 0.96, 53.41, 38.12, and 0.20, respectively, for the flexural capacity model, and 0.95, 39.47, 28.77, and 0.22, respectively, for the shear capacity model. In addition to these fitness criteria, the performance of the proposed models was also assessed by making a comparison with the American design code of concrete structures ACI 318-19. The ACI model reported R^2, $RMSE$, MAE, and $RRSE$ to be 0.88, 101.86, 51.74, and 0.39, respectively, for flexural capacity, and 0.87, 238.74, 183.66, and 1.35, respectively, for shear capacity outputs. The comparison depicted a better performance and higher accuracy of the proposed models as compared to that of ACI 318-19.

Keywords: reinforced concrete columns; lateral load carrying capacity; bearing capacity; flexural capacity; shear capacity; gene expression programming

1. Introduction

1.1. Background

Columns are the primary structural elements intended to transfer the load of the overlying structure to the foundations underneath through axial compression [1]. Based on the material, they can be classified into several types. However, reinforced concrete (RC) columns are widely used in buildings and bridge structures. They are generally considered as the most important RC frame members of these structures in terms of their seismic

response [2,3]; this is because most of the recent structural failures during earthquakes are reported to occur because of poor column performances [4]. A failure normally triggered through the columns progressively results in the global failure of the overall structure. The modern generation of earthquake engineering aims to equip structural engineers with advanced techniques to design structures based on their performance rather than on conventional strength-based methods [5–8]. To effectively apply this approach for the seismic design of RC rectangular columns, failure modes and ultimate load-carrying capacities are required to be determined [9]. Enough literature is already available on the failure mode classification of RC columns. According to Acun and Sucuoglu [10], failure in columns occurs because of insufficient strength or deformation capacity. Based on this principle, RC column failures can be classified into three categories [11], i.e., shear failure, flexural failure, and flexural-shear failure. Shear failure occurs in a brittle manner with a sudden loss of load carrying capacity before the yielding of longitudinal reinforcement, and therefore also termed as brittle shear failure [12]. Flexural failure occurs in a ductile fashion, with the column enduring a large deformation without any significant loss of load-carrying capacity [12]. This type of failure is also known as a ductile flexural failure. An intermediate failure stage exists between pure shear and flexural failure, i.e., flexural-shear failure. In this type of failure, columns usually suffer material damage or flexural yielding just before experiencing pure shear failure [12]. This is also termed as ductile shear failure.

Many researchers have worked on quantifying these failure modes of RC columns. Some have presented pure analytical techniques for classifying these failure modes based on either geometrical or strength parameters, whereas others have presented mixed analytical-experimental approaches. According to Feng et al. [13], the failure mode classification is normally based on two simple indices, i.e., aspect ratio (A_r) and shear demand to the capacity ratio. If A_r (i.e., the ratio of column length to the cross-sectional height) is less than 2, the column failure would be categorized as a pure shear failure, whereas if it is greater than 4, the failure would be classified as a pure flexural failure. If $2 \leq A_r \leq 4$, the failure is said to be a flexural-shear failure. Similarly, if the shear demand to capacity ratio is greater than 1, the column would suffer pure shear failure, whereas if it is less than 0.7, the failure would be a pure flexural failure. However, if the shear demand to capacity ratio is between 0.7 and 1, combined flexural-shear failure is expected to occur for the given RC column [13].

Once the failure mode is classified, the next step is determining the lateral load-carrying capacity (i.e., bearing capacity) of the RC columns. This can be defined as the maximum lateral load taken by the column before undergoing failure during seismic events. The bearing capacity of an RC column further depends on its flexural and shear capacities, the minimum of which is regarded as the bearing capacity of the overall column. The flexural capacity is normally based on the plane section assumptions and can be determined accurately using the famous flexural formula ($f = MY/I$). However, the determination of shear capacity is relatively complicated because of multiple influencing factors and complex shear transfer mechanisms. Many experimental and finite element studies have predicted both these capacities of RC columns with higher accuracy [1,11,14], but these traditional capacity determination techniques are costly, time-consuming, and laborious. The solution to the problem lies in the development of analytical models [13,15–20]. Up till now, many empirical models have been developed for the prediction of bearing capacity of RC columns, e.g., failure mode and bearing capacity prediction model [13], strut and tie model [21,22], modified compression field theory model [23,24], softened truss model [25,26], damage model [27,28], etc. However, most of these models are based on simple regression analysis incorporating limited experimental datasets [13]. Therefore, their applicability in a broader aspect is always questionable. Moreover, the models with relatively larger datasets have considered only a few parameters in their model development. Therefore, their accuracy is also not ideal. Artificial intelligence (AI) techniques are generally recommended to overcome these issues using relatively larger databases developed over the past few

decades [29]. AI is a "black box" that relates the inputs and outputs in a simplified manner while avoiding complex mathematical derivations. It has better performance as compared to the traditional analytical models and therefore can be employed to predict the bearing capacity of RC columns quickly, accurately, and robustly.

The term AI was first introduced in 1956 [30] to represent the ability of a machine to imitate intelligent human behaviour in order to solve complicated problems based on human-inspired algorithms [31]. However, its use was limited then in the civil and structural engineering community because of the lack of innovative techniques and advanced computer technology. The recent progress in AI techniques and modern computer systems has enhanced the interest of structural engineers in developing precise and consistent models to solve complex structural problems [32]. AI procedures are generally based on machine learning (ML), pattern recognition (PR) and deep learning (DL), which further consist of artificial neural networks (ANNs), fuzzy logic, genetic programming (GP), etc. [16,17,31,33–50]. The application of these AI techniques in structural engineering are traced back to the early 1980s, where they were first used in the compliance checking of design codes [51,52] and expert interactive design of concrete columns (EIDOCC) [53,54]. In addition to these basic works, an AI technique, i.e., ML, was also applied to predict the location and extent of the damage to various structural systems subjected to varying loading conditions [55–57]. Multiple AI techniques, including ANN, logistic and linear regressors, lasso and supper vector machines (SVMs), have been adopted to predict the fragility behaviour of RC structural systems [58–61]. Some recent works have also employed AI techniques to predict the shear strength of RC beams [62], shear strength of RC joints [63], punching shear capacity of RC slabs [64], failure mode and capacity prediction of RC columns, etc. [12,65,66]. However, the methodology adopted in all these works requires a bulk of memory to accommodate a large number of hidden neurons involved in the model development. Moreover, using these techniques, it is also difficult to represent the relationship between inputs and outputs in a practically simplified form, which requires the researchers and other stakeholders to shift towards more efficient, simple, and practically applicable techniques.

Another AI method, i.e., GP, provides an effective solution to the above-discussed practicality issue. This was first introduced by Cramer in 1985 [67] and further improved by Koza in 1992 [68]. It is an evolutionary AI algorithm intended to produce high-quality solutions to complicated problems through analogues to computer-based genetic operations, e.g., deletion, reproduction, duplication, mutation, and crossover. It is a domain-independent technique based on Darwin's evolutionary theory [69]. GP starts with the given initial population of programs consisting of multiple functions and terminals (Figure 1). The functions describe the standard procedure to execute the program, whereas terminals outline the arguments for each function [70,71]. The functions may include arithmetical operations, program manoeuvres, and mathematical equations. The performance of a certain program is evaluated based on its effectiveness in explaining the output while exploiting the given inputs. Evaluating the initial population, worst performing programs are deleted, whereas a new generation of programs is created based on the top performing programs through the operations of reproduction, mutation, duplication, and crossover [68]. The mutation involves replacing an arbitrary part of a program with an alternative part randomly picked from another program of the existing population. In addition to the reproduction of the programs through mutation and crossover, some programs are simply copied from the previous generation to the new generation through a process known as duplication. These processes eventually result in the production of a new generation of programs having a better fitness than their parent population [68]. This loop of genetic operations is repeated again and again until an ultimate termination condition is reached. Some common termination conditions of GP may include maximum number of iterations, maximum allowable computational time, minimum level of difference between the populations of successive iterations or a solution satisfying the set criteria. In case either of these conditions is reached, the program is terminated to provide a final solution [68].
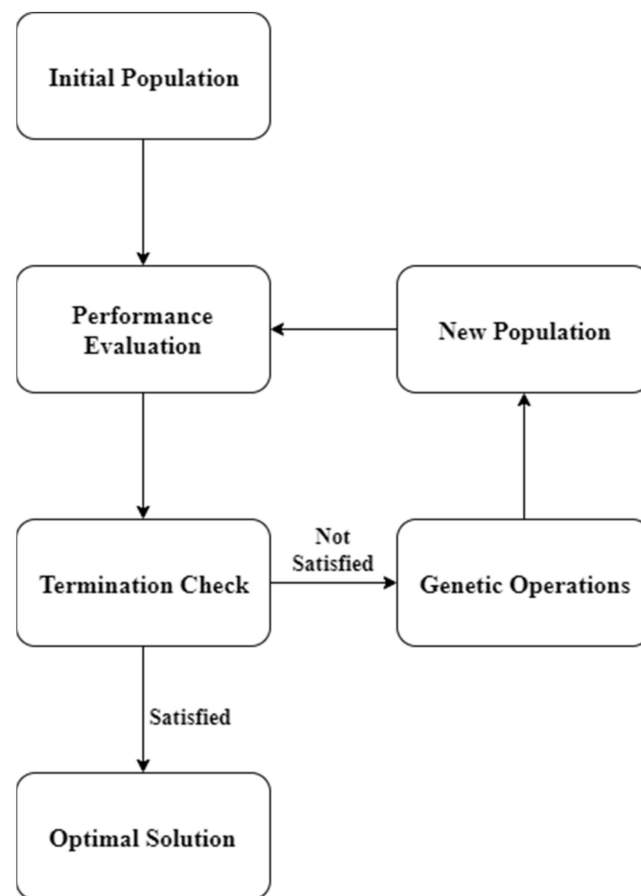
Figure 1. Working Procedure of GP Algorithms.

Recently, GP has been successfully used as an automated problem-solving technique in many practical applications, e.g., engineering design, traffic routing, curve fitting, data modelling, failure classification, capacity prediction, symbolic regression, etc. [72]. However, along with all the positive utilities it offers, there are certain limitations associated with it. It is computationally very expensive in the case of high dimensional real-world problems which require a huge amount of time to complete simulation to get the optimum results [73]. It also tends to destructive mutation, i.e., once the best fit chromosome is generated, it can be reconstructed into a lesser fit chromosome of the new generation, reducing its overall efficiency [73]. The effectiveness of GP is always a matter of debate in the case of decision problems, where there is the least probability for the convergence of a solution [73]. It can also sometimes result in an exponential increase in the population size during mutation and crossover operations, making this technique difficult to apply to complex design problems [73]. Another drawback of GP lies in its inability to have an independent genome. Therefore, its nonlinear structure has to act both as a genotype and phenotype [74], which makes it incapable of producing simple and robust expressions. To successfully deal with all these limitations, an extension of GP, i.e., gene expression programming (GEP), is often recommended. GEP further improves the effectiveness of this technique by making the model adaptive to different conditions, like living organisms. It uses the simple chromosomes of defined length to encode a computer program for the given model [74]. GEP has the ability to describe the output by a simple mathematical equation, which can be recommended for practical purposes with higher prediction accuracy [75]. Recently, GEP has been considered an alternative to the conventional analytical methodologies, particularly in civil and structural engineering [76].

Therefore, this research aims to predict the lateral load carrying capacity of RC rectangular columns subjected to earthquake loading utilizing GEP. Moreover, as according to the

authors' knowledge, no equation has been proposed in the past research articles until now to directly predict the lateral load carrying capacity of RC columns, this research also aims to provide simplified empirical relationships for future design purposes while assuring the universal nature of the proposed model. It is further extended to provide a comparison on the lateral load carrying capacity of RC columns as recommended by the proposed model and international design codes.

1.2. Gene Expression Programming

GEP was proposed by Candida Ferreira [77] in 2001 as an extension of GP. It belongs to the family of evolutionary AI algorithms that create complex tree structure computer programs, often known as parse trees. Multiple parse trees exist in each chromosome of these computer programs, which are usually created by the multigenic system of GEP. Therefore, they are also known as expression trees [78,79]. These computer programs are supposed to explore the environment and adapt to the given conditions by altering their sizes, shapes, and composition. The computer programs are generally encoded using simple linear chromosomes of fixed length as inherited from genetic algorithms (GA), resembling the genetic model of living organisms. GEP incorporates these linear chromosomes as genotype, and the parse trees as phenotype of their complex system of genomes and phenomes [79]. The genomes are intended to keep and transfer genetic information, whereas phenomes are responsible for exploring and adapting the programs to the given environment. Despite their fixed sizes, genomes are used to encode the expression trees of varying sizes and shapes, allowing the adaptation and evolution of phenomes to occur smoothly. An expression tree generally consists of a phenome expression and the linear strings of GEP genes [79]. These linear string genes are often known as k-expressions which can be read directly from top to bottom and left to the right in the expression tree structure of GEP. The k-expressions are normally used to represent the region of genes that are being expressed, which in turn epitomises the validity of computer programs and expressions. GEP genes can be further expressed as a function of two domains, i.e., a head and a tail [77]. The head is normally used to encode all the input variables and functions chosen to get the required solution, whereas the tail provides a pool of terminals to ensure the error-free nature of the GEP programs. Following the gene structure of living organisms, GEP genes are also associated with the homeotic genes, which controls the interaction between different main program modules [79]. Homeotic genes possess an identical structure to the normal GEP genes as they are the product of the same process. They are also comprised of a head and a tail domain. The only difference is that they contain linking functions and special terminals called genic terminals. The cellular system in GEP allows the unrestricted evolution of linking functions as well as the recursive use of encoded programs. The head-tail domains of GEP and homeotic genes are the fundamental units of all GEP algorithms [79]. However, GEP also tries to explore additional genes with an extra domain. This extra domain usually encodes random constants that are persistently fined tuned by the program to arrive at optimum solution.

The basic GEP algorithm involves four fundamental steps, i.e., the creation of initial random population, execution of GEP programs, verification of termination conditions and the generation of subsequent populations or the representation of the final optimum solution in the form of a simple mathematical expression (Figure 2). The initial random population of GEP programs is usually created with the help of functions and terminal sets. GEP only requires the population units to be expressed in the form of simple linear chromosomes of fixed length without wondering about their structural reliability [79]. This is because GEP expressions always result in syntactically sound programs. After creating the initial random population, all the GEP programs are executed according to predefined fitness functions. These fitness functions normally depend on the type of problem under consideration. Based on prediction outputs, the overall GEP problems can be classified into three main categories, i.e., the regression problems, classification problems and Boolean logic problems [79]. For the regression problems, the output is usually a numeric value.

Therefore, the fitness function for the evaluation of their model performance is based on the error between the model output and actual output. The most frequently used fitness functions for regression problems include the coefficient of determination (R^2), root mean squared error (*RMSE*), mean absolute error (*MAE*), mean absolute percentage error (MAPE), number of loop iterations, etc. [79].
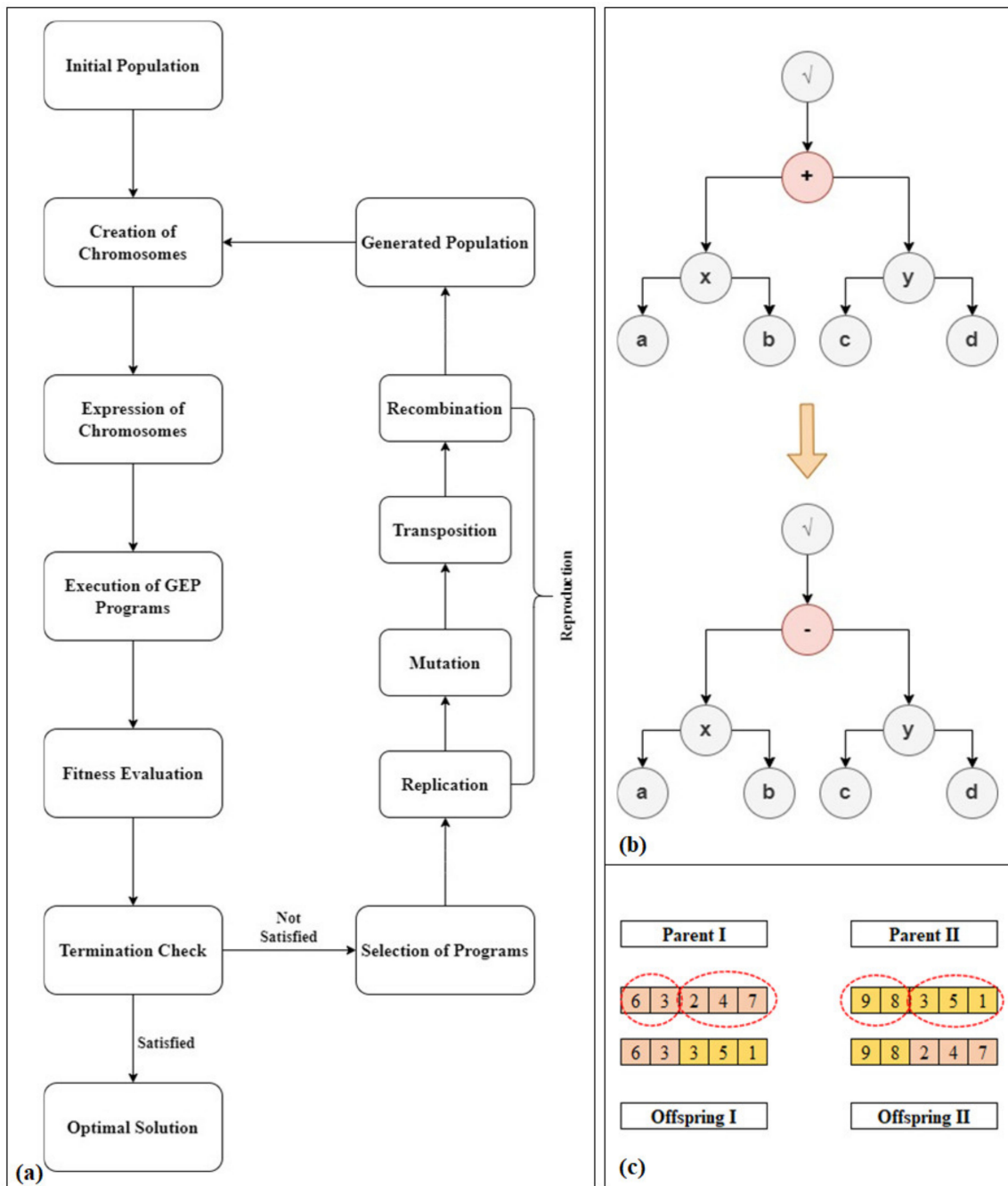
Figure 2. (**a**) GEP Algorithm (**b**) Mutation Process (**c**) Crossover Process.

The output of classification and Boolean logic problems generally cannot be expressed in the form of numeric values; therefore, their performance evaluation functions are based on a matrix counting the number of correct and incorrect predictions, often known as a confusion matrix. Some common fitness functions based on this confusion matrix include sensitivity, precision, Jaccard similarity, F-measure, etc. [79]. These fitness functions are quite sophisticated and are adequate for efficient solution of most of the classification

and Boolean logic problems. However, any complexity encountered while dealing with these problems can be solved effectively by introducing an additional fitness function capable of exploring the overall model structure, distribution of outputs and classifier margins [79]. Once the GEP programs are executed, their performance is evaluated based on the already defined above-discussed fitness functions. If the evaluation results satisfy the termination criteria, the GEP loop algorithm is terminated, and the final optimum solution is presented in the form of a simple mathematical expression. However, if the termination criteria are not satisfied, the best-fit programs are selected to reproduce the subsequent population of GEP programs, and the loop is repeated [77]. The reproduction process of the GEP algorithm involves the four basic genetic operations, i.e., replication, mutation, recombination, and transposition. During replication, all the genomes of the best performing programs of the previous population are copied to the new population. The sub-structural elements of these genomes are then replaced with one another during the mutation process of genetic reproduction [77]. After the basic operations of replication and mutation, the new chromosomes of the next generation are produced by combining different parts of the parent chromosomes. To ensure the soundness and reliability of the generated programs, a transposition insertion sequence is introduced in the heads of genes while preserving the chromosome length and gene structure [79]. This loop algorithm of GEP is repeated again and again until an optimum solution is obtained (Figure 2).

A significant characteristic of GEP is that only the genome is transferred to the successive population of computer programs instead of the global parent structure, which enhances the overall algorithm efficiency [77]. There also exists a direct relationship between the chromosomes and corresponding functions or terminals which makes the genetic variation of GEP simple and easy to understand [32]. Another advantage of GEP lies in its ability to provide the optimum solution with simple mathematical equations [17]. These equations can be easily manipulated in practical works, especially in civil and structural engineering. Considering all these positive features, this research aims to employ GEP to predict the lateral load carrying capacity of RC rectangular columns subjected to earthquake loading.

2. Methods

The development of an analytical model for predicting the bearing capacity of RC rectangular columns involves three fundamental steps i.e., the preparation of a comprehensive universal database, selection of model parameters, and the implementation of the GEP algorithm. According to Mundfrom et al. [80], the minimum sample size in the database must be 3–20 times the involved parameters for general engineering problems. However, it is still a controversial topic and is often decided according to logical judgment criteria [13]. The selection of parameters for the model development can be based on linear regression or other simpler analysis determining the relationship between two or more variables. Once an integrated and comprehensive database is established, the GEP algorithm can be implemented to obtain the required analytical model for the given problem.

2.1. Database

An inclusive structural performance database assembled by the PEER centre [29] was adopted to develop an analytical model for predicting RC rectangular columns' bearing capacity. This was prepared by the researchers based on the experimental work at the National Institute of Standards and Technology. It contains all the data required to evaluate the seismic performance of RC columns. The overall information provided in the database can be divided into two main categories, i.e., the key column properties (e.g., material properties, geometric properties, confinement details and test configuration) and experimental test results (e.g., failure modes, bearing capacities, force–displacement relationship, axial load effect, and damage pattern). The database is currently available on the worldwide websites of the University of Washington [81] and the PEER centre [82]. However,

PEER website provides additional information about the experimental work, i.e., structural drawings, testing images, etc., compared to the University of Washington's website.

The distribution of some important column properties, i.e., depth, aspect ratio, axial load ratio, longitudinal reinforcement ratio and transverse reinforcement ratio, was examined to evaluate the comprehensiveness and universal nature of the model database (Table 1). From the detailed statistical analysis, the database was found to be normally distributed about a mean value of 319 mm with respect to depth, 3.6 with respect to the aspect ratio, and 2.39% with respect to longitudinal reinforcement ratio. However, it was found to be skewed more towards the lesser values of aspect ratio, longitudinal reinforcement ratio and axial load ratio. The analysis has also shown that approximately 80% of the columns have depths ranging between 200 and 500 mm, whereas 65% have an axial load ratio of between 0 and 0.3. The database was also found to be weighted around a mean value of 2% of the transverse reinforcement ratio. However, no characterizable distribution was observed. Analysing the distribution characteristics of all the above discussed column properties, the database can be said to possess population samples from one extreme to the other following a specific distribution, more often a normal distribution, which indicates its inclusiveness and universality.

Table 1. Database Characteristics.

Property	Unit	Statistical Parameters		
		Mean	**STD**	**COV**
Depth	mm	319	117	0.37
Aspect Ratio	Decimals	3.58	1.46	0.41
Axial Load Ratio	Decimals	0.27	0.19	0.70
Longitudinal Reinforcement Ratio	%	2.39	0.96	0.40
Transverse Reinforcement Ratio	%	2.01	1.22	0.61

STD: Standard Deviation, COV: Coefficient of Variations.

2.2. Parametric Selection

The structural performance database assembled by the PEER centre provides more than 40 parameters describing all aspects of the seismic response of RC rectangular columns. The selection of output variables was made conferring to the set objective (i.e., determining the bearing capacity of RC rectangular columns) according to which flexural capacity and shear capacity outputs were selected. However, for selecting input variables against the established outputs, a detailed stepwise procedure was adopted. A linear regression analysis was performed using Microsoft Excel while considering all the available input variables in the first step. The analysis results showed a strong relationship between the inputs and outputs with a coefficient of determination (R^2) of 0.9408 for flexural capacity output and 0.9339 for shear capacity output. However, the number of variables employed in the regression analysis was found to be too much to be expressed by a simple mathematical expression. Therefore, reducing the number of inputs was decided while ensuring minimum damage to the statistical relationship between the inputs and outputs. To reduce the number of input parameters, all the groups of similar variables were replaced by their corresponding single universal variable (e.g., instead of width and height, area parameter was used, which is the combined/universal variable for the preceding two variables), based on the fundamental principles of civil and structural engineering. This utilization of a single universal variable instead of a number of default variables reduced the number of input parameters to 12, almost one-third of their initial quantity, i.e., 35. Linear regression analysis was performed again to assess the damage caused by the reduction of input variables to the statistical relationship between the inputs and outputs. However, the analytical results showed no significant impairment as the coefficient of determination was found to be 0.8912 in the case of flexural capacity output, and 0.8784 in the case of shear capacity output.

A combined approach of the linear regression and cosine amplitude methods was employed to further reduce the number of input variables for the model development. Linear regression along with analysis of variance (ANOVA) was performed in Microsoft Excel to calculate the *p*-value, whereas the cosine amplitude method was used to compute the relationship coefficient *R* for each parameter (Table 2). The *R*-value was obtained according to Equation (1), where X_1 and X_2 represent the parameters, whose relationship is to be determined, and "*m*" epitomises the sample size. For the acceptability of each of the above discussed modified parameters, a maximum threshold of 0.05 for the *p*-value, and a minimum threshold of 0.75 for the *R*-value, were established. Any parameter satisfying either of these criteria for the flexural or shear capacity output was selected as the model input for predicting the bearing capacity of RC rectangular columns. This combined approach of linear regression and cosine amplitude method reduced the number of input variables to approximately half of their previous quantity, i.e., seven (Table 3). Linear regression analysis was performed again to evaluate the damage caused by this reduction of input variables to the statistical relationship between the inputs and outputs. However, the analysis results showed no substantial loss as the coefficient of determination was 0.8896 for flexural capacity output, and 0.8760 for shear capacity output, which is similar to that calculated earlier employing 12 variables.

$$R - Value = \frac{|\sum_{k=1}^{m} X_{1k} \cdot X_{2k}|}{\sqrt{\left(\sum_{k=1}^{m} X_{1k}^2\right) \left(\sum_{k=1}^{m} X_{2k}^2\right)}}, \quad (1)$$

Table 2. Parametric Selection Based on Linear Regression and Cosine Amplitude.

Parameters	Flexural Capacity Output		Shear Capacity Output	
	p-Value	*R*-Value	*p*-Value	*R*-Value
Column Length	0.2209	0.8101	1.04×10^{-44}	0.7199
Cross Sectional Area	6.72×10^{-12}	0.9342	1.59×10^{-11}	0.8889
Long. Rein. Ratio	0.0005	0.6239	0.0102	0.6985
Long. Rein. Yield Strength	0.4242	0.6818	0.5124	0.7348
Long. Rein. Ultimate Strength	0.8822	0.6417	0.2295	0.6391
Trans. Rein. Ratio	0.1707	0.5013	0.1874	0.5181
Trans. Rein. Yield Strength	0.7121	0.5861	0.0921	0.6881
Trans. Rein. Ultimate Strength	0.7703	0.5797	0.0995	0.6186
Concrete Compressive Strength	0.0083	0.5741	0.0718	0.6585
Applied Axial Load	9.25×10^{-18}	0.8053	3.64×10^{-21}	0.7979
Design Axial Load	0.0081	0.9068	9.63×10^{-05}	0.8839
Clear Cover	0.0027	0.7468	0.3666	0.7707

Long.: Longitudinal, Rein.: Reinforcement, Trans.: Transverse.

Table 3. Statistical Characteristics of Model Parameters.

Parameters	Symbol	Unit	Type	Minimum	Maximum	Mean	STD
Column Length	L	m	Input	0.08	2.34	1.095	0.5485
Cross Sectional Area	A	cm^2	Input	64	4180.64	1021.6	777.88
Long. Rein. Ratio	ρ	Decimal	Input	0.007	0.0603	0.024	0.0101
Concrete Comp. Strength	f'_c	MPa	Input	16	118	51.91	29.244
Applied Axial Load	P_A	KN	Input	0	8000	1238.33	1350.28
Design Axial Load	P_D	KN	Input	109.51	7359.6	2424.26	1421.61
Clear Cover	C_c	cm	Input	0	6.51	2.395	1.0855
Flexural Capacity	M_F	KN-m	Output	2	1680	264.75	264.55
Shear Capacity	V_S	KN	Output	23	1339	207.78	176.55

Comp.: Compressive, P_D: Function of Geometric and Material Properties.

2.3. GEP Modelling

The modelling of flexural and shear capacity models for predicting the overall bearing capacity of RC rectangular columns was carried out using an extremely flexible GEP data modelling software, GeneXproTools 5.0. This is an easy and resourceful data mining software especially designed for simple functional regression, logistic regression, classification, time series prediction and logic synthesis. A well-organized Microsoft Excel database consisting of seven input variables against the established outputs was imported into GeneXproTools to initialize the modelling process. For effective and robust model development, all the 250 population samples were regarded as both the training and validation datasets [75,83–85]. GeneXproTools exhibits a distinct characteristic where it allows the user to specify the key modelling parameters, i.e., the number of chromosomes, head size, number of genes, linking function, constant per gene, model functions, etc., for the preparation of the predictive model. Exploiting this characteristic of GeneXproTools, multiple GEP models were created using different combinations of modelling parameters. A brief summary of the modelling characteristics of these GEP models is given in Table 4. The performance of all these models was then evaluated based on the four most commonly used fitness measures, i.e., coefficient of determination (R^2), root mean squared error (*RMSE*), mean absolute error (*MAE*), and root relative squared error (*RRSE*). A higher R^2, and lower *RMSE*, *MAE*, and *RRSE*, usually represent a better fit model. There also exists another performance measure, i.e., the correlation coefficient. However, this was ignored because of its insensitivity towards the multiplication and division of outputs to a constant [86]. Based on these performance indicators, the best fitted model was proposed for future practical purposes.

- Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(P_i - T_i)^2}{\sum_{i=1}^{m}(P_i - \overline{T})^2}, \tag{2}$$

- Root Mean Squared Error (*RMSE*)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m}(P_i - T_i)^2}{m}}, \tag{3}$$

- Mean Absolute Error (*MAE*)

$$MAE = \frac{\sum_{i=1}^{m}|P_i - T_i|}{m}, \tag{4}$$

- Root Relative Squared Error (*RRSE*)

$$RRSE = \sqrt{\frac{\sum_{i=1}^{m}(P_i - T_i)^2}{\sum_{i=1}^{m}(T_i - \overline{T})^2}}, \tag{5}$$

Table 4. Summary of GEP Models for Predicting the Bearing Capacity of RC Rectangular Columns.

GEP Models	Model Details								Performance Indicators			
	Chromosomes	Head Size	Genes	Linking Function	Fitness Function	Model Functions	Input Variables	Variables Used	R^2	*RMSE*	*MAE*	*RRSE*
Flexural Capacity Models (FCM)												
FCM 1	30	8	3	+	*RMSE*	+, −, *, /	7	7	0.9228	73.42	47.21	0.2781
FCM 2	30	8	3	−	*RMSE*	+, −, *, /	7	5	0.9275	71.10	48.03	0.2693
FCM 3	30	8	3	*	*RMSE*	+, −, *, /	7	6	0.9448	62.17	41.82	0.2355
FCM 4	30	8	3	/	*RMSE*	+, −, *, /	7	7	0.9454	61.94	45.67	0.2346
FCM 5	30	8	3	Average	*RMSE*	+, −, *, /	7	7	0.9233	74.09	47.03	0.2806
FCM 6	30	8	3	Minimum	*RMSE*	+, −, *, /	7	7	0.9221	74.33	46.89	0.2815
FCM 7	30	8	3	Maximum	*RMSE*	+, −, *, /	7	6	0.9156	88.17	60.35	0.3340

Table 4. *Cont.*

GEP Models	Model Details								Performance Indicators			
	Chromosomes	Head Size	Genes	Linking Function	Fitness Function	Model Functions	Input Variables	Variables Used	R^2	*RMSE*	*MAE*	*RRSE*
FCM 8	80	8	3	/	*RMSE*	+, −, *, /	7	7	0.9496	59.68	41.90	0.2260
FCM 9	50	8	3	/	*RMSE*	+, −, *, /	7	7	0.9614	53.41	38.12	0.2023
FCM 10	50	12	4	/	*RMSE*	+, −, *, /	7	7	0.9376	66.01	42.78	0.2500
FCM 11	50	5	2	/	*RMSE*	+, −, *, /	7	5	0.9298	70.13	43.56	0.2656
FCM 12	50	8	3	/	*RMSE*	+, −, *, /, $\sqrt{}$	7	6	0.9362	66.97	45.06	0.2536
FCM 13	50	8	3	/	*RMSE*	+, −, *, /, ln	7	7	0.9264	71.63	43.96	0.2713
Shear Capacity Models (SCM)												
SCM 1	30	8	3	+	*RMSE*	+, −, *, /	7	6	0.9233	48.81	32.94	0.2770
SCM 2	30	8	3	−	*RMSE*	+, −, *, /	7	7	0.9012	55.40	39.05	0.3144
SCM 3	30	8	3	*	*RMSE*	+, −, *, /	7	6	0.9138	51.80	37.88	0.2940
SCM 4	30	8	3	/	*RMSE*	+, −, *, /	7	6	0.9246	49.11	35.69	0.2787
SCM 5	30	8	3	Average	*RMSE*	+, −, *, /	7	7	0.9032	56.41	38.40	0.3201
SCM 6	30	8	3	Minimum	*RMSE*	+, −, *, /	7	7	0.9133	53.36	40.25	0.3029
SCM 7	30	8	3	Maximum	*RMSE*	+, −, *, /	7	7	0.9234	48.88	35.24	0.2774
SCM 8	80	8	3	/	*RMSE*	+, −, *, /	7	6	0.9268	47.71	35.05	0.2708
SCM 9	50	8	3	/	*RMSE*	+, −, *, /	7	7	0.9481	40.23	29.13	0.2283
SCM 10	50	12	4	/	*RMSE*	+, −, *, /	7	7	0.9052	54.30	33.45	0.3082
SCM 11	50	5	2	/	*RMSE*	+, −, *, /	7	5	0.8942	58.46	44.75	0.3318
SCM 12	50	8	3	/	*RMSE*	+, −, *	7	7	0.9512	39.47	28.77	0.2240
SCM 13	50	8	3	/	*RMSE*	+, −, *, /, $\sqrt{}$	7	5	0.9287	47.21	34.01	0.2679

Here P represents the predicted values, T represents the tested values or experimental values, $\overline{T}$ represents the mean tested value, and m represents the number of population samples.

3. Results

The nominal failure in RC rectangular columns subjected to earthquake loading occurs when either of their shear or flexural capacity is reached. This is because shear strength and flexural strength are the only parameters triggering the failure of RC rectangular columns in seismic conditions. Based on this fact, two types of capacity prediction models, i.e., the flexural capacity model and shear capacity model, are required to explain their overall bearing capacity comprehensively. Based on the results obtained from both these models, the ultimate bearing capacity or lateral load carrying capacity (V_u) of the aforementioned columns can be obtained using Equation (6).

$$V_u = \min\left(\frac{M_F}{L}, V_S\right), \tag{6}$$

3.1. Flexural Capacity Model

To develop the analytical model for the prediction of the flexural capacity of RC rectangular columns, multiple GEP models were created in GeneXproTools with different combinations of modelling parameters, a summary of which is shown in Table 4. The performance of all these models was assessed based on the already defined fitness indicators, i.e., R^2, *RMSE*, *MAE*, and *RRSE*. Based on the evaluation results, the best performing GEP model, i.e., the model with the highest value of R^2 and least value of *RMSE*, *MAE*, and *RRSE*, was selected and proposed for predicting the flexural capacity of RC rectangular columns for future practical purposes. The GEP expression tree was extracted from the modelling software GeneXproTools as shown in Figure 3, to ensure the effective and easy utilization of the proposed model in real-world problems. This expression tree was then decoded to be presented in the form of simple mathematical expressions as given in Equations (7)–(10). Here, Equation (7) can be used to predict the flexural capacity of RC rectangular columns, the input parameters of which can be calculated using Equations (8)–(10). The rest of the variables used in all these equations are already defined in Table 3. However, attention must be paid to the measuring units.

$$M_F = \frac{X_1}{X_2 \cdot X_3}, \tag{7}$$

$$X_1 = P_D - [\{\rho L(P_D - P_A)\} \cdot \{C_c - 10.963L\}], \tag{8}$$

$$X_2 = \rho + \frac{1}{\rho P_A + 11.8583}, \tag{9}$$

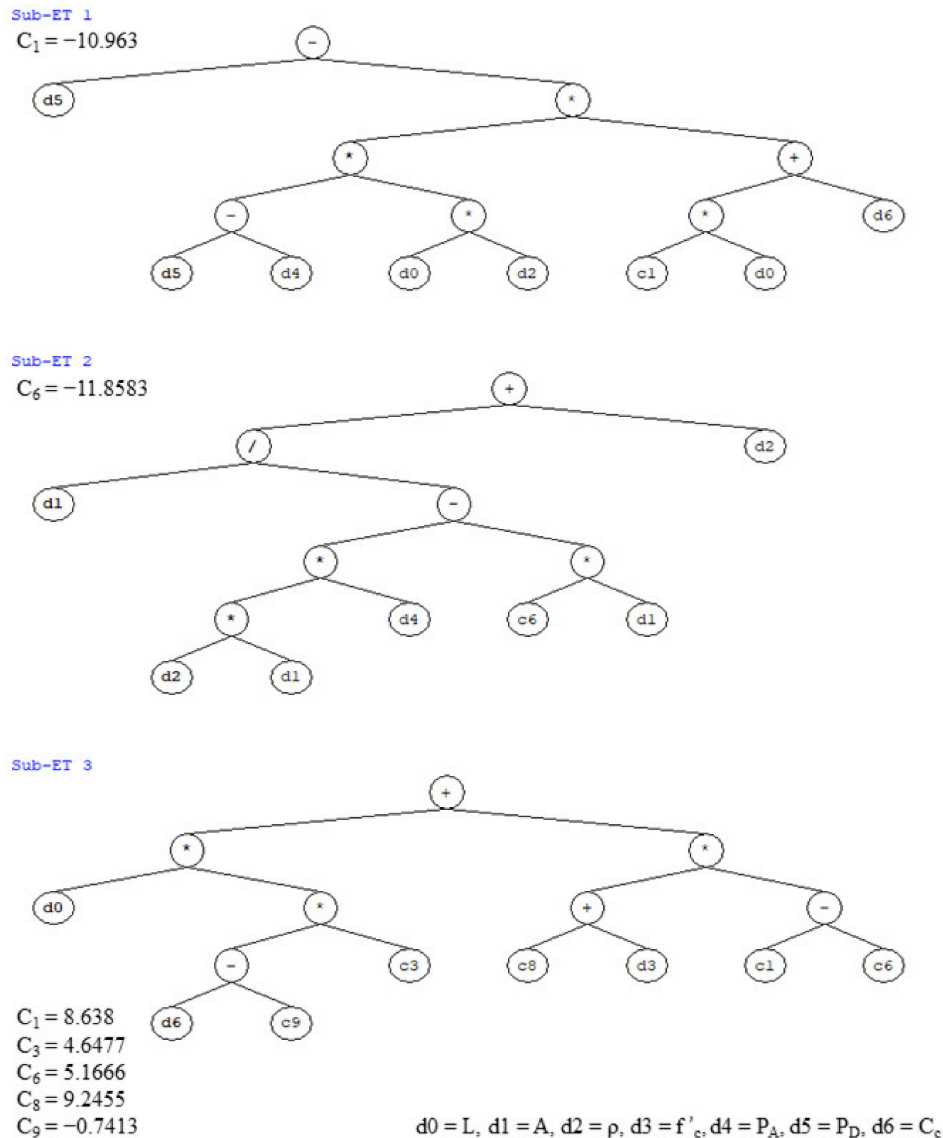$$X_3 = 4.6477L\,(C_c + 0.7413) + 3.4714f'_c + 32.0948, \tag{10}$$

Figure 3. GEP Expression Tree for Flexural Capacity Model.

3.2. Shear Capacity Model

A similar procedure was adopted to develop the analytical model for predicting shear capacity of RC rectangular columns as used for the development of the flexural capacity model. The best performing GEP model from all those shown in Table 4 was selected based on the same performance criteria as described earlier in the flexural capacity model. The expression tree of this model (Figure 4) was extracted from GeneXproTools, which was then decoded and expressed in terms of simplified mathematical expression. Here, Equation (11) was proposed to predict the shear capacity of RC rectangular columns. The parameters involved in this equation can be calculated using Equations (12)–(14).

$$V_S = \frac{X_A}{X_B \cdot X_C}, \tag{11}$$

$$X_A = 1.4668A\,(P_A + 400.271) + (f'_c + P_D)^2, \tag{12}$$

$$X_B = A + \left[\{2L\,(C_c + 6.9691)\}\,.\,\{Lf'_c + 60.6393\}\right], \tag{13}$$

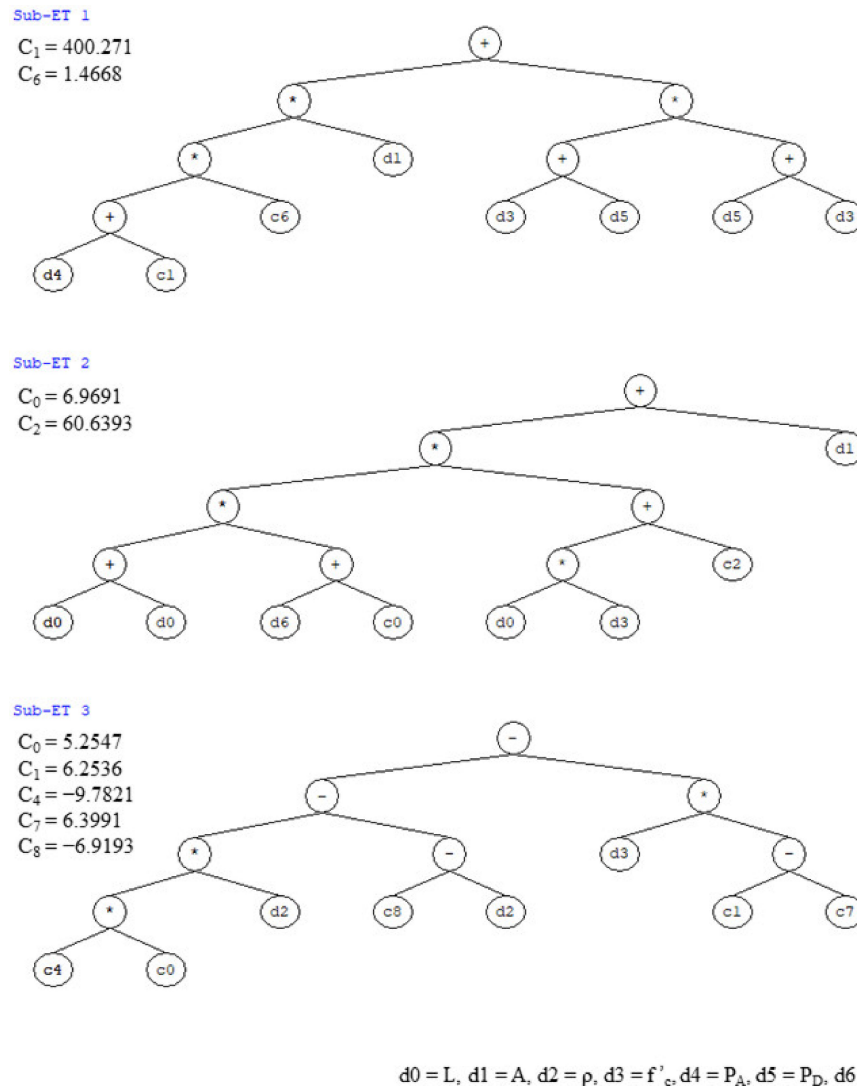$$X_C = 0.1455f'_c - 50.402\rho + 6.9193, \tag{14}$$

Figure 4. GEP Expression Tree for Shear Capacity Model.

3.3. Model Validation

In AI-based data modelling, it is often recommended to conduct a variety of statistical analyses to ensure the proposed models' robustness and universality [67]. Two types of such analysis, i.e., the sensitivity analysis and the parametric analysis, were employed in this research to investigate whether the proposed models represent the actual physical phenomenon or not [85]. The sensitivity analysis is normally used to explore input variables' relative contribution (RLC) towards the overall model development. The mathematical expressions for conducting the sensitivity analysis are given in Equations (15) and (16) [67]. Here, $f_{max}\,(x_i)$ and $f_{min}\,(x_i)$ represent the maximum and minimum predicted output corresponding to the ith input variable in the database while maintaining all other variables at their average value. From the sensitivity analysis results, P_D was found to be the most significant input variable contributing around 50% towards the development of flexural and shear capacity models of RC rectangular columns (Figure 5). The rest of six variables

i.e., L, f'_c, P_A, ρ, C_c and A cumulatively account for the remaining 50% of the overall model development.

$$RLC\ (\%) = \frac{N_i}{\sum_{i=1}^{n} N_i} . 100, \tag{15}$$

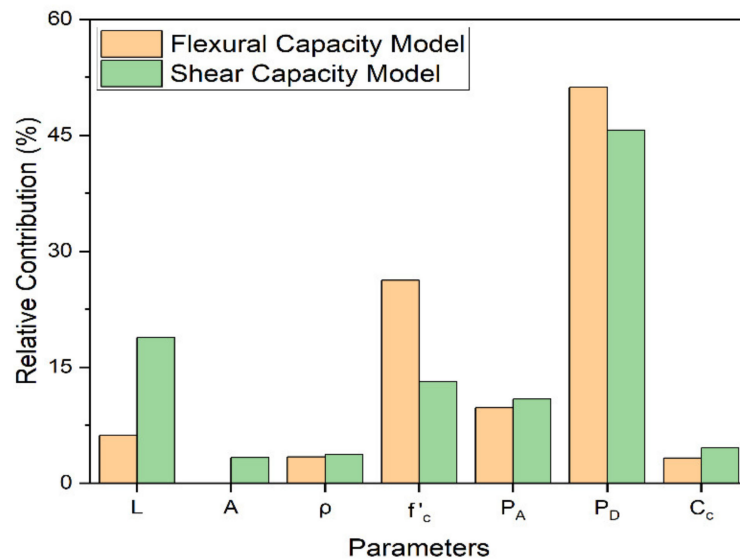$$N_i = f_{max}(x_i) - f_{min}(x_i), \tag{16}$$

Figure 5. Relative Contribution of Input Parameters to the Model Development.

3.3.1. Impact Assessment of Model Influencing Parameters

Parametric analysis was conducted to further validate the efficiency of the proposed models to capture behind-scenes real-world phenomena. In the parametric analysis, the trend of the output variable to a certain change in a specific input variable was observed while keeping all the other variables constant at their average value. A detailed description of the parametric analysis results is given in the following sections.

Geometric Parameters

Two geometric parameters, i.e., L and A were selected as the input variables for developing GEP models for predicting the bearing capacity of RC rectangular columns. The impact of these variables on the overall bearing capacity of the aforementioned columns was investigated based on parametric analysis, the results of which are shown in Figure 6. The figure shows an increase in the flexural and shear capacity of RC rectangular columns along with the decrease in length, and the increase in the cross-sectional area of columns. The inverse relationship between the length and flexural or shear capacity of the RC rectangular columns is because of the probability of buckling phenomenon in the long columns. Similar results were also presented in the nonlinear finite element-based study conducted by Mahmood and Ghulam [14], which ensures the strength and effectiveness of the proposed models.

Material Parameters

f'_c was the only material characteristic considered in developing the proposed models. Based on the parametric analysis, its impact on the overall bearing capacity of RC rectangular columns is presented in Figure 6. This figure shows a direct relationship between the flexural and shear capacity of RC rectangular columns and f'_c. A similar relationship was also observed by the finite element-based research program of Mahmood and Ghulam [14]. This relationship can also be validated by Equation (18) of ACI 318-19 [87].
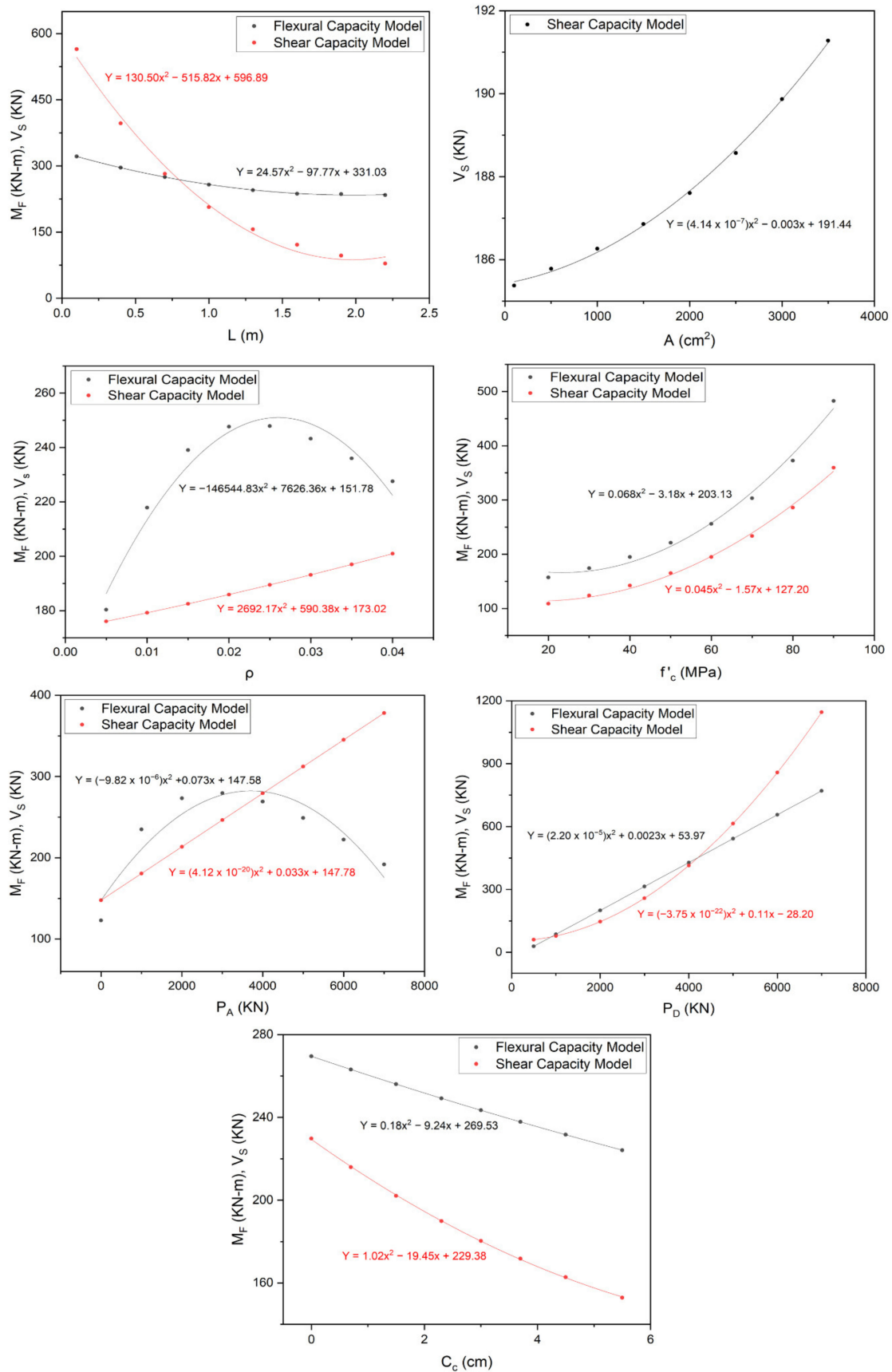
Figure 6. Parametric Analysis Results for Proposed Capacity Prediction Models.

Structural Parameters

Four structural parameters of RC rectangular columns, i.e., ρ, P_A, P_D, and C_c, were considered in developing their capacity prediction GEP models. The impact of these parameters on the overall bearing capacity of RC rectangular columns was analyzed based on parametric analysis, the results of which are presented in Figure 6. The figure shows that the flexural and shear capacity of the RC rectangular column increases along the increase in ρ, P_A, and P_D, whereas it decreases along with the increase in C_c. Moreover, flexural capacity was found to be declining after reaching a certain threshold of ρ and P_A. These observations were found to be well in agreement with the non-linear finite element-based investigations of Mahmood and Ghulam [14] and the design guidelines of ACI 318-19 [87]. The trend of predicted output corresponding to a certain change in these parameters was also found to be consistent with the experimental results in the database which further authenticates the validity of the proposed models.

3.4. Performance Evaluation

The performance of both the proposed capacity prediction models was evaluated based on four key fitness indicators, i.e., R^2, *RMSE*, *MAE*, and *RRSE*, as described earlier in the "GEP Modelling" section. For this evaluation, some fundamental statistical analyses were employed, which reported these performance indicators to be 0.9614, 53.41, 38.12, and 0.2023, respectively, for the flexural capacity model, and 0.9512, 39.47, 28.77, and 0.2240, respectively, for the shear capacity model (The details of these evaluation results are shown in Figure 7). From these performance indicators, it can be observed that there exists a strong correlation between the predicted and experimental output with the least possible error (Figure 8), which indicates the higher prediction accuracy of the proposed models. In addition to these fitness criteria, the performance of the proposed models was also assessed by comparing with the prevailing design code of concrete structures in the country (Pakistan) i.e., ACI 318-19 [87]. (Explanation: The comparison was made only with ACI 318-19 [87] because almost all the design codes provide similar guidelines for the capacity prediction of discussed RC members. Moreover, no specific mathematical expressions were found by the author in past research articles, on the basis of which comparison could have been made.) For the comparison, the flexural and shear strength of RC rectangular columns present in the database were computed in accordance with the guidelines of sections 22.4 and 22.5 of ACI 318-19 [87], the mathematical expressions for which are also given in Equations (17)–(20). All the parameters involved in these equations are defined clearly in Chapter 2 of the given design code. The performance of ACI model was also evaluated by employing the same statistical analysis. The results of these analyses reported R^2, *RMSE*, *MAE*, and *RRSE* to be 0.8849, 101.86, 51.74, and 0.3858, respectively, for flexural capacity predictions, whereas 0.8737, 238.74, 183.66, and 1.35 respectively for shear capacity predictions (Figure 9). The comparison depicted a better performance and higher accuracy of the proposed models than that of ACI 318-19 [87]. Following the model validation and performance evaluation results, given models can be recommended for computing the lateral load-carrying capacity of RC rectangular columns for future practical purposes.

$$V_S = V_c + V_{st}, \tag{17}$$

$$V_c = \left[0.66\lambda(\rho_w)^{\frac{1}{3}}\sqrt{f'_c} + \frac{N_u}{6A_g} \right] b_w d, \tag{18}$$

$$V_{st} = \frac{A_v f_{yt} d}{S}, \tag{19}$$
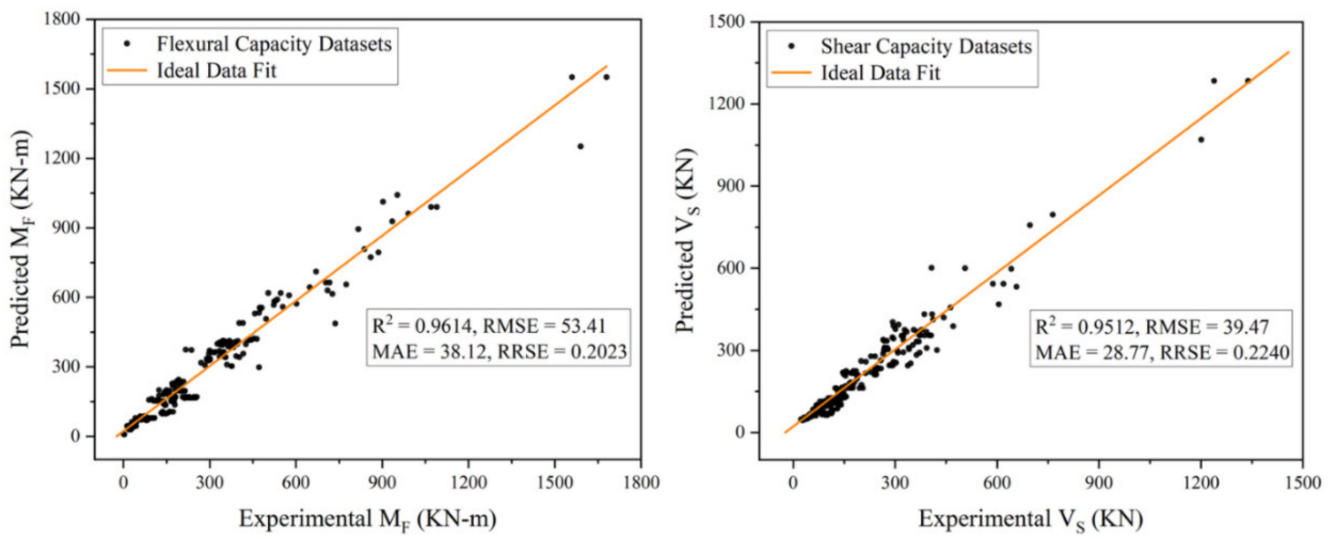
$$V_S = \frac{2M_F}{L}, \tag{20}$$

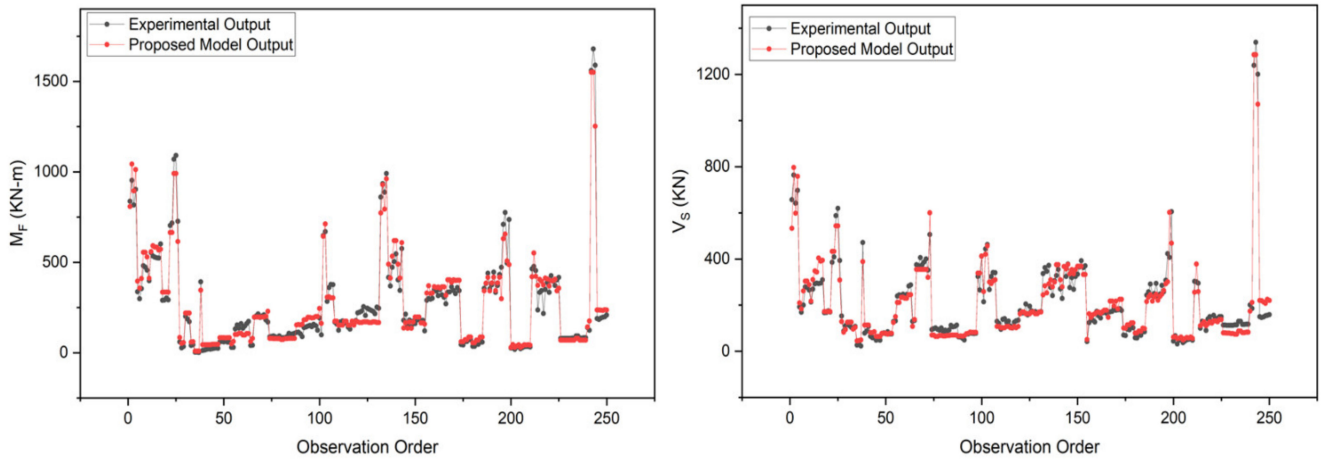**Figure 7.** Predicted Versus Experimental Values Plot for Proposed Models.

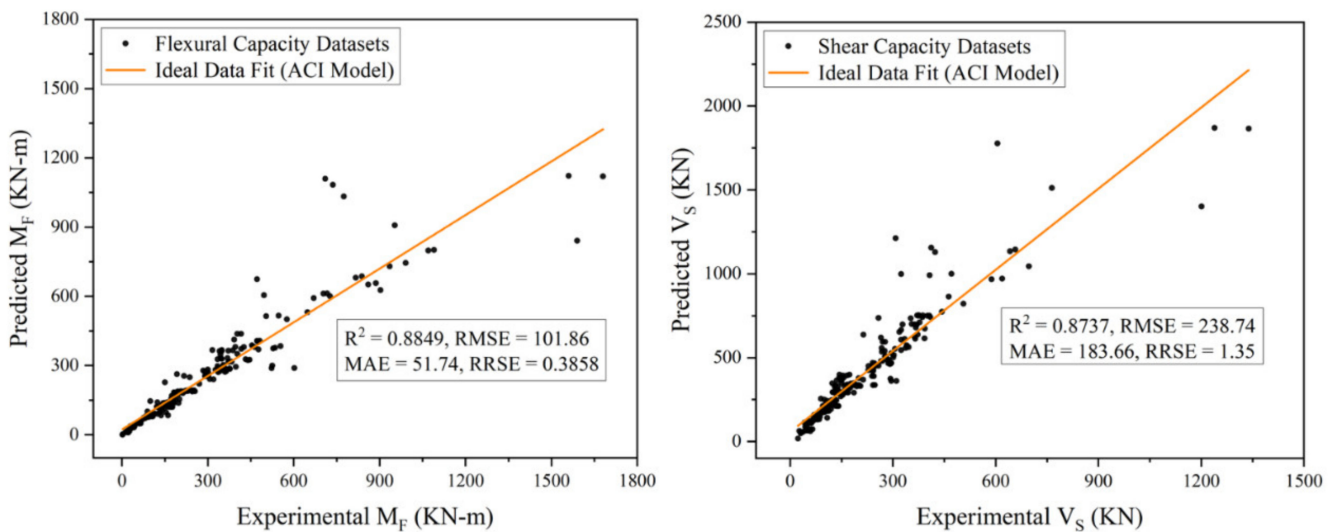**Figure 8.** Comparison of Experimental and Proposed Model Results.

Figure 9. Predicted Versus Experimental Values Plot for ACI Models.

4. Conclusions

This research presented a novel AI approach to GEP for predicting the ultimate lateral load-carrying capacity of RC rectangular columns subjected to earthquake loading. For this, an experimental database assembled by the PEER centre consisting of 250 cyclic tests of RC rectangular columns was utilized. Based on statistical GEP modelling results as presented in the earlier sections, the following conclusions can be drawn:

- The proposed AI technique provides an alternative method for the determination of lateral load carrying capacity of RC rectangular columns while avoiding complicated structural and mathematical computations. Moreover, it is also simpler and easier to be implemented in practical applications.
- The proposed capacity prediction models were found to exhibit better accuracy when compared to that of the ACI model. The major performance indicator, i.e., R^2, was found to be 0.9614 and 0.9512 in the proposed flexural and shear capacity model, respectively, and 0.8849 and 0.8737 in the case of flexural and shear capacity models ACI, respectively.
- Design axial load (P_D) was found to be the most significant input variable, contributing around 50% towards the development of both the proposed models. The rest of the six input variables were observed to cumulatively account for the remaining 50% of the overall model development.
- From the parametric analysis results of the proposed models, the trend of output variables corresponding to most of the input variables was found to be consistent with the experimental results in the database, which validates the ability of the proposed models to capture behind the scenes real-world phenomena.

The main incentive of the proposed models is that they are based on a "black box" model that relates the inputs and outputs in a simplified manner while avoiding complex mathematical derivations. It has better performance as compared to the traditional analytical models and therefore can be employed to predict the lateral load carrying capacity of RC rectangular columns quickly, accurately, and robustly. Moreover, the proposed GEP models can be expressed in terms of simple mathematical expressions, which is not possible in some of the other conventional and modern analytical techniques. Apart from all its advantages, it is associated with certain limitations. The proposed models are unable to explain the failure mode, as well as the involved mechanism in predicting the lateral load carrying capacity of RC rectangular columns. Moreover, as it was not feasible to consider every involved parameter in the model development, the applicability of the proposed models in special cases is questionable. Considering all the pros and cons of the proposed capacity prediction models, future researchers are recommended to work on the best AI algorithm, either individual or ensembled [88–96], that considers all the possible aspects and explains the mechanism involved.

Author Contributions: Conceptualization, R.A. (Raheel Asghar) and M.F.J.; methodology, R.A. (Raid Alrowais); software, R.A. (Raheel Asghar); validation, A.M. and N.I.V.; formal analysis, A.M.M.; investigation, A.K.; data curation, R.A. (Raid Alrowais); writing—original draft preparation, R.A. (Raheel Asghar); writing—review and editing, A.K.; supervision, M.F.J. 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: The data presented in this study can be obtained on request from the corresponding author.

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

References

1. Isleem, H.F.; Abid, M.; Alaloul, W.S.; Shah, M.K.; Zeb, S.; Musarat, M.A.; Javed, M.F.; Aslam, F.; Alabduljabbar, H. Axial Compressive Strength Models of Eccentrically-Loaded Rectangular Reinforced Concrete Columns Confined with FRP. *Materials* **2021**, *14*, 3498. [CrossRef] [PubMed]
2. Buka-Vaivade, K.; Sliseris, J.; Serdjuks, D.; Pakrastins, L.; Vatin, N.I. Rational use of HPSFRC in multi-storey building. *Mag. Civ. Eng.* **2018**, *84*, 3–14.
3. Ahmad, S.; Hasnain, G. FEA of reinforced concrete beam-column joint with steel fibers for cyclic loading. *Int. J. Struct. Integr.* **2021**, *12*, 670–687. [CrossRef]
4. Li, H.; Xiao, S.; Huo, L. Damage investigation and analysis of engineering structures in the Wenchuan earthquake. *J. Build. Struct.* **2018**, *4*, 10–19.
5. Deierlein, G.; Moehle, J. A Framework Methodology for Performance-Based Earthquake Engineering. In Proceedings of the 13th World Conference on Earthquake Engineering, Vancouver, BC, Canada, 1–6 August 2004; p. 14.
6. Lao, X.-C.; Han, X.-L. Performance index limits of high reinforced concrete shear wall components. *J. Cent. South Univ. Technol.* **2011**, *18*, 1248–1255. [CrossRef]
7. Asghar, R.; Shahzad, A.; Amjad, S.U.; Akhtar, A. Comparative Study on the Seismic Performance of Bare Frame and Infilled Frame RC Structures with Brick Masonry and Low Strength Concrete Block Masonry Infills. In Proceedings of the 2nd Conference on Sustainability in Civil Engineering CSCE20, Islamabad, Pakistan, 12 August 2020; 2020; p. 8.
8. Asghar, R.; Shahzad, A.; Amjad, S.U.; Akhtar, A. Experimental Determination of the Mechanical Properties of Brick Masonry and Low Strength Concrete Block Masonry. In Proceedings of the 2nd Conference on Sustainability in Civil Engineering CSCE20, Islamabad, Pakistan, 12 August 2020; p. 6.
9. Qi, Y.-L.; Han, X.-L.; Ji, J. Failure mode classification of reinforced concrete column using Fisher method. *J. Cent. South Univ.* **2013**, *20*, 2863–2869. [CrossRef]
10. Acun, B.; Sucuoğlu, H. The effect of displacement history on the performance of concrete columns in flexure. In *Advances in Performance-Based Earthquake Engineering*; Springer: Dordrecht, The Netherlands, 2010; pp. 373–382.
11. Zhu, L.; Elwood, K.J.; Haukaas, T. Classification and Seismic Safety Evaluation of Existing Reinforced Concrete Columns. *J. Struct. Eng.* **2007**, *133*, 1316–1330. [CrossRef]
12. Mangalathu, S.; Jeon, J.-S. Machine Learning–Based Failure Mode Recognition of Circular Reinforced Concrete Bridge Columns: Comparative Study. *J. Struct. Eng.* **2019**, *145*, 04019104. [CrossRef]
13. Feng, D.-C.; Liu, Z.-T.; Wang, X.-D.; Jiang, Z.-M.; Liang, S.-X. Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm. *Adv. Eng. Inform.* **2020**, *45*, 101126. [CrossRef]
14. Abbas, R.M.; Awazli, A.G. Behavior of Reinforced Concrete Columns Subjected to Axial Load and Cyclic Lateral Load. *J. Eng.* **2017**, *23*, 21–40.
15. Lazarevska, M.; Cvetkovska, M.; Knežević, M.; Gavriloska, A.T.; Milanovic, M.; Murgul, V.; Vatin, N. Neural Network Prognostic Model for Predicting the Fire Resistance of Eccentrically Loaded RC Columns. In *Applied Mechanics and Materials*; Trans Tech Publications Ltd.: Bäch, Switzerland, 2014; Volume 627, pp. 276–282.
16. Khan, S.; Khan, M.A.; Zafar, A.; Javed, M.F.; Aslam, F.; Musarat, M.A.; Vatin, N.I. Predicting the Ultimate Axial Capacity of Uniaxially Loaded CFST Columns Using Multiphysics Artificial Intelligence. *Materials* **2021**, *15*, 39. [CrossRef] [PubMed]
17. Iqbal, M.; Zhang, D.; Jalal, F.E.; Javed, M.F. Computational AI prediction models for residual tensile strength of GFRP bars aged in the alkaline concrete environment. *Ocean Eng.* **2021**, *232*, 109134. [CrossRef]
18. 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]
19. Iqbal, M.; Zhang, D.; Jalal, F.E. Durability evaluation of GFRP rebars in harsh alkaline environment using optimized tree-based random forest model. *J. Ocean Eng. Sci.* **2021**, *in press*. [CrossRef]
20. Kaloop, M.R.; Samui, P.; Iqbal, M.; Hu, J.W. Soft computing approaches towards tensile strength estimation of GFRP rebars subjected to alkaline-concrete environment. *Case Stud. Constr. Mater.* **2022**, *16*, e00955. [CrossRef]
21. Hwang, S.-J.; Lee, H.-J. Strength Prediction for Discontinuity Regions by Softened Strut-and-Tie Model. *J. Struct. Eng.* **2002**, *128*, 1519–1526. [CrossRef]
22. Sezen, H.; Moehle, J.P. Shear Strength Model for Lightly Reinforced Concrete Columns. *J. Struct. Eng.* **2004**, *130*, 1692–1703. [CrossRef]
23. Vecchio, F.J.; Collins, M.P. The modified compression-field theory for reinforced concrete elements subjected to shear. *ACI J.* **1986**, *83*, 219–231.
24. Bentz, E.C.; Vecchio, F.J.; Collins, M.P. Simplified Modified Compression Field Theory for Calculating Shear Strength of Reinforced Concrete Elements. *ACI Struct. J.* **2006**, *103*, 614.
25. Hsu, T.T. Softened truss model theory for shear and torsion. *Struct. J.* **1988**, *85*, 624–635.
26. Pang, X.-B.D.; Hsu, T.T. Fixed angle softened truss model for reinforced concrete. *Struct. J.* **1996**, *93*, 196–208.
27. Feng, D.-C.; Wu, G.; Sun, Z.-Y.; Xu, J.-G. A flexure-shear Timoshenko fiber beam element based on softened damage-plasticity model. *Eng. Struct.* **2017**, *140*, 483–497. [CrossRef]
28. Feng, D.-C.; Ren, X.-D.; Li, J. Softened Damage-Plasticity Model for Analysis of Cracked Reinforced Concrete Structures. *J. Struct. Eng.* **2018**, *144*, 04018044. [CrossRef]

29. Berry, M.; Parrish, M.; Eberhard, M. *PEER Structural Performance Database User's Manual (Version 1.0)*; University of California: Berkeley, CA, USA, 2004.

30. Russell, S.J.; Norvig, P.; Canny, J.F.; Malik, J.M.; Edwards, D.D. *Artificial Intelligence: A Modern Approach. Vol. 2*; Prentice Hall: Upper Saddle River, NJ, USA, 2003.

31. Salehi, H.; Burgueño, R. Emerging artificial intelligence methods in structural engineering. *Eng. Struct.* **2018**, *171*, 170–189. [CrossRef]

32. Gholampour, A.; Gandomi, A.; Ozbakkaloglu, T. New formulations for mechanical properties of recycled aggregate concrete using gene expression programming. *Constr. Build. Mater.* **2017**, *130*, 122–145. [CrossRef]

33. Lim, J.C.; Karakus, M.; Ozbakkaloglu, T. Evaluation of ultimate conditions of FRP-confined concrete columns using genetic programming. *Comput. Struct.* **2016**, *162*, 28–37. [CrossRef]

34. Perera, R.; Arteaga, A.; De Diego, A. Artificial intelligence techniques for prediction of the capacity of RC beams strengthened in shear with external FRP reinforcement. *Compos. Struct.* **2010**, *92*, 1169–1175. [CrossRef]

35. Aval, S.B.; Ketabdari, H.; Gharebaghi, S.A. Estimating Shear Strength of Short Rectangular Reinforced Concrete Columns Using Nonlinear Regression and Gene Expression Programming. In *Structures*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 12, pp. 13–23.

36. Jalal, M.; Ramezanianpour, A.A.; Pouladkhan, A.R.; Tedro, P. RETRACTED ARTICLE: Application of genetic programming (GP) and ANFIS for strength enhancement modeling of CFRP-retrofitted concrete cylinders. *Neural Comput. Appl.* **2013**, *23*, 455–470. [CrossRef]

37. HIsleem, H.F.; Tayeh, B.A.; Alaloul, W.S.; Musarat, M.A.; Raza, A. Artificial Neural Network (ANN) and Finite Element (FEM) Models for GFRP-Reinforced Concrete Columns under Axial Compression. *Materials* **2021**, *14*, 7172. [CrossRef]

38. Ebadi-Jamkhaneh, M.; Ahmadi, M. Comprehensive Investigations of the Effect of Bolt Tightness on Axial Behavior of a MERO Joint System: Experimental, FEM, and Soft Computing Approaches. *J. Struct. Eng.* **2021**, *147*, 04021197. [CrossRef]

39. Bardhan, A.; GuhaRay, A.; Gupta, S.; Pradhan, B.; Gokceoglu, C. A novel integrated approach of ELM and modified equilibrium optimizer for predicting soil compression index of subgrade layer of Dedicated Freight Corridor. *Transp. Geotech.* **2022**, *32*, 100678. [CrossRef]

40. Kardani, N.; Bardhan, A.; Samui, P.; Nazem, M.; Zhou, A.; Armaghani, D.J. A novel technique based on the improved firefly algorithm coupled with extreme learning machine (ELM-IFF) for predicting the thermal conductivity of soil. *Eng. Comput.* **2021**, 1–20. [CrossRef]

41. Bardhan, A.; Kardani, N.; GuhaRay, A.; Burman, A.; Samui, P.; Zhang, Y. Hybrid ensemble soft computing approach for predicting penetration rate of tunnel boring machine in a rock environment. *J. Rock Mech. Geotech. Eng.* **2021**, *13*, 1398–1412. [CrossRef]

42. Bardhan, A.; Kardani, N.; Alzoùbi, A.; Roy, B.; Samui, P.; Gandomi, A.H. Novel integration of extreme learning machine and improved Harris hawks optimization with particle swarm optimization-based mutation for predicting soil consolidation parameter. *J. Rock Mech. Geotech. Eng.* **2022**, *in press*. [CrossRef]

43. Asteris, P.G.; Skentou, A.D.; Bardhan, A.; Samui, P.; Pilakoutas, K. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cem. Concr. Res.* **2021**, *145*, 106449. [CrossRef]

44. Taghizadeh-Mehrjardi, R.; Emadi, M.; Cherati, A.; Heung, B.; Mosavi, A.; Scholten, T. Bio-Inspired Hybridization of Artificial Neural Networks: An Application for Mapping the Spatial Distribution of Soil Texture Fractions. *Remote Sens.* **2021**, *13*, 1025. [CrossRef]

45. Kardani, N.; Bardhan, A.; Samui, P.; Nazem, M.; Asteris, P.G.; Zhou, A. Predicting the thermal conductivity of soils using integrated approach of ANN and PSO with adaptive and time-varying acceleration coefficients. *Int. J. Therm. Sci.* **2021**, *173*, 107427. [CrossRef]

46. Tavoosi, J.; Zhang, C.; Mohammadzadeh, A.; Mobayen, S.; Mosavi, A.H. Medical Image Interpolation Using Recurrent Type-2 Fuzzy Neural Network. *Front. Neuroinform.* **2021**, *15*, 667375. [CrossRef]

47. Zha, T.H.; Castillo, O.; Jahanshahi, H.; Yusuf, A.; Alassafi, M.O.; Alsaadi, F.E.; Chu, Y.M. A fuzzy-based strategy to suppress the novel coronavirus (2019-NCOV) massive outbreak. *Appl. Comput. Math.* **2021**, *20*, 160–176.

48. Zhao, T.; Khan, M.I.; Chu, Y. Artificial neural networking (ANN) analysis for heat and entropy generation in flow of non-Newtonian fluid between two rotating disks. *Math. Methods Appl. Sci.* **2021**. [CrossRef]

49. Sadeghiravesh, M.H.; Khosravi, H.; Abolhasani, A.; Ghodsi, M.; Mosavi, A. Fuzzy Logic Model to Assess Desertification Intensity Based on Vulnerability Indices. *Acta Polytech. Hung.* **2021**, *18*, 7–24. [CrossRef]

50. Manshadi, M.D.; Ghassemi, M.; Mousavi, S.; Mosavi, A.; Kovacs, L. Predicting the Parameters of Vortex Bladeless Wind Turbine Using Deep Learning Method of Long Short-Term Memory. *Energies* **2021**, *14*, 4867. [CrossRef]

51. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.

52. Rosenman, M.A.; Gero, J.S.; Oxman, R. An expert system for design codes and design rules. In *Applications of Artificial Intelligence in Engineering Problems*; Springer: Berlin/Heidelberg, Germany, 1986; pp. 745–758.

53. Sacks, R. 'EIDOCC' Expert Interactive Design of Concrete Columns. Ph.D. Thesis, Massachusetts Institute of Technology, Department of Civil Engineering, Cambridge, MA, USA, 1985.

54. Sacks, R.; Büyüköztürk, O. Expert Interactive Design of R/C Columns Under Biaxial Bending. *J. Comput. Civ. Eng.* **1987**, *1*, 69–81. [CrossRef]

55. Butcher, J.; Day, C.; Austin, J.; Haycock, P.; Verstraeten, D.; Schrauwen, B. Defect Detection in Reinforced Concrete Using Random Neural Architectures. *Comput. Civ. Infrastruct. Eng.* **2013**, *29*, 191–207. [CrossRef]
56. Gui, G.; Pan, H.; Lin, Z.; Li, Y.; Yuan, Z. Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection. *KSCE J. Civ. Eng.* **2017**, *21*, 523–534. [CrossRef]
57. Vitola, J.; Pozo, F.; Tibaduiza, D.A.; Anaya, M. A sensor data fusion system based on k-nearest neighbor pattern classification for structural health monitoring applications. *Sensors* **2017**, *17*, 417. [CrossRef]
58. Lagaros, N.D.; Fragiadakis, M. Fragility Assessment of Steel Frames Using Neural Networks. *Earthq. Spectra* **2007**, *23*, 735–752. [CrossRef]
59. Jalayer, F.; Ebrahimian, H.; Miano, A.; Manfredi, G.; Sezen, H. Analytical fragility assessment using unscaled ground motion records. *Earthq. Eng. Struct. Dyn.* **2017**, *46*, 2639–2663. [CrossRef]
60. Mangalathu, S.; Jeon, J.-S.; Desroches, R. Critical uncertainty parameters influencing seismic performance of bridges using Lasso regression. *Earthq. Eng. Struct. Dyn.* **2018**, *47*, 784–801. [CrossRef]
61. Mahmoudi, S.N.; Chouinard, L. Seismic fragility assessment of highway bridges using support vector machines. *Bull. Earthq. Eng.* **2016**, *14*, 1571–1587. [CrossRef]
62. Lee, S.; Lee, C. Prediction of shear strength of FRP-reinforced concrete flexural members without stirrups using artificial neural networks. *Eng. Struct.* **2014**, *61*, 99–112. [CrossRef]
63. JJeon, J.-S.; Shafieezadeh, A.; DesRoches, R. Statistical models for shear strength of RC beam-column joints using machine-learning techniques. *Earthq. Eng. Struct. Dyn.* **2014**, *43*, 2075–2095. [CrossRef]
64. Vu, D.-T.; Hoang, N.-D. Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach. *Struct. Infrastruct. Eng.* **2016**, *12*, 1153–1161. [CrossRef]
65. SMangalathu, S.; Jeon, J.-S. Classification of failure mode and prediction of shear strength for reinforced concrete beam-column joints using machine learning techniques. *Eng. Struct.* **2018**, *160*, 85–94. [CrossRef]
66. Ahmadi, M.; Naderpour, H.; Kheyroddin, A. A Proposed Model for Axial Strength Estimation of Non-compact and Slender Square CFT Columns. *Iran. J. Sci. Technol. Trans. Civ. Eng.* **2018**, *43*, 131–147. [CrossRef]
67. 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]
68. Koza, J.R.; Koza, J.R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*; MIT Press: Cambridge, MA, USA, 1992; Volume 1.
69. Gao, W. A comprehensive review on identification of the geomaterial constitutive model using the computational intelligence method. *Adv. Eng. Inform.* **2018**, *38*, 420–440. [CrossRef]
70. Gandomi, A.H.; Babanajad, S.K.; Alavi, A.H.; Farnam, Y. Novel Approach to Strength Modeling of Concrete under Triaxial Compression. *J. Mater. Civ. Eng.* **2012**, *24*, 1132–1143. [CrossRef]
71. Gandomi, A.H.; Yun, G.J.; Alavi, A.H. An evolutionary approach for modeling of shear strength of RC deep beams. *Mater. Struct.* **2013**, *46*, 2109–2119. [CrossRef]
72. Genetic Programming. Wikipedia. 11 December 2021. Available online: https://en.wikipedia.org/w/index.php?title=Genetic_programming&oldid=1059826574 (accessed on 28 January 2022).
73. Genetic Algorithm. Wikipedia. Available online: https://en.wikipedia.org/w/index.php?title=Genetic_algorithm&oldid=1068402434 (accessed on 28 January 2022).
74. Ferreira, C. *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 21.
75. 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]
76. Jafari, S.; Mahini, S.S. Lightweight concrete design using gene expression programing. *Constr. Build. Mater.* **2017**, *139*, 93–100. [CrossRef]
77. Ferreira, C. Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. *arXiv* **2001**, arXiv:cs/0102027.
78. Mohammadzadeh, S.D.; Kazemi, S.F.; Mosavi, A.; Nasseralshariati, E.; Tah, J.H. Prediction of compression index of fine-grained soils using a gene expression programming model. *Infrastructures* **2019**, *4*, 26. [CrossRef]
79. Gene Expression Programming. Wikipedia. 30 October 2021. Available online: https://en.wikipedia.org/w/index.php?title=Gene_expression_programming&oldid=1052685265 (accessed on 28 January 2022).
80. Mundfrom, D.J.; Shaw, D.G.; Ke, T.L. Minimum Sample Size Recommendations for Conducting Factor Analyses. *Int. J. Test.* **2005**, *5*, 159–168. [CrossRef]
81. UW Civil & Environmental Engineering Home Page | Civil & Environmental Engineering. UW Civil & Environmental Engineering. Available online: https://www.ce.washington.edu/node/947 (accessed on 28 January 2022).
82. PEER Structural Performance Database. Available online: https://nisee.berkeley.edu/spd/ (accessed on 28 January 2022).
83. 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. In *Structures*; Elsevier: Amsterdam, The Netherlands, 2020; Volume 25, pp. 212–228.

84. Javed, M.F.; Amin, M.N.; Shah, M.I.; Khan, K.; Iftikhar, B.; Farooq, F.; Aslam, F.; Alyousef, R.; Alabduljabbar, H. Applications of Gene Expression Programming and Regression Techniques for Estimating Compressive Strength of Bagasse Ash based Concrete. *Crystals* **2020**, *10*, 737. [CrossRef]
85. Shah, M.I.; Javed, M.F.; Abunama, T. Proposed formulation of surface water quality and modelling using gene expression, machine learning, and regression techniques. *Environ. Sci. Pollut. Res.* **2021**, *28*, 13202–13220. [CrossRef]
86. 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]
87. ACI Committee 318. *Building Code Requirements for Structural Concrete (ACI 318-19) and Commentary (ACI 318R-19)*; American Concrete Institute (ACI) Committee 318: Farmington Hills, MI, USA, 2019. [CrossRef]
88. Iqbal, A.; Wang, Y.; Miah, M.; Osman, M.S. Study on Date–Jimbo–Kashiwara–Miwa Equation with Conformable Derivative Dependent on Time Parameter to Find the Exact Dynamic Wave Solutions. *Fractal Fract.* **2021**, *6*, 4. [CrossRef]
89. Chu, H.-H.; Zhao, T.-H.; Chu, Y.-M. Sharp bounds for the Toader mean of order 3 in terms of arithmetic, quadratic and contraharmonic means. *Math. Slovaca* **2020**, *70*, 1097–1112. [CrossRef]
90. Zhao, T.-H.; Wang, M.-K.; Chu, Y.-M. A sharp double inequality involving generalized complete elliptic integral of the first kind. *AIMS Math.* **2020**, *5*, 4512–4528. [CrossRef]
91. Dehghani, E.; Ranjbar, S.H.; Atashafrooz, M.; Negarestani, H.; Mosavi, A.; Kovacs, L. Introducing Copula as a Novel Statistical Method in Psychological Analysis. *Int. J. Environ. Res. Public. Health* **2021**, *18*, 7972. [CrossRef] [PubMed]
92. Ebrahimi-Khusfi, Z.; Taghizadeh-Mehrjardi, R.; Roustaei, F.; Ebrahimi-Khusfi, M.; Mosavi, A.H.; Heung, B.; Soleimani-Sardo, M.; Scholten, T. Determining the contribution of environmental factors in controlling dust pollution during cold and warm months of western Iran using different data mining algorithms and game theory. *Ecol. Indic.* **2021**, *132*, 108287. [CrossRef]
93. Ahmadi, M.; Kheyroddin, A.; Dalvand, A.; Kioumarsi, M. New empirical approach for determining nominal shear capacity of steel fiber reinforced concrete beams. *Constr. Build. Mater.* **2020**, *234*, 117293. [CrossRef]
94. Bui, D.T.; Nhu, V.-H.; Hoang, N.-D. Prediction of soil compression coefficient for urban housing project using novel integration machine learning approach of swarm intelligence and Multi-layer Perceptron Neural Network. *Adv. Eng. Informatics* **2018**, *38*, 593–604.
95. Koopialipoor, M.; Fallah, A.; Armaghani, D.J.; Azizi, A.; Mohamad, E.T. Three hybrid intelligent models in estimating flyrock distance resulting from blasting. *Eng. Comput.* **2019**, *35*, 243–256. [CrossRef]
96. Khan, M.U.A.; Shukla, S.K.; Raja, M.N.A. Load-settlement response of a footing over buried conduit in a sloping terrain: A numerical experiment-based artificial intelligent approach. *Soft Comput.* **2022**, 1–18. [CrossRef]