ORIGINAL ARTICLE

Fuzzy logic model for pullout capacity of near-surface-mounted FRP reinforcement bonded to concrete

Kourosh Nasrollahzadeh¹ · Solmaz Afzali¹

Received: 7 August 2017 / Accepted: 23 June 2018 / Published online: 9 July 2018 - The Natural Computing Applications Forum 2018

Abstract

The application of fiber-reinforced polymer (FRP) strips or rods in the form of near-surface-mounted (NSM) reinforcement has become an attractive solution to strengthen the existing buildings and bridges. It is of interest to engineers to have an accurate estimate of the bond capacity of this technique. In this paper, fuzzy logic approach is utilized to propose an alternative method of determining the pullout strength of NSM FRP strips/rods which are bonded to the concrete block. Two types of fuzzy logic models, namely Mamdani and Takagi–Sugeno, are developed. With the aim of enhancing the interpretability of the fuzzy model, the rule base of Mamdani model is extracted from the classification decision tree, and the membership functions corresponding to the linguistic concepts are built by uniform partitioning the range of variables. On the other hand, in order to arrive at closed-form equations for pullout capacity, the subtractive clustering algorithm is employed to deduce the rule base and membership functions of Takagi–Sugeno model (first order), and its consequent part is tuned by the least square optimization using training dataset. Several fuzzy logic models of both types with different numbers of rules are developed and compared in terms of different error measures. To train and validate the fuzzy models, a large database of 384 direct pullout tests on NSM FRP bonded to concrete is assembled from the literature. The results reveal that both of the proposed Mamdani and Takagi–Sugeno models demonstrate good accuracy against the experimental data and outperform the published models. A parametric study indicates that the proposed fuzzy models can predict the maximum effective bond length, and thus, they are able to capture the underlying mechanics of the problem.

Keywords Fuzzy logic · Data mining · Decision tree · Bond strength · Fiber-reinforced polymer · Concrete

1 Introduction

Retrofit of the existing concrete buildings and bridges is an urgent need worldwide. The retrofit is conducted to improve the seismic resistance to meet newly developed design codes or to upgrade the load carrying capacity against the increased gravity load. Fiber-reinforced polymer (FRP) composites have been extensively applied in retrofitting of structures and are known as a superior technique comparing other traditional strengthening methods due to their lightweight, high strength, easy installation, and excellent durability. In order to enhance flexural or shear strength of reinforced concrete beams and

slabs by means of FRP, there are two methods of application as follows. First, the FRP sheets/plates are bonded to the external surface of the concrete, that is, called the externally bonded reinforcement (EBR) [\[1](#page-27-0)]. Second, FRP rods/strips are mounted into the groove cut near the surface of the concrete and filled with the resin. The latter is briefly called as near-surface-mounted (NSM) FRP reinforcement [\[2](#page-27-0)], and its geometric properties are illustrated in Fig. [1.](#page-1-0) The NSM technique is newer than EBR method and is less investigated. The advantages of NSM over EBR are considered as follows: the protection of FRP bars against vandalism or environmental effects as a result of its embedment in the concrete cover [\[3](#page-27-0)] and the removal of need for preparation of concrete surface prior to FRP installation [[4\]](#page-27-0). The bond capacity and failure mode of NSM FRP are affected by variety of factors including mechanical properties of FRP, epoxy, and concrete; surface

 \boxtimes Kourosh Nasrollahzadeh nasrollahzadeh@kntu.ac.ir

¹ Department of Civil Engineering, K. N. Toosi University of Technology, P.O. Box 15875-4416, Tehran, Iran

treatment and dimensions of FRP rod/strip; and groove depth and width [[5\]](#page-27-0).

It is of importance to engineers to accurately estimate the bond capacity of NSM FRP so as to take full advantage of materials and to avoid debonding failure. The bond capacity here refers to the pullout strength of NSM FRP bonded to a concrete block. Unlike the EBR method, very few models have been reported in the literature to predict the bond strength of NSM FRP technique. The existing models on the pullout capacity of NSM FRP have been developed by regression of experimental results combined with a fracture mechanics-based approach or outcomes of finite element models [\[6–8](#page-27-0)]. The available models on NSM differ from each other in the way they incorporate a number of parameters such as concrete strength or groove dimensions. Besides, the mechanical properties of bonding agent have not been accounted for in some models. Moreover, some of the models were developed only for FRP strips and thus are not applicable to FRP rods.

With the aim of developing an alternative and reliable method to predict the pullout strength of NSM FRP rods and strips, the fuzzy logic approach is employed in this paper. The approach of fuzzy logic has not yet been applied to the present topic. The main advantage of fuzzy logic

system stems from its interpretability [[9,](#page-27-0) [10](#page-27-0)], which may not be pronounced in other neural computing techniques. Neuro-fuzzy methods have been applied to the field of concrete [[11,](#page-27-0) [12](#page-27-0)] and FRP-wrapped concrete [[13\]](#page-27-0). The Takagi–Sugeno fuzzy inference system has been already employed by the author $[14, 15]$ $[14, 15]$ $[14, 15]$ in other applications of FRP in concrete structures. The Mamdani type of fuzzy logic model has been used in the field of concrete [[16\]](#page-27-0) and FRP applications [[17,](#page-27-0) [18\]](#page-27-0). Also, fuzzy logic expert systems have been utilized recently to evaluate other design issues related to NSM FRP such as deflections and cracks of concrete beams strengthened by NSM FRP [[19,](#page-27-0) [20](#page-27-0)]. In the present work, the two available types of fuzzy logic systems, namely Mamdani and Takagi–Sugeno (or Sugeno for short), are investigated. Classification decision tree is used to develop Mamdani model, whereas subtractive clustering is employed to derive Takagi–Sugeno model. Both fuzzy models are inferred from 384 experimental data on the direct pullout tests of NSM FRP, which have been collected from the open literature. Detailed procedures to develop each of the fuzzy models are explained, and the outcomes of the proposed fuzzy models are compared with the predictive equations published in the literature.

2 Experimental database

The experimental data, which are employed in this paper to develop fuzzy models, contain 384 direct pullout tests on NSM FRP rods or strips bonded to concrete blocks. The test data were assembled from several experimental campaigns reported in the open literature and were conducted by different researchers [\[3–7](#page-27-0), [21–39](#page-27-0)]. Since all experimental data used here are extracted from papers published in the well-known journals, it is believed that the data are reliable and the instruments are accurate enough.

As shown in Fig. [1,](#page-1-0) the pullout test contains a prismatic concrete block in which a groove is cut. The FRP strip/rod is inserted in the groove which is later filled with the resin. After the resin hardens, the FRP rod/strip is subjected to the tensile force exerted by a jack, while the concrete block is restrained by appropriate supports at a face perpendicular to the direction of the applied force. The concrete block may be placed horizontally, or alternatively, vertically in the test setup during the loading test. The main characteristics of the pullout test (as illustrated in Fig. [1\)](#page-1-0) are common among the experimental campaigns employed here. Thus, the testing conditions for determining the pullout capacity are similar among the studies considered herein.

The detailed information on the experimental data including the geometrical/mechanical properties of FRP bars/strips, groove, epoxy, and concrete as well as other testing conditions such as the experimental pullout capacity and bonded length is presented in ''Appendix [3](#page-18-0).'' The FRP reinforcements used in the specimens were different in cross-sectional shape (round bar, and square or narrow strip), surface treatment (smooth, sand coated, or spirally wound), and type of the fiber (carbon, glass, or basalt). The majority of NSM FRP in the test samples was of CFRP type. In the case that some samples were duplicated, only one test sample with the minimum pullout capacity among the replica samples was taken into account. By putting aside the duplicate samples, the whole database decreased from the original 384 experiments to 222 test specimens. Besides, 78 specimens out of 222 samples were FRP strips,

and the rest were rounded bars (i.e., rods). The statistics of properties for the experimental database are listed in Table 1.

3 Selection of input variables for fuzzy logic models

Selection of input parameters is a major stage in developing a fuzzy logic model. According to the NSM FRP bond models currently available in the literature [\[7](#page-27-0), [8](#page-27-0)], the parameters assumed to have an influence on the bond capacity are as follows: the bonded length (L) , concrete compressive strength (f'_c) , the axial rigidity of FRP rod or strip $(A_f E_f)$, and the groove depth-to-width ratio (D_g/W_g) . In the case of NSM strips, another parameter representing the length of debonding failure plane, which is defined as summation of three side lengths of the groove [\[8](#page-27-0)] or sum of twofold of strip width and its thickness plus 4 mm [\[7](#page-27-0)], is also considered in some models. Besides, the FRP tensile strength, though not affecting the bond capacity, is used in certain models to separate the FRP rupture failure from the bond failure mode. Furthermore, the rod/strip surface treatment has been considered as a test variable in experimental papers [\[24](#page-27-0), [33\]](#page-27-0). However, the surface treatment, due to its qualitative characteristics, is not included in the mathematical expressions of the existing bond models. Another important factor is related to the mechanical properties of the adhesive used between NSM FRP and concrete. Among different properties of adhesive, its tensile strength (f_e) has been reported in majority of the experimental campaigns.

It is also noted that some of the existing NSM bond models have been developed only for FRP strips and not for rods, whereas the present paper is aimed at proposing a model which is applicable to both cases of strip and rod. In order to distinguish between strip and rod, a quantitative term is introduced herein as the ratio of the normalized perimeter to the normalized cross-sectional area of the NSM reinforcement, denoted as $P_{\text{norm}}/A_{\text{norm}}$, in which the

Table 1 Statistics of the experimental data

| Parameter | L (mm) | $A_{\iota}E_{\iota}$ (kN) | f_c' (MPa) | f_e (MPa) | $P_{\text{norm}}/A_{\text{norm}}$ | $D_{\rm g}/W_{\rm g}$ | D_{σ} (mm) | f_u (MPa) | P_f (kN) |
|---------------------------------|----------|---------------------------|--------------|-------------|-----------------------------------|-----------------------|-------------------|-------------|------------|
| Minimum | 30 | 1300 | 15 | b | 0.90 | | 10 | 749 | 7.17 |
| Maximum | 508 | 16.966 | 64.8 | 90.7 | 5.45 | 6.89 | 40 | 3100 | 100 |
| Mean | 206.9 | 7409.6 | 34.9 | 39.6 | l.93 | 1.77 | 18.4 | 1954 | 46.1 |
| Median | 228 | 7647.4 | 30 | 30 | . 46 | | 17.1 | 2068 | 46.8 |
| Coefficient of variation $(\%)$ | 49.4 | 55.6 | 38.5 | 60 | 67.8 | 81.9 | 31.2 | 29.5 | 45.3 |

 D_g is groove depth, W_g is groove width, P_{norm} is normalized perimeter of rod/strip, A_{norm} is normalized sectional area of rod/strip, f_e is epoxy tensile strength, f_c' is concrete compressive strength, A_jE_f is the FRP axial rigidity, L is bond length, and P_f is bond capacity

perimeter and cross-sectional area are divided by their own maximum values. In the experimental database used in the present paper, the maximum value for the rod/strip perimeter and area is 45.5 mm and 132.7 mm², respectively. It is obvious that the value of $P_{\text{norm}}/A_{\text{norm}}$ for the rod shape is smaller than that for the strip shape.

Regarding the above-mentioned explanations, six independent variables are selected as inputs for the proposed fuzzy models, as follows: the bond length (L) , concrete compressive strength (f_c') , the axial rigidity of FRP rod/ strip $(A_f E_f)$, the groove depth-to-width ratio (D_e/W_g) , tensile strength of epoxy (f_e) , and the ratio of $P_{\text{norm}}/A_{\text{norm}}$. The statistical descriptors of these variables for the experimental database employed in the current paper are listed in Table [1](#page-2-0).

4 Overview of fuzzy logic-based models

In order to derive a relationship between the input variables (namely, L, f_c' , $A_f E_f$, D_g/W_g , f_e , and $P_{\text{norm}}/A_{\text{norm}}$) and the output (i.e., the pullout capacity, P_f), fuzzy logic approach is utilized in this paper. The objective is to develop fuzzy logic models inferred from the available experimental data to predict the pullout capacity of NSM FRP rod/strip bonded to concrete. A fuzzy logic model consists of several rules; each has an antecedent together with a consequent part. The antecedent is composed of some membership functions; each assigned to an input variable and are interconnected by use of logical operators such as AND or OR [\[40](#page-27-0)]. The role of membership function is to map the value of an input variable to a number in the range of 0–1, expressing the fuzzy nature of the model. Depending on the format of the consequent part, there are two types of fuzzy models, namely Takagi–Sugeno fuzzy models [[41](#page-28-0)] and Mamdani fuzzy models [[42\]](#page-28-0).

In the present paper, both Mamdani and Takagi–Sugeno (or Sugeno for short) fuzzy models are applied to the problem in hand. The merit of Sugeno model, as developed herein, lies in its closed-form presentation of the pullout capacity. On the other hand, Mamdani model is proposed here with the aim of providing highly interpretable rules. In order to develop models and run the related codes, MATLAB is used in the present paper. The different steps in building both fuzzy models are explained in the followings.

4.1 Development of Mamdani fuzzy model based on decision tree

In the current paper, the main purpose from developing a Mamdani-type model is to enhance the interpretability of the fuzzy logic system. Interpretability is known as the advantage of fuzzy modeling as compared to other soft computing techniques such as neural network [\[9](#page-27-0), [10](#page-27-0)]. Interpretability of fuzzy model means that an expert can easily understand the rule base of the model and modify or extend it if needed. It is particularly important when the model is inferred from data to ensure that the reasoning is understandable to the user, as opposed to the case that the automatic rule induction through supervised learning may result in the models containing meaningless coefficients obtained only on the basis of minimizing error between experimental output and model prediction. The interpretability can be achieved at two levels: membership function and rule induction technique, which will be discussed as follows.

4.1.1 Selection of membership functions by uniform partitioning the input/output space

A simple approach to reach an interpretable membership function is to divide the practical range of the variable into several equal partitions. The number of partitions is the same as the number of membership functions whose coordinates are determined based on the length of each partition, as will be discussed later. The practical range of each variable is presented in Table [1](#page-2-0) according to the available experimental database. In here, the range of all input variables is divided into five equal intervals except for the variable D_{g}/W_{g} , which is partitioned into three equal intervals. The number of partitions can be decided by looking at the histograms for distribution of experimental data, as demonstrated in Fig. [2](#page-4-0) for D_g/W_g and f_c' as examples of input parameters. For instance, according to Fig. [2](#page-4-0), there are gaps in the experimental data points in the range of D_g/W_g , suggesting that three divisions would be sufficient to cover the whole range of this input variable, whereas for f_c' , five partitions are considered due to a relatively uniform distribution of data points within the range of f_c' . It goes without saying that increasing the number of partitions may result in a more complex model. In the next step, membership functions, which correspond to different linguistic concepts, are constructed. In the case of three partitions for D_g/W_g , the linguistic classes of low (L), medium (M), and high (H) are considered as illustrated in Fig. [3](#page-5-0). Also shown in Fig. 3, for f_c' with five partitions, the linguistic classes of very low (VL), low (L), medium (M), high (H), and very high (VH) are defined. As for the shape of membership functions, trapezoids are assigned to the first and last classes, while triangles are allocated to the intermediate classes (see Fig. [3](#page-5-0)). Triangular and trapezoidal shapes are the most common membership functions reported in the literature [\[9](#page-27-0), [10,](#page-27-0) [17\]](#page-27-0). It is also noted that the right corner and top and left corner of a triangular membership function lie in the midpoint of three consecutive

Fig. 2 The statistical distributions of D_g/W_g (top) and f_c' (bottom) in the experimental database

partitions, respectively, resulting in overlapping parts between two successive membership functions. As for the trapezoidal functions, the top base spans between the midpoint of the first (or last) partition and the minimum (or maximum) value in the range of the considered input variable. In order to build the membership functions for the model output (i.e., P_f), a similar procedure is employed except that six intervals are considered in order to achieve a higher accuracy in prediction of the bond capacity. The linguistic terms for P_f are designated as very low (VL), low (L), almost low (AL), almost high (AH), high (H), and very high (VH), as demonstrated in Fig. [3](#page-5-0). The coordinates of each membership function are presented in Tables [2](#page-5-0) and [3](#page-6-0) for the inputs and output, respectively. The membership functions, as described above, are constructed only by use of the given range of each variable without any need for

optimization or other a priori assumptions, leading to enhance the interpretability of the model.

4.1.2 Rule induction by classification decision tree and its application to the present problem

Another aspect of interpretability of fuzzy model is related to the rule induction method. As a technique to induce fuzzy rules, the classification decision tree has been suggested by several researchers [\[10](#page-27-0), [43\]](#page-28-0). In this paper, the rule base of the proposed fuzzy model is extracted from the binary decision tree, which is employed for classification of the experimental observations in terms of the linguistic labels considered for the variables. First, each variable should be converted to a linguistic label, i.e., the so-called fuzzification process. Given a value for an input or output variable, the linguistic label corresponding to that is determined based on its value falling in which partition of the variable range. It is recalled that the range of each variable is partitioned into a number of equal intervals. For instance, if a given value of f_c is within the third partition out of five partitions of the experimental range of f'_c , then the linguistic label of medium (M) is assigned. In other words, among the considered membership functions for f'_c , the one denoted by medium (M) yields the largest degree of membership for the given f_c' .

The experimental database used in this study contains 222 numerical data points, as mentioned earlier. However, if these data are expressed in terms of the relevant linguistic labels, 165 unique data remain, while the rest are duplicate. Next, the algorithm of decision tree [[44\]](#page-28-0) is applied to the 165 labeled experimental data. It is also noticed that the membership labels (and not the numeric values) of the experimental data are employed in developing the decision tree and that no parameter is tuned during the model development. The idea is to select an input variable and choose one of its linguistic labels as a splitting level, which divides the database into two parts: One is larger or equal than the splitting level, and the other is less than that. The criterion for selection of the best variable and its splitting level is to maximize the information gain (IG) in a node, defined as follows:

$$
IG(node) = G(node) - GL \times wL - GR \times wR
$$
 (1)

where a node refers schematically to a point in the decision tree which shows the selected input variable with its splitting level and G stands for Gini index and is calculated for either a node [i.e., G (node)] or the left/right branch of the considered node (i.e., G_L or G_R) with the following formula:

$$
G(\text{node}) = 1 - \sum_{i=1}^{6} \left(\frac{n_i}{n}\right)^2 \tag{2}
$$

Table 2 The coordinates of membership functions for the input variables

VL and VH (L and H for D_g/W_g) are trapezoidal membership functions, and L, M and H (M for D_g/W_g) are triangular ones The numbers inside each bracket refer to the abscissas of the trapezoid/triangle clockwise

where n_i is the number of experimental observations whose outputs' level is the ith linguistic label. In the current paper, $i = 1, 2, ..., 6$, which refer to the linguistic labels of VL, L, AL, AH, H, and VH, respectively. Also, n is the total number of the observations in the node, that is, $= \sum_{i=1}^{6} n_i$. The Gini index [\[45](#page-28-0)] is a measure of impurity of classification in a node, meaning that a node, in which all observations belong to only one class of output, will have a Gini index of equal to zero, that is, a pure node. In Eq. (1), w_L (or w_R) refers to the ratio of the number of observations in the left (or right) branch to that in the original node $(w_L + w_R = 1).$

The algorithm of the classification decision tree starts by sorting the experimental observations based on the

Table 3 The coordinates of membership functions for the output variable

| Linguistic labels for the output variable | | | | | | | | | |
|---|--------------|-----------------|------------------|------------|-----------------|--|--|--|--|
| Very low [VL] | Low [L] | Almost low [AL] | Almost high [AH] | High [H] | Very high [VH] | | | | |
| [7 7 15 30] | [15, 30, 46] | [30 46 62] | [46 62 78] | [62 78 93] | [78 93 100 100] | | | | |
| | | | | | | | | | |

VL and VH are trapezoidal membership functions, and L, AL, AH, and H are triangular ones

The numbers inside each bracket refer to the abscissas of the trapezoid/triangle clockwise

ascending order of the linguistic labels of each input variable. For example, the linguistic labels of f_c' in the ascending order are as follows: VL, L, M, H, and VH. Then, any linguistic label is examined as a potential splitting point. Thereby, IG from Eq. (1) is calculated for each sorted input variable and each linguistic label as a potential splitting point. The final decision on selection of the best input as well as the best linguistic label for the splitting level is made on the basis of the maximum value of IG which is obtained. Once an optimum input with the corresponding splitting label (such as, f_c' as an input and [M] as the splitting level) is selected, the database is divided into two parts designated by the left and right branches. As an example, a branch designated by f_c - \langle [M] refers to the values of f_c' belonging to linguistic labels of either [VL] or [L], whereas the branch on opposite side is allocated to $f_c' \geq [M]$, referring to the linguistic labels of either [M], [H], or [VH]. Then, the above algorithm is repeated for the experimental observations belonging to the left (or right) branch. Again, in each branch, an optimum input variable with the best splitting label is determined by maximizing the value of IG. By proceeding the algorithm, the number of branches and eventually depth of the tree increase. The procedure continues till one of the following thresholds is met: (a) The number of observations related to a node shall be greater than 5 and that related to the left or right branch shall be at least 1; (b) the algorithm can no longer improve the value of IG.

The decision tree derived for classification of the experimental database of the current paper is illustrated in Fig. [4](#page-7-0). The uppermost node in the decision tree of Fig. [4](#page-7-0) is called as the root node, whereas the lowest nodes are known as leaves which correspond to the linguistic labels of the output variable (i.e., P_f). For instance, a leaf showing the class of [H] means that the majority of experimental observations falling in this particular branch possess the linguistic label of HIGH for P_f . By following consecutive branches starting from the root node down to a leaf, a fuzzy rule is established. The rule coming from a decision tree is not necessarily complete, meaning that its antecedent may not include all the input variables. Finally, the classification decision tree shown in Fig. [4](#page-7-0) is equivalent to a Mamdanitype fuzzy logic model, in which the number of fuzzy rules is equal to the number of leaves in the decision tree, that is, 35 rules. The rules derived from a decision tree are clear and easy to understand, leading to enhance the interpretability of the fuzzy model.

In order to assess the misclassification error of a node, denoted as E (node), the following relation is used:

$$
E(\text{node}) = 1 - \frac{n_j}{n} \tag{3}
$$

where the index j refers to the output class label whose number of experimental observations (i.e., n_i) is the largest one among the observations of different classes existing in the considered node. The total number of observations in a node is denoted by n. As an example, if the majority of observations in a node or a leaf have an output belonging to the label of almost high (i.e., AH), the value of j becomes 4 since [AH] is the fourth linguistic label of P_f . Furthermore, if the number of observations related to [AH] is 20 out of the total observations of 50, then $n_i = 20$ and $n = 50$, and the error of the node is $E = 0.6$ according to Eq. (3). It is recalled that in a leaf, the final class of the output is decided based on the value of n_i , meaning that the class of a leaf is the linguistic label with the largest occurrences in the experimental observations. If the error in a leaf node, evaluated by Eq. (3), exceeds a user-specified threshold, then that leaf can be eliminated by merging the leaves in that level with a node located just in one level higher than it. Repeating the process of merging nodes of lower level with the node of higher level (i.e., the so-called pruning procedure) can help decrease the depth of a decision tree and eventually reduce the number of the corresponding fuzzy rules.

By pruning the decision tree of Fig. [4](#page-7-0) at several con-secutive levels, the simpler tree shown in Fig. [5](#page-7-0) is obtained, which is finally chosen as the basis to generate the Mamdani model with 17 fuzzy rules. The formats of all rules derived from the decision tree of Fig. [5](#page-7-0) are listed in "Appendix [1.](#page-18-0)" An example of one rule is as follows: IF L is Low or Medium, and $A_f E_f$ is not Very Low, and f_c ' is Medium or High or Very High, and f_e is Medium or High or Very High, and $P_{\text{norm}}/A_{\text{norm}}$ is Very Low, THEN P_f is Almost Low. It is observed that a rule may contain a set of

Fig. 4 The original classification decision tree (35 rules)

Fig. 5 The pruned decision tree (17 rules)

membership functions (i.e., linguistic variables) interconnected by Boolean operators of OR and NOT, which are defined as the minimum and complement, respectively. For example, the membership function for not Very Low can be constructed as complement of that for Very Low. Similarly, the membership function for the combined variable of Medium or High is built by taking into account the minimum of both membership functions for Medium as well as High. The schematic graphs for the combined membership functions are illustrated in Fig. 6.

Fig. 6 Membership functions combined by logical operators OR and NOT (VL, L, M, H, VH stand for linguistic variables of very low, low, medium, high, and very high, respectively)

4.2 Development of Sugeno fuzzy model based on subtractive clustering

The only difference between the rule base of a Sugeno fuzzy model and that of a Mamdani-type fuzzy model lies in the consequent part. In Sugeno model, the consequent part is either a constant value (i.e., the so-called zeroth order) or a linear combination of all input variables plus a constant term (i.e., the so-called first order). The unknown coefficients of the consequent part in Sugeno fuzzy model are commonly determined following an optimization procedure through training against the known experimental data points. However, this is considered as a drawback in the sense that optimization reduces the interpretability of the fuzzy model. Another issue is the definition of membership functions for the input variables and the induction of rules in Sugeno model. In the current paper, a wellknown procedure called as subtractive clustering [[46\]](#page-28-0) is employed to generate the rule base of the fuzzy model. The application of subtractive clustering-based Sugeno model to the field of FRP composites has been already reported by the author [\[14](#page-27-0), [15\]](#page-27-0). Another widely used technique for clustering of numerical data is grid partitioning. The disadvantage of grid partitioning is that the number of generated rules may increase a lot when the number of input variables increases. For instance, in the case of the current problem, where there are six input variables, if only two membership functions are assigned to each input, then the number of generated rules becomes 64 (i.e., 2 to the power of 6). Increased numbers of rules yield more complex models with larger number of parameters to be tuned, which in turn reduce the model generalization ability. On the other hand, if subtractive clustering is utilized, then the number of rules, the number of clusters, and the number of membership functions would be all the same. For instance, by considering three membership functions for each input variable, the number of rules still remains as low as three. This is the main reason that subtractive clustering method is employed in this paper so that the simpler models with less number of rules can be created. In the following, a brief overview of the method of subtractive clustering is explained.

4.2.1 Overview of subtractive clustering algorithm

The subtractive clustering algorithm is a technique to extract cluster centers from an assembly of numerical data points (i.e., the experimental database) such that wherever more numbers of the data are concentrated, a cluster center is found. To do so, a potential of being a cluster center, denoted by P_i , is assigned to each of the data (i.e., the *i*th experimental datum), as follows:

$$
P_i = \sum_{j=1}^n \exp\left(\frac{-4}{r_a^2} \cdot \left\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\right\|^2\right) \tag{4}
$$

where $\|$ $\|$ stands for the Euclidean distance between two of data points, which are expressed as vectors whose coordinates contain input and output values, n is the total number of data points, and r_a is a user-specified value (in the range of 0–1) for the radius of a cluster. In the problem in hand, $\mathbf{x}_i = [L, A_j E_f, f_c', f_e, P_{norm}/A_{norm}, D_g/W_g, P_f]_i$, and $\hat{\mathbf{x}}_i$ is normalized with respect to the range of each coordinate. The datum with the maximum potential value is selected as the first cluster center. In order to find the next cluster centers, an amount of $P_k^* \cdot \exp\left(\frac{-4}{r_b^2} \cdot \left\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_k^*\right\| \right)$ $\left(\frac{-4}{r^2} \cdot ||\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_k^*||^2 \right)$ is subtracted from the potentials of all data points, where P_k^* and $\hat{\mathbf{x}}_k^*$ are the potential and the vector of normalized coordinates for the kth cluster center and r_b is a multiplier of r_a . Thereafter, the point with the highest remaining potential is the next cluster center. The procedure ends if $P_k^* < \underline{\varepsilon} P_1^*$, where ε is a user-defined value, and the current cluster center (i.e., the kth center) is not accepted. Besides, if $\underline{\varepsilon}P_1^* < P_k^* < \overline{\varepsilon}P_1^*$, where $\overline{\varepsilon}$ is another user-defined value, then the following condition is evaluated: When $\frac{d_{\min}}{r_a} + \frac{P_k^*}{P_1^*} \ge 1$, where d_{\min} is the shortest of the distances between the kth cluster center and all of the other cluster centers, then the current data point is accepted as the kth cluster center and the procedure continues; otherwise, the current cluster center is rejected and P_k^* is set to 0, and the data point with the largest remaining potential is chosen as the new kth cluster center, and the above process is repeated. In the above-mentioned algorithm, there are several user-defined parameters including r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$, which affect the number and coordinates of the generated cluster centers. In this paper, varying effects of each of them are examined, and the best values are selected based on the minimum error between the model output and experimental results.

4.2.2 Membership function selection and rule induction by subtractive clustering and its application to the present problem

Each of the identified cluster centers forms a fuzzy rule as follows: ''IF an input variable is close to the cluster center THEN its output is close to the cluster center's output.'' The level of truth for this rule depends on Euclidean distance between the given variable and the cluster center, meaning that the degree of truth increases as the distance decreases. The following relation is used to determine the level of truth (or weight), w_i , of the *i*th fuzzy rule:

$$
w_{i} = \exp\left\{\frac{-4}{r_{a}^{2}} \left[\left(\frac{L - L_{i}^{*}}{L_{\max} - L_{\min}}\right)^{2} + \left(\frac{A_{f}E_{f} - \left(A_{f}E_{f}\right)_{i}^{*}}{\left(A_{f}E_{f}\right)_{\max} - \left(A_{f}E_{f}\right)_{\min}}\right)^{2} + \left(\frac{f_{c}^{'} - f_{ci}^{*}}{f_{c, \max}^{'} - f_{c, \min}^{'}}\right)^{2} + \left(\frac{f_{c} - f_{ci}^{*}}{f_{c, \max}^{'} - f_{c, \min}^{'}}\right)^{2} + \left(\frac{\frac{P_{\text{norm}}}{A_{\text{norm}}} - \left(\frac{P_{\text{norm}}}{A_{\text{norm}}}\right)_{i}^{*}}{\left(\frac{P_{\text{norm}}}{A_{\text{norm}}}\right)_{\max} - \left(\frac{P_{\text{norm}}}{A_{\text{norm}}}\right)_{\min}}\right)^{2} + \left(\frac{\frac{D_{g}}{W_{g}} - \left(\frac{D_{g}}{W_{g}}\right)_{i}^{*}}{\left(\frac{D_{g}}{W_{g}}\right)_{\max} - \left(\frac{D_{g}}{W_{g}}\right)_{\min}}\right)^{2} \right\}
$$
\n(5)

where the denominators are the range of input variables in the experimental database, and in the numerators, the variables without star refer to the input coordinates of a given variable, and those with a star refer to the input coordinates of the i th cluster center. The index i varies between 1 and c , where c is the number of rules that is equal to the number of cluster centers. Equation (5) implies an antecedent with Gaussian membership functions interconnected by AND operators defined as multiplication. The consequent part of the ith rule in Sugeno fuzzy model (first order) is as follows:

$$
P_{f,i} = a_i L + b_i (A_f E_f) + c_i f_c' + d_i f_e + e_i \left(\frac{P_{\text{norm}}}{A_{\text{norm}}}\right) + f_i \left(\frac{D_g}{W_g}\right) + g_i
$$
\n
$$
(6)
$$

The final output (i.e., the pullout capacity P_f) of the fuzzy model is calculated as:

$$
P_f = \sum_{i=1}^{c} \frac{w_i}{\sum_{j=1}^{c} w_j} P_{f,i}
$$
\n(7)

where a_i , b_i , c_i , d_i , e_i , f_i , and g_i ($i = 1,..., c$) are unknown coefficients to be determined by minimizing the error between the model output and experimental value for P_f .

As mentioned earlier, in the developed Sugeno model, which is based on subtractive clustering algorithm, the membership function is of Gaussian type (i.e., exponential) as proposed by Chiu [\[46](#page-28-0)]. The reason for selecting the Gaussian membership function stems from the method of subtractive clustering, in which the membership value of a data point in a cluster is determined based on an exponential function of the Euclidean distance between the given point and the center of the cluster under study, as presented in Eq. (5). In fact, each input coordinate of each cluster center forms a Gaussian membership function, where the mean is the cluster center coordinate and the standard deviation is proportional to the range of the values of the given input variable. The membership functions considered for Sugeno models are illustrated in Fig. [7](#page-10-0) (bottom part). The controlling parameters of the membership function are the same as the parameters affect the

Fig. 7 The diagram of fuzzy rules for the proposed Mamdani model (top) with 17 rules and Sugeno model (bottom) with 3 rules (the experimental value = 30.46 kN)

subtractive clustering algorithm, namely r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$. The effects of varying values of these parameters are discussed in the following parts of the paper.

The first step to derive a fuzzy model is to decide the values for the clustering parameters (i.e., r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$) which result in identifying the cluster centers within the domain of the experimental database. Then, the IF-part of the rule base is materialized by use of the cluster centers (see Eq. 5), while the THEN-part is formed after optimizing the coefficients of Eq. (6) in terms of the model error evaluated on experimental datasets.

For developing soft computing models which are inferred from experimental data, it is required to use only a portion of data for training process, whereas the rest should be employed to check or validate the generalization ability of the developed model. In this regard, the total database of experiments is divided randomly into two equal portions called as training and checking datasets. The training dataset is used for tuning the parameters, whereas the checking dataset is employed to assess the model generalization on an unfamiliar data to avoid overfitting. Thus, the checking dataset is not used in the training process. For division of data into training and checking datasets, the data points which are assigned to a certain cluster are divided randomly into two equal portions. It is noted that the cluster center which is the closest to a data point is considered as the cluster assigned to that data point. Table 4 presents several error measures obtained on training, checking, and total experimental datasets for several Takagi–Sugeno fuzzy models derived using some selected values of r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$. The error measures considered here include: coefficient of variation (CoV) and mean of experiment-to-prediction ratio, as well as the mean absolute percentage error (MAPE) between the experiment and model prediction. Employment of checking dataset, which is here generated randomly, is essential to prove that the developed models are insensitive to the used training dataset considering that one of the criteria to select the

proposed models in this study is to assure that the error on training dataset is similar to that on checking dataset (see Table 4) so as to avoid overfitting. As a result of this, the proposed models are indeed the best models which can be extracted from the current database.

Each set of values for the clustering parameters (i.e., r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$) results in a particular Sugeno fuzzy model. In order to examine the effect of varied values of these userdefined parameters and to select the best values, each of the clustering parameters is varied with a small increment within its possible range so that it is assured that all possible values for each clustering parameter are generated. Then, all possible combinations of the values for the clustering parameters are considered, and as a result, numerous sets of values for r_a , r_b , $\underline{\varepsilon}$, and $\overline{\varepsilon}$ are obtained. In the next step, each set of the parameters is used to build a Sugeno fuzzy model by subtractive clustering algorithm. Finally, the best fuzzy model is considered the one whose errors on both training and checking datasets are similar to each other (to ensure the generalization ability of the model) and also are minimum among the other prospective models. Besides, under the same performance in terms of errors, a simpler model with less number of rules is preferred. The above-mentioned procedure is implemented in a MATLAB code. For instance, as seen in Table 4, five different Sugeno models with the same number of rules (i.e., four rules) are built by varying the values of clustering parameters. The models presented in Table 4 serve as just demonstration and are chosen from many models built in this study by varying the clustering parameters. The finally selected set of parameter values for the case of fuzzy model containing four rules is listed in Table [5](#page-12-0), which is decided based on the minimum error between the model output and experimental results while developing a high ability of generalization, as explained above. By following this procedure, the most accurate fuzzy models with different numbers of rules are chosen and listed in Table [5,](#page-12-0) which also reports the associated clustering parameters, the

Table 4 The effects of varying values of clustering parameters on the generated Sugeno fuzzy models

| Clustering parameters | | No. of rules | Training dataset | | | | Checking dataset | | Total dataset | | | | |
|-----------------------|---------------------|----------------------------|-------------------|----------------|------|-----------|------------------|------|---------------|--------------|------|--------------|--------------|
| r_a | $\bar{\varepsilon}$ | $\boldsymbol{\varepsilon}$ | $\frac{r_b}{r_a}$ | | Mean | $CoV(\%)$ | MAPE $(\%)$ | Mean | CoV $(\%)$ | MAPE $(\%)$ | Mean | CoV $(\%)$ | MAPE $(\%)$ |
| 0.695 | 1.15 | 0.6° | 0.25 | $\overline{4}$ | 1.00 | 19.2 | 15.7 | 1.06 | 54.6 | 32.6 | 1.03 | 39.6 | 23.4 |
| 0.703 | 1.15 | 0.6° | 0.25 | 4 | 1.00 | 22.9 | 20.9 | 1.07 | 52.5 | 21.8 | 1.03 | 39.3 | 21.3 |
| 0.706 | 1.15 | 0.6° | 0.25 | $\overline{4}$ | 1.01 | 23.9 | 18.1 | 0.98 | 26.0 | 29.9 | 0.99 | 24.9 | 23.6 |
| 0.7095 | 1.15 | 0.6° | 0.25 | 4 | 1.00 | 20.3 | 16.1 | 0.97 | 25.8 | 22.7 | 0.99 | 23.0 | 22.4 |
| 0.7095 | 1.15 | 0.3 | 0.25 | $\overline{4}$ | 1.01 | 24.4 | 17.7 | 0.98 | 25.9 | 29.5 | 0.99 | 25.1 | 23.2 |
| 0.7095 | 1.15 | 0.3 | 0.15 | -8 | 1.00 | 19.1 | 13.1 | 0.96 | 77.2 | 35.3 | 0.98 | 54.4 | 23.4 |
| 0.7095 | 1.17 | 0.3 | 0.15 | -8 | 1.01 | 19.5 | 13.3 | 1.01 | 83.9 | 35.1 | 1.01 | 58.9 | 23.4 |

Mean values are calculated for the ratios of the experimental values to the predicted values

Table 5 The coordinates of cluster centers and coefficients in consequent part of several Sugeno models with different clustering parameters

Table 5 The coordinates of cluster centers and coefficients in consequent part of several Sugeno models with different clustering parameters

coordinates of the cluster centers, and the constants of consequent part.

5 Results and discussion

5.1 Comparison of predictions by fuzzy logic models and experiments

The performances of different fuzzy logic models developed in this paper are evaluated in Table 6 in terms of various error measures in comparison with the experimental data. There are three Mamdani-type fuzzy logic models developed here including the one with 17 rules in addition to two others containing 13 and 21 rules. All these Mamdani models are derived from the same classification decision tree having 35 leaves, as explained earlier (see Fig. [4](#page-7-0)), but by pruning at different levels. As for Sugenotype fuzzy logic models, five models with 1 up to 5 rules are built herein, as appeared in Table [5](#page-12-0) with their characteristics. According to Table 6, the Mamdani model with 17 rules develops definitely less errors in terms of CoV, MAPE, RMSE, and MAE, which are evaluated on the total experimental database, as compared to Mamdani models with 13 or 21 rules. Thus, the Mamdani model with 17 rules is finally chosen. As for Sugeno models reported in Table 6, the difference in one error type among the models with 3, 4, and 5 rules is hardly noticeable. Thereby, all these models may be considered as a candidate for the final selection. However, selection of the best model is a tradeoff between the reduced error measures and the simplicity of the model. In other words, if a model with less numbers of rules demonstrates error measures being more or less similar to another model which contains more numbers of rules, the model with less numbers of rules is of more interest since it is simpler. As a result, the Mamdani model with 17 rules and the Sugeno model with 3 rules are finally selected.

The schematic diagrams of the rule base and membership functions for Mamdani model (17 rules) and Sugeno model (3 rules) are illustrated in Fig. [7.](#page-10-0) As a numerical example, also shown in Fig. [7](#page-10-0), for an input variable with $L = 200$ mm, A_f $E_f = 2053$ kN, $f_c' = 52.8$ MPa, f_e . = 16 MPa, $P_{\text{norm}}/A_{\text{norm}} = 5.28$, and $D_g/W_g = 3.77$, the value of P_f is equal to 30.2 kN in Mamdani model and 30.5 kN in Sugeno model as compared to its experimental value of 30.46 kN. In this specific work, the accuracies offered by both Mamdani and Sugeno models are almost similar. However, a general comparison of these two types of fuzzy logic models is beyond the scope of the current paper. It should be mentioned that although the number of rules in the proposed Sugeno model is much less than that in Mamdani model, the number of parameters need to be tuned in Sugeno model with 3 rules is equal to 21 unknowns in the consequent part (i.e., 7 constants per each rule). As stated earlier, tuning these parameters against experimental data needs an optimization procedure, which in turn reduces the interpretability of the fuzzy logic model. On contrary, in development of Mamdani model based on the classification decision tree, there is no need to optimization, resulting in a highly interpretable model. Furthermore, with reference to Fig. [7](#page-10-0), it is observed that for the proposed Sugeno model, the membership functions, which are obtained by subtractive clustering, have no linguistic or qualitative meaning as opposed to the case of the proposed Mamdani model, where the membership functions are essentially attributed to the linguistic concepts. This aspect also contributes to the interpretability of the

| Models | Error measures | | | | | | | | | | | |
|---------------------|----------------|-----------|--------------------------------|---------------|------------|----------------------------------|-----------------|--------------|-----------|----------|--|--|
| | | | Total database (222 specimens) | | | Strip FRP dataset (78 specimens) | | | | | | |
| | Mean | CoV (%) | MAPE $(\%)$ | $RMSE$ (kN) | MAE (kN) | | Mean CoV $(\%)$ | MAPE $(\%)$ | RMSE (kN) | MAE (kN) | | |
| Mamdani (13 rules) | 0.99 | 34.0 | 33.7 | 15.2 | 11.72 | 1.08 | 33.8 | 32.8 | 18.2 | 14.85 | | |
| Mamdani (17 rules) | 1.04 | 26.6 | 21.5 | 11.9 | 9.15 | 1.09 | 22.9 | 17.4 | 12.8 | 9.14 | | |
| Mamdani (21 rules) | 0.94 | 27.2 | 28.8 | 12.2 | 9.19 | 1.07 | 21.6 | 16.1 | 11.6 | 8.20 | | |
| Sugeno (1 rule) | 1.01 | 29.4 | 28.20 | 12.79 | 9.96 | 1.09 | 24.6 | 19.2 | 13.1 | 10.08 | | |
| Sugeno (2 rules) | 1.01 | 24.5 | 22.71 | 11.02 | 8.29 | 1.07 | 22.0 | 16.3 | 11.0 | 8.61 | | |
| Sugeno (3 rules) | 1.00 | 23.1 | 21.90 | 10.66 | 7.88 | 1.04 | 20.3 | 15.9 | 10.5 | 8.03 | | |
| Sugeno (4 rules) | 0.99 | 23.0 | 22.42 | 10.48 | 7.77 | 1.02 | 19.9 | 16.5 | 10.7 | 8.25 | | |
| Sugeno (5 rules) | 0.99 | 22.9 | 22.14 | 10.22 | 7.57 | 1.02 | 20.0 | 15.6 | 10.4 | 7.93 | | |
| Zhang et al. $[8]$ | 1.05 | 36.4 | 33.86 | 16.27 | 12.11 | 0.96 | 30.2 | 29.5 | 16.4 | 12.74 | | |
| Seracino et al. [7] | - | | | | | 1.54 | 39.3 | 33.0 | 22.3 | 18.05 | | |

Table 6 Error measures of several fuzzy models with different numbers of rules and comparison with the existing models

Mean values are calculated for ratios of the experimental values to the predicted values

Models

Table 7 The expressions of the existing models for bond capacity of NSM FRP

 P_f is bond capacity, G_f is interfacial fracture energy, A_f is FRP bar cross-sectional area, E_f is FRP elasticity modulus, L_{per} is cross-sectional contour of the failure surface, f_{uf} is FRP tensile strength, L is bond length, L_e is effective bond length, τ_{max} is maximum bond stress of local stress–slip curve, D_g is groove depth, W_g is groove width, f_c' is concrete compressive strength, φ_f is height-to-width ratio of failure contour, and t_f and h_f are FRP strip thickness and height

proposed Mamdani model. As mentioned previously, interpretability is an important feature of fuzzy logic-based modeling when compared to the other soft computing paradigms. The rule base of the proposed Mamdani model is presented in ''Appendix [1](#page-18-0).'' On the other hand, the Sugeno model, in the format developed herein by use of subtractive clustering approach, can be presented as closedform relations, making it of particular interest to practicing engineers. The closed-form equations of the proposed Sugeno model appear in ''Appendix [2](#page-18-0).''

5.2 Comparison between the proposed fuzzy models and existing equations for pullout strength of NSM FRP bonded to concrete

In the literature, there are two models developed by Seracino et al. [\[7](#page-27-0)] and Zhang et al. [[8\]](#page-27-0) to predict the pullout capacity of NSM FRP bonded to concrete blocks. The equations for both models are presented in Table 7. It is noted that Seracino's model is applicable only for NSM strips (i.e., not including NSM rods). The error measures of both models are evaluated on the experimental database, and the results are reported in Table [6](#page-13-0). The data on NSM strips consist of 78 specimens out of 222 experiments of the total database and are employed in Table [6](#page-13-0) so as to compare Seracino's model with the other considered models. According to Table [6](#page-13-0), both of the proposed fuzzy logic models (i.e., Mamdani type and Sugeno type) demonstrate superior performance compared to Seracino's and Zhang's models. For instance, the CoVs for the ratios of the experimental value to model prediction for the proposed Mamdani model (17 rules) and Sugeno model (3 rules) are equal to 22.9 and 20.3% on the strip database, respectively, as compared to CoVs for Zhang's and Seracino's models, which are equal to 30.2 and 39.3%, respectively. Also, other error measures such as MAPE and RMSE decrease for the proposed fuzzy models comparing with the existing equations. Besides, the bond capacity predictions by selected fuzzy models and Zhang's model versus 222 experimental results are depicted in Fig. [8,](#page-15-0) which shows that dispersion of the results by Zhang's model is larger than that by the fuzzy logic models.

As a result, the performance of the proposed fuzzy models in terms of errors reported in Table [6](#page-13-0) and Fig. [8](#page-15-0) is better than the other existing models in the literature. Meanwhile, it is worth emphasizing that this paper seeks not only reducing the errors but also enhancing the interpretability of the proposed fuzzy models. In many cases, very low errors may be achieved but at the cost of reducing the interpretability issues. For instance, the adaptive network-based fuzzy inference system (ANFIS) is expected to provide higher accuracy since the parameters of the membership functions are iteratively adjusted by the training algorithm. On contrary, in both of the fuzzy models used here (namely subtractive clustering-based Sugeno model and classification decision tree-based Mamdani model), the membership functions of input variables are set fixed in the beginning and not altered during training. This feature, itself, enhances the interpretability of the model but may not yield the results as accurate as ANFIS does. Another point which should be noticed is that increasing the number of parameters to be tuned requires a larger database so as to assure that overfitting can be avoided.

Fig. 8 Predictions by different fuzzy models and Zhang's model [\[8\]](#page-27-0) for bond (i.e., pullout) capacity of NSM FRP versus experimental data

Although the fuzzy models proposed here (see ''Appendices [1,](#page-18-0) [2](#page-18-0)'') are more accurate than the existing regression-based models reported in the literature, they appear to be more complex than Zhang's or Seracino's models (see Table [7\)](#page-14-0). However, this is a common issue for most of soft computing models as compared to regressionbased models in the sense that a regression-based formula is so simple to be applied. On the other hand, a fuzzy

Fig. 9 Parametric study on the effective bond length using different fuzzy models (common details: $A_f E_f = 7115.5$ kN, $f_c' = 36.4$ MPa, $f_e = 39.7 \text{ MPa}, P_{\text{norm}}/A_{\text{norm}} = 2.14, D_g/W_g = 2.01$

model, though looks more complex, is not a black box and its advantage lies on its rule base which makes an expert easily understand the reasoning behind the model, and extend or modify it when needed. This is apparently opposed to a regression-based formula containing some coefficients suffering from lack of sound meaning.

5.3 Parametric study on the bond length of NSM FRP

In order to assess whether the proposed fuzzy models are able to capture the underlying mechanics of the problem in hand, the varying effects of the bond length L on the bond capacity P_f are examined in Fig. 9, where the other input variables are kept unchanged and equal to their mean values. According to the literature [\[8](#page-27-0)], there exists an effective bond length, beyond which the bond capacity no longer increases. As seen in Fig. 9, the proposed Mamdani model (with 17 rules) and Sugeno model (with 3 rules) are able to develop the concept of the effective bond length, meaning that after a certain amount of the bond length of NSM FRP, the pullout capacity remains constant. It is also observed from Fig. 9 that the amount of the effective bond length identified by the proposed Mamdani model is almost the same as that by the proposed Sugeno model. The other fuzzy models illustrated in Fig. 9 can predict, to some extent, the existence of an effective bond length except for the one-rule Sugeno model due to the fact that just one rule is most likely to be insufficient to describe the full behavior of the problem.

The simultaneous variations of both bond length, L, and concrete compressive strength, f_c' , across their own ranges, and their effects on the pullout capacity, P_f , are shown in three-dimensional diagrams of Fig. [10](#page-17-0) for the proposed

Mamdani and Sugeno models. The other input variables are kept unchanged. According to Fig. [10,](#page-17-0) the general trend of variation in the proposed Mamdani model is similar to that in the proposed Sugeno model. This conclusion is important in the sense that the rule bases of the two proposed fuzzy models are initiated from two different sources: The former comes from classification decision tree, whereas the latter stems from subtractive clustering algorithm. More specifically, it is revealed from Fig. [10](#page-17-0) that the proposed Sugeno model endorses the Mamdani model's rules, which are extracted from the decision tree of Fig. [5,](#page-7-0) such as: (1) Given very low values of L, the value of P_f remains very low or low, and (2) if L and f_c' are not very low, while $A_f E_f$ is not very low as it is the case in Fig. [10](#page-17-0), then P_f is not in very low and low ranges.

6 Conclusions

As an alternative method of determining the pullout capacity of NSM FRP rods or strips bonded to the concrete block, two fuzzy logic models are proposed in the current paper. The first fuzzy model is of Mamdani type, in which the rule base is deduced from a classification decision tree. The second one is of Sugeno type (first order), which is developed by the subtractive clustering algorithm. Several fuzzy logic models of each type with different numbers of rules are built. Among them, a Mamdani model with 17 rules and a Sugeno model with 3 rules are finally selected as they demonstrate the least errors on model predictions. The conclusions are as follows:

- 1. The decision tree-based Mamdani fuzzy model of the pullout strength, which contains the membership functions corresponding to linguistic labels, provides a highly interpretable rule base as listed in ''Appendix [1](#page-18-0).'' The proposed Sugeno model based on subtractive clustering serves as closed-form formulations for the pullout capacity of NSM FRP, as presented in ''Appendix [2.](#page-18-0)'' With reference to the mentioned Appendices, the researchers and engineers can employ the fuzzy logic models proposed in this study.
- 2. Comparison with a large experimental database including 386 direct pullout tests on NSM FRP, which are collated from the open literature, indicates that the predictions by the proposed fuzzy models are in good agreement with the experimental results. The prediction accuracies of both of the proposed fuzzy models are similar in this particular study. However, a general comparison of Mamdani versus Sugeno fuzzy models is beyond the scope of this paper.
- 3. The proposed fuzzy models outperform the existing models reported in the literature. Unlike some of the

Fig. 10 Prediction of P_f with respect to simultaneous variations of L and f_c' by the proposed Mamdani fuzzy model (top figure) and Sugeno model (bottom figure); (common details: $A_f E_f = 9533$ kN, $f_e = 35$ MPa, $P_{\text{norm}}/$ $A_{\text{norm}} = 1.55, D_g/W_g = 1$

published models which have been developed only for FRP strips, the proposed fuzzy models are applicable to both rods and strips. The coefficient of variation (CoV) for the ratio of the experiment to prediction in the proposed fuzzy models is almost 10% lower than that in Zhang's model on the total data and about 16% lower than that in Seracino's model on the strip data.

4. By conducting a parametric study on the bonded length of NSM FRP, it is revealed that the proposed fuzzy models are able to capture the underlying mechanics of the problem in hand since they can develop the concept of the effective bond length, that is, a bond length beyond which the bond capacity no longer increases.

The results obtained in the present work endorse the efficiency of fuzzy logic approach in predicting the pullout strength of NSM FRP bonded to concrete. It is, however, noted that the fuzzy model predictions are valid within the

considered range of input variables. Future work may deal with applications of the fuzzy logic approach to other design aspects of NSM FRP reinforcement such as flexural or shear strengthening of the concrete beams.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix 1: List of fuzzy rules in the proposed Mamdani model for the pullout capacity of NSM FRP

| IF | THEN |
|---|---------------------|
| 1. $L = [VL] \bigwedge f_c' = [VL]$ | $P_f = [VL]$ |
| 2. $L = [VL] \wedge f_c' \neq [VL]$ | $P_f = [L]$ |
| 3. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VL] $\bigwedge D_g / W_g \neq$ [H] $\bigwedge P_{\text{norm}}/$ $A_{\text{norm}} = [VL]$ | $P_f = [\text{AL}]$ |
| 4. $L \neq$ [VL] $\bigwedge A_j E_j =$ [VL] $\bigwedge D_g/W_g \neq$ [H] $\bigwedge P_{\text{norm}}/$ $A_{\text{norm}} = [L]$ | $P_f = [L]$ |
| 5. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VL] $\bigwedge D_g / W_g \neq$ [H] $\bigwedge P_{\text{norm}}/$ $A_{\text{norm}} = [M]$ | $P_f = [AH]$ |
| 6. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VL] $\bigwedge D_g / W_g \neq$ [H] $\bigwedge P_{\text{norm}}/$ $A_{\text{norm}} = ([H] \vee [VH])$ | $P_f = [L]$ |
| 7. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VL] $\bigwedge D_g / W_g =$ [H] | $P_f = [AH]$ |
| 8. $L \neq$ [VL] $\bigwedge A_f E_f = (L)$ v | $P_f = [AL]$ |
| [M] \vee [H]) $\bigwedge f_c' = ([VL] \vee [L]) \bigwedge f_e = [VL]$ 9. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VH] $\bigwedge f_c' =$ ([VL] v [L]) $\bigwedge f_e =$ [VL] | $P_f = [L]$ |
| 10. $L \neq$ [VL] $\bigwedge A_f E_f \neq$ [VL] $\bigwedge f_c' =$ [VL] \bigwedge $f_e \neq$ [VL] | $P_f = [AL]$ |
| 11. $L \neq [VL] \wedge A_f E_f = ([L] \vee [M] \vee$ [H]) $\bigwedge f_c' = [L] \bigwedge f_e \neq [VL]$ | $P_f = [AL]$ |
| 12. $L \neq$ [VL] $\bigwedge A_f E_f =$ [VH] $\bigwedge f_c' =$ [L] $\bigwedge f_e \neq$ [VL] | $P_f = [AH]$ |
| 13. $L = ([L] \vee [M]) \wedge A_i E_f \neq [VL] \wedge f_c' = ([M] \vee$ [H] \vee [VH]) $\bigwedge f_e = ([\nabla L] \ \vee \ [L]) \bigwedge P_{\text{norm}}/$ $A_{\text{norm}} = [VL]$ | $P_f = [H]$ |
| 14. $L = ([L] \ v [M]) \wedge A_f E_f \neq [VL] \wedge f_c' = ([M] \ v$ [H] \vee [VH]) $\bigwedge f_e = ([M] \vee [H] \vee [VH]) \bigwedge$ $P_{\text{norm}}/A_{\text{norm}} = [\text{VL}]$ | $P_f = [AL]$ |
| 15. $L = ([L] \vee [M]) \wedge A_f E_f \neq [VL] \wedge f_c' = ([M] \vee$ [H] v [VH]) \wedge $P_{\text{norm}}/A_{\text{norm}} \neq$ [VL] | $P_f = [AL]$ |
| 16. $L = ([H] \ v [VH]) \wedge A_f E_f \neq [VL] \wedge f_c' = [M]$ | $P_f =$ [VH] |
| 17. $L = ([H] \vee [VH]) \wedge A_i E_f \neq [VL] \wedge f_c' = ([H])$ v [VH]) | $P_f = [H]$ |

[[]L], [M], [H], [VL], [VH], [AL], and [AH] stand for low, medium, high, very low, very high, almost low, and almost high, respectively (see Fig. [4\)](#page-7-0)

Appendix 2: Closed-form relations for the proposed Sugeno model (with three rules) for the pullout capacity of NSM FRP

$$
w_1 = \exp\left\{\frac{-4}{0.9935^2}\left[\left(\frac{L-250}{508-30}\right)^2 + \left(\frac{A_fE_f - 5740.3}{16966-1300}\right)^2 + \left(\frac{f_c' - 28.5}{64.8-15}\right)^2\right.\newline + \left(\frac{f_c - 35.2}{90.7-6}\right)^2 + \left(\frac{A_{\text{norm}}}{3\text{mm}} - 1.5557\right)^2 + \left(\frac{B_e}{6.89-1}\right)^2\right]\right\}
$$
\n
$$
w_2 = \exp\left\{\frac{-4}{0.9935^2}\left[\left(\frac{L-200}{508-30}\right)^2 + \left(\frac{A_fE_f - 2114.5}{16966-1300}\right)^2 + \left(\frac{f_c' - 41.8}{64.8-15}\right)^2\right.\newline + \left(\frac{f_c - 16}{90.7-6}\right)^2 + \left(\frac{P_{\text{norm}}}{5.45-0.9}\right)^2 + \left(\frac{P_{\text{norm}}}{6.89-1}\right)^2\right]\right\}
$$
\n
$$
w_3 = \exp\left\{\frac{-4}{0.9935^2}\left[\left(\frac{L-36}{508-30}\right)^2 + \left(\frac{A_fE_f - 14017}{16966-1300}\right)^2 + \left(\frac{f_c' - 30.8}{64.8-15}\right)^2\right.\newline + \left(\frac{f_c - 30}{90.7-6}\right)^2 + \left(\frac{P_{\text{norm}}}{5.45-0.9}\right)^2 + \left(\frac{A_fE_f - 14017}{16966-1300}\right)^2 + \left(\frac{f_c' - 30.8}{64.8-15}\right)^2\right.\newline + \left(\frac{f_c - 30}{90.7-6}\right)^2 + \left(\frac{P_{\text{norm}}}{5.45-0.9}\right)^2 + \left(\frac{B_e - 1}{6.89-1}\right)^2\right]\right\}
$$
\n
$$
P_{f1} = -0.0188L + 0.0030A_fE_f + 0.6019f_c' - 0.3151f_e
$$

$$
-15.9532 \frac{P_{\text{norm}}}{A_{\text{norm}}} + 19.1137 \frac{D_g}{W_g} + 42.5840
$$

$$
P_{f2} = 0.0757L + 0.0209A_f E_f + 0.3489f_c - 0.1097f_e
$$

$$
+ 17.2851 \frac{P_{\text{norm}}}{A_{\text{norm}}} + 0.9352 \frac{D_g}{W_g} - 139.2611
$$

$$
P_{f3} = 0.1041L + 0.0030A_f E_f + 0.1635f_c' + 0.1917f_e
$$

$$
- 5.6615 \frac{P_{\text{norm}}}{A_{\text{norm}}} + 2.6405 \frac{D_g}{W_g} - 53.1304
$$

$$
P_f = \frac{w_1 P_{f1} + w_2 P_{f2} + w_3 P_{f3}}{w_1 + w_2 + w_3}
$$

Example An experimental data point with $L = 200$ mm, $A_f E_f = 2053 \text{ kN}, \quad f_c' = 52.8 \text{ MPa}, \quad f_e = 16 \text{ MPa}, \quad P_{\text{norm}}/$ $A_{\text{norm}} = 5.28$, and $D_g/W_g = 3.77$ is considered. The details of calculations are as follows: $P_{f1} = 59.5475 \text{ kN}, P_{f2}$ $= 30.2443$ kN, $P_{f3} = -34.3894$ kN, $w_1 = 0.0064$, $w_2 = 0.8182$, and $w_3 = 0.0003$. Thereby, the proposed Sugeno fuzzy model yields $P_f = 30.54$ kN, which agrees with the experimental value of 30.45 kN (see Fig. [7\)](#page-10-0).

Appendix 3: The detailed information on experimental data

This appendix includes a table listing the geometrical/ mechanical properties of FRP bars/strips, groove, epoxy, and concrete as well as other testing conditions such as the experimental pullout capacity and bonded length. The table is as follows:

 $D_{\rm b}$ is the diameter of FRP bars, whereas t_f and w_f are the dimensions of FRP strips

References

- 1. Chen JF, Teng JG (2001) Anchorage strength models for FRP and steel plates bonded to concrete. J Struct Eng 127(7):784–791
- 2. Kara IF, Ashour AF, Köroğlu MA (2016) Flexural performance of reinforced concrete beams strengthened with prestressed nearsurface-mounted FRP reinforcements. Compos Part B Eng 91:371–383
- 3. Kalupahana WKKG, Ibell TJ, Darby AP (2013) Bond characteristics of near surface mounted CFRP bars. Constr Build Mater 43:58–68
- 4. Sharaky IA, Torres L, Baena M, Vilanova I (2013) Effect of different material and construction details on the bond behaviour of NSM FRP bars in concrete. Constr Build Mater 38:890–902
- 5. Bilotta A, Ceroni F, Di Ludovico M, Nigro E, Pecce M, Manfredi G (2011) Bond efficiency of EBR and NSM FRP systems for strengthening concrete members. J Compos Constr 15(5):757–772
- 6. Seo SY, Feo L, Hui D (2013) Bond strength of near surfacemounted FRP plate for retrofit of concrete structures. Compos Struct 95:719–727
- 7. Seracino R, Raizal Saifulnaz MR, Oehlers DJ (2007) Generic debonding resistance of EB and NSM plate-to-concrete joints. J Compos Constr 11(1):62–70
- 8. Zhang SS, Teng JG, Yu T (2013) Bond strength model for CFRP strips near-surface mounted to concrete. J Compos Constr. [https://](https://doi.org/10.1061/(ASCE)CC.1943-5614.0000402) [doi.org/10.1061/\(ASCE\)CC.1943-5614.0000402](https://doi.org/10.1061/(ASCE)CC.1943-5614.0000402)
- 9. Guillaume S (2001) Designing fuzzy inference systems from data: an interpretability-oriented review. IEEE Trans Fuzzy Syst 9(3):426–443
- 10. Mikut R, Jäkel J, Gröll L (2005) Interpretability issues in databased learning of fuzzy systems. Fuzzy Set Syst 150(2):179–197
- 11. Dilmaç H, Demir F (2013) Stress-strain modeling of highstrength concrete by the adaptive network-based fuzzy inference system (ANFIS) approach. Neural Comput Appl 23(1):385–390
- 12. Cevik A, Ozturk S (2009) Neuro-fuzzy model for shear strength of reinforced concrete beams without web reinforcement. Civ Eng Environ Syst 26(3):263–277
- 13. Cevik A (2011) Modeling strength enhancement of FRP confined concrete cylinders using soft computing. Expert Syst Appl 38(5):5662–5673
- 14. Nasrollahzadeh K, Basiri MM (2014) Prediction of shear strength of FRP reinforced concrete beams using fuzzy inference system. Expert Syst Appl 41(4):1006–1020
- 15. Nasrollahzadeh K, Nouhi E (2018) Fuzzy inference system to formulate compressive strength and ultimate strain of square concrete columns wrapped with fiber-reinforced polymer. Neural Comput Appl 30(1):69–86
- 16. Subaşı S, Beycioğlu A, Sancak E, Şahin İ (2013) Rule-based Mamdani type fuzzy logic model for the prediction of compressive strength of silica fume included concrete using non-destructive test results. Neural Comput Appl 22(6):1133–1139
- 17. Doran B, Yetilmezsoy K, Murtazaoglu S (2015) Application of fuzzy logic approach in predicting the lateral confinement coefficient for RC columns wrapped with CFRP. J Eng Struct 88(1):74–91
- 18. Hoque N, Jumaat MZ, Shukri AA (2017) Critical curtailment location of EBR FRP bonded RC beams using dimensional analysis and fuzzy logic expert system. Compos Struct 166:87–95
- 19. ud Darain KM, Jumaat MZ, Hossain MA, Hosen MA, Obaydullah M, Huda MN, Hossain I (2015) Automated serviceability prediction of NSM strengthened structure using a fuzzy logic expert system. Expert Syst Appl 42(1):376–389
- 20. ud Darain KM, Shamshirband S, Jumaat MZ, Obaydullah M (2015) Adaptive neuro fuzzy prediction of deflection and

cracking behavior of NSM strengthened RC beams. Constr Build Mater 98:276–285

- 21. Sharaky IA, Torres L, Baena M, Miàs C (2013) An experimental study of different factors affecting the bond of NSM FRP bars in concrete. Compos Struct 99:350–365
- 22. De Lorenzis L, Rizzo A, La Tegola A (2002) A modified pull-out test for bond of near-surface mounted FRP rods in concrete. Compos Part B Eng 33(8):589–603
- 23. De Lorenzis L, Nanni A (2002) Bond between near-surface mounted fiber-reinforced polymer rods and concrete in structural strengthening. ACI Struct J 99(2):123–132
- 24. Lee D, Cheng L, Yan-Gee Hui J (2012) Bond characteristics of various NSM FRP reinforcements in concrete. J Compos Constr 17(1):117–129
- 25. Ceroni F, Pecce M, Bilotta A, Nigro E (2012) Bond behavior of FRP NSM systems in concrete elements. Compos Part B Eng 43(2):99–109
- 26. Bilotta A, Ceroni F, Nigro E, Pecce M (2014) Strain assessment for the design of NSM FRP systems for the strengthening of RC members. Constr Build Mater 69:143–158
- 27. Seracino R, Jones NM, Ali MS, Page MW, Oehlers DJ (2007) Bond strength of near-surface mounted FRP strip-to-concrete joints. J Compos Constr 11(4):401–409
- 28. Rashid R, Oehlers DJ, Seracino R (2008) IC debonding of FRP NSM and EB retrofitted concrete: plate and cover interaction tests. J Compos Constr 12(2):160–167
- 29. Oehlers DJ, Haskett M, Wu C, Seracino R (2008) Embedding NSM FRP plates for improved IC debonding resistance. J Compos Constr 12(6):635–642
- 30. Soliman SM, El-Salakawy E, Benmokrane B (2010) Bond performance of near-surface-mounted FRP bars. J Compos Constr 15(1):103–111
- 31. Novidis DG, Pantazopoulou SJ (2008) Bond tests of short NSM-FRP and steel bar anchorages. J Compos Constr 12(3):323–333
- 32. Shield C, French C, Milde E (2005) The effect of adhesive type on the bond of NSM tape to concrete. In: ACI SP230: 7th international symposium on fiber-reinforced polymer (FRP) reinforcement for concrete structures
- 33. Lee D, Hui J, Cheng L (2012) Bond characteristics of NSM reinforcement in concrete due to adhesive type and surface configuration. In: 6th International conference on FRP composites in civil engineering (CICE2012). Rome, Italy
- 34. Palmieri A, Matthys S, Taerwe L (2012) Double bond shear tests on NSM FRP strengthened members. In: 6th international conference on FRP composites in civil engineering (CICE2012)
- 35. Novidis D, Pantazopoulou SJ, Tentolouris E (2007) Experimental study of bond of NSM-FRP reinforcement. Constr Build Mater 21(8):1760–1770
- 36. Palmieri A, Matthys S, Barros JA, Costa I, Bilotta A, Nigro E, Ceroni F, Szambo Z, Balazs G (2012) Bond of NSM FRP strengthened concrete: round robin test initiative. In: 6th international conference on FRP composites in civil engineering (CICE2012)
- 37. De Lorenzis L, Lundgren K, Rizzo A (2004) Anchorage length of near-surface mounted fiber-reinforced polymer bars for concrete strengthening-experimental investigation and numerical modeling. ACI Struct J 101(2):269–278
- 38. El-Gamal S, Al-Salloum Y, Alsayed S, Aqel M (2012) Performance of near surface mounted glass fiber reinforced polymer bars in concrete. J Reinf Plast Compos 31(22):1501–1515
- 39. Peng H, Hao H, Zhang J, Liu Y, Cai CS (2015) Experimental investigation of the bond behavior of the interface between nearsurface-mounted CFRP strips and concrete. Constr Build Mater 96:11–19
- 40. Zadeh LA (1965) Fuzzy sets. Inform Control 8(3):338–353
- 41. Takagi T, Sugeno M (1985) Fuzzy identification of systems and its applications to modeling and control. IEEE Trans Syst Man Cybern 1:116–132
- 42. Mamdani EH, Assilian S (1975) An experiment in linguistic synthesis with a fuzzy logic controller. Int J Man Mach Stud 7(1):1–13
- 43. Yuan Y, Shaw MJ (1995) Induction of fuzzy decision trees. Fuzzy Set Syst 69(2):125–139
- 44. Quinlan JR (1986) Induction of decision trees. Mach Learn 1(1):81–106
- 45. Brieman L, Friedman JH, Olshen R, Stone C (1984) Classification and regression trees. Wadsworth, Belmont
- 46. Chiu SL (1994) Fuzzy model identification based on cluster estimation. J Intell Fuzzy Syst 2(3):267–278