Predicting Bond Strength between FRP Plates and Concrete Substrate: Applications of GMDH and MNLR Approaches

Seyed Mahmood Hamze-Ziabari, Amir Yasavoli


**Related Papers** Click to Download full PDF!

**Experimental Study on Bond-Strengthening Hooks for RC Beams**
Yuichi Sato, Katsuhiro Nagatomo, Yoshifumi Nakamura

*Journal of Advanced Concrete Technology*, volume 7 (2009), pp. 263-271

**A Design Proposal for Concrete Cover Separation in Beams Strengthened by Various Externally Bonded Tension Reinforcements**
Dawei Zhang, Tamon Ueda, Hitoshi Furuuchi


**Investigation of the Bond Behaviour between PBO-FRCM Strengthening Material and Concrete**
Chanh Thai Minh Tran, Boonchai Stitmannaithum, Tamon Ueda

*Journal of Advanced Concrete Technology*, volume 12 (2014), pp. 545-557

Click to Submit your Papers

Japan Concrete Institute  [http://www.j-act.org]
Predicting Bond Strength between FRP Plates and Concrete Substrate: Applications of GMDH and MNLR Approaches

Seyed Mahmood Hamze-Ziabari and Amir Yasavoli

Abstract

Debonding of the fiber-reinforced polymer (FRP) reinforcement due to shear stresses is a very significant issue in design of concrete structures. Several experimental and theoretical investigations have been carried out to produce a relationship between the shear bond strength and the governing variables. However, existing empirical models do not provide an accurate prediction due to the complexity of the debonding process. In the present study, group method of data handling (GMDH) network as a novel machine learning approach was employed to predict the externally bond strength between FRP composites and concrete structures. The GMDH model was developed based on a reliable database including 342 experimental tests obtained from literature. The GMDH results were compared to the most common existing equations and also to the regression approaches developed in this study through statistical error parameters. Furthermore, some correction factors for four well-known equations were suggested based on regression approaches to improve their accuracy. Results indicated that the developed GMDH model outperformed the existing equations and also the developed regression-based equations in terms of both accuracy and safety aspects. Finally, parametric and sensitivity analyses were performed for further verification of the developed GMDH model in capturing the underlying physical behaviors of bond strength.

1. Introduction

During the last few decades, external bonding of fiber reinforced polymer (FRP) plates or sheets have been employed as an excellent option for strengthening of reinforced concrete (RC) structures (Teng et al. 2003; Bakis et al. 2002). Several advantages of FRP materials such as high strength and resistance to corrosion, high durability, non-magnetic and having higher strength to weight ratio which leads to reduce the self-weight of RC structure, and high fatigue resistance make them as a viable alternative materials for retrofitting seismically deficient structures and also structures suffering from corrosion related problems. The externally bonded FRP plates can also be used to improve confinement in compression members and increase the moment capacity of flexural members. The mechanical performance of the interface between FRP and concrete is often affecting on the behavior of the RC structures. The bond interface is often one of the weakest links in RC structures, and debonding at the interface between FRP and concrete substrate is one of the critical failure modes in these structures. Therefore, investigating the bond behavior of the externally bonded FRP is a very sensitive issue in design of RC structures, and it has widely been scrutinized in recent years.

In general, the bond characteristics of externally bonded FRP have been investigated through three methods: (a) direct measurements by using experimental studies such as pull-off tests or the single-lap shear tests (e.g. Chajes et al. 1996; Maeda et al. 1997; De Lorenzis et al. 2001; Nakaba et al. 2001; Wu et al. 2001; Sharma et al. 2006; Ferracuti et al. 2007; Kolluru et al. 2007; Mazzotti et al. 2008; Yun et al. 2008; Carloni and Subramaniam 2010; Malena et al. 2014); (b) numerical methods (e.g. Lu et al. 2005; Carrara et al. 2013); and (c) hybrid model extended by applying experimental test results and analytical solutions (e.g. Dai et al. 2005,2006). The first method is the most popular and widely studied by many researchers. In pull-off test, the slip and local bond stress are directly computed by varying the strain. The strain is measured by closely spaced strain gauges located on top of the FRP laminate (Wu et al. 2002; Yuan et al. 2004; Dai et al. 2005,2006), as shown in Fig. 1. The stress state of the interface is similar to that in a pull-off test in various debonding failure modes. As a result, many studies, both theoretical and experimental have been performed on pull tests of FRP bonded to concrete substrate (Chen and Teng 2001; Yuan et al. 2004). Through the test, the maximum transferable load ($P_{\text{max}}$) and the strain profiles along the bonded length are usually assessed.

Understanding of the behavior of FRP-to-concrete interface is required to be extended for the economic and safe design of externally bonded FRP systems. Investigations indicated that the main failure mode of FRP-to-concrete joints in shear tests is cracking of concrete under shear, occurring commonly at a few millimeters...
from the adhesive-concrete interface (Chen and Teng 2001). A reliable local bond-slip model for the interface and maximum transferable load is very significant issue for accurately modeling and understanding of debonding failures in FRP-strengthened RC structures.

Various design equations have been developed to predict the maximum transferable load based on empirical analysis of experimental data obtained from pull-off test (e.g. Van Gemert 1980; Tanaka 1996; Yoshizawa and Wu 1997; Maeda et al. 1997; Neubauer and Rostasy 1997; Khalifa et al. 1998; Adhikary and Mutsuyoshi 2001; Chen and Teng 2001; De Lorenzis et al. 2001; Yang et al. 2001; JCI 2003; Dai et al. 2005; Lu et al. 2005; Camli and Binici 2007). Many of these equations used to estimate the bond strength between FRP and concrete substrates are highly empirical, and their predictive abilities are limited by the corresponding data sets from which they were derived, and do not provide a reliable prediction of maximum transferable load (Yuan et al. 2012; Diab and Farghal 2014; D’Antino and Pellegrino 2014). Furthermore, the bond mechanism is complex because it depends on different parameters including the concrete strength, the bond length, the FRP plate axial stiffness, the FRP-to-concrete width ratio, the adhesive stiffness, and the adhesive strength (Lu et al. 2005). The mentioned formulae do not completely consider the nonlinearity between input and output parameters and do not check different combinations of input parameters for estimating bond strength. Most of the existing prediction models also neglect the adhesive material properties (Diab and Farghal 2014; Ueno et al. 2014).

Recently, machine learning approaches such as Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS) as the most common soft computing methods have been employed to overcome these limitations. Mashrei et al. (2013) investigated the application of Back-Propagation Neural Network (BPNN) model for predicting the bond strength of FRP-to-concrete joints. Mansouri and Kisb (2015) also evaluated the applications of neuro fuzzy and neural network approaches for estimation of debonding strength for masonry elements retrofitted with FRP composites. Results showed that ANN and ANFIS approaches can be successfully used to predict the bond strength. However, these methods do not give enough insight into the generated models and are not as easy to use as the empirical formulas. Among soft computing methods, the GMDH network is known as a self-organized method to model and discover the behaviors of unknown or complicated systems based on given input-output data points (Ivakhnenko 1971; Ivakhnenko and Ivakhnenko 2000). The main objective of this study is to investigate the efficiency of the GMDH network for predicting of bond strength between FRP and concrete substrate. The main advantage of GMDH method in comparison with ANN method is that the dependencies between input parameters and output parameter are represented in parametric form as an equation while these dependencies are hidden within neural network structures in ANN method. Furthermore, the ANN methods need an essential time for learning and, consequently, it is difficult to be applied for modeling and forecasting in real time systems.

In order to generate GMDH models, a comprehensive database from literature containing 342 experimental was collected. The developed GMDH model related the bond strength to the width of FRP plate, elastic modulus of FRP, concrete cylinder compressive strength, the effective thickness of concrete contributing to shear deformation, the width of concrete prism, the bond length, the elastic modulus of adhesive, and the thickness of adhesive layer. The developed GMDH results were compared to the most common existing equations and also to the multiple linear regression (MLR) and multiple nonlinear regression (MNLR) approaches developed in this study through statistical error indicators. Some correction factors for Maeda et al. (1997), Neubauer and Rostasy (1997), JCI (2003), and Lu et al. (2005) equations as the most accurate predictive equations from current dataset were also suggested based on MNLR approach to improve their predictive abilities. Results confirmed that the developed GMDH model outperformed the existing equation and the developed regression-based equations in terms of both accuracy and safety aspects. The relative importance of significant parameters dealing with bond strength is also investigated through sensitivity analysis. The robustness of proposed GMDH model was also verified through a parametric analysis.

The remaining sections of this paper are organized as follows. In Section 2, a background of Shear bond strength between FRP and concrete substrate is presented. GMDH and MNLR algorithms are outlined in Section 3. Fourth section states to the involved input parameters and modeling process. Section 5 evaluates the proposed model. Furthermore, sensitivity and parametric analyses are conducted to evaluate the robustness of GMDH model, and the reliability and safety of the GMDH model and modified equations through demerit point’s classifications are evaluated in the penult of this section. At the end, the paper is concluded in Section 6.

2. Shear bond strength between FRP and concrete substrate

Debonding of the FRP reinforcement due to shear

Fig. 1 Schematic diagram of single pull-off test for FRP to concrete.
stresses is often one of the crucial failure modes of composite elements, which externally bonded with FRP. The externally bond strength between FRP and concrete substrate have been tested and analyzed generally on the basis of a single pull-off test. In the test, the direct tensile force \( P_{\text{app}} \) is applied to the FRP plate bonded to concrete substrate in order to determine the maximum transferable load \( (P_{\text{max}}) \). A sketch of single pull-off test for FRP to concrete is presented in Fig. 1. The shear bond strength is generally estimated as the maximum transferable load by applying a coupled model of a shear stress-tangential slip law and normal stress-displacement law (Yuan et al. 2012).

The maximum transferable load is related to the mechanical characteristics of the materials and the interfacial fracture energy. Generally, \( P_{\text{max}} \) for FRP externally bonded reinforcement on brittle substrate is defined as:

\[
P_{\text{max}} = b_f \int_0^\infty \tau(x)dx
\]

where \( b_f \) is the width of FRP plate, \( x \) is the longitudinal axis, and \( \tau(x) \) is the bond shear stress distribution along the interface. The interfacial fracture energy is defined as the work done by the interfacial shear stress until the complete separation of the FRP plate from concrete substrate. According to many studies in literature, the maximum transferable load is directly proportional to the square root of the interfacial fracture energy \( G_f \), regardless of the shape of the bond-slip curve (Ferracuti et al. 2007; Kashyap et al. 2012). The interfacial fracture energy is defined as:

\[
G_f = \int_0^\infty \tau(s)ds
\]

where \( s \) represents the FRP local slip. The general expression for \( P_{\text{max}} \) (debonding load) is derived as a function of \( G_f \) and material properties as follows:

\[
P_{\text{max}} = b_f \sqrt{2E_f t_f G_f}
\]

where \( E_f \) and \( t_f \) are elastic modulus and thickness of the FRP, respectively. The fracture energy can be estimated as (Ferracuti et al. 2007):

\[
G_f = k_s k_c \sqrt{f_s f_u}
\]

in which \( k_s \) is a width factor coefficient for considering the effect of reinforcement width on \( P_{\text{max}} \), \( k_c \) is a calibration parameter, which is determined through experimental results of bond tests, \( f_s \) is the compressive strength and \( f_u \) is the tensile strength of the concrete substrate.

### 2.1 Existing bond strength models

Since the 1980s many researchers have developed various equations for predicting bond strength between FRP and concrete substrate, including van Gemert (1980), Tanaka (1996), Yoshizawa and Wu (1997), Maeda et al. (1997), Neubauer and Rostasy (1997), Khalifa et al. (1998), Adhikary and Mutusuyoshi (2001), Chen and Teng (2001), De Lorenzis et al. (2001), Yang et al. 2001, JCI (2003), Dai et al. (2005), Lu et al. (2005), and Camli and Binici (2007). Furthermore, the international federation for structural concrete (fib) (2001) and CNR-DT200/2004 (2004) design guidelines also present different formulae for estimating the bond strength of FRP-to-concrete elements. These equations are presented in Table 1. These relations are mainly based on either fracture mechanics based theory or on derived empirical equation calibrated with experimental datasets or combinations of the two. Most of these equations neglect the influence of the adhesive material properties. The accuracy of these empirical and combined theoretical and empirical models seems to be limited. These formulae were developed mainly based on regression analysis on experimental datasets and do not completely consider the nonlinearity between input and output parameters. Recently, the machine learning approaches have been widely applied in engineering problems. The main advantages of these approaches can be mentioned as: they consider all possible relations between input and output parameters and also unlike the regression approaches, check different combinations of input parameters for estimating output parameter.

### 3. Methodology

The data mining approaches such as Artificial Neural Networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS) have been employed as useful tools for modeling and forecasting complex engineering problems. Among these methods, the group method of data handling (GMDH) network has not yet been widely applied in structural engineering. The main advantage of GMDH method in comparison with ANN method is that the dependencies between input parameters and responses are represented in a parametric form as an equation while these dependencies are hidden within neural network structures in ANN method. Besides that ANN methods need an essential time of learning and therefore it is difficult to be applied for modeling and forecasting in real time systems. The description of the GMDH network and multiple nonlinear regression methods (MNLR) are discussed in below.

#### 3.1 Principle of GMDH network

GMDH network is known as a self-organized machine learning approach that models the nonlinear relationship between input and response variables based on the polynomial theory of complex systems (Ivakhnenko 1971). It is a self-organized neural network because the most important input parameters, number of layers, number of neurons in the middle layers, and also the optimal topology design of the network are automatically defined by the algorithm. The relationship between any sets of input–response variables in the GMDH algorithm can be estimated by Volterra-Kolmogorov-Gabor (VKG) polynomial (Ivakhnenko 1971):
Table 1: Existing equations for predicting maximum transferable load.

<table>
<thead>
<tr>
<th>Model</th>
<th>Equation</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen and Teng</td>
<td>$P_{\text{max}} = 0.315B_yB_p \sqrt{f_{\text{t},p} b_p L_y}$</td>
<td></td>
</tr>
<tr>
<td>De Lorenzis et al.</td>
<td>$P_{\text{max}} = b_p \sqrt{\frac{2E_p t_p G_f}{f}}$</td>
<td>$G_f$: the fracture energy per unit area of the joint, assumed equal to 1.43 $\text{Nm} / \text{mm}^2$.</td>
</tr>
<tr>
<td>TYong et al.</td>
<td>$P_{\text{max}} = \left(0.5 + 0.08 \sqrt{\frac{E_p t_p}{1000}}\right) b_p L_y 0.5 f_{\text{cm}}$</td>
<td></td>
</tr>
<tr>
<td>Izumo (in JCI 2003)</td>
<td>$P_{\text{max}} = (3.8 f_{\text{ck}}^{2/3} + 15.2) L_y b_p E_p t_p \times 10^{-3}$</td>
<td></td>
</tr>
<tr>
<td>Iso (in JCI 2003)</td>
<td>$P_{\text{max}} = b_p L_y 0.93 f_{\text{ck}}^{0.44}$</td>
<td>$L_y = 0.125 (E_p t_p)^{0.57}$</td>
</tr>
<tr>
<td>Sato (in JCI 2003)</td>
<td>$P_{\text{max}} = (b_p + 7.4) L_y 2.68 f_{\text{ck}}^{0.5} E_p t_p \times 10^{-3}$</td>
<td>$L_y = 1.89 (E_p t_p)^{0.4}$</td>
</tr>
<tr>
<td>Van Gemert</td>
<td>$P_{\text{max}} = 0.5 b_p L_y 0.5 f_{\text{cm}}$</td>
<td></td>
</tr>
<tr>
<td>Tanaka</td>
<td>$P_{\text{max}} = (6.13 - L y L_y) b_p L_y$</td>
<td></td>
</tr>
<tr>
<td>Yoshizawa and Wu</td>
<td>$P_{\text{max}} = b_p L_y (5.88 L_y^{0.6})$</td>
<td></td>
</tr>
<tr>
<td>Maeda et al.</td>
<td>$P_{\text{max}} = 110.2 \times 10^{-9} E_p t_p b_p L_y$</td>
<td>$L_y = e^{0.13 - 0.58 \ln L_y}$</td>
</tr>
<tr>
<td>Neubauer and Rostasy</td>
<td>$P_{\text{max}} = 0.64 K_p b_p \sqrt{f_{\text{cm}} E_p t_p}$ $L_y$ if $L_y \geq L_e$</td>
<td>$K_p = \sqrt{\frac{2 - b_p}{b_p}}$ $L_y &lt; L_e$</td>
</tr>
<tr>
<td></td>
<td>$P_{\text{max}} = 0.64 K_p b_p \sqrt{f_{\text{cm}} E_p t_p} L_y$ $\left(2 - L_y / L_e\right)$ if $L_y &lt; L_e$</td>
<td></td>
</tr>
<tr>
<td>Khalifa et al.</td>
<td>$P_{\text{max}} = 110.2 \times 10^{-6} \left(\frac{f_{\text{ck}}}{42}\right)^{2/3} E_p t_p b_p L_y$</td>
<td>$L_y = e^{0.13 - 0.58 \ln L_y}$</td>
</tr>
<tr>
<td>Adhikary and Matsuyoshi</td>
<td>$P_{\text{max}} = b_p L_y \left(0.25 f_{\text{ck}}^{2/3}\right)$</td>
<td></td>
</tr>
<tr>
<td>Dai et al.</td>
<td>$P_{\text{max}} = (b_p + 7.4) \sqrt{2E_p t_p G_f}$</td>
<td>$G_f = 0.514 f_{\text{ck}}^{0.236}$</td>
</tr>
<tr>
<td>Lu et al.</td>
<td>$P_{\text{max}} = B_y b_p \sqrt{2E_p t_p G_f}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$L_y = a + \frac{1}{2\lambda_e} \ln \left(\lambda_1 + \lambda_2 \tan(\lambda_2 a)\right)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\lambda_1 = \frac{\tau_{\text{max}}}{S_y E_p t_p}$ \hspace{1cm} $\lambda_2 = \frac{\tau_{\text{max}}}{(S_y - S_{\lambda}) E_p t_p}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$a = \frac{1}{\lambda_2} \arcsin \left[\frac{0.99}{S_y}\sqrt{\frac{S_y - S_{\lambda}}{S_y}}\right]$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\tau_{\text{max}} = \alpha_3 \beta_{\text{f,y}} f_{\text{y}}$ $\alpha_3 = 1.5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$G_f = 0.308 B_y^{0.5} \sqrt{f_{\text{ck}}}$ $S_y = 0.0195 \beta_{\text{f,y}} f_{\text{y}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$B_y = \sqrt{\frac{2.25}{b_p} + \frac{b_p}{1.25 + \frac{b_p}{b_p}}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$B_y = \sin \left[\frac{\pi L_y}{2L_e}\right]$</td>
<td></td>
</tr>
</tbody>
</table>
\[
P_{\text{max}} = \sqrt{\tau_f \sigma_f} \sqrt{E_p f_p} b_p \tanh \left( \frac{\theta L_o}{L_o} \right)
\]

where \( X = (x_1, x_2, \ldots, x_n) \) is the vector of input variables, \( w \) is the vector of weight coefficients, and \( \hat{y} \) is the predicted output. The general form of VKG can be simplified as a partial quadratic polynomials consisting of only two variables (\( x_i \) and \( x_j \)):

\[
P_{\text{max}} = 0.64 \alpha \kappa K_p b_p \sqrt{f_{\text{con}} E_p f_p} \left( 2 - \frac{L_o}{L_e} \right) \quad \text{if} \quad L_o \geq L_e
\]

\[
P_{\text{max}} = 0.64 \alpha \kappa K_p b_p \sqrt{f_{\text{con}} E_p f_p} \left( \frac{L_o}{L_e} \right) \quad \text{if} \quad L_o < L_e
\]

\[
K_\theta = 0.03
\]

\[
K_i = \begin{cases} 
1 & \text{if} \quad L \geq L_e \\
\frac{L_o}{L_e} \left( 2 - \frac{L_o}{L_e} \right) & \text{if} \quad L < L_e
\end{cases}
\]

\[
L_e = \sqrt{\frac{E_p f_p}{2f_c}}
\]

\[
K_k = \begin{cases} 
\frac{2 - 0.33}{1 + \frac{b_p}{b_c} / 400} & \text{if} \quad b_p / b_c < 0.33 \\
\frac{2 - 0.33}{1 + \frac{b_p}{b_c}} & \text{if} \quad b_p / b_c \geq 0.33
\end{cases}
\]

\[
\hat{y} = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j + \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} w_{ijk} x_i x_j x_k + \ldots
\]

\[
\hat{y} = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 x_1^2 + w_5 x_2^2
\]

It should be noted that there are different types of polynomial forms such as bilinear, quadratic, tri-quadratic, and third order which can be employed to design the GMDH system. The application of tri-quadratic and third order can lead to a remarkably more complicated network in comparison with quadratic polynomial. On the other hand, using bilinear form may lead to a simpler model in comparison with quadratic; however, its accuracy can be very limited. Therefore, in this study, the quadratic form is applied.

The structure of the GMDH network is configured thorough the training stage with polynomial model which produces the minimum error between the predicted values (\( \hat{y} \)) and observed ones (\( y \)). The weighting coefficients of the VKG function are calculated in a way that the output obtained for the whole set of input-output data pairs optimally fit to observed ones based on the following error criterion:

where \( y_i \) is the actual output and \( \hat{y}_i \) is the estimated one based on VKG. To obtain the best fit, the value \( E \) should be minimized, in other words, the partial derivatives of Eq. (7) with respect to each constant \( w_i \) are taken and set equal to zero:

\[
\frac{\partial E}{\partial w} = 0
\]

Solving Eq. (8) leads to a system of equations that are solved by training set of data:

\[
Y = \begin{bmatrix} 1 & x_1 & x_2 \end{bmatrix}
\]

\[
X = Y^T Y
\]

\[
w = \begin{bmatrix} w_0 & w_1 & w_2 & w_3 & w_4 & w_5 \end{bmatrix}
\]
Following equation can be obtained as:

\[ b = (yY)^T \]  

(13)

Therefore, the following system of equation must be solved:

\[ \sum_{i=1}^{n} wX = \sum_{i=1}^{n} b \]  

(14)

In this study, the GMDH network is improved using a back propagation algorithm. This method includes two main steps: (1) the weighting coefficients of the quadratic polynomial are determined using the least squares method from the input layer to output layer in the form of a forward path; and (2) the weighting coefficients were updated using a back-propagation algorithm in a backward path. This procedure may be continued until the error of the training network \( E \) is minimized.

### 3.2 Multiple nonlinear regression (MNLR)

Multiple non-linear regression (MNLR) approach is one of the classic statistical analysis that finds a suitable nonlinear relationship between input and response variables. To explain this method, let assume that \( y \) is a dependent variable and have a nonlinear relation with \( n \) independent variables as \( x_1, x_2, ..., x_n \). The nonlinear relation between them can be expressed as:

\[ y = a_0 x_1^a x_2^b ... x_n^c \]  

(15)

By applying the logarithmic transformation, the following equation can be obtained as:

\[ \log y = \log a_0 + a_1 \log x_1 + a_2 \log x_2 + ... + a_n \log x_n \]  

(16)

where the coefficients \( a_0, a_1, ..., a_n \) can be determined by applying least squares method in similar to multi linear regression (MLR) method.

### 4. Data set and modeling

#### 4.1 Influential parameters used in simulation

Several parameters can effect on the bond strength between FRP and concrete substrate based on many theoretical and experimental studies. For example according to Lu et al. (2005), the following six parameters are known to be the most effective parameters on bond strength between FRP and concrete substrate: (a) the concrete strength (b) the bond length; \( E_b \) is elastic modulus of adhesive; \( t_p \) and \( t_a \) are the thicknesses of the FRP sheets and adhesive layer, respectively.

### 4.2 Data collection

To drive a reliable model for predicting the externally bond strength between fiber-reinforced polymer (FRP) composites and concrete structures, a comprehensive database of laboratory testing from Chajes et al. (1996), Maeda et al. (1997), Tåljenst (1997), Ueda et al. (1999), Zhao et al. (2000), Adhikary and Matsuyoshi (2001), Fu-quan et al. (2001), Nakaba et al. (2001), Wu et al. (2002), Tan (2002), Ren (2003), Dai et al. (2005), Yao et al. (2005), Sharma et al. (2006), Toutanji et al. (2007) and Ueno et al. (2014), including 342 sub assemblages are collected from published literature as presented in Table 2. It should be noted that the adhesive material properties for some datasets were not provided. Therefore, the following assumptions according to studies of Ueno et al. (2014) and Kaw (2006) were considered to add these properties in analyzing process: (i) the poisson’s ratio is assumed to be 0.4 (ii) the thickness and Young modulus of adhesive material are assumed as 1 mm and 3.792 MPa.

### 4.3 Model development using GMDH network

To develop new models for predicting shear bond strength between FRP sheets and concrete substrate using GMDH algorithm, the available database were randomly divided into training and testing subsets. The training data were taken for the learning of the algorithm. The testing dataset was used to specify the generalization capability of the model to new data they did not train with. In other words, the testing data were applied to measure the performance of the models obtained by GMDH algorithm when applied to dataset which played no role in building the models. Of the 342 tests, 273 data vectors (80%) were taken for the training process. The remaining 69 data vectors (20%) were used for the testing of the models.

Following data division, training dataset was presented to the GMDH for model training. The GMDH returned the following selective polynomials for prediction of shear bond strength between FRP and concrete substrate as:

**Layer #1**

\[ L_s = 1.05 - 1.08 \times 10^{-5} E_b + 0.0714 b_p - 0.607 t_p + 8.07 \times 10^{-7} b_p E_b + 7.44 \times 10^{-5} t_p E_b + 0.186 t_p b_p + 3.37 \times 10^{-11} E_b^2 - 0.000669 b_p^2 - 2.62 t_p \]  

(18-a)

**Layer #2**

\[ L_s = 15.4 + 0.757 L_c - 0.21 b_p - 0.0912 b_p - 0.00248 b_p L_c + 0.0027 b_p L_c + 0.0014 b_p b_p + 0.00135 L_c^2 - 3.03 \times 10^{-5} b_p^2 - 0.000121 b_p^2 \]  

(18-b)
Layer #3

\[ L_s = -14.9 + 0.539L_a + 0.047L_b + 0.194t_\epsilon + 0.0018L_aL_b^2 + 0.000595L_a^2 - 0.000317t_\epsilon L_a - 0.000126L_b^2 - 4.35 \times 10^{-6}L_b^2 - 0.000536t_\epsilon^{0.5} \]

(18-c)

Layer #4

\[ P_{\text{max}}(kN) = -2.9 + 1.1L_s + 1.89 \times 10^{-5}E_p + 0.0281f'_{\epsilon} + 2.91 \times 10^{-7}E_pL_s - 0.00732f'_{\epsilon}L_s + 2.2 \times 10^{-7}f'_{\epsilon}E_p^2 + 0.0092L_a^2 - 7.38 \times 10^{-11}E_p^2 + 0.0016f'_{\epsilon}^2 \]

(18-d)

The GMDH algorithm generated the first network layer by examining each possible pairs of 9 input variables based on 9(9-1)/2 = 36 neurons. According to Eq. (18-a), the layer one consists of two neurons based on three input variables (i.e. \(E_p\), \(t_\epsilon\), and \(L_s\)). For more illustration, the evolved structure of generalized GMDH network for estimation of \(P_{\text{max}}\) is shown in Fig. 2. As shown, the second layer is created using inputs from the first layer (\(L_s\), \(b_n\), and \(b_e\)) and so on. Note that selecting the input parameters and neurons are based on their con-
contributions to the output parameter. According Eq. (18), the first network layer consists of a quadratic function of the input variables, the second layer includes fourth degree polynomial, the third layer involves eighth degree polynomials, and etc. It should be noted that the developed network structure is fully self-determined by the algorithm itself.

To further evaluate the proposed GMDH model against the most common regression approaches, two different equations were developed using multiple linear regression (MLR) and MNLR as:

**MLR:**

\[
P_{\text{max}}(kN) = 0.0013E_{\text{c}} + 0.0612b_2 - 0.0729t_3 + 14.84t_5 + 0.17b_6 + 0.01L_8 - 17.21\]  

(19)

To develop the MNLR model, first the input and output parameters are transformed to logarithmic space as:

\[
\log P_{\text{max}} = \log a_0 + a_1 \log E_{\text{c}} + a_2 \log t_3 + a_3 \log b_5 + a_4 \log t_5 + a_5 \log t_4 + a_6 \log L_8 + a_7 \log f_{\text{c}} + a_8 \log E_{\text{c}} + a_9 \log t_5 + a_{10} \log b_5
\]

(20)

Then, the coefficients of \(a_0, a_1, \ldots, a_{10}\) are determined based on the MLR method. After determining the coefficients, the \(P_{\text{max}}\) parameter is converted from logarithmic space to linear space as the following equation:

\[
\log P_{\text{max}} = \log a_0 + a_1 \log E_{\text{c}} + a_2 \log t_3 + a_3 \log b_5 + a_4 \log t_5 + a_5 \log t_4 + a_6 \log L_8 + a_7 \log f_{\text{c}} + a_8 \log E_{\text{c}} + a_9 \log t_5 + a_{10} \log b_5
\]

(21)

In the developed models (Eqs. (18), (19), and (21)), the units of \(E_{\text{c}}\), \(E_{\text{c}}\) and \(f_{\text{c}}\) are MPa and the units of \(t_3, t_5, b_5, b_5, \) and \(L_8\) are mm.

To evaluate the performances of the developed models, the following statistical error parameters were used: BIAS, mean absolute error (MAE), root mean square error (RMSE), correlation coefficient \((R)\) and coefficient of determination \((R^2)\):

\[
\text{BIAS} = \frac{\sum_{i=1}^{N}(P_i - O_i)}{N}
\]

(22)

\[
\text{MAE} = \frac{\sum_{i=1}^{N}|P_i - O_i|}{N}
\]

(23)

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N}(P_i - O_i)^2}
\]

(24)

\[
R = \frac{\sum_{i=1}^{N}(P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N}(P_i - \bar{P})^2} \sqrt{\sum_{i=1}^{N}(O_i - \bar{O})^2}}
\]

(25)

\[
R^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2}
\]

(26)

where \(O_i\) is the measured value, \(P_i\) stands for prediction values; \(N\) is the number of data points, \(O_{\text{m}}\) is the mean value for observation and \(P_{\text{m}}\) is the mean value of prediction.

**5. Results and discussion**

The results of GMDH network and existing traditional equations for prediction of bond strength between FRP and concrete substrate are presented in this section. Moreover, some traditional equations were modified through different correction factors developed by MNLR method. The performances of GMDH and modified equations were evaluated through externally validation criteria. In addition; influences of important parameters on bond strength have been investigated through a parametric analysis. The most important predictive parameters were determined through sensitivity analysis in prediction of bond strength. Finally, reliability analysis was done by using demerit point’s classifications.

**5.1 Comparison of GMDH model and existing equations**

The performance of GMDH model in prediction of the maximum bond strength between FRP sheet and concrete substrate for training and testing databases are presented in Fig. 3, which illustrates the scattering between measured and predicted bond strength around the optimal line of equality. As shown, there is a little scatter around the optimal line between predicted and measured values of maximum bond strength in both training and testing sets. The same performance of the GMDH model on the training and testing database indicates that it can have both good predictive ability and generalization performances for the input ranges used.

The performances of the developed GMDH model are further confirmed through analytical analyses in Table 3, which contains four different performance measures including coefficient of correlation, \(R\), coeffi-

![Fig. 3 Comparison between measured and predicted maximum transferable load \((P_{\text{max}})\) by GMDH model for training and testing dataset.](image-url)
Table 3 Performance of GMDH and regression models in predicting bond strength for training and testing datasets.

<table>
<thead>
<tr>
<th>Subsets</th>
<th>Model</th>
<th>$BIAS$</th>
<th>$RMSE$</th>
<th>$R$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>GMDH</td>
<td>$8.35 \times 10^{-15}$</td>
<td>2.7394</td>
<td>0.9514</td>
<td>0.9052</td>
</tr>
<tr>
<td></td>
<td>MNLR</td>
<td>-0.7167</td>
<td>3.3093</td>
<td>0.9345</td>
<td>0.8616</td>
</tr>
<tr>
<td></td>
<td>MLR</td>
<td>-10.8270</td>
<td>12.5592</td>
<td>0.7137</td>
<td>-0.9936</td>
</tr>
<tr>
<td>Testing</td>
<td>GMDH</td>
<td>0.2079</td>
<td>2.8820</td>
<td>0.9481</td>
<td>0.8983</td>
</tr>
<tr>
<td></td>
<td>MNLR</td>
<td>-0.0881</td>
<td>3.4673</td>
<td>0.9235</td>
<td>0.8527</td>
</tr>
<tr>
<td></td>
<td>MLR</td>
<td>-10.0737</td>
<td>12.3950</td>
<td>0.6480</td>
<td>-0.8819</td>
</tr>
<tr>
<td>Total</td>
<td>GMDH</td>
<td>0.0419</td>
<td>2.7688</td>
<td>0.9507</td>
<td>0.9037</td>
</tr>
<tr>
<td></td>
<td>MNLR</td>
<td>-0.5899</td>
<td>3.3418</td>
<td>0.9313</td>
<td>0.8598</td>
</tr>
<tr>
<td></td>
<td>MLR</td>
<td>-10.6751</td>
<td>12.5262</td>
<td>0.6992</td>
<td>-0.9704</td>
</tr>
</tbody>
</table>

Table 4 Performance of existing equations for prediction of maximum bond strength.

<table>
<thead>
<tr>
<th>Model</th>
<th>$BIAS$</th>
<th>$RMSE$</th>
<th>$R$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaka</td>
<td>-7.45</td>
<td>10.94</td>
<td>0.4395</td>
<td>-0.5043</td>
</tr>
<tr>
<td>Maeda et al.</td>
<td>-0.72</td>
<td>4.83</td>
<td>0.8446</td>
<td>0.7061</td>
</tr>
<tr>
<td>Neubauer and Rostásy</td>
<td>2.08</td>
<td>4.40</td>
<td>0.9057</td>
<td>0.7563</td>
</tr>
<tr>
<td>Khalifa et al.</td>
<td>-1.76</td>
<td>5.86</td>
<td>0.7839</td>
<td>0.5677</td>
</tr>
<tr>
<td>Adhikary and Mutsuyoshi</td>
<td>14.69</td>
<td>28.08</td>
<td>0.5024</td>
<td>-8.90</td>
</tr>
<tr>
<td>Chen and Teng</td>
<td>-5.49</td>
<td>7.30</td>
<td>0.9087</td>
<td>0.3295</td>
</tr>
<tr>
<td>De Lorenzis et al.</td>
<td>6.00</td>
<td>8.39</td>
<td>0.8645</td>
<td>0.1155</td>
</tr>
<tr>
<td>Yang et al.</td>
<td>-1.70</td>
<td>5.54</td>
<td>0.8070</td>
<td>0.6141</td>
</tr>
<tr>
<td>Iso</td>
<td>1.25</td>
<td>5.18</td>
<td>0.8652</td>
<td>0.6625</td>
</tr>
<tr>
<td>Sato</td>
<td>16.65</td>
<td>40.75</td>
<td>0.7763</td>
<td>-19.85</td>
</tr>
<tr>
<td>Dai et al.</td>
<td>7.01</td>
<td>8.97</td>
<td>0.8728</td>
<td>-0.01</td>
</tr>
<tr>
<td>Lu et al.</td>
<td>-2.10</td>
<td>4.41</td>
<td>0.9143</td>
<td>0.7551</td>
</tr>
<tr>
<td>Camli and Binici</td>
<td>2.14</td>
<td>7.15</td>
<td>0.7619</td>
<td>0.3573</td>
</tr>
<tr>
<td>fib model</td>
<td>2.07</td>
<td>4.39</td>
<td>0.9057</td>
<td>0.7570</td>
</tr>
<tr>
<td>CNR-DT200</td>
<td>-4.38</td>
<td>6.42</td>
<td>0.8779</td>
<td>0.4819</td>
</tr>
</tbody>
</table>

cient of determination (or efficiency), $R^2$, root mean squared error, $RMSE$, and $BIAS$. These performance measures were employed to evaluate the performances of the GMDH, MLR and MNLR models. The $R$ parameter shows the correlation between predicted and measured values. Smith (1986) showed that if the $R$ value is more than 0.8, there is strong correlation between measured and predicted values. Nonetheless, $R$ may not necessarily indicate a better performance due to the tendency of the model to deviate toward higher or lower values, particularly when the data range is very wide and most of the data are distributed around their mean. Thus, the coefficient of determination, $R^2$, can be used as a better error indicator because of its unbiased estimation of errors. Furthermore, the $BIAS$ and $RMSE$ parameters should be at the minimum for having precise results. The error indicators of the GMDH, MLR, and MNLR models are presented in Table 3. The developed GMDH increases the accuracy of the optimal MNLR by 16.5% and 5.1% with respect to $RMSE$ and $R^2$ criteria. It should be noted that in spite of MLR, the derived MNLR model in this study can also be successfully used in predicting deboning strength.

Table 4 compares the performance of the developed GMDH model in prediction of bond strength between FRP sheet and concrete substrate with those of the available empirical models and design codes against a set of statistical error indicators. The newly developed GMDH models showed the highest prediction accuracy and lowest prediction errors in respect to all models for predicting maximum bond strength, with $R^2$ of 0.90 and $RMSE$ of 2.77. The best empirical equation is that of Neubauer and Rostásy (1997), which had an $R^2$ of 0.76 and $RMSE$ of 4.40. However, the proposed GMDH model outperforms Neubauer and Rostásy (1997) equation by improving $R^2$ value by 19.4% and $RMSE$ value by 37.2%. Thus, the proposed GMDH model can predict the target values of the maximum bond strength with acceptable accuracy and less error than available empirical models over a wide range of input parameters.

5.2 Developing new ideas for accuracy improvement

Errors of an appropriate model should be independent or less sensitive to the variation of the input parameters involved in the process of the event studied. Otherwise, it can be interpreted that those input parameters either aren’t correctly incorporated or should be included in that model. Figure 4 shows the variation of discrepancy ratio between measured and predicted maximum bond strength ($DR=P_{\text{max,measured}}/P_{\text{max,predicted}}$) as a function of $f_c$, $L_b$, and $E_{\text{pf}}$ parameters, which were directly incorporated in each model, for different equations. To observe the general trend of errors, the linear regression between $DR$ values of each model and input parameters are also shown in Fig. 4. The $DR$ values of Maeda et al. and Lu et al. increase with increasing $f_c$ while $DR$ values of Neubauer and Rostásy (1997) (N&R) and JCI (2003)
decrease. However, according to Fig. 4, the developed GMDH model had a better performance than other models in this aspect. The sensitivity of DR values of Maeda et al. (1997) and Lu et al. (2005) equations to change of bond length is more remarkable than other models and similar trends for $E_{d_p}$ can be observed. N&R and Iso errors are less sensitive to variations of $L_b$ and $E_{d_p}$ than other models. However, the GMDH errors have the least sensitivity to change of $L_b$ and $E_{d_p}$ parameters amongst other empirical equations.

Most of the existing prediction models neglected the adhesive material properties (Ueno et al. (2014)) and the FRP-to-concrete width ratio. Therefore, the variations of DR values for different models as a function of these properties are separately depicted in Fig. 5. As shown, the DR values of empirical models are very sensitive to change of $E_a$ parameter. The similar trend for $(b_p/b_c)$ can be observed with less intensity. Just like the other input parameters, the proposed GMDH model errors are independent of $E_a$ and $b_p/b_c$ parameters. It can be concluded that some effective parameters, which aren’t directly incorporated, should be added to the empirical equations with regarding their meaningful errors tendency to change of these input parameters. Therefore, some correction factors were suggested to improve the accuracy of Maeda et al. (1997), Neubauer and Rostaşy (1997), JCI (2003), and Lu et al. (2005) equations. The correction factor for each model was developed based on MNLR regression approaches by employing the log (input parameters which weren’t considered in that model) as input and log (DR=Measured/predicted) as target parameters:

$$\log(K_{corr}) = \log(\text{Measured/Predicted}) = a_1 + a_2 \log(E_a) + a_3 \log(b_c)$$  \hspace{1cm} (27)

After transformation:

$$K_{corr} = (E_a)^{a_1} b_c^{a_2}$$ \hspace{1cm} (28)

where $a_1$, $a_2$, $a_3$, $a'_1$, $a'_2$, and $a'_3$ are constant and were determined by applying MNLR method. The following correction factors are suggested for mentioned equations as:

$$K_{Iso} = 0.62(E_a)^{0.11}$$ \hspace{1cm} (29)

$$K_{Maeda et al.} = 0.47(E_a)^{0.10} f_c^{0.17}$$ \hspace{1cm} (30)

$$K_{Neubauer and Rostaşy} = 0.94(E_a)^{0.09}$$ \hspace{1cm} (31)

$$K_{Lu et al.} = 0.78(E_a)^{0.10} f_c^{0.13}$$ \hspace{1cm} (32)

Table 5 presents the statistical error parameters of these empirical models after applying proposed correction factors. The Maeda et al. (1997) model performances are slightly improved by employing its suggested correction factor but the improvements for the other three

<table>
<thead>
<tr>
<th>Model</th>
<th>BIAS</th>
<th>RMSE</th>
<th>$R$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maeda et al.</td>
<td>-0.87 (+22%)</td>
<td>4.59 (-4.7%)</td>
<td>0.8653 (+2.4%)</td>
<td>0.7344 (+4%)</td>
</tr>
<tr>
<td>Neubauer and Rostaşy</td>
<td>-0.45 (-78%)</td>
<td>3.62 (-17.6%)</td>
<td>0.9188 (+1.4%)</td>
<td>0.8350 (+10.4%)</td>
</tr>
<tr>
<td>Iso</td>
<td>0.22 (-81%)</td>
<td>4.25 (-17.9%)</td>
<td>0.8934 (+3.2%)</td>
<td>0.7731 (+16.6%)</td>
</tr>
<tr>
<td>Lu et al.</td>
<td>-0.60 (-71.6%)</td>
<td>3.61 (-18.0%)</td>
<td>0.9194 (+0.5%)</td>
<td>0.8359 (+10.7%)</td>
</tr>
</tbody>
</table>

models are remarkable. The RMSE values of Neubauer and Rostásy (1997), JCI (2003), and Lu et al. (2005) equations with respect to their original equations are decreased by 17.6%, 17.9% and 18%, respectively. Their $R^2$ values also increased by 10.4%, 16.6%, and 10.7%.

5.3 Evaluation of GMDH model and corrected equations

Tropsha et al. (2003) introduced new validation criteria for checking models based on their performances in testing data subsets. It is recommended that at least one of the gradients of the regression lines ($k$ or $k'$) through the origin for the predicted versus measured values, or for the measured versus predicted values, should be close to 1.

$$ k = \sum_{i=1}^{n} (O_i \times P_i) / P_i^2 \quad \text{or} \quad k' = \sum_{i=1}^{n} (O_i \times P_i) / O_i^2 $$

(33)

Also, the coefficient of determination for the regression line $(m, n)$ through the origin should be less than 0.1.

$$ m = \frac{R^2 - R^2_r}{R^2} $$

(34)

![Graph](image1)

**Fig. 5** Variation of discrepancy ratio (DR) between measured and predicted maximum bond strength as a function of (a) FRP-to-concrete width ratio (b) $E_{fr}$. 

![Graph](image2)
modified equations are shown in measures of the developed GMDH models and the tiles. The black edges are also the 10th and 90th percentiles. The edges of the box are the 25th and 75th value of them specifies the central mark (red line) of the boxes from largest to smallest values and then the median is displayed. It displays the variation in samples of a statistical population without making any assumptions of the underlying distribution. To generate a box plot, the DR values for entire database are sorted from largest to smallest values and then the median value of them specifies the central mark (red line) of the box. The edges of the box are the 25th and 75th percentiles. The black edges are also the 10th and 90th percentiles. The spacing between the different parts of the box indicates the degree of dispersion (spread) or skewness in the database. Based on Tables 3 and 4, the GMDH predictions are more reliable than the other empirical models. However, it is impossible to identify the uncertainty/safety factor incorporated in these models. Therefore, the exact degree of uncertainties in the mentioned models cannot be determined. To overcome this limitation, a safety factor based on the acceptable level of risk for each model can be considered.

Figure 6 presents the box plot and safety factors of different equations. It can be seen that the Neubauer and Rostásy (1997), Adhikary and Mutsuyoshi (2001), De Lorenzis et al. (2001), JCI (2003), Dai et al. (2005) and Camli and Binici (2007) generally overestimate the bond strength and are completely non-conservative. Conversely, the predictions of Tanaka (1996), Khalifa et al. (1998), and Yang et al. (2001) are conservative. In this aspect, Maeda et al. (1997) and JCI (2003) equations had the best performances amongst the other empirical equations. However, the proposed GMDH model shows good predictions close to one with less scatter (smaller box plots) than other existing equations. For having a precise and accurate predictive model, the error distribution of DR values should be symmetrical around their mean value and close to one. Wider distribution generally leads to more uncertainty. In addition, the safety factor of GMDH model is generally smaller than other equations. For example, if 10% risk is acceptable, the prediction of GMDH model should be divided by 1.4 while this factor for Maeda et al. (1997) and JCI (2003) equations are 1.8 and 1.7, respectively.

5.4 Sensitivity and parametric analysis
A sensitivity analysis was employed to identify the most effective involved parameters in predicting the bond strength between FRP and concrete substrate. To investigate sensitivity, each predictive parameter was eliminated one by one in order to observe how the accuracy of developed model based on GMDH algorithm would change. Statistically error measures are presented in Table 7. This table shows that $P_{max}$ are mostly affected by mechanical $(E_0)$ and geometrical $(b_p$ and $t_p$) properties of the FRP sheet. According to this table, the thickness of FRP sheet was the most effective parameter. The next important parameters were $b_p$, $E_p$, $L_p$, $f_c$, $E_{as}$, $b_s$, $t_s$, and $t_r$, respectively. The robustness of the developed GMDH model can be verified by examining how closely the model predictions agree with the physical behavior of the studied

### Table 6 External validation statistical measures for different approaches.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R(R&gt;0.8)$</th>
<th>$K (0.85&lt;k&lt;1.15)$</th>
<th>$K’ (0.85&lt;k&lt;1.15)$</th>
<th>$m (m&lt;0.1)$</th>
<th>$n (n&lt;0.1)$</th>
<th>$R_m (R_m&gt;0.5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMDH</td>
<td>0.9507</td>
<td>0.9973</td>
<td>0.9775</td>
<td>-0.1064</td>
<td>-0.1043</td>
<td>0.6236</td>
</tr>
<tr>
<td>Maeda et al. (new)</td>
<td>0.8653</td>
<td>1.0591</td>
<td>0.8813</td>
<td>-0.3137</td>
<td>-0.2636</td>
<td>0.3859</td>
</tr>
<tr>
<td>N&amp;R (new)</td>
<td>0.9188</td>
<td>1.0450</td>
<td>0.9173</td>
<td>-0.1731</td>
<td>-0.1536</td>
<td>0.5215</td>
</tr>
<tr>
<td>Iso (new)</td>
<td>0.8934</td>
<td>0.9482</td>
<td>0.9950</td>
<td>-0.2406</td>
<td>-0.2528</td>
<td>0.4484</td>
</tr>
<tr>
<td>Lu et al. (new)</td>
<td>0.9194</td>
<td>1.0508</td>
<td>0.9130</td>
<td>-0.1690</td>
<td>-0.1488</td>
<td>0.5258</td>
</tr>
</tbody>
</table>

$$n = \frac{R^2 - R_m^2}{R^2}$$  \quad (35)

where the squared correlation coefficient (through the origin) between predicted and measured values $(R_n^2)$ and the coefficient between measured and predicted values $(R_m^2)$ are calculated as:

$$R_n^2 = 1 - \frac{\sum_{i=1}^{n} (P_i - \bar{P})^2}{\sum_{i=1}^{n} (P_i - \bar{P})^2}$$  \quad (36)

$$R_m^2 = 1 - \frac{\sum_{i=1}^{n} (O_i - \bar{O})^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2}$$  \quad (37)

Additionally, the condition of cross validation should satisfy

$$R_n = R^2 \times \left(1 - \sqrt{R^2 - R_m^2}\right) > 0.5$$  \quad (38)

The validation criteria and relevant performance measures of the developed GMDH models and the modified equations are shown in Table 6. Models are considered valid for prediction if they satisfy some or all of the required conditions. It can be observed that the developed GMDH model satisfies all of the mentioned criteria; thus it has strong prediction power and is not a random correlation. Moreover, most of criteria were also satisfied by modified equations except the condition of cross validation $(R_n)$ for JCI (2003) and Maeda et al. (1997) models.

A reliable estimation of bond strength between FRP and concrete substrate is vital for engineers to have a reliable, technically correct and economic design. The box plot of discrepancy ratio $(DR)$ can be used to assess the reliability of the empirical equations and the proposed GMDH model. Box plot can be employed as a convenient graphical way for illustrating data points through their quartiles. It displays the variation in samples of a statistical population without making any assumptions of the underlying distribution. To generate a box plot, the DR values for entire database are sorted from largest to smallest values and then the median value of them specifies the central mark (red line) of the box. The edges of the box are the 25th and 75th percentiles. The black edges are also the 10th and 90th percentiles. The spacing between the different parts of the box indicates the degree of dispersion (spread) or skewness in the database. Based on Tables 3 and 4, the GMDH predictions are more reliable than the other empirical models. However, it is impossible to identify the uncertainty/safety factor incorporated in these models. Therefore, the exact degree of uncertainties in the mentioned models cannot be determined. To overcome this limitation, a safety factor based on the acceptable level of risk for each model can be considered.

### Table 7 External validation statistical measures for different approaches.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R(R&gt;0.8)$</th>
<th>$K (0.85&lt;k&lt;1.15)$</th>
<th>$K’ (0.85&lt;k&lt;1.15)$</th>
<th>$m (m&lt;0.1)$</th>
<th>$n (n&lt;0.1)$</th>
<th>$R_m (R_m&gt;0.5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMDH</td>
<td>0.9507</td>
<td>0.9973</td>
<td>0.9775</td>
<td>-0.1064</td>
<td>-0.1043</td>
<td>0.6236</td>
</tr>
<tr>
<td>Maeda et al. (new)</td>
<td>0.8653</td>
<td>1.0591</td>
<td>0.8813</td>
<td>-0.3137</td>
<td>-0.2636</td>
<td>0.3859</td>
</tr>
<tr>
<td>N&amp;R (new)</td>
<td>0.9188</td>
<td>1.0450</td>
<td>0.9173</td>
<td>-0.1731</td>
<td>-0.1536</td>
<td>0.5215</td>
</tr>
<tr>
<td>Iso (new)</td>
<td>0.8934</td>
<td>0.9482</td>
<td>0.9950</td>
<td>-0.2406</td>
<td>-0.2528</td>
<td>0.4484</td>
</tr>
<tr>
<td>Lu et al. (new)</td>
<td>0.9194</td>
<td>1.0508</td>
<td>0.9130</td>
<td>-0.1690</td>
<td>-0.1488</td>
<td>0.5258</td>
</tr>
</tbody>
</table>
parametric analysis was carried out with the aim to deeper understand of the bond strength between FRP and concrete substrate. The parametric analysis investigates the response of the predicted bond strength from the GMDH model to a set of hypothetical input data generated over the ranges of the minimum and maximum data used for the model training. The methodology is based on the change of only one input variable at a time while the other variables are kept constant at their average values in entire data sets. A set of synthetic data for the single varied parameter is generated by incrementally increasing the mentioned parameter. These inputs are presented to the prediction equation and the bond strength is calculated. This procedure is repeated using another variable until the model response is tested for all input variables. It should be noted that the average values of $E_p$, $f'c$, $tc$, $tp$, $bc$, $bp$, and $Lb$ are considered 196 GPa, 36 MPa, 128 mm, 0.48 mm, 137 mm, 52 mm, and 183 mm, respectively.

In Fig. 7, the values of $P_{max}$ predicted by the developed GMDH model are plotted as functions of the main design parameters including $f'c$, $Lb$, $E_p$,$p$, and $bp/bc$. From Fig. 7(a), it can be seen that the sensitivity of the predicted bond strength by GMDH model to the change of concrete compressive strength is very low. This observation can be explained by the fact that debonding cracks develop when the maximum transferable load reaches the concrete tensile strength that is a function of $f'c^{0.5}$ and as a result, compression strength can slightly impact on the bond strength (Mashrei et al. 2013).

Figure 7(b) presents the variations of bond strength with bond length ($L_b$). As shown, the bond strength increases with increasing the bond length up to an effective bond length and then, a slight decrease occurs beyond this length. This behavior is also in agreement with experimental database as shown in Fig. 7(b). Generally, in members externally bonded to concrete substrate, no matter how long the adhesion length may increase, bond strength could not absolutely reach to the tensile strength of FRP. However, the more increase of bond strength, the more increase of the ductility in the fracture process (Woo and Lee 2010). It should be noted

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>RMSE</th>
<th>$R$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>In absence of $E_p$</td>
<td>1.9688</td>
<td>2.7660</td>
<td>0.9504</td>
<td>0.9033</td>
</tr>
<tr>
<td>In absence of $t_p$</td>
<td>1.9705</td>
<td>2.7263</td>
<td>0.9519</td>
<td>0.9061</td>
</tr>
<tr>
<td>In absence of $b_p$</td>
<td>1.9705</td>
<td>2.7263</td>
<td>0.9519</td>
<td>0.9061</td>
</tr>
<tr>
<td>In absence of $f'_c$</td>
<td>1.9593</td>
<td>2.6966</td>
<td>0.9529</td>
<td>0.9081</td>
</tr>
<tr>
<td>In absence of $b_c$</td>
<td>2.0428</td>
<td>2.8007</td>
<td>0.9491</td>
<td>0.9009</td>
</tr>
<tr>
<td>In absence of $t_c$</td>
<td>4.0232</td>
<td>5.9538</td>
<td>0.5520</td>
<td>0.7429</td>
</tr>
<tr>
<td>In absence of $b_p$</td>
<td>2.9467</td>
<td>3.9017</td>
<td>0.8987</td>
<td>0.8076</td>
</tr>
<tr>
<td>In absence of $L_b$</td>
<td>2.1197</td>
<td>2.9374</td>
<td>0.9439</td>
<td>0.8909</td>
</tr>
<tr>
<td>In absence of $E_p$</td>
<td>2.4733</td>
<td>3.3805</td>
<td>0.9250</td>
<td>0.8556</td>
</tr>
</tbody>
</table>

Fig. 6 Box plot of different equations.
that the sensitivity of bond strength to change of bond length is more than to variation of concrete compressive strength, which this result is also in agreement with sensitivity analysis obtained by GMDH model based on the Table 7.

In Fig. 7(c), the variation of bond strength with both $E_p$ and $L_b$ as a function of axial rigidity of FRP reinforcement ($E_p L_b$) is shown. It is clear that the sensitivity of bond strength to change of $E_p L_b$ is more significant than other parameters. It can be seen that the bond strength almost linearly increased with increase of $E_p L_b$. This behavior can be expected and also appears in several existing equations. The sensitivity of bond strength predictions to variation of both $b_p$ and $b_p$ parameters as a function of FRP-to-concrete width ratio ($b_p/b_c$) is also illustrated in Fig. 7(d). It is clear that the bond strength increases with increasing the width ratio, as expected, but at a decreasing rate. From these observations, it can be concluded that the developed GMDH model is in good agreement with the physical characteristics of debonding phenomenon and previous findings.

5.5 Reliability analysis using demerit points classifications

Collins (2001) introduced a new scale to evaluate and classify the reliability of design codes. This scale is known as Demerit Points Classification (DPC), which consider the safety, accuracy and scattering of design codes as a function of the ratio between the ultimate resistances in experimental tests and the estimated theoretically load capacity. In this study, the measured ultimate bond strength ($P_{\text{max,exp}}$) and predicted bond strength by existing formulae and GMDH algorithm ($P_{\text{max,predicted}}$) are the ultimate resistances in experimental tests and the estimated theoretically load capacity, respectively. Table 8 presents an adaption made in the present study to the original values proposed by Collins. According to bond strength predicted by each formula, a demerit point for each prediction for 342 data points is attributed to that formula based on Table 8. Then, the general value of demerit of each formula is calculated by the sum of the products of the number of specimens in each interval and their corresponding demerit penalty. The lower the value of total sum indicates that the considered formula is more reliable.

![Fig. 7](image_url)

Fig. 7 Maximum bond strength parametric analysis in the GMDH-based model for different ranges of: (a) $f'_c$, (b) $L_b$ (c) $E_p L_b$ (d) $b_p/b_c$.
658

Table 8 Classification by demerit points.

<table>
<thead>
<tr>
<th>$P_{\text{exp}}/P_{\text{predicted}}$</th>
<th>Classification</th>
<th>Demerit points</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;0.50$</td>
<td>Extremely dangerous</td>
<td>10</td>
</tr>
<tr>
<td>$[0.50-0.85)$</td>
<td>Dangerous</td>
<td>5</td>
</tr>
<tr>
<td>$[0.85-1.15)$</td>
<td>Appropriate and safe</td>
<td>0</td>
</tr>
<tr>
<td>$[1.15-2.00)$</td>
<td>Conservative</td>
<td>1</td>
</tr>
<tr>
<td>$\geq2.00$</td>
<td>Extremely conservative</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 9 Classification of GMDH and modified equations according to the criteria of Collins.

<table>
<thead>
<tr>
<th>Model</th>
<th>DR $&lt;0.5$</th>
<th>0.5 $\leq$ DR $&lt;0.85$</th>
<th>0.85 $\leq$ DR $&lt;1.15$</th>
<th>1.15 $\leq$ DR $&lt;2$</th>
<th>DR $&gt;2$</th>
<th>Demerit points</th>
</tr>
</thead>
<tbody>
<tr>
<td>GMDH (present study)</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>397</td>
</tr>
<tr>
<td>Maeda et al. (new)</td>
<td>-</td>
<td>67</td>
<td>213</td>
<td>62</td>
<td>-</td>
<td>458</td>
</tr>
<tr>
<td>N&amp;R (new)</td>
<td>1</td>
<td>63</td>
<td>189</td>
<td>89</td>
<td>-</td>
<td>414</td>
</tr>
<tr>
<td>Iso (new)</td>
<td>1</td>
<td>85</td>
<td>137</td>
<td>119</td>
<td>-</td>
<td>554</td>
</tr>
<tr>
<td>Lu et al. (new)</td>
<td>1</td>
<td>63</td>
<td>191</td>
<td>87</td>
<td>-</td>
<td>412</td>
</tr>
<tr>
<td>fib model (N&amp;R)</td>
<td>2</td>
<td>171</td>
<td>140</td>
<td>29</td>
<td>-</td>
<td>904</td>
</tr>
<tr>
<td>CNR-DT200</td>
<td>-</td>
<td>21</td>
<td>46</td>
<td>249</td>
<td>26</td>
<td>406</td>
</tr>
</tbody>
</table>

Table 9 presents the evaluation of GMDH and modified equations as a function of the adapted criterion from Collins (2001). According to this criterion, fib model, which is actually a modification of N&R model, presents higher total demerit points (904 points) than other models, with 50% of the values in the second classification range (between 0.5 and 0.85), unfavorable in terms of safety. CNR-DT200/2004 model presents the lowest total demerit points amongst the other modified equations. However, 72.8% of its prediction values are in the fourth classification range (between 1.15 and 2), which is classified as conservative in terms of safety. The existing equations usually suffer from either having limited accuracy or being too conservative. Generally, the GMDH model with lowest total demerit points (397) and having 62% of its prediction in the range of the third classification (appropriate and safe) had the best performance amongst the other equations in terms of safety.

6. Summary and conclusion

The applicability of group method of data handling (GMDH) as an alternative approach was investigated and assessed in prediction of the externally bond strength between fiber-reinforced polymer (FRP) composites and concrete substrates. The database used for the development of GMDH model was collected from the literature and comprised a series of 342 experimental tests of bond strength between FRP and concrete substrate. The developed GMDH model related the bond strength to the width of FRP plate, elastic modulus of FRP, concrete cylinder compressive strength, the effective thickness of concrete contributing to shear deformation, the width of concrete prism, the bond length, the elastic modulus of adhesive, and the thickness of adhesive layer. The predictive ability of GMDH model was examined by comparing its predictions with those obtained from the most common formulae and regression approaches including MLR and MNLR approaches. Some correction factors for Maeda et al. (1997), Neubauer and Rostásy (1997), JCI (2003), and Lu et al. (2005) are suggested by applying MNLR method. The most important outcomes of this study can be summarized as follows:

• Among the existing equations in literature, there is a good agreement between measured and predicted maximum bond strength using Neubauer and Rostásy (1997) equation. However, most of its predictions were categorized in unsafe region based on Collins criteria. In contrast, the CNR-DT200 model represented the safest predictions according to this criterion while its accuracy is remarkably limited.

• The newly proposed GMDH model outperformed the most common existing equations in terms of both accuracy and safety. The errors of developed GMDH model also showed a symmetrical and predictable behavior.

• The results of statistical measures show that the proposed GMDH model outperformed the other equations in literature. The proposed equations yield an $R^2=0.90$ and RMSE=2.76 kN, this represents a 37.2% improvement in RMSE and 19.4% improvement in terms of $R^2$ in respect to Neubauer and Rostásy (1997) equation as the best equation among the other design equations.

• Some correction factors were also suggested for Neubauer and Rostásy (1997), JCI (2003), and Lu et
al. (2005) equations to improve their accuracy. Results indicated that the RMSE values of modified equations with respect to their original equations were improved by 17.6%, 17.9% and 18%, respectively

- The sensitivity analysis of GMDH model indicated that the maximum bond strength are mostly affected by mechanical (\(E_s\)) and geometrical (\(l_p\) and \(t_p\)) properties of the FRP sheet. The thickness of FRP plate or sheet was the most important input variables.

- The robustness of GMDH model in capturing the underlying physical behaviors of bond strength was verified through parametric study. Results of parametric analysis were also confirmed the outcomes of the sensitivity analysis.

In general, GMDH can be successfully used as a reliable alternative approach to predict maximum bond strength between FRP sheets and concrete substrates.

References


plates bonded to concrete.” Engineering Structures, 27(4), 564-575.


