Evaluation of ultimate conditions of FRP-confined concrete columns using genetic programming

Jian C. Lim, Murat Karakus *, Togay Ozbakkaloglu

School of Civil, Environmental and Mining Engineering, University of Adelaide, South Australia 5005, Australia

A R T I C L E   I N F O

Article history:
Received 13 February 2015
Accepted 22 September 2015

Keywords:
Genetic programming (GP)
Fiber reinforced polymer (FRP)
Confinement
Concrete
Compressive strength
Ultimate axial strain

A B S T R A C T

A large database consisting of 832 axial compression tests results of fiber reinforced polymer (FRP)-confined concrete specimens was assembled. Using the test database, existing conventional and evolutionary algorithm models developed for FRP-confined concrete were then assessed. New genetic programming (GP) models for predicting the ultimate condition of FRP-confined concrete were developed. The predictions of the proposed models suggest that more accurate results can be achieved in explaining and formulating the ultimate condition of FRP-confined concretes by GP. The model assessment also illustrates the influences of the size of the databases and the selected parameters used in the GP models.

1. Introduction

It is now well understood that the confinement of concrete with fiber reinforced polymer (FRP) composites can substantially enhance concrete strength and deformability. A large number of studies undertaken to date have produced over 3000 test results on FRP-confined concrete and resulted in the development of over 90 models. The conventional models that were developed using regression analysis can be classified into two broad categories namely the design-oriented models and the analysis-oriented models. In predictions of the ultimate condition of FRP-confined concrete, the design-oriented models use closed-form empirical expressions that were derived directly from database results. The analysis-oriented models, on the other hand, use a combination of empirical and theoretical expressions through an incremental procedure to consider the interaction mechanism between the external FRP jacket and the internal concrete core. The analysis-oriented models are built on the assumption of stress path independence, which assumes that the axial stress and axial strain of FRP-confined concrete at a given lateral strain are the same as those of the same concrete actively confined with a constant confining pressure equal to that supplied by the FRP shell.

As indicated by the assessment results of these models using a comprehensive experimental database, the performances of a large proportion of the conventional models were compromised when they were assessed against a large database covering parametric ranges that are much wider than the original databases used to develop these models [1,2]. This can be attributed to the limited capability of the design-oriented models in handling uncertainties in complex experimental database, whereas the assumption adopted by the analysis-oriented models has recently been shown to be inaccurate [3,4]. In addition, the development of these existing conventional models is often based on the expressions proposed by Richart et al. [5], with refinement subsequently applied to those earlier expressions to incorporate new research findings. Given the dependencies of the conventional models on the base expressions and the gradual refinement process, an efficient alternative approach is therefore required. Recently, new models based on artificial neural network, genetic programming, stepwise regression, and fuzzy logic algorithms have been proposed by many researchers [6–18]. Models in this category could handle complex databases containing large number of independent variables, identify the sensitivity of input parameters, and provide mathematical solutions between dependent and independent variables. However, the complexity of modeling frameworks and the dependency of these models on computers have significantly reduced their versatility in design applications. The computer generated statistical solutions have also compromised the physical significance unfolding the structural behavior of FRP-confined concrete. In addition, several common modeling issues identified from the assessment of these models include: (i) limited size of database results, (ii) overfitted with redundant test parameters that cause unreliable prediction beyond their original observation range,

* Corresponding author. Tel.: +61 8 8313 6471; fax: +61 8 8313 4359.
E-mail address: murat.karakus@adelaide.edu.au (M. Karakus).

http://dx.doi.org/10.1016/j.compstruc.2015.09.005
0045-7949/© 2015 Elsevier Ltd. All rights reserved.
and (iii) lack of consideration of important test parameters, including the ultimate rupture strain of FRP jacket.

With proper treatments given to the aforementioned modeling issues, genetic programming (GP) can be a potential candidate to address these shortcomings. GP is an evolutionary algorithm attempting to find key variables for a problem in a given search space, and generates mathematical expressions to explain the relations between the variables. By using GP based on the principles of symbolic regression (SR) analysis, the relationships between the dependent and independent variables of a complex database involving uncertainties and variability can be solved. SR is a process of finding a mathematical expression by minimizing the errors between given finite data set as well as providing a method of function identification [19]. Symbolic regression, as opposed to other regression techniques, discovers both the form of the model and its parameters from the search space. In other words, a measured dataset is fitted to an appropriate mathematical formula by a fitness function. Determination or identification of key variables and variable combinations, providing comprehension of developed models are among the benefits of symbolic regression analysis. In this research, SR analysis was conducted using GP approach which fits well to wide range of engineering problems.

In the recent years, the use of GP for optimum solution and function identification of engineering problems have been gaining acceleration as the approach is capable of dealing with complex database that contains a large number of parameters. GP progressively improves solution while it maintains the versatility of the model in closed-form expressions (e.g. [20–25]). For example, Johari et al. [26] has successfully applied GP for the prediction of soil–water characteristic curve. Baykasoglu et al. [27] applied multi expression programming (MEP), gene expression programming (GEP) and linear genetic programming (LGP) to estimate compressive and tensile strength of limestone for the first time with good predictions. Javadi et al. [28] introduced a new technique based on GP for the determination of liquefaction induced lateral spreading. Cabalar and Cevik [29] applied GP for the prediction of peak ground acceleration using strong-ground-motion data from Turkey. The use of GP for the prediction of axial compressive strength of FRP-confined concrete has also been demonstrated by Cevik et al. [11], Cevik and Cabalar [6], and Cevik et al. [8]. Results from these studies [6–8] proved that highly accurate prediction models can be developed by using genetic programming.

In the present study, the GP approach was used to establish models to predict the ultimate conditions of FRP-confined concrete columns under concentric compression. Based on a comprehensive experimental database that was carefully assembled using a set of selection criteria to ensure the reliability and consistency of the database, three closed-form expressions are proposed for the predictions of the compressive strength, ultimate axial strain and hoop rupture strain of FRP-confined concrete. This is the first study in the literature to establish expressions for the ultimate axial strain and hoop rupture strain of FRP-confined concrete on the basis of evolutionary algorithms. Details of the adopted approach are discussed in Section 2. A summary of the experimental database is provided in Section 3. The selection process of independent variables, functions and fitness rule, together with the proposed expressions are presented in Section 4. The predictions statistics of the proposed and the existing models with experimental results are presented in Section 5.

2. Overview of genetic programming

In this section, the GP paradigm is discussed and the essentials of GP are highlighted. Further concepts and terminology behind GP can be found from the inventor of this paradigm [19]. However, it is advised that the genetic algorithm concept developed in 1975 by Holland [30] and work from his student, Goldberg [31], can also be visited for further insight.

2.1. Basic concepts of GP

Genetic programming is an extension of the conventional genetic algorithm (GA), generating novel solutions to complex problems, developed by Koza [19]. Unlike GA which uses a string of numbers to represent the solution, the GP automatically creates several computer programs (CP) with a parse tree structure to solve the problem considered. The process of solving the problems with GP is equivalent to searching a space of possible computer programs for the fittest individual computer program [19]. The generated CP is based on Darwinian concept of survival and reproduction of the fittest as well as appropriate mating of CPs. As illustrated in the flow chart in Fig. 1, the problem will be solved using the Darwinian genetic operators such as reproduction, crossover, and mutation. Initial population consists of randomly generated CPs, which is composed of functions and terminals appropriate to the characteristics of the problem. If the functions and terminals selected are not appropriate for the problem, the desired solution cannot be achieved. A basic flow chart of the genetic programming paradigm is shown in Fig. 1.

As stated earlier, CP is composed of functions and terminals. The functions can be standard arithmetic operators, trigonometric log-arithmetic or power, e.g. \( f = \{ /, \times, -, +, \sin, \cos, \log, 2, \text{power}, \ldots \} \) and/or any mathematical functions, logical functions, as well as user-defined operators. Depending on the nature of the problem investigated, the computer program might be Boolean-valued, integer-valued, real-valued, complex-valued, vector-valued, symbolic-valued, or multiple-valued [19]. A typical expression tree, representing the computer program is shown in Fig. 2. In this example, the function set \( F \) is composed of multiplication, division,......
addition, subtraction and the sine function, $F = \{+, -, \sin\}$. The terminal set ($T$) is composed of $N = 3$ variable as $T = \{x, y, z\}$. The functions and terminals must fulfill two important properties in order to solve the problem with an appropriate representation [19]. These parameters are closure property and sufficiency property. The closure property includes protection of the function set and the terminal set against all possible argument values, e.g. protection of negative square root. Sufficiency property is the selection of the appropriate functions and terminals to solve the problem at hand.

2.2. Genetic operations

Genetic operations used in GP are composed of: reproduction, crossover, and mutation. Reproduction operation involves selecting, in proportion to fitness, a computer program from the current population of programs, and allowing it to survive by copying it into the new population. Several different types of reproduction operations such as fitness proportionate reproduction or roulette wheel algorithm, tournament selection and lexicographic parsimony pressure selection are commonly used in GP. In this study Lexicographic parsimony pressure selection was used, which is a multi-objective method similar to tournament selection. This particular method optimizes both fitness and parse tree size. The shortest individual, the tree with fewer nodes, is selected as the fittest when two individuals are equally fit. Silva and Almeida [32] reported that this technique is effective in controlling the bloat which is a phenomenon consisting of an excessive code growth without the corresponding improvement in fitness. The theory of Parsimony Pressure are discussed in detail by Poli and McPhee [33]. The standard method of controlling bloat is to set up a maximum depth on trees in the proposed GP model.

Crossover operation involves choosing random nodes in two parent trees and swapping respective branches creating two new offspring. Fig. 3 illustrates the crossover operation. Mathematical expressions for parent I and parent II are as follows before crossover:

Parent I : $\frac{xy}{\sin(z)} + (x - z)$  Parent II : $\frac{x}{2} + e^y - z^2$

Resulting offspring after crossover operation are:

Offspring I : $\frac{x}{2} + (x - z)$  Offspring II : $\frac{xy}{\sin(z)} + e^y - z^2$

As can be seen from the above expressions, crossover between S-expressions consists of swapping the randomly selected sub S-expressions.

In mutation operation, “a random node is selected from the parent tree and substituted by a new random tree created with the terminals and the functions available” as described by Silva [32]. This is known as tree mutation. However, Koza [19] stated that mutation plays a minor role in GP. Therefore, it can be disregarded in most cases.

3. Database of FRP-confined concrete

The database of FRP-confined concrete was assembled through an extensive review of the literature that covered 3042 test results from 253 experimental studies published between 1991 and 2013. The suitability of the results was then assessed using a set of carefully established selection criteria to ensure the reliability and consistency of the database. Only monotonically loaded circular specimens with unidirectional fibers orientated in the hoop direction and an aspect ratio ($H/D$) of less than three were included in the database. Specimens containing internal steel reinforcement or partial FRP confinement were not included. This resulted in a final database size of 832 datasets collected from 99 experimental studies. The complete database of experimental results used in the present study can be found in Ozbakkaloglu and Lim [34]. The database consists of specimens confined by five main types of FRP materials (carbon FRP (CFRP), high-modulus carbon FRP (HM CFRP), ultra-high-modulus carbon FRP (UHM CFRP), S- or E-glass FRP (GFRP), and aramid FRP (AFRP)) and two confinement techniques (wraps and tubes). The carbon FRPs were categorized into three subgroups on the basis of their elastic modulus of fibers ($E_f$) i.e., carbon FRP with $E_f \leq 270$ GPa is categorized as CFRP; followed by $270 < E_f < 440$ GPa as HM CFRP; and $E_f > 440$ GPa as UHM CFRP). 755 specimens in the database were FRP-wrapped, whereas 77 specimens were confined by FRP tubes. 495 of the specimens were confined by CFRP; 206 by GFRP; 79 by AFRP; 40 by HM CFRP; and 12 by UHM CFRP. The unconfined concrete strength ($f_{cu}$) and strain ($\varepsilon_{cu}$), as obtained from concrete cylinder tests, varied from 6.2 to 55.2 MPa and 0.14% to 0.70%, respectively. The diameters of the specimens ($D$) included in the test database varied between 47 and 600 mm, with the majority of the specimens having a diameter of 150 mm. The hoop rupture strain ($\varepsilon_{h,rup}$) of the FRP-confined concrete specimens varied from 0.09% to 3.21%. The actual confinement ratio, defined as the ratio of the ultimate confining pressure of the FRP jacket at rupture to the compressive strength of an unconfined concrete specimen ($f_{cu}$), varied from 0.02 to 4.74. Figs. 4 and 5, respectively, show the relationships of the strength enhancement ratio ($f_{cudf} / f_{cu}$) and strain enhancement ratio ($\varepsilon_{edf} / \varepsilon_{cu}$) versus the actual confinement ratio ($f_{hu} / f_{cu}$), established from the database.

3.1. Parameter selection

The GP analysis is based on three main stages. Firstly, the terminal sets are selected. The term terminal refers to independent variables used to approximate dependent variables. Parameters that were identified to be non-influential were either eliminated or combined with several other parameters until a strong trend of influence was observed. One of such input parameters that combines several other input parameters is the confinement stiffness of FRP jacket ($K_0$), defined by Eq. (2), which was presented later in Section 4.1. In the experimental database [34], it has already been observed that strong exponential or power relations exist between the independent variable $K_0$ and the dependent variables $f_{cc}$ and $\varepsilon_{w}$. With the use of $K_0$ as an independent variable, GP also generated similar exponential and/or power formulations which underscore the experimental observations. Other influential independent parameters selected for the GP analyses include the unconfined strength of concrete ($f_{cu}$), elastic modulus of fiber ($E_f$), hoop rupture strain of FRP jacket ($\varepsilon_{h,rup}$), and ultimate tensile strength of fiber ($\sigma_f$). A summary of the selected input parameters.

![Fig. 2. A typical expression tree.](image-url)
related to FRP-Confined concrete and GP parameters are given in Table 1.

In the second stage, a set of functions which are related to the nature of the problem or data set was determined. As can be seen from the Table 1, additional to the basic arithmetic operations (+, −, , ), exponential (exp), square root (sqrt) and power functions are also included. We have also included logarithmic functions which did not improve the fitness. Therefore it was removed from the function set.

In the third stage, a symbolic expression (S-expression), a list or an atom in LISP, was generated by GP in terms of the fitness func-

Fig. 3. Crossover operation in genetic programming.

Fig. 4. Variation of strength enhancement ratio (f'cc/fo) of FRP-confined concrete with confinement ratio (flu,a/fo).

Fig. 5. Variation of strain enhancement ratio (εcu/εco) of FRP-confined concrete with confinement ratio (flu,a/fo).
term adopted. The S-expression is the only syntactic form of the LISP programming language. For example, \((+(-3 4)5)\) is a LISP S-expression. In this S-expression the atoms or individuals \((3\text{ and }4)\) are subtracted first and the result \((-1)\) is added to 5 yielding in value 4. S-expressions were chosen according to the lower fitness value. The lower the fitness value, the better the model is. According to the selected fitness \((\text{AAE})\), the lowest value indicates a small error between the measured and predicted data. The genetic programming will run until the termination criterion \((\text{stopping condition})\) is satisfied. This can be done either by determining a maximum generation limit or a tolerated error limit. The program can also be terminated by the variation in fitness observed during the run. We used maximum generation limit as the termination criterion for the all GP runs.

Some of the mathematical functions included in the GP are protected against zero division or negative square root. In the division operation, if the denominator is equal to zero then the results returns to the numerator. In the power operation \(x^0\) returns zero if \(x^0\) is not a number \((\text{NaN})\) or infinity \((\infty)\), or has an imaginary part, otherwise it returns \(x^0\). In the function square root, \(\sqrt{x}\) returns zero if \(x \leq 0\) and \(\sqrt{x}\) otherwise. Table 1 summarizes the parameters adopted in GP analysis.

### 4. Proposed model for ultimate condition of FRP-confined concrete

The ultimate condition of FRP-confined concrete is often characterized as the compressive strength and the corresponding axial strain of concrete and hoop strain recorded at the rupture of the FRP jacket. This makes the relationship between the ultimate axial stress \(f_{cc}\), ultimate axial strain \(e_{cu}\) and hoop rupture strain \(e_{rup}\) an important one. Using the comprehensive experimental database \([34]\), a model consisting of three expressions for the predictions of the compressive strength \(f_{cc}\), ultimate axial strain \(e_{cu}\) and hoop rupture strain \(e_{rup}\) was developed using GP and is presented in this section. The model is applicable to FRP-confined concrete with unconfined concrete strength up to 55 MPa. Table 1 summarizes the parameters that were used in GP analysis. As GP analysis does not directly reveal the underlying physical relationships of a given dataset, the search of a physically meaningful model structure relies on users’ engineering knowledge of FRP-confined concrete systems. This includes the understanding of experimentally influential parameters for implementation in GP formulation, and the careful selection of population size, functions and model structure that will result in practical closed-form expressions. These are no easy tasks since both the structure and parameters of the physical model must be determined \([25]\). With the help of GP approach, a large number of potential model components and structures can be tested, while the best parts of these structures and combinations can be retained to produce new and possibly better expressions. In addition, the expressions resulting from GP formulation process are often useful in revealing pertinent aspects that are physically meaningful. Hence, with carefully selected functions for these pertinent aspects, an accurate and physically meaningful model can be established. In our trial runs of this approach, larger population sizes up to 1500 were tested. However, the use of the larger population sizes tends to create bloated/large number of nodes leading to extremely complex formulation. Therefore, we gradually reduced the population size to 300 in the actual analysis. Mutation and crossover probabilities are not fixed but random. In tree crossover, random nodes are chosen from both parent trees, and the respective branches are swapped creating two offspring. It should be noted that not all the datasets included in the database contained all the relevant details required for the model development. As a result, out of 832 results, 753, 511, and 325 were used in the development of the expressions of the compressive strength \(f_{cc}\), ultimate axial strain \(e_{cu}\) and hoop rupture strain \(e_{rup}\) respectively. The experimental values of the compressive strength \(f_{cc}\) and the corresponding axial strain \(e_{cu}\) of unconfined concrete were based on the test results of the unconfined cylinders. In order to avoid over-fitting random data division method was used, in which 70% of the total data were randomly selected for training and the remaining 30% were used for testing purposes. This approach was previously used by a number of existing studies \([23,36,37]\).

#### 4.1. Compressive strength of FRP-confined concrete

As was discussed earlier, the majority of the conventional models are based on the expression forms proposed by Richart et al. \([5]\). GP, on the other hand, initiates the formulation of expressions from a few random seeds, of which the evolution took place artificially and terminated when a robust solution was found. In the GP formulation process, the selection of independent variables, functions and model structure for a given terminal set relies significantly on the users’ knowledge of the FRP-confined concrete systems. Overfitting of model with redundant independent variables and functions results in a complex mathematical solution rather than a physically meaningful model structure, whereas underfitting reduces the model accuracy. In this process, finding the potential combination of several input parameter to form a single representative input parameter, such as the confinement stiffness of FRP jacket \((K_i)\) used in this study, is important. As presented in Table 1, only parameters that had been observed experimentally to influence the terminal sets \([2]\) were selected as independent variables for GP formulation. In addition, only simple model structure yielding practical closed-formed expressions suitable for engineering application were adopted by the authors. The converged expression for the prediction of the compressive strength of FRP-confined concrete \(f_{cc}\) is presented in Eq. (1).

\[
f_{cc} = f_{cc0} + K_i e_{rup} + K_1^* e_{rup}^2 + a \quad \text{where} \quad a = \sqrt{K_i - \frac{f_{cc0}}{e_{rup}}} \geq 0
\]  

(1)

where \(f_{cc0}\) is the unconfined concrete strength, \(K_i\) is the confinement stiffness of the FRP jacket, to be calculated using Eq. (2), and \(e_{rup}\) is the hoop rupture strain of the FRP jacket to be calculated using Eq. (8) presented later in Section 4.3. \(K_i\) and \(f_{cc0}\) are in MPa.

\[
K_i = \frac{2E_f t_f}{D}
\]  

(2)

\(E_f\) is the elastic modulus of fibers, \(t_f\) is the total nominal fiber thickness of the FRP jacket, \(D\) is the diameter of the concrete specimen.

To demonstrate an example of the GP formulation process, Fig. 6 shows the relationship between accuracy and complexity
of the GP formulation of the compressive strength expression. As illustrated by the figure, the fitness or accuracy of the expression improves significantly after the first few generations. After which, the complexity of the expression represented by the level and nodes in Fig. 6 continue to increase with marginal improvements in its fitness in the proceeding generations. At the 28th generation, where a low level and a small number of nodes were found at lower fitness, the final expression is selected and presented as Eq. (1). Level is the depth of parse tree which controls the number of nodes. It is used to avoid bloating, an excessive code growth without corresponding improvement in fitness.

Fig. 7 shows the comparison of the strength enhancement ratio predictions ($f_{cc}'/f_{cc}$) of the proposed expression (Eq. (1)) with the 30% testing datasets of the experimental database. The comparison indicates that the model predictions are in close agreement with the test results, which are quantified through the use of statistical indicators: average absolute error (AAE) to establish overall model accuracy; mean ($M$) to establish average overestimation or underestimation of the model; and standard deviation ($SD$) to establish the magnitude of the associated scatter of the model prediction. These indicators are defined by Eqs. (3)–(5). As evident from the figure, 10.5% of AAE, 100.9% of $M$, and 13.5% of $SD$ were achieved.

$$\text{AAE} = \frac{\sum_{i=1}^{n} |\text{model}_i - \text{exp}_i|}{n}$$

$$M = \frac{\sum_{i=1}^{n} \text{model}_i}{n}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (\text{model}_i - M)^2}{n-1}}$$

4.2. Ultimate axial strain of FRP-confined concrete

Based on the GP formulation process discussed earlier, the expression for the prediction of the ultimate axial strain ($\epsilon_{cu}$) of FRP-confined concrete is established and presented in Eq. (6).

$$\epsilon_{cu} = (\epsilon_{co} + b)(\epsilon_{co} + \frac{K_i}{f_{cc}}(2\epsilon_{co} + b))$$

where $b = \epsilon_{h_rup} - \frac{\epsilon_{co}}{\epsilon_{h_rup}}$ and $c = \frac{f_{cc}(\epsilon_{co} + \epsilon_{h_rup} + \epsilon_{e^aexp})}{k_i}$

where $\epsilon_{co}$ is the axial strain corresponding to the compressive strength of unconfined concrete, to be calculated using Eq. (7) proposed by Lim and Ozbakkaloglu [38], $\epsilon_{h_rup}$ is the hoop rupture strain of the FRP jacket, to be calculated using Eq. (8) presented in the next section, and $K_i$ is the confinement stiffness of the FRP jacket, to be calculated using Eq. (2).

$$\epsilon_{co} = \frac{f_{in}^{0.225}}{1000k_i}$$

where $k_i = \left(\frac{152}{D}\right)^{0.1}$ and $k_s = \left(\frac{2D}{H}\right)^{0.13}$

where $f_{cc}$ is in MPa, and $k_i$ and $k_s$, respectively, are the coefficients to allow for the specimens size and specimen aspect ratio.

Fig. 8 shows comparisons of the strain enhancement ratio predictions ($\epsilon_{cal}/\epsilon_{cu}$) of the proposed model with testing datasets. The comparison indicates that the model predictions are in close agreement with the test results, of which AAE, $M$, and $SD$ of 25.7%, 101.1%, and 33.1%, respectively, were achieved.

4.3. Hoop rupture strain of FRP jacket

The hoop rupture strains ($\epsilon_{h_rup}$) of FRP jackets are commonly reported to be lower than the ultimate tensile strain of the fiber material ($\epsilon_f$) [39–43]. As previously reported in Ozbakkaloglu and Akin [44], an increase in the compressive strength of concrete, which alters the concrete cracking pattern from heterogenic microcracks to localized macroracks, has an adverse influence on the hoop rupture strain ($\epsilon_{h_rup}$) of the FRP jacket. An increase in the elastic modulus of fibers ($E_f$) of the FRP jacket was also reported.

---

**Fig. 6.** Accuracy versus complexity for the compressive strength expression. Please note that each point on the fitness and nodes curves correspond to the lowest fitness values.

**Fig. 7.** Comparison of model predictions of strength enhancement ratios ($f_{cc}'/f_{cc}$) with experimental results.

**Fig. 8.** Comparison of model predictions of strain enhancement ratios ($\epsilon_{cal}/\epsilon_{cu}$) with experimental results.
to decrease the hoop rupture strain ($\varepsilon_{h,rup}$) [45]. To account for such dependencies of the hoop rupture strain on the FRP and concrete materials, the elastic modulus of fibers ($E_f$), ultimate tensile strain of fibers ($\varepsilon_f$), and unconfined concrete strength ($f_{co}$) were considered as separate input variables for the development of the hoop rupture strain expression ($\varepsilon_{h,rup}$). The expression established using the GP formulation process for the prediction of the hoop rupture strain of FRP ($\varepsilon_{h,rup}$) is presented in Eq. (8).

$$\varepsilon_{h,rup} = \frac{E_f}{f_{co}} \left( \frac{E_f}{f_{co}} + d \right)^{0.5}$$

To demonstrate an example of overfitted expression, the GP formulation process was allowed to continue until a next expression (Eq. (9)) with a slightly higher accuracy. Fig. 9(a) and (b) show the comparisons of the results from testing dataset with the hoop rupture strains ($\varepsilon_{h,rup}$) predicted using Eqs. (8) and (9), respectively. The comparison shows that the model predictions are in close agreement with the test results, of which AAE, $M$, and SD of 22.7%, 109.6%, and 34.2% were achieved in the first expression, while 22.5%, 112.2%, and 33.1% were achieved in the second expression. With only a small margin of improvement achieved, the form of the expression of Eq. (9) became relatively complex in comparison to the earlier form shown in Eq. (8). On this basis, Eq. (8) is recommended for its simplicity. In addition, as Eq. (9) was overfitted for the current size of test database, its performance is likely to degrade at out-of-range prediction, due to its increased sensitivity to the parametric ranges of the current database.

$$\varepsilon_{h,rup} = \left( \frac{E_f}{f_{co}} \right)^{0.5}$$

where $E_f$ and $f_{co}$ are in MPa.

5. Model validation and comparisons with existing models

To establish the relative performance of the proposed model, its prediction statistics were compared with those of the 10 best performing conventional models identified in a recent comprehensive review study reported in Ozbakkaloglu et al. [34]. In addition, the model was also compared with seven artificial intelligence (AI) models currently available in the literature that were developed using evolutionary programming techniques, including neural network (NN) [7,9,11,13,14], genetic programming (GP) [6,8,11], and

<table>
<thead>
<tr>
<th>Model</th>
<th>Model category</th>
<th>Prediction of $f_{co}/f_{uo}$</th>
<th>Test data</th>
<th>Average absolute error (%)</th>
<th>Mean (%)</th>
<th>Standard deviation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed model</td>
<td>GP</td>
<td>753</td>
<td>10.8</td>
<td>102.1</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td>Proposed model [30% testing dataset]</td>
<td>GP</td>
<td>226</td>
<td>10.5</td>
<td>100.9</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Ozbakkaloglu and Lim [34]</td>
<td>DO</td>
<td>753</td>
<td>11.2</td>
<td>99.6</td>
<td>13.7</td>
<td></td>
</tr>
<tr>
<td>Teng et al. [47]</td>
<td>AO</td>
<td>753</td>
<td>11.8</td>
<td>98.8</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Lam and Teng [48]</td>
<td>DO</td>
<td>753</td>
<td>12.4</td>
<td>99.4</td>
<td>15.3</td>
<td></td>
</tr>
<tr>
<td>Wu and Zhao [49]</td>
<td>DO</td>
<td>753</td>
<td>12.4</td>
<td>102.1</td>
<td>15.2</td>
<td></td>
</tr>
<tr>
<td>Wu and Wang [50]</td>
<td>DO</td>
<td>753</td>
<td>12.7</td>
<td>101.4</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Al-Salloum and Siddiqui [51]</td>
<td>DO</td>
<td>753</td>
<td>12.7</td>
<td>101.7</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td>Wei and Wu [52]</td>
<td>DO</td>
<td>753</td>
<td>12.7</td>
<td>101.5</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td>Realfonzo and Napoli [53]</td>
<td>DO</td>
<td>753</td>
<td>12.7</td>
<td>102.2</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td>Bissy et al. [54]</td>
<td>DO</td>
<td>753</td>
<td>12.8</td>
<td>101.9</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td>Jiang and Teng [1]</td>
<td>AO</td>
<td>753</td>
<td>12.9</td>
<td>93.9</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Cevik 1 [11]</td>
<td>NN</td>
<td>753</td>
<td>15.6</td>
<td>96.2</td>
<td>18.3</td>
<td></td>
</tr>
<tr>
<td>Elsanadedy et al. [13]</td>
<td>NN</td>
<td>753</td>
<td>18.4</td>
<td>100.4</td>
<td>31.0</td>
<td></td>
</tr>
<tr>
<td>Cevik and Cabalar [6]</td>
<td>GP</td>
<td>753</td>
<td>19.6</td>
<td>115.1</td>
<td>22.6</td>
<td></td>
</tr>
<tr>
<td>Cevik 2 [11]</td>
<td>SW</td>
<td>753</td>
<td>20.1</td>
<td>95.7</td>
<td>29.9</td>
<td></td>
</tr>
<tr>
<td>Cevik et al. 1 [8]</td>
<td>GP</td>
<td>753</td>
<td>23.9</td>
<td>108.0</td>
<td>33.7</td>
<td></td>
</tr>
<tr>
<td>Cevik et al. 2 [8]</td>
<td>SW</td>
<td>753</td>
<td>33.3</td>
<td>108.8</td>
<td>42.2</td>
<td></td>
</tr>
<tr>
<td>Naderpour et al. [9]</td>
<td>NN</td>
<td>753</td>
<td>40.2</td>
<td>127.6</td>
<td>50.3</td>
<td></td>
</tr>
<tr>
<td>Cevik and Guzelbey [7]</td>
<td>NN</td>
<td>753</td>
<td>75.6</td>
<td>154.3</td>
<td>110.7</td>
<td></td>
</tr>
<tr>
<td>Jalal and Ramezanianpour [14]</td>
<td>NN</td>
<td>753</td>
<td>79.2</td>
<td>178.3</td>
<td>79.1</td>
<td></td>
</tr>
</tbody>
</table>

stepwise regression (SW) [8,11]. It should be noted that the models proposed in various studies that only report the architecture of their modeling framework without complete expressions for the prediction of the ultimate condition of FRP-confined concrete were not assessed (e.g., [10,12,15,17,18]). The prediction statistics for the strength and strain enhancement ratios \( f_{cc}/f_{co} \) and \( \varepsilon_{cu}/\varepsilon_{co} \) are given in Tables 2 and 3, respectively. In addition to the results shown in Tables 2 and 3, the graphical comparison of the AAEs for the conventional and the AI models are given in Figs. 10 and 11, respectively. In calculating model predictions of the peak strain ratios \( \varepsilon_{cu}/\varepsilon_{co} \), the \( \varepsilon_{co} \) values were determined using the expressions given in the original publication when available. If no expression was not specified in the original publication, the \( \varepsilon_{co} \) values were then based on the experimental values obtained from cylinder tests. If an experimental \( \varepsilon_{co} \) value was not available from a given dataset, Eq. (7) proposed by Tasdemir et al. [46] is used to determine the \( \varepsilon_{co} \) value. For the proposed model, two sets of prediction statistics are presented in Tables 2 and 3. The first set was based on the full datasets of 753 and 511 results in the experimental database for strength and strain enhancement ratios \( f_{cc}/f_{co} \) and \( \varepsilon_{cu}/\varepsilon_{co} \), while the second set was based the 30% testing datasets.

As evident from Tables 2 and 3 and Figs. 10 and 11, the proposed model provides improved predictions of the strength enhancement ratios \( f_{cc}/f_{co} \) compared to all of the existing models. On the other hand, the performance of the model developed by Ozbakkaloglu and Lim [34] for the predictions of the strain enhancement ratios \( \varepsilon_{cu}/\varepsilon_{co} \) is slightly better than that of the current proposed model, given that the influence of axial strain instrumentation methods was considered in the model [34]. For such consideration, a numerical input such as the gauge length of the instrument which is necessary for GP, however, was not available from the experimental database [34]. As a result, the minimum AAE of the predicted strain enhancement ratios \( \varepsilon_{cu}/\varepsilon_{co} \) by the proposed model is closer to the natural scatter of the database of 23%. It is also worthwhile noting that, none of the reviewed AI models proposed an expression for the prediction of the ultimate axial strain \( \varepsilon_{cu} \) of FRP-confined concrete. This paper presents the first expression established for the prediction of the ultimate axial strain \( \varepsilon_{cu} \) on the basis of GP (Eq. (6)). As evident from Fig. 8 and Table 3, the proposed expression provides reasonable predictions of the test results, which can be further improved in the future through accurate modeling of the specimen instrumentation methods.

**Table 3**

Statistics of strain enhancement ratio \( \varepsilon_{cu}/\varepsilon_{co} \) predictions of best performing models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model category</th>
<th>Prediction of ( \varepsilon_{cu}/\varepsilon_{co} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test data</td>
<td>Average absolute error (%)</td>
</tr>
<tr>
<td>Ozbakkaloglu and Lim [34]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Proposed model</td>
<td>GP</td>
<td>511</td>
</tr>
<tr>
<td>Proposed model (30% testing dataset)</td>
<td>GP</td>
<td>153</td>
</tr>
<tr>
<td>Tamuz et al. [55]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Wei and Wu [52]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Binici [56]</td>
<td>AO</td>
<td>151</td>
</tr>
<tr>
<td>Jiang and Teng [57]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Youssef et al. [58]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Teng et al. [59]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Fahmy and Wu [60]</td>
<td>DO</td>
<td>511</td>
</tr>
<tr>
<td>Teng et al. [47]</td>
<td>AO</td>
<td>151</td>
</tr>
<tr>
<td>De Lorenzis and Tepfers [61]</td>
<td>AO</td>
<td>151</td>
</tr>
</tbody>
</table>

A close examination of the assessment results has led to a number of important findings on factors influencing the performances of existing models including the size of database, parameters considered, ability to handle uncertainties, dependency on assumptions, and the architecture of their modeling frameworks. For the conventional models, only the statistics of the model performances are presented in Tables 2 and 3. However, a detailed review of factors influencing the performances of these models was previously discussed by Ozbakkaloglu et al. [2]. In this section, we only discussed the factors influencing the performances of AI models. As evident from Fig. 11, the proposed model outperformed existing AI models by a significant margin. This improvement was achieved through consideration of a wide range of parameters that was covered by the comprehensive experimental database and careful selection of influential input parameters in model development. In the better performing models illustrated in Fig. 11, including the NN models proposed by Cevik [11] and Elsanadedy et al. [13], and the GP model proposed by Cevik and Cabalar [6], it was found that the sizes of the databases used in their development were generally larger than those of the underperforming counterparts, but the architecture of their modeling frameworks was not necessarily more complex. On the other hand, excessive complexity of modeling framework due to overfitting with redundant variables, as evident from some of the underperforming models (e.g., [9,14]), significantly undermined the modeling accuracy. On these bases, it is recommended that comprehensive experimental databases should be used and selection of key parameters should be carefully implemented in the future development of AI models for FRP-confined concrete.

6. Conclusions

A comprehensive experimental test database that consisted of 832 test results of FRP-confined concrete has been assembled from the published literature. Using the test database, the performances of a number of existing empirical, theoretical, and artificial intelligence models developed for FRP-confined concrete were then assessed. A close examination of the results from model assessments has led to a number of important findings on factors influencing the strengths and weaknesses of models in each category. These findings have been summarized and discussed in detail in the paper. On the basis of the experimental database, a new model for evaluating the ultimate condition of FRP-confined concrete was developed using genetic programming and has been presented in this paper. The model is the first to establish the ultimate axial strain and hoop rupture strain expressions for FRP-confined concrete on the basis of evolutionary algorithms. Comparisons with experimental test results show that the predictions from the proposed model are in good agreement with the test results of the database. The proposed models provide improved predictions compared to the existing artificial intelligence models. Genetic programming proved that more accurate results can be achieved in explaining and formulating the ultimate condition of FRP confined concretes. The model assessment presented in this study clearly illustrates the important influences of the size of the test databases and the selected test parameters used in the development of artificial intelligence models on their overall performances.

References


