Engineering Structures 195 (2019) 524–535

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

Predicting reinforcing bar development length using polynomial chaos expansions

Zaher Mundher Yaseen[a], Behrooz Keshtegar[b], Hyeon-Jong Hwang[c], Moncef L. Nehdi[d,*]

[a] Sustainable Developments in Civil Engineering Research Group, Faculty of Civil Engineering, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
[b] Department of Civil Engineering, Faculty of Engineering, University of Zabol, Zabol, Iran
[c] Department of Architecture and Civil Engineering, Hunan University, Changsha, Hunan, China
[d] Department of Civil and Environmental Engineering, Western University, London, Ontario, Canada

ARTICLE INFO

Keywords:
Polynomial chaos expansion
Model
Prediction
Rebar
Development length
Bond stress
Concrete cover
Compressive strength
Design code

ABSTRACT

The bond stress of a reinforcing bar in a cementitious matrix varies along the bar length and is difficult to quantify. Thus, design code provisions refer to the concept of bar development length and rely on statistical analysis of rebar-pull-out test results. In the present study, a novel data-driven predictive model based on Polynomial Chaos Expansions (PCE) was developed to predict the reinforcing bar development length using 534 experimental results of simple pull-out tests on short unit bar lengths. The predictive capability of PCE was compared to that of other data-driven models, namely the Response Surface Method (RSM) and Artificial Neural Networks (ANN). Moreover, predictions of the PCE, RSM and ANN were further compared with calculations of three commonly used design code formulas (i.e., ACI 318-14, ACI 408R-03, and Eurocode 2) and predictions of two existing empirical models (i.e. Model Code 2010 and Hwang et al. model). A parametric study was conducted to explore the sensitivity of the proposed model to influential input parameters. It was found that the Polynomial Chaos Expansions model offers a powerful predictive tool for reinforcing bar bond strength. The model was able to capture trends that differ from that of existing models that assume unrealistic uniform bond stress along the rebar. This flexible and data intensive model for predicting rebar bond stress and full embedment length could offer an intelligent platform for accommodating new bar materials, new test data, and calibrating existing design provisions to keep design codes relevant.

1. Introduction

Reinforced concrete technology has been undergoing transformative changes in recent decades. For instance, various types of reinforcing steel rebar are being used, including carbon steel with various deformed shapes, galvanized, epoxy-coated, stainless, and high-strength steel rebar [1,2]. Moreover, new composite material bars with different engineering properties and surface texture have emerged as contenders for concrete reinforcement, including carbon, aramid, glass and basalt rebar.

Likewise, the cementitious matrices in which various rebar can be embedded have become diverse. Indeed, different types of concrete have become mainstream with inherently different bond to and compatibility with reinforcing rebar. These include conventional concrete, mass concrete, cellular concrete, fiber-reinforced concrete, high-strength concrete, reactive powder concrete, ultrahigh-performance concrete, polymer-modified concrete, lightweight concrete, pervious concrete, shotcrete, rubberized concrete, roller-compacted concrete, self-consolidating and anti-washout concrete, concrete using calcium aluminate or calcium sulfo-aluminate cements, just to name the most common types. Moreover, sustainability considerations have brought about eco-efficient concrete types such as geo-polymer concrete, concrete using alkali-activated binders, and other alternative cementitious matrices.

Such developments pose challenges to the design engineer. Classical design equations reflecting old age wisdom acquired on traditional concrete and rebar materials have, in many cases, become obsolete. Thus, substantial experimental testing is often needed to develop more reliable design provisions applicable to the emerging new materials, so that design codes can be appropriately revised.

Furthermore, the bond of steel rebar to a cementitious matrix has become the subject of considerable research and controversy. Various

* Corresponding author.
E-mail address: mnehdi@uwo.ca (M.L. Nehdi).

https://doi.org/10.1016/j.engstruct.2019.06.012
Received 8 December 2018; Received in revised form 5 May 2019; Accepted 5 June 2019
Available online 17 June 2019

parameters influence the bond strength of rebar in concrete. These include mutual adhesion between the rebar and concrete interfaces; friction interlock between bar deformations or projections of rebar and the cementitious matrix; gripping effect imparted for instance by shrinkage stresses; the tensile, compressive and stiffness properties of the cementitious medium; mechanical anchorage of bar ends; diameter, shape, surface condition, and stress-strain behavior of the rebar; concrete cover; rebar embedment length; and rebar confinement and transverse reinforcement.

The rebar development length, l_d, is generally defined as the shortest length of rebar through which the bond stress can increase from zero to the bar yield strength, f_y. This implies that when the distance from the end of rebar to a point where the stress equals f_y is less than the development length, then the rebar will not yield, but rather pull out of the concrete. Discussion of bond development of reinforcing rebar and basic rebar development length has been extensively addressed elsewhere (e.g., [3]).

The concepts of bond stress and rebar development length are fundamental in reinforced concrete design codes and guidelines, such as the ACI 318-14, ACI 408R-03, and Eurocode 2. Discussion of existing design methods for bar development length has been provided by others (e.g., [4]) and will not be covered further herein. Such design provisions generally specify the rebar development length empirically based on existing test results. The development of a more flexible and data intensive framework for predicting rebar average bond stress and full embedment length could offer an intelligent platform for accommodating new data, calibrating existing design provisions and keeping design codes relevant.

The polynomial chaos expansions (PCE) approach proposed by Wiener (1983) can capture stochastic relations for complex nonlinear engineering events using homogeneous orthogonal polynomials basis functions [5]. Non-parametric analysis-based modelling approach using PCE has been considered in the past for multiple complex engineering problems characterized by uncertainty. For instance, quantification of uncertainties in numerical simulation by PCE is a technique, which has been recently used in various problems and applications. It can also be used in global sensitivity analysis through approximation of sensitivity indices. The PCE-based modelling has indeed been widely utilized in numerus engineering problems. These include structural reliability analysis [6,7], prediction of vehicle dynamics [8], fluid mechanics [9], structural dynamics [10], and fluid-structure interaction problems [11]. The PCE-based model was implemented for uncertainties assessment of hydrologic model under randomness [12,13]. Sochala and Le Maître used the PCE-based surrogate model to represent uncertainties of input parameters in simulating subsurface flows [14]. Kewlani et al. [8] employed the PCE for evaluating vehicle dynamics under uncertainty. The PCE was also utilized to predict the dissolved oxygen (DO) concentration in river water and structured using water quality input variables [15]. However, it has not previously been adopted to solve data intensive problems such as modeling the bar bond stress in reinforced concrete design.

In the present study, a novel modelling approach based on PCE is proposed for using multivariate orthogonal polynomial (MOP) functions with high-order terms. The high-order cross Hermite PCE functions based on the input data are applied to predict the average reinforcing bar bond strength (ARBS). The accuracy of the PCE-based model in predicting ARBS is appraised via two data-driven models, first using the response surface method (RSM) second-order polynomial functions with cross terms, and then using artificial neural networks (ANN). The accuracy of predicting the average reinforcing bar bond strength by the three heuristic modeling approaches (PCE, RSM and ANN) was assessed. First such predictions were compared to that of two existing empirical models (Model Code 2010 and the Hwang et al. 2017 model). Second, it was compared with predictions of three commonly used reinforced concrete design codes (i.e. ACI 318-14, ACI 408R-03, and Eurocode 2) [4,16–19].

2. Experimental database

The development of a data-driven model for predicting the average bond stress and full development length of reinforcing bar in concrete based on Polynomial Chaos Expansions requires a comprehensive database capturing the key parameters influencing the overall behavior. In this study, such a database was retrieved from the open literature. In total, 534 data sets from 27 different studies based on the lap-splice test method, which is described in the Joint ACI-ASCE Committee 408, were used. The excellent database is compiled and presented in detail elsewhere [4,20]. Since this database is readily accessible, there is no need to reproduce it herein. The database includes the lap splice length l_s, ranging from 76 to 2311 mm, the reinforcing bar diameter d_b, ranging from 9.5 to 43.0 mm, the concrete compressive strength f_c', ranging from 12.6 to 113.0 MPa, and the yield strength of the reinforcing bar f_y, ranging from 345 to 830 MPa. For all 534 experimental cases, the database presents the experimental tensile strength (f_s) of the bar splices according to the splice length. The database was randomly divided into training (80%) and validation (20%) sets to build the models, as discussed below.

3. Methodology and model development

3.1. Polynomial chaos Expansions

The polynomial chaos expansion (PCE) was generally formulated by homogeneous and orthogonal Hermite polynomial functions to build stochastic models. The PCE model to predict the tensile strength of the bar splices ($\hat{f}_s$) is presented using series expansions as follows [21]:

$$\hat{f}_s(\xi) = \sum_{i=1}^{P} w_i \prod_{j=1}^{M} H_i(\xi_j) \tag{1}$$

where w_i is the unknown coefficients, which provides a connection between the predicted $\hat{f}_s(\xi)$ and PCE functions. ξ_i are the normal standard input variables, which are given by input variable (X) as $\xi = \frac{X-\mu}{\sigma}$ (in which, μ and σ are the mean and standard deviation of input variable X, respectively). P represents the total number of PCE terms, which is given by MOP basis functions $\Psi_i(\xi)$ as follows:

$$\Psi_i(\xi) = \prod_{j=1}^{M} H_i(\xi_j) \tag{2}$$

where M is the dimension of MOP, which is given by the input data using each Hermite polynomials H_i, which in turn is given by the following relations [22]:

$$H_O(\xi) = 1,\ H_1(\xi) = \xi,\ H_P(\xi) = \xi H_{P-1}(\xi) - (p-1)H_{P-2}(\xi) \quad p = 2, 3, 4, \cdots. \tag{3}$$

The multivariate orthogonal polynomials are given by homogeneous chaos functions H_0-H_p with dimension M in the basic predicted relation of PCE in Eq. (1). As can be observed in Eq. (3), the basic functions, $\Psi_i(\xi)$ are increased exponentially by increasing the total expansion order, p and dimension of the problem. High-order degree for polynomial basis chaos functions may be provided based on the order chaos polynomial factor, p. Consequently, the input chaos data points are given based on an order chaos polynomial factor of $p = 5$ to build the multi-dimensional Hermite polynomial functions (Ψ)in the current study. The predicted rebar stress is given based on the Hermite polynomial functions as follows:

$$\hat{f}_s(\xi_i) = ([\Psi^T\Psi]^{-1}\Psi^T f_{train}) \prod_{j=1}^{M} H_i(\xi_j) \tag{4}$$

where f_{train} is the rebar stress data in the training phase. The matrix Ψ_{NxP} is MOPs determined as follows:

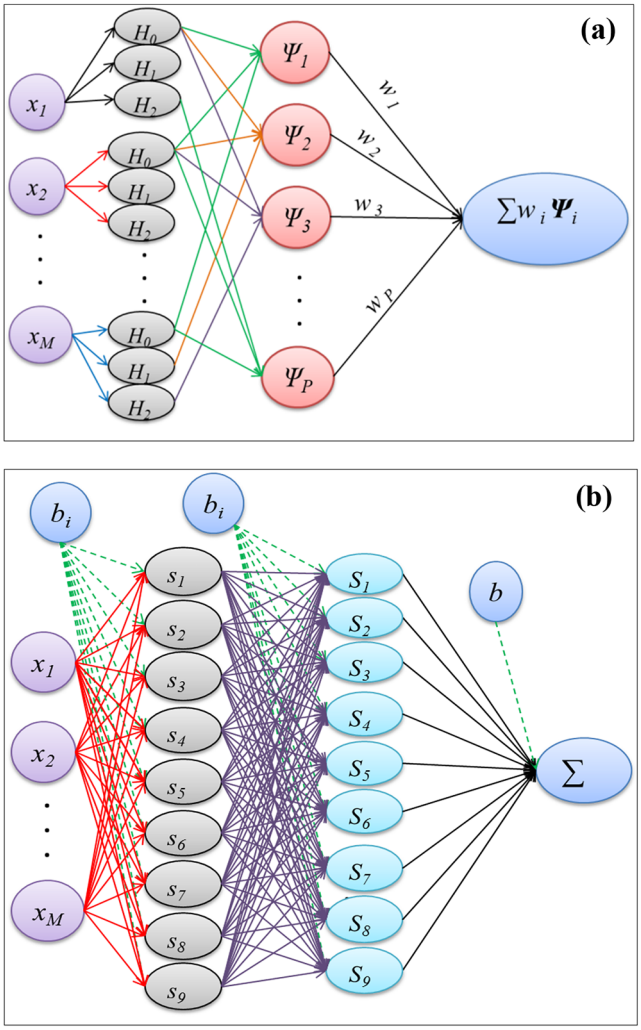

Fig. 1. (a) Shamanic view of PCE structure for $p = 2$, (b) Schematic structure of ANN with 2-hidden layer.

$$\Psi = \begin{bmatrix} \Psi_1(\xi_1)\Psi_2(\xi_1) & \cdots & \Psi_P(\xi_1) \\ \vdots & \ddots & \vdots \\ \Psi_1(\xi_N)\Psi_2(\xi_N) & \cdots & \Psi_P(\xi_N) \end{bmatrix} \tag{5}$$

The number of training data points (N) is selected to be larger than the number of orthogonal polynomials (P), i.e. $N > P$. Accordingly, a total number of data used in the training phase, i.e. 427 data points, were randomly selected and represent about 80% of the entire database (534 data points). The schematic view of the PCE model based on the input variables x_1-x_M and $p = 2$ is plotted in Fig. 1a. The PCE model represents the input variables by MOPs functions-based Hermite polynomials of each input data using mapping $\xi = \frac{X-\mu}{\sigma}$ for prediction of the average rebar bond stress. These MOPs basis functions for training and testing data points can be extracted from the basic input variables in the modelling process by basic Hermite polynomials. For example, the MOPs are determined for two input data (ξ_1 andξ_2) and $p = 5$ using basic Hermite polynomials of $H_0(\xi) = 1$; $H_1(\xi) = \xi$; $H_2(\xi) = \xi^2 - 1$; $H_3(\xi) = \xi(\xi^2 - 3)$; $H_4(\xi) = \xi^4 - 6\xi^2 + 3$; $H_5(\xi) = \xi(\xi^4 - 10\xi^2 + 15)$; asfolllows:

$$\begin{aligned} \Psi_1(\xi) &= 1,\ \Psi_2(\xi) = \xi_1,\ \Psi_3(\xi) = \xi_1^2 - 1,\ \Psi_4(\xi) = \xi_1(\xi_1^2 - 3),\ \Psi_5(\xi) \\ &= \xi_1^4 - 6\xi_1^2 + 3,\ \Psi_6(\xi) = \xi_1(\xi_1^4 - 10\xi_1^2 + 15),\ \Psi_7(\xi) = \xi_1\xi_2\Psi_8(\xi) \\ &= \xi_2(\xi_1^2 - 1),\ \Psi_9(\xi) = \xi_2\xi_1(\xi_1^2 - 3),\ \Psi_{10}(\xi) \\ &= \xi_2(\xi_1^4 - 6\xi_1^2 + 3),\ \Psi_{11}(\xi) = \xi_2^2 - 1,\ \Psi_{12}(\xi) \\ &= (\xi_2^2 - 1)(\xi_1^2 - 1),\ \Psi_{13}(\xi) = \xi_1(\xi_2^2 - 1)(\xi_1^2 - 3),\ \Psi_{14}(\xi) \\ &= \xi_2(\xi_2^2 - 3),\ \Psi_{15}(\xi) = \xi_2(\xi_2^2 - 3)(\xi_1^2 - 1),\ \Psi_{16}(\xi) \\ &= \xi_2^4 - 6\xi_2^2 + 3,\ \Psi_{17}(\xi) = \xi_1(\xi_2^4 - 6\xi_2^2 + 3),\ \Psi_{17}(\xi) \\ &= \xi_2^4 - 6\xi_2^2 + 3,\ \Psi_{18}(\xi) = \xi_2(\xi_2^4 - 10\xi_2^2 + 15) \end{aligned} \tag{6}$$

It can be noted that a large variety of Hermite PCE is applied to predict f_s. The database using MOPs basis functions is the connecting tool between the input data and predicted average rebar bond stress.

3.2. Response surface method

The response surface method (RSM) is among the most efficient and simple modeling tools using second-order polynomial functions, as follows [23]:

$$f_s(X) = a_o + \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j \tag{7}$$

where $f_s(X)$ is the predicted tensile strength; n is the number of input variables; and a_o, a_i and a_{ij} are unknown coefficients. The basic polynomial basic function is directly computed using input data without chaos form in the RSM. The unknown coefficients in Eq. (7) can be calibrated using the least square method via the following direct relation in Eq. (8) [24,25]:

$$a = [P(X)^T P(X)]^{-1} P(X)^T f_{train} \tag{8}$$

where a represents the unknown coefficients vector; f_{train} are the experimental rebar stress data in the training phase; $P(X)$ is the polynomial basic function which is computed with n-input data as $P(\boldsymbol{X}) = \{P(X_1), P(X_2), \ldots P(X_n)\}^T$; N denotes the number of data points in the training phase; and $P(X_n)$ is the input data by second-order function, which are computed by X_N as follows:

$$\begin{aligned} &P(X) \\ &= \{1, x_1, x_2, \cdots, x_n, x_1^2, x_1, x_2, x_3, \cdots, x_1, x_n, x_2^2, x_2, x_3, \cdots, x_2, x_n, x_3^2, x_3, x_4, \\ &\quad \cdots, x_3, x_n, \cdots, x_n^2\} \end{aligned} \tag{9}$$

3.3. Artificial neural network

Artificial Neural Networks (ANN) are robust statistical optimization frameworks that mimic the biological nervous system. They can establish logical models comprised of several interconnected neurons in a computing platform [26–28]. ANNs are employed in finding solutions to complex modeling tasks, such as estimating, classifying, or performing pattern recognition. The development of any ANN demands three modeling steps. ANNs for regression or classification tasks are mainly classified into two main forms: supervised and unsupervised. Supervised ANNs are trained via a regulation of the inter-neuronal weight values, making it possible to predict the output values after introducing several input data from the previous experiments. In unsupervised ANNs, there are no set target values during the introduction of the input. In the multi-layer perceptron (MLP) feedforward ANN, which is one of the most common algorithms for training of optimization frameworks, there is usually one or more hidden layers where the selection of the appropriate input data mainly depends on the experience of the analyst, as well as on the nature of the problem to be solved. The input, in the feedforward backpropagation algorithms traverses the network, and at the end, the network output is matched with a set values for estimating the error level [29,30]. Back-propagation learning rule ensures the establishment of an input-output relationship, which is

usually determined by randomly allocating initial weights to the input data before updating it. The randomly weighted input data are updated by comparing the outcome of the iteration process to the desired output.

Most studies that utilized neural computations implemented several transfer functions based on the nature of the problem (complexity and nonlinearity). In recent engineering tasks, the two most common transfer functions often utilized for the hidden and output layers, respectively, are the tangent sigmoid and the linear functions. The tangent sigmoid transfer functions are usually applied to the hidden neurons to ensure that the input-output behavior of the system is significantly improved when the updated weights are slightly varied. In Fig. 1b, a multi-layer ANN comprising two hidden layers having nine neurons is depicted.

During the training of a network, the termination criterion is reached when an error evaluation criterion such as the mean squared errors (*MSE*) and correlation coefficient (*R*), has reached an acceptable limit, or after passing a pre-adjusted number of training epochs. These error measures (*MSE* and *R*) are defined as follows [31]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_{sp} - f_{so})^2 \tag{10}$$

$$R = \frac{\sum_{i=1}^{n}(f_{so} - \overline{f_{so}})(f_{sp} - \overline{f_{sp}})}{\sqrt{\sum_{i=1}^{n}(f_{so} - \overline{f_{so}})^2 \sum_{i=1}^{n}(f_{sp} - \overline{f_{sp}})^2}} \tag{11}$$

where f_{so} represents the experimental data value, f_{sp} is the predicted value, and n represents the number of samples in the data set. $\overline{f_{so}}$ and $\overline{f_{sp}}$ are the mean values of f_{so} and f_{sp}, respectively.

4. Model application, results and discussion

Three commonly used design code formulas, namely ACI 318-14, ACI 408R-03, and Eurocode 2, along with two empirical mathematical models, and three data-driven modeling approaches using PCE, RSM and ANN were applied for the prediction of the average reinforcing rebar bond strength in concrete. The PCE and RSM were coded in MATLAB, while the ANN was applied based on a MATLAB toolbox with one input layer, two hidden layers with nine neurons, and one output layer as M-9-9-1, where M is the number of input data. In this study, two input scenarios were used to compare the different predicated models of the average bond strength of reinforcing rebar. Scenario I applied four elements(l_s, d_b, f_c', f_y), while scenario II used the parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$(where, $c = c_{min} + d_b/2$, $w = 0.1(c_{max}/c_{min}) + 0.9 \le 1.25$, $K_{atr} = 6\sqrt{f_c'}A_{tr}t_d/(nS_t)$, $c_{max} = \max(c_b, c_s)$, $c_{min} = \min(c_b, c_s)$, $c_s = \min(c_{so}, c_{si} + 6.4)$, S_t = center-to-center distance of the transverse bars, n = the number of bars being developed, A_{tr} = total cross-sectional transverse bar area within spacing S_t, c_{si} = one-half of the center-to-center bar spacing, c_b = thickness of bottom cover concrete, and c_{so} = thickness of side cover concrete)

The PCE is structured using the MOPs basis functions with a high-order of $p = 5$. The highly nonlinear function can provide suitable flexibly for the model to consider high cross-correlations between all input data and the average reinforcing rebar bond strength. Fig. 2 illustrates the effect of the order, p of the multivariate orthogonal polynomial function on the RMSE value for the training and testing phases. It can be observed that the RMSE decreased with increasing p in the training phase, which means that high-order polynomials in calibrating the PCE model can improve its training prediction. However, for the testing phase, no significant reduction in RMSE could be observed beyond $p = 5$. In fact, RMSE for $p = 8$ was higher than that for $p = 5$. This justifies why the authors limited the order p of the PCE model in this study to 5, which was perceived as optimal. This is also advantageous since it reduces the computational burden of the PCE model.

The database was randomly divided into two sets. About 80% of the

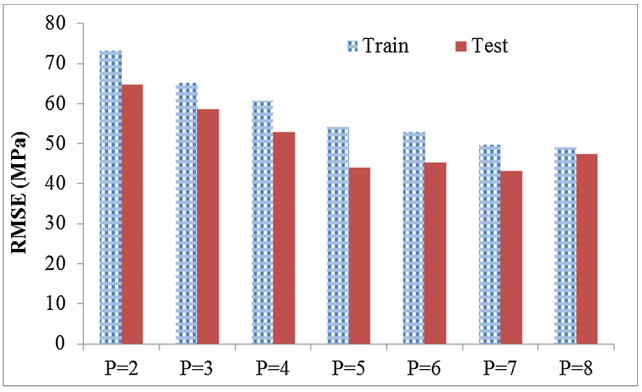

Fig. 2. Effect of the order of the multivariate orthogonal polynomial function, p, on the RMSE value for the training and testing phases.

experimental data points were utilized for the training stage of the models, while about 20% of the data was used for the model validation phase. The accuracy of model predictions and agreement of predicted data amongst these nonlinear models were compared based on several comparative statistics, including the root mean square error (*RMSE*), mean bias error (*MBE*), mean absolute error (*MAE*), modified agreement index (*md*) and modified Nash and Sutcliffe efficiency (*mNSE*) statistics [32–34], which are expressed below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(f_{so})_i - (f_{sp})_i]^2} \tag{12}$$

$$MBE = \frac{1}{N}\sum_{i=1}^{N}\frac{(f_{so})_i - (f_{sp})_i}{(f_{so})_i} \tag{13}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(f_{so})_i - (f_{sp})_i| \tag{14}$$

$$md = 1 - \frac{\sum_{i=1}^{N}|(f_{so})_i - (f_{sp})_i|}{\sum_{i=1}^{N}|(f_{so})_i - \bar{f_{so}}| + |(f_{sp})_i - f_{so}|}, \quad < md \le 1 \tag{15}$$

$$mNSE = 1 - \frac{\sum_{i=1}^{N}|(f_{so})_i - (f_{sp})_i|}{\sum_{i=1}^{N}|(f_{so})_i - (\bar{f}_{sp})_i|}, \quad -\infty < mNSE \le 1 \tag{16}$$

When *RMSE*, *MBE*, and *MAE* approach zero, the model provides near perfect prediction and close agreement with experimental data, with *md* and *mNSE* values that tend to one.

Figs. 3 and 4 with the statistical presentation provided in Tables 1 and 2 compare the model results for predictions of the proposed models (RSM, ANN, and PCE) in the training phase (i.e., 80% of the database) and testing phase (i.e., 20% of the database). Fig. 3 and Table 1 show predictions for scenario I, which considers the lap-splice length l_s, reinforcing bar diameter d_b, concrete compressive strength f_c', and yield strength of the reinforcing bar f_y as design parameters. It can be observed that each of the various models exhibited similar level of accuracy for its predictions in the training and testing phases. The RSM and ANN models showed large variance for predictions for the normal steel rebar stress levels (i.e., less than 500 MPa), and tended to overestimate tensile stresses for the bar splices in the high stress levels (i.e., higher than 500 MPa). On the other hand, the PCE model predicted experimental test results better than both the RSM and ANN models. This can be explained by the excellent ability of the PCE model to abstract the physical mechanism and interrelations between the independent and dependent variables. This was also reported in several previous studies (e.g., [7,35–36]). Hence, it would be interesting to revisit previous studies which exploited ANN modeling (e.g. 37–40) to explore whether PCE modeling could provide superior performance.

Fig. 4 and Table 2 illustrate predictions for scenario II, which additionally considers the effects of the concrete cover and transverse

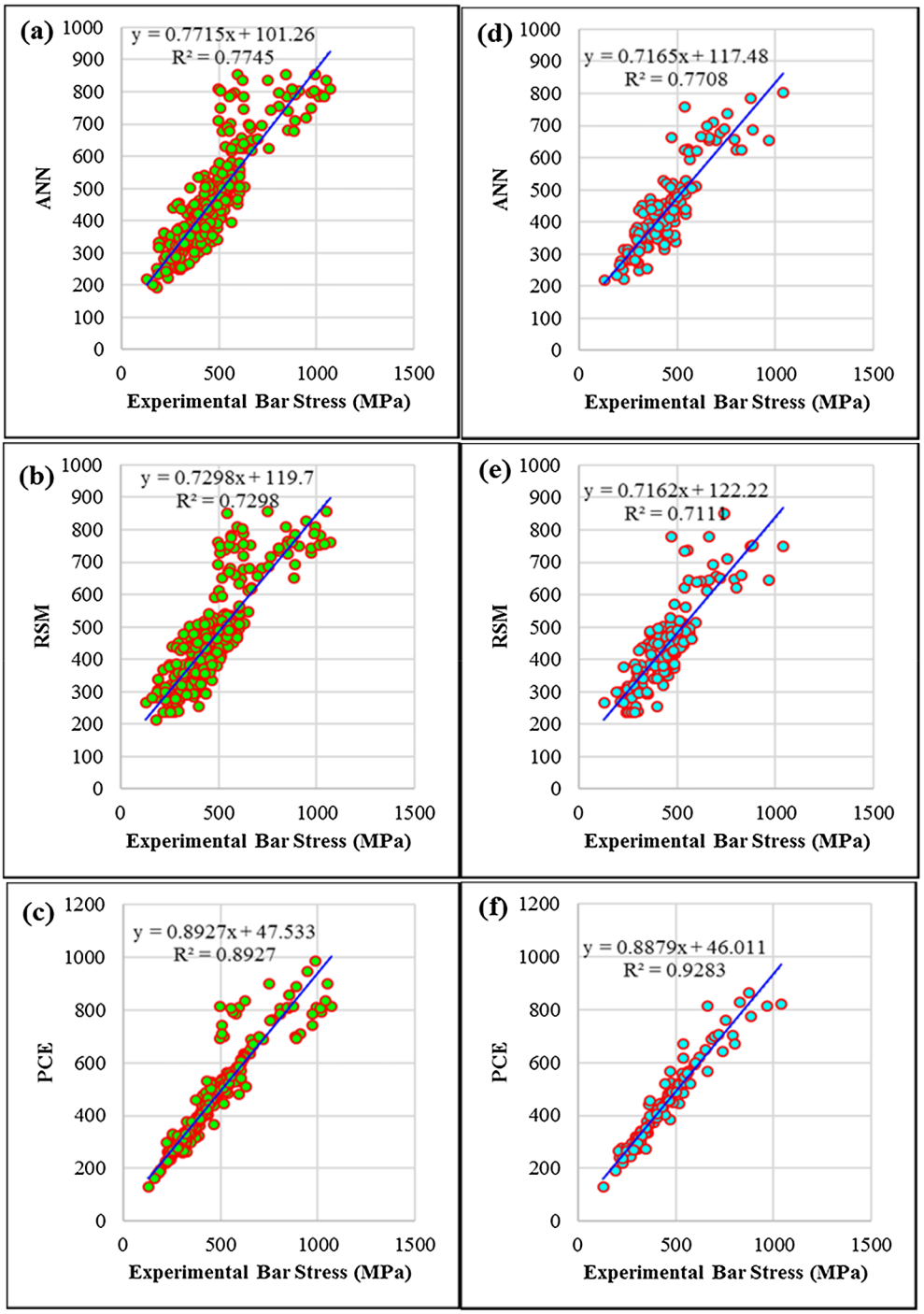

Fig. 3. Scatterplot (Scenario I) of (a, b and c) the applied modeling strategies (ANN, RSM and PCE) over the training phase (80% all data selected randomly), and (d, e, f) testing phase (20% all data selected randomly).

reinforcement, i.e.$\left[\frac{(cw + K_{atr})}{d_b}\right]$ on the bond strength. Compared to scenario I, the proposed models in scenario II better predicted the experimental test results. This is because the bond strength of the bar splices showing splitting failure is significantly affected by the concrete cover and transverse reinforcement. The trend of model predictions of the training phase was similar to that of the testing phase. In addition, the data-driven models are essentially based on regression analyses in which non-stationary and non-linear features can better be captured (related variables as predictors).

Figs. 5–6 and Table 3 visualize the calculations of various design code empirical equations versus the actual experimental rebar stress values. The experimental dataset was split into 80% training and 20% validation phases, in harmony with the models developed in the current study (training and testing phases), so that the validation can be performed in a more reliable manner. It can be observed that the predictions of the ACI 408R-03 [19] and Hwang et al. [4] models correlated well with the test results, regardless of the stress level of the rebar splice.

Fig. 7 compares the experimental rebar stress test results to corresponding calculations of current design codes and that of the models proposed in the present study. For direct comparison, safety factors were neglected in the calculations of design codes. ACI 318-14 [16], Eurocode 2 [17], and Model Code 2010 [18] generally showed ratios of

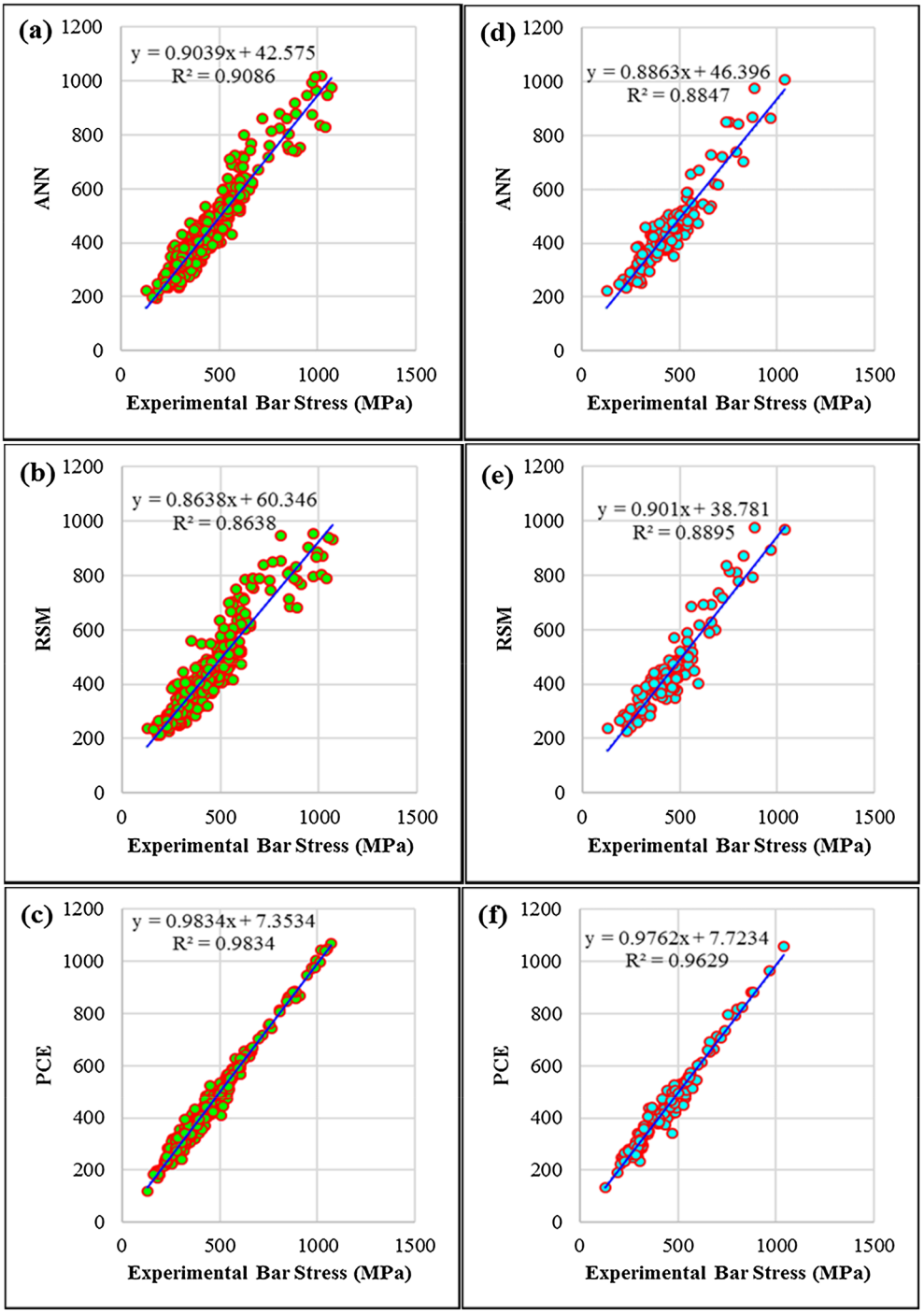

Fig. 4. Scatterplot (Scenario II) of (a, b and c) the applied modeling strategies (ANN, RSM and PCE) over the training phase (80% of all data selected randomly), and (d, e, f) testing phase (20% all data selected randomly).

Table 1
Statistical comparative indicators of different modeling strategies for scenario (I) over the training and testing phases.

Models	MAE	RMSE	MBE	md	mNSE	MAE	RMSE	MBE	md	mNSE
	Training phase					*Testing phase*				
ANN	56.785	78.687	0.030	0.761	0.541	59.476	78.686	0.015	0.739	0.515
RSM	66.320	86.142	0.037	0.714	0.464	68.007	87.182	0.028	0.705	0.445
PCE	21.549	54.282	0.008	0.912	0.826	24.300	44.085	0.000	0.899	0.802

Table 2
Statistical comparative indicators of different modeling strategies for scenario (II) over the training and testing phases.

Models	MAE	RMSE	MBE	md	mNSE	MAE	RMSE	MBE	md	mNSE
	Training phase					*Testing phase*				
ANN	37.595	50.106	0.014	0.909	0.845	44.720	55.162	0.010	0.808	0.635
RSM	46.281	61.163	0.019	0.864	0.809	42.605	54.149	0.008	0.820	0.652
PCE	14.698	21.350	0.003	0.983	0.940	22.779	31.419	−0.002	0.906	0.814

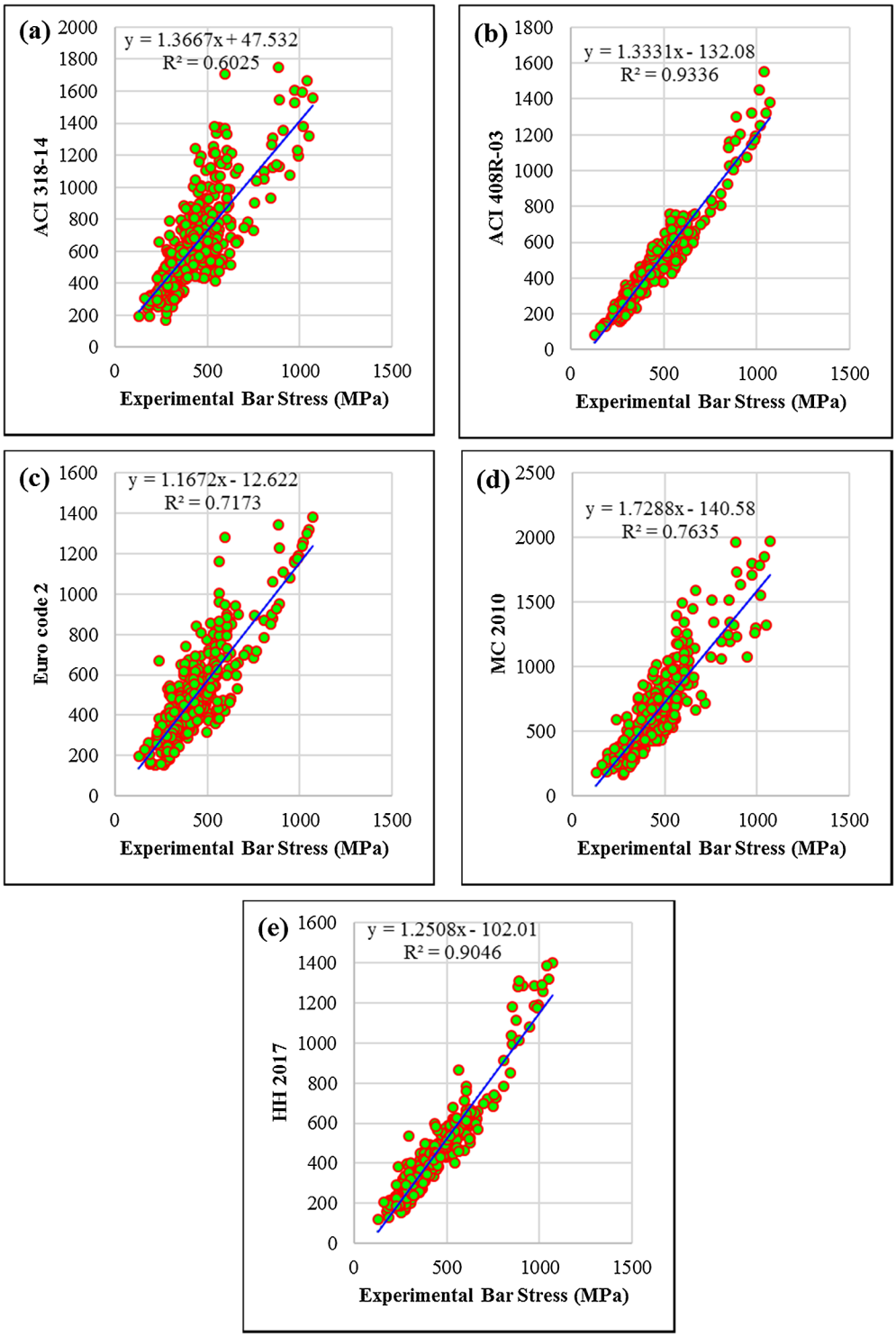

Fig. 5. Scatterplot of empirical design code methods over the training phase (80% all data points randomly) (a) ACI 318-14, (b) ACI 408R-03, (c) Eurocode 2, (d) Model Code 2010, and (e) Hwang et al. 2017 model.

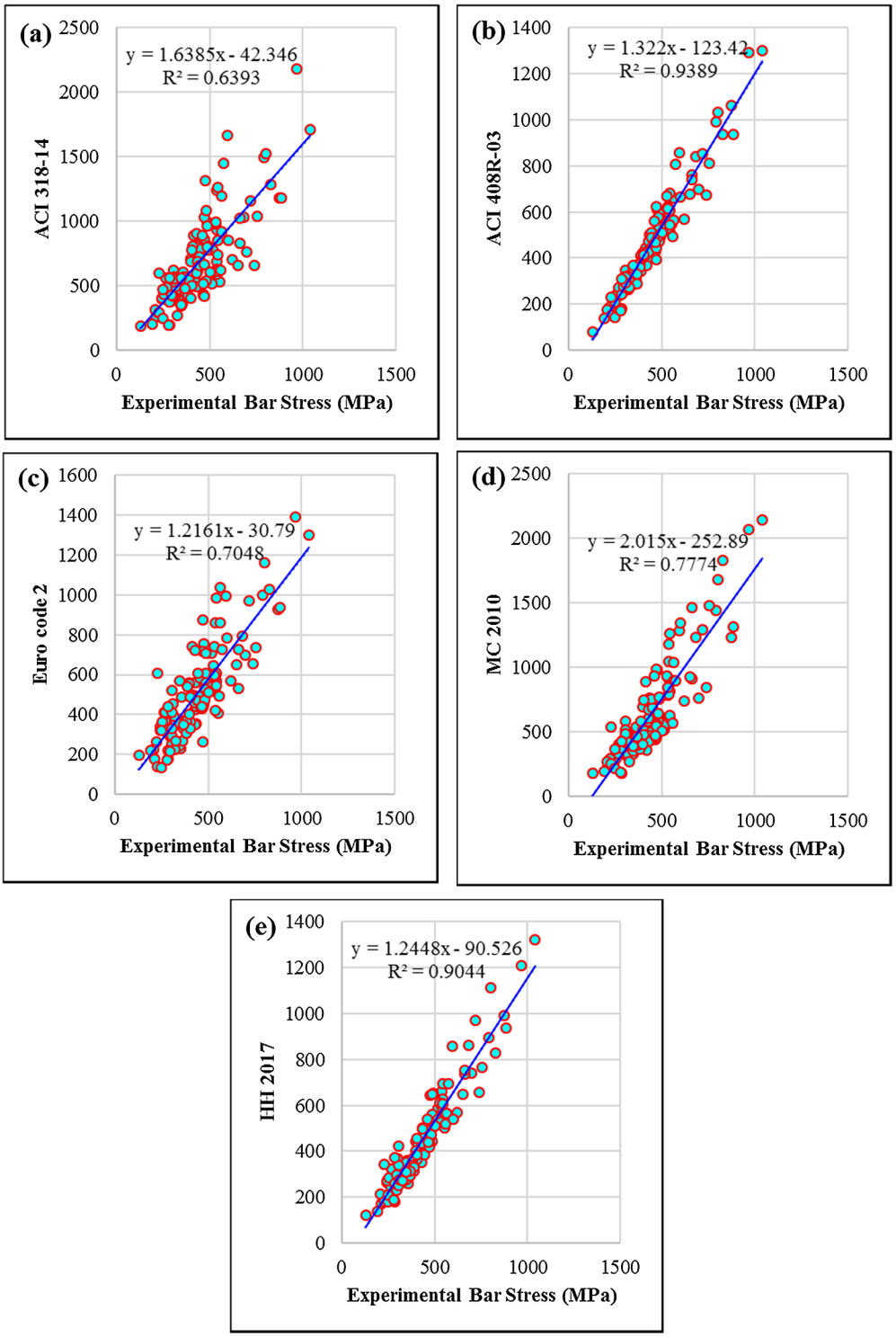

Fig. 6. Scatterplot of empirical/code methods over the testing phase (20% all data points which are selected randomly) (a) ACI 318-14, (b) ACI 408R-03, (c) Eurocode 2, (d) Model Code 2010, and (e) Hwang et al. 2017 model.

f_{test}/f_s greater than 1.0 in the small l_s/d_b range. This indicates that the design code methods (without safety factor) underestimated the tensile strength of bar splices related to the pull-out failure. Moreover, the design codes considered showed lower accuracy in their predictions. Conversely, the ACI 408R-03 [19] and the Hwang et al. [4] model predicted the test results with reasonable accuracy, showing an average f_{test}/f_s ratio = 1.002, and 1.005, and COV. = 0.151 and 0.155, respectively.

The predictions of the models proposed in this study, including RSM, ANN, and PCE were better than those of the existing methods, showing an average f_{test}/f_s ratio = 1.016, 1.013, and 1.005, and COV. = 0.143, 0.124, and 0.064, respectively. In particular, the PCE model showed the best accuracy of predictions amongst all methods considered in this study. While predicting well the test results, the proposed methods also tended to somewhat overestimate the experimental test results. Accordingly, a safety factor is necessary to consider the proposed methods for design applications.

To satisfy the design safety level, the value of "average –

Table 3
Statistical comparative indicators of different empirical/code formulations over the same data sets.

Models	MAE	RMSE	MBE	md	mNSE	MAE	RMSE	MBE	md	mNSE
	Training phase					*Testing phase*				
ACI 318-14	216.706	285.700	0.484	0.427	−0.752	248.388	329.904	0.534	0.387	−1.028
ACI 408R-03	52.805	82.179	0.000	0.815	0.573	50.773	78.106	0.010	0.823	0.586
Euro code 2	102.158	138.882	0.139	0.650	0.174	104.538	147.486	0.144	0.647	0.147
MC 2010	188.979	270.593	0.378	0.494	−0.528	205.149	311.433	0.393	0.467	−0.675
HH 2017	50.908	79.641	0.000	0.815	0.588	54.404	78.795	0.020	0.805	0.556

$K \times$ standard deviation" (where $K = 1.645$ for 5% fractile (quantile) with 90% confidence level of 534 specimens) should be greater than 1.0, so that the likelihood of failure would be less than 5%. For the proposed methods, and without the safety factor, the values are 0.777 (RSM), 0.806 (ANN), and 0.897 (PCE). Thus, in the case of the PCE method, a safety factor$\phi = 0.9$ can be considered for 5% fractile (quantile) with 90% confidence level. Although the RSM and ANN models attained good results in the testing phase, they require lower safety factors (e.g., about 0.75 to 0.8) considering their relatively lower accuracy compared to that of the PCE method. Lower safety factors imply that the lap splice length should be increased for design safety. Thus, the use of the PCE method can improve design accuracy and cost efficiency compared to all other methods investigated.

5. Parametric study

To explore the sensitivity of the PCE model to key parameters affecting the rebar bond strength, a parametric study was conducted whereby each influential parameter was varied within its practical range, while the other input parameters were maintained constant, and the model was then asked to predict the rebar bond strength.

Fig. 8-a illustrates the results of the parametric study for the effect of the concrete compressive strength on the bar bond strength for a ratio of bar splice length to bar diameter (l_s/d_b) of 10, 20, 40 and 70. The yield strength of the bar was maintained at 830 MPa and the $\left[\frac{(cw + K_{atr})}{d_b}\right]$ parameter was kept equal to 2. It can be observed that as expected, when the l_s/d_b ratio increased, the rebar bond strength also increased. Increasing the l_s/d_b ratio seemed to be most effective for increasing the maximum tensile stress in the bar, f_s. As the compressive strength increased, f_s increased linearly. This linear relationship is very interesting since until now, existing models and design codes have generally considered the effect of concrete strength as $(f_c')^{0.25}$ to $(f_c')^{0.5}$. It should be noted that, unlike existing models and design codes, the PCE model is data-driven and reflects the actual effect of the 534 points in the database.

Fig. 8-b portrays the effect of the bay yield strength on the maximum bar stress. The concrete compressive strength was maintained at 30 MPa and the $\left[\frac{(cw + K_{atr})}{d_b}\right]$ parameter was maintained equal to 2. The l_s/d_b ratio was 10, 20, 40 and 70. In general, existing models do not consider the effect of the bar yield strength on the bar bond strength. Only Model Code 2010 (factor η4 in clause 6.1.3.2), considers this aspect where the bond strength decreases as the yield strength increases. However, the present parametric study based on polynomial chaos expansions shows an opposite relationship. Except in the proximity of the rebar yielded range, a linear relationship can be defined. While this result emulates from the data drive model, it is rather difficult to explain the mechanism controlling this behavior.

Fig. 8c captures the effect of the parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$ on the maximum tensile stress in the bar f_s. The concrete compressive strength was maintained at 30 MPa and the l_s/d_b ratio was 10, 20, 40 and 70. The maximum tensile stress in the bar, f_s increased with increased$\left[\frac{(cw + K_{atr})}{d_b}\right]$. Except in the proximity of the bar yield stress, this increase can be approximated by a linear function, which is similar to existing models and design code equations. One discrepancy observed is that since $f_y = 830$ MPa was used, f_s should be less than f_y. However, in the case of $l_s/d_b = 70$, f_s was greater than f_y. Nonetheless, this model is data driven and by examining the database, a significant number of experimental outputs justifies the appearance of the above noted discrepancy.

Fig. 9 exhibits results of the parametric study whereby the l_s/d_b ratio was increased for various levels of concrete compressive strength (Fig. 9a), bar yield strength (Fig. 9b), and parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$ (Fig. 9c). It can be observed in Fig. 9a that as l_s/d_b increased, f_s increased non-linearly. This is unlike in design codes in which this increase is rather linear due to the common assumption of uniform bond stress distribution.

In Fig. 9b, the effect of the l_s/d_b ratio on f_s is shown when the bar yield stress varied from 350, 550, 750 and 950 MPa. Again, the concrete compressive strength was maintained at 30 MPa and the parameter was kept equal to $\left[\frac{(cw + K_{atr})}{d_b}\right]$ 2. As the l_s/d_b ratio increased, f_s also increased, and this increase was more significant the higher the bar yield stress. Except in the proximity of the bar yield stress, the relationship between l_s/d_b and f_s can be approximated to be linear, which is generally similar to existing models and design codes.

Fig. 9c captures the relationship between l_s/d_b and f_s when the parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$is varied between 1, 2, 3 and 4. The concrete compressive strength was maintained at 30 MPa and f_y was kept at 830 MPa. The parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$ had an influential effect on f_y as observed earlier. As the l_s/d_b ratio increased, f_s also increased. Again; this increase can be considered linear, except near the bar yield stress. The observed discrepancy of f_s exceeding f_y was discussed earlier.

It is worth exploring the computational effort for computing the unknown coefficients (weights) for the established PCE, RSM and ANN models. Scenario I was selected as an example to quantify the computational cost for each model and the results are reported in Table 4. The ANN needs a feedforward backpropagation algorithm to define best connections between neurons in the training process. Conversely, the PCE and RSM models are directly computed using a least square estimator. On the other hand, the PCE model needs to compute the multivariate orthogonal polynomial as input data, which requires higher computational effort for higher order multivariate orthogonal polynomials. It can be observed in Table 4 that, at the optimal order of $p = 5$, the PCE model was much more computationally effective than the ANN model. However, when increasing the order p to 8, the PCE model became more computationally intensive than the ANN model. The RSM was the most computationally effective. However, using high-order polynomial functions would also make the RSM computationally intensive, but this was not investigated herein.

Finally, the practical relevance of the PCE model needs to be clarified. First, the proposed PCE model could be readily used in design subject to the application of a safety factor. Owing to its superior predictive capability, this safety factor would be smaller than that in existing design provisions, hence averting overly conservative design and its associated cost. Moreover, the proposed PCE model is highly flexible. This intelligent data intensive framework for predicting rebar average bond stress and full embedment length can better accommodate new

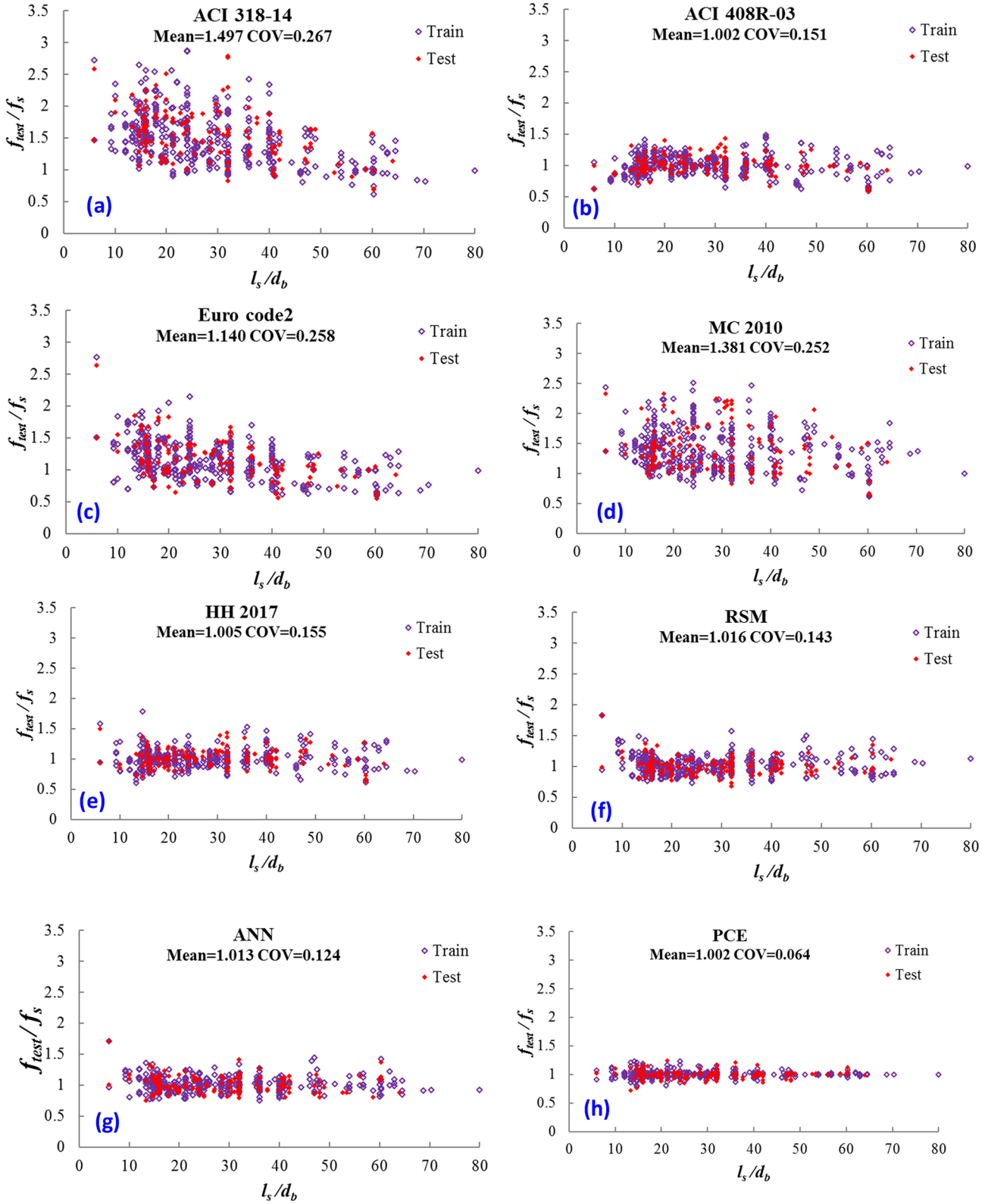

Fig. 7. The f_{test}/f_s ratio against l_s/d_b ratio: (a) ACI 318-14, (b) ACI 408R-03, (c) Eurocode 2, (d) MC 2010, (e) Huang et al. 2017, (f) RSM, (g) ANN, and (h) PCE.

data and new rebar materials and configurations for calibrating existing design provisions and keeping design codes relevant. Moreover, we are entering an era of the fourth industrial revolution and witnessing a fusion of technologies that is blurring boundaries between the physical, digital and biological spheres. Emerging technology breakthroughs such as robotics, artificial intelligence, nanotechnology, quantum computing, biotechnology, the Internet of Things, decentralized consensus, additive manufacturing, etc., are changing the way we do things. Hence, intelligent models such as the proposed PCE tool could find a place in automated and semi-automated design platforms of the future [37–40].

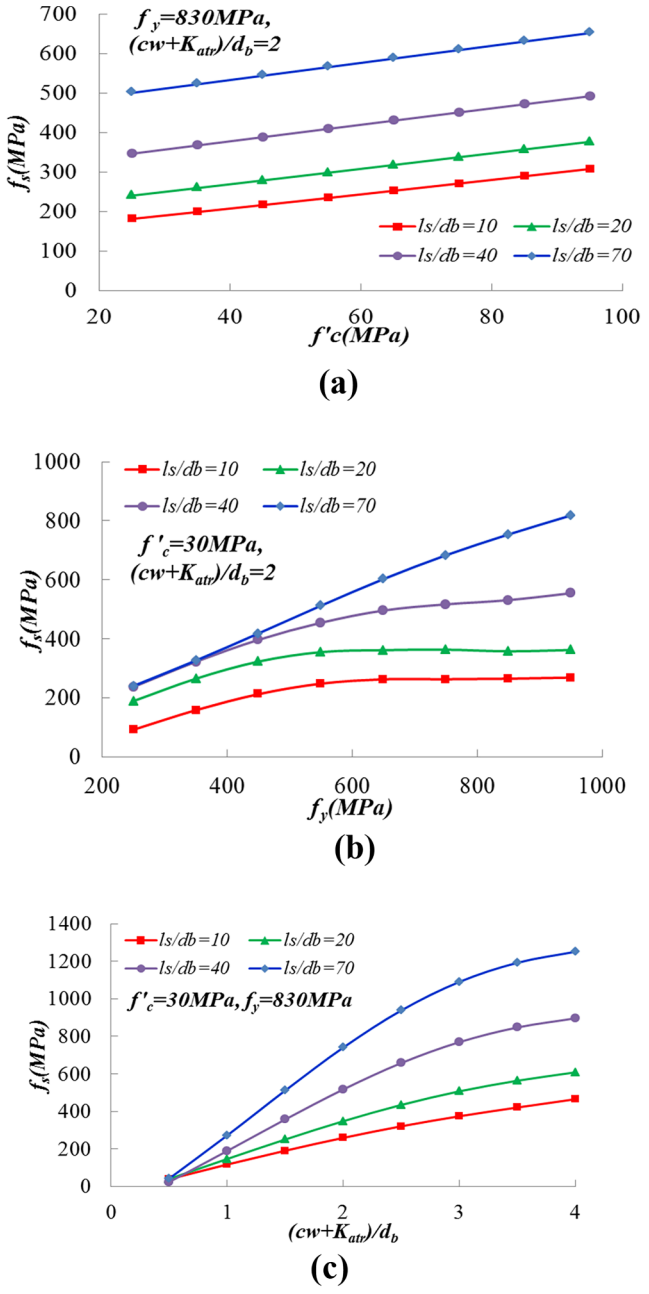

Fig. 8. Bond strength of bar splices with respect to: (a) concrete compressive strength f_c', (b) yield strengths of the reinforcing bar f_y, and (c) parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$ for different lap splice length to reinforcing bar diameters (l_s/d_b) of 10, 20, 40 and 70.

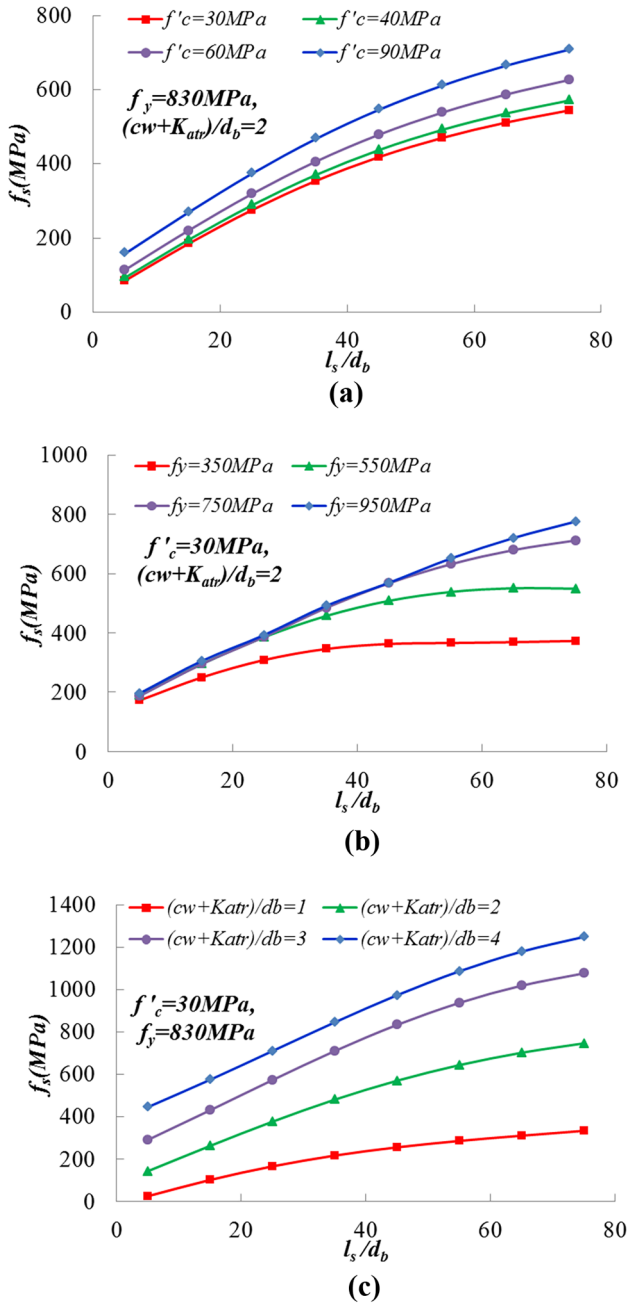

Fig. 9. Effects of lap splice length-to-reinforcing bar diameter (l_s/d_b) on average bond strength of bar splices for (a) different concrete compressive strengths f_c', (b) different yield strengths of reinforcing bar f_y, and (c) different values of parameter $\left[\frac{(cw + K_{atr})}{d_b}\right]$.

6. Summary and concluding remarks

The reinforcing bar bond stress and rebar development length are fundamental concepts in reinforced concrete design codes and guidelines. However, the bond stress of a reinforcing bar anchored in a cementitious matrix varies along the bar length and cannot be easily quantified. Hence, design codes rather refer to the concept of bar development length and average bond stress. Moreover, design provisions for bond strength are generally based on statistical analysis of bar pullout test results. In the present study, a novel statistical predictive model based on Polynomial Chaos Expansions (PCE) was developed to predict the reinforcing bar development length using 534 experimental results of simple pullout tests on short unit bar lengths retrieved from

Table 4
Computational effort for computing the unknown coefficients (weights) of the developed PCE, RSM and ANN models.

Models	PCE					RSM	ANN
	P = 3	P = 4	P = 5	P = 6	P = 7		
Times to build models (seconds)	0.11	0.12	0.47	2.77	14.68	0.07	10.82

the open literature. Its predictions were compared to calculations of commonly used design code methods (i.e., ACI 318-14 [16], ACI 408R-03 [19], and Eurocode 2 [17]) and bond models including [4,16–19]. A parametric study was conducted using the proposed PCE model to explore its sensitivity to key input parameters. The following conclusions can be drawn:

- Among existing methods, predictions of the ACI 408R-03 [19] and the Hwang et al. [4] model better correlated with the experimental test results, regardless of the stress level of the rebar splice.
- The capability of the novel model based on Polynomial Chaos Expansions proposed in this study in predicting experimental rebar stress outperformed all existing methods considered in this study, as well as the response surface model and the artificial neural network model.
- The proposed model showed adequate sensitivity to influential input parameters. While its trends generally agreed with that of existing design provisions, it captured certain features not identified by previous models which assume unrealistic uniform bond stress along the reinforcing bar.
- The proposed model could be used in design upon the application of a safety factor. Considering the superior predictive capability of the novel model, this safety factor can be smaller than that of existing methods. This can prevent overly conservative design and should thus lead to more cost effective design.
- The proposed PCE model is highly flexible. It can offer an intelligent data intensive framework for predicting rebar average bond stress and full embedment length, with ability to accommodate new data and new rebar materials for calibrating existing design provisions and keeping design codes relevant.

References

[1] Saleem MA, Mirmiran A, Xia J, Mackie K. Development length of high-strength steel rebar in ultrahigh performance concrete. J Mater Civ Eng 2013;25:991–8. https://doi.org/10.1061/(ASCE)MT.1943-5533.0000571.

[2] Alam MS, Moni M, Tesfamariam S. Seismic overstrength and ductility of concrete buildings reinforced with superelastic shape memory alloy rebar. Eng Struct 2012;34:8–20. https://doi.org/10.1016/j.engstruct.2011.08.030.

[3] Wight JK, MacGregor JG. Reinforced Concrete, Mechanics and Design, Section 6–3; 2012.

[4] Hwang HJ, Park HG, Yi WJ. Nonuniform bond stress distribution model for evaluation of bar development length. ACI Struct J 2017;114:839–49. https://doi.org/10.14359/51689446.

[5] Wiener N. The homogeneous chaos. Am J Math 1938;60:897–936.

[6] Hu C, Youn BD. Adaptive-sparse polynomial chaos expansion for reliability analysis and design of complex engineering systems. Struct Multidiscip Optim 2011;43:419–42. https://doi.org/10.1007/s00158-010-0568-9.

[7] Sudret B. Global sensitivity analysis using polynomial chaos expansions. Reliab Eng Syst Saf 2008;93:964–79. https://doi.org/10.1016/j.ress.2007.04.002.

[8] Kewlani G, Crawford J, Iagnemma K. A polynomial chaos approach to the analysis of vehicle dynamics under uncertainty. Veh Syst Dyn 2012;50:749–74. https://doi.org/10.1080/00423114.2011.639897.

[9] Najm HN. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Annu Rev Fluid Mech 2009;41:35–52. https://doi.org/10.1146/annurev.fluid.010908.165248.

[10] Jacquelin E, Adhikari S, Sinou JJ, Friswell MI. Polynomial chaos expansion in structural dynamics: accelerating the convergence of the first two statistical moment sequences. J Sound Vib 2015;356:144–54. https://doi.org/10.1016/j.jsv.2015.06.039.

[11] Witteveen JAS, Sarkar S, Bijl H. Modeling physical uncertainties in dynamic stall induced fluid-structure interaction of turbine blades using arbitrary polynomial chaos. Comput Struct 2007;85:866–78. https://doi.org/10.1016/j.compstruc.2007.01.004.

[12] Fajraoui N, Ramasomanana F, Younes A, Mara TA, Ackerer P, Guadagnini A. Use of global sensitivity analysis and polynomial chaos expansion for interpretation of nonreactive transport experiments in laboratory-scale porous media. Water Resour Res 2011;47.

[13] Laloy E, Rogiers B, Vrugt JA, Mallants D, Jacques D. Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion. Water Resour Res 2013;49:2664–82.

[14] Sochala P, Le Maître OP. Polynomial Chaos expansion for subsurface flows with uncertain soil parameters. Adv Water Resour 2013. https://doi.org/10.1016/j.advwatres.2013.10.003.

[15] Keshtegar B, Heddam S, Hosseinabadi H. The employment of polynomial chaos expansion approach for modeling dissolved oxygen concentration in river. Environmental Earth Sciences 2019;78:34.

[16] ACI Committee 318. Building Code Requirements for Structural Concrete (ACI 318-14) and Commentary (ACI 318R-14). 2014. doi: http://doi.org/10.1016/0262-5075(85)90032-6.

[17] BS EN 1992-1-1. Eurocode 2: Design of concrete structures - Part 1-1 : General rules and rules for buildings. British Standards Institution 2004;1:230. doi:[Authority: The European Union Per Regulation 305/2011, Directive 98/34/EC, Directive 2004/18/EC].

[18] Code M. fib Model Code for Concrete Structures 2010. Struct Concr 2010;14:1–402. https://doi.org/10.1002/9783433604090.

[19] Joint ACI-ASCE Committee 408. Bond and Development of Straight Reinforcing Bars in Tension (ACI 408R-03). Detroit, Michigan, US: American Concrete Institute; 2002.

[20] Orangun CO, Jirsa JO, Breen JE. A revaluation of test data on development length and splices. J Am Concr Instit 1977. https://doi.org/10.14359/10993.

[21] Wang S, Huang GH, Baetz BW, Huang W. A polynomial chaos ensemble hydrologic prediction system for efficient parameter inference and robust uncertainty assessment. J Hydrol 2015;530:716–33.

[22] Lasota R, Stocki R, Tauzowski P, Szolc T. Polynomial chaos expansion method in estimating probability distribution of rotor-shaft dynamic responses. Bull Polish Acad Sci: Tech Sci 2015;63:413–22. https://doi.org/10.1515/bpasts-2015-0047.

[23] Keshtegar B, Kisi O. RM5Tree: radial basis M5 model tree for accurate structural reliability analysis. Reliab Eng Syst Saf 2018. https://doi.org/10.1016/j.ress.2018.06.027.

[24] Keshtegar B, Heddam S. Modeling daily dissolved oxygen concentration using modified response surface method and artificial neural network: a comparative study. Neural Comput Appl 2017:1–12. https://doi.org/10.1007/s00521-017-2917-8.

[25] Keshtegar B, Kisi O, Asce M. Modified response-surface method : new approach for modeling pan evaporation. J Hydrol Eng 2017;22:1–14. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001541.

[26] Morfidis K, Kostinakis K. Approaches to the rapid seismic damage prediction of r/c buildings using artificial neural networks. Eng Struct 2018;165:120–41. https://doi.org/10.1016/j.engstruct.2018.03.028.

[27] Pathirage CSN, Li J, Li L, Hao H, Liu W, Ni P. Structural damage identification based on autoencoder neural networks and deep learning. Eng Struct 2018;172:13–28. https://doi.org/10.1016/j.engstruct.2018.05.109.

[28] Allahyari H, Nikbin IM, Rahimi S, Heidarpour A. A new approach to determine strength of Perfobond rib shear connector in steel-concrete composite structures by employing neural network. Eng Struct 2018. https://doi.org/10.1016/j.engstruct.2017.12.007.

[29] Kurt H, Kayfeci M. Prediction of thermal conductivity of ethylene glycol-water solutions by using artificial neural networks. Appl Energy 2009;86:2244–8. https://doi.org/10.1016/j.apenergy.2008.12.020.

[30] Hemmat Esfe M, Afrand M, Yan W-M, Akbari M. Applicability of artificial neural network and nonlinear regression to predict thermal conductivity modeling of Al2O3–water nanofluids using experimental data. Int Commun Heat Mass Transfer 2015;66:246–9. https://doi.org/10.1016/j.icheatmasstransfer.2015.06.002.

[31] Durairaj M, Thamilselvan P. Applications of artificial neural network for IVF data analysis and prediction. J Eng Comput Appl Sci (JEC&AS) 2013;2:11–5.

[32] Ghorbani MA, Deo RC, Karimi V, Yaseen ZM, Terzi O. Implementation of a hybrid MLP-FFA model for water level prediction of Lake Egirdir, Turkey. Stoch Env Res Risk Assess 2017:1–15. https://doi.org/10.1007/s00477-017-1474-0.

[33] Yaseen ZM, Ramal MM, Diop L, Jaafar O, Demir V, Kisi O. Hybrid adaptive neuro-fuzzy models for water quality index estimation. Water Resour Manage 2018. https://doi.org/10.1007/s11269-018-1915-7.

[34] Yaseen ZM, Awadh SM, Sharafati A, Shahid S. Complementary data-intelligence model for river flow simulation. J Hydrol 2018;567:180–90. https://doi.org/10.1016/j.jhydrol.2018.10.020.

[35] Ernst OG, Mugler A, Starkloff H-J, Ullmann E. On the convergence of generalized polynomial chaos expansions. ESAIM. Math Model Numer Anal 2012;46:317–39. https://doi.org/10.1051/m2an/2011045.

[36] Ghosh D, Ghanem R. Stochastic convergence acceleration through basis enrichment of polynomial chaos expansions. Int J Numer Meth Eng 2008;73:162–84. https://doi.org/10.1002/nme.2066.

[37] Suleiman AR, Nehdi ML. Modeling self-healing of concrete using hybrid genetic algorithm – artificial neural network. Materials 2017;10: 135:15 p.

[38] Nehdi ML, Soliman AM. Artificial intelligence model for early-age autogenous shrinkage of concrete. ACI Mater J 2012;109(3):353–61.

[39] Shahriar A, Nehdi ML. Modelling rheological properties of oilwell cement slurries using artificial neural networks. J Mater Civ Eng ASCE 2011;23(12):1703–10.

[40] Nehdi ML, El Chabib H, Said AM. Predicting effect of stirrups on shear strength of reinforced NSC and HSC slender beams using artificial intelligence. Can J Civ Eng 2006;33(8):933–44.