ASCE

Data-Driven Approach to Predict the Plastic Hinge Length of Reinforced Concrete Columns and Its Application

De-Cheng Feng, A.M.ASCE[1]; Barbaros Cetiner[2]; Mohammad Reza Azadi Kakavand[3]; and Ertugrul Taciroglu, M.ASCE[4]

Abstract: Inelastic response of reinforced concrete columns to combined axial and flexural loading is characterized by plastic deformations localized in small regions, which are idealized as plastic hinges. Under extreme events such as earthquakes, the load-carrying and deformation capacities of reinforced concrete beam/columns are highly dependent on the accuracy of this idealization for which the plastic hinge length is a key parameter. From a design perspective, a reinforced concrete column can only attain the ductility characteristics prescribed by its performance level if it is provided with sufficient confinement along the length of its plastic hinge zones. From an analysis standpoint, an efficient, nonlocalized, and objective finite-element simulation of column behavior requires accurate plastic hinge length definitions. This paper presents a novel data-driven model for predicting the plastic hinge length of reinforced concrete columns and its implementation in force-based fiber beam-column elements. The model is based on an ensemble machine learning algorithm named adaptive boosting (AdaBoost) and is trained using the results of 133 reinforced concrete column tests conducted in the period from 1984 to 2013. The performance of the model is assessed using the 10-fold cross-validation technique. It is shown that the prediction accuracy achieved using the proposed method is considerably higher than those of state-of-the-art empirical relationships and several other highly effective machine learning base models. Furthermore, numerical experiments reveal that the force-based beam-column models using plastic hinge length predictions of the developed model closely resemble the monotonic and cyclic behavior observed in laboratory experiments. **DOI: 10.1061/(ASCE)ST.1943-541X.0002852.**

Author keywords: Plastic hinge length; RC columns; Machine learning; Adaptive boosting; Fiber element analysis.

Introduction

Because a RC column experiences lateral deformations, regions of the column subjected to moments exceeding the member's yielding capacity undergo yielding of reinforcement and crushing of concrete over finite lengths around these regions. Under extreme loading conditions, these finite lengths, referred to as plastic hinge zones (or plastic hinges), sustain large inelastic curvatures, typically assumed constant throughout the plastic hinge. Thus, to ensure ductile structural behavior, the extent of plastic hinge zones are determined, and extensive confinement is provided in these regions.

The extent of a plastic hinge zone can be defined in terms of plastic hinge zone length (PZL) or plastic hinge length (PHL). PZL refers to the full length of a plastic hinge zone, whereas for a column, the PHL marks the location where the plastic rotations along a plastic hinge zone can be concentrated into a hinge. PHL is measured from member ends, and is typically 0.5 PZL. In the present paper, PZL is henceforth assumed to be calculated in terms of PHL when relevant.

Performance assessment of RC columns also requires an understanding of the PHL. Inelastic displacement demand on a structural member can be calculated by integrating its curvature field only if the length and nonlinear response of its plastic hinge zones are known. Given the dependency of inelastic finite-element models appropriate for determining plastic rotations on PHL, analyzing column performance as well is tightly coupled with the knowledge of PHL. For obtaining computationally efficient and mesh-independent numerical solutions for inelastic beam-column behavior, the most common approach is using distributed plasticity beam-column elements based on the force-based fiber section formulation (Neuenhofer and Filippou 1997; Hjelmstad and Taciroglu 2003). To accommodate strain-softening or strain-hardening responses, these types of elements require numerically consistent regularization based on a realistic knowledge of PHL (Scott and Hamutçuoğlu 2008; Feng et al. 2016; Feng and Ren 2017).

In the last 4 decades, numerous empirical PHL estimation relationships have been developed, e.g., by Park et al. (1982), Priestley and Park (1987), Paulay and Priestley (1992), Sheikh and Khoury (1993), Sheikh et al. (1994), Panagiotakos and Fardis (2001), Lu et al. (2005), Berry et al. (2008), Bae and Bayrak (2008), Elmenshawi et al. (2012), and Ning and Li (2015). In general terms, these relationships were derived by fitting a statistical model to experimental results, where the model parameters are selected in

[1]Associate Professor, Key Laboratory of Concrete and Prestressed Concrete Structures of the Ministry of Education, Southeast Univ., Nanjing 211189, China. ORCID: https://orcid.org/0000-0003-3691-6128. Email: dcfeng@seu.edu.cn

[2]Dept. of Civil and Environmental Engineering, Univ. of California, Los Angeles, CA 90095. ORCID: https://orcid.org/0000-0002-9726-8120

[3]Unit of Strength of Materials and Structural Analysis, Institute of Basic Sciences in Engineering Sciences, Univ. of Innsbruck, Innsbruck 6020, Austria.

[4]Professor, Dept. of Civil and Environmental Engineering, Univ. of California, Los Angeles, CA 90095 (corresponding author). ORCID: https://orcid.org/0000-0001-9618-1210. Email: etacir@ucla.edu

Note. This manuscript was submitted on August 30, 2019; approved on July 14, 2020; published online on November 30, 2020. Discussion period open until April 30, 2021; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Structural Engineering***, © ASCE, ISSN 0733-9445.

recognition of the mechanisms involved in plastic hinge development. In reality, the spreading of yielding over a plastic hinge zone is a complicated phenomenon dependent on a combination of factors including moment gradient, tension shift, and strain penetration (Ning and Li 2015). As a result, it is difficult to capture the processes leading to plastic hinge formation using a limited number—typically five or fewer—of parameters, as in the PHL relationships mentioned previously.

Selecting model parameters is another source of complication. There is a lack of consensus among the research community over which one of the geometric, material, and section properties govern plastic hinge formation for columns. For instance, Priestley and Park (1987) and Mendis (2002) concluded that the PHL is insensitive to the axial load ratio, whereas Paultre et al. (2001) and Pam and Ho (2009) reported that the PHL increases with axial load ratio. A direct effect of this difference in understanding is a high variation of model parameters, and hence the wide variability in PHL predictions made by the existing models.

With the growing advancements in computational capabilities, availability of more comprehensive data sets, and better algorithms, machine learning (ML) techniques offer a compelling alternative to traditional statistical modeling for predicting PHL. ML provides an algorithmic approach for regression problems where a preconceived relationship between predictors and outcomes cannot otherwise be determined. Irrespective of the complexity of data-generating processes, ML focuses on prediction by using multipurpose learning algorithms to find patterns in the data set at hand. Several successful regression-type applications of ML exist in the literature for structural engineering problems. To name a few, Siddique et al. (2011), Pham et al. (2015), and Omran et al. (2016) used regression tree models, support vector machines (SVM), and artificial neural networks (ANN) to predict the compressive strength of different types of concrete matrices. Jeon et al. (2014) developed a shear strength model for RC beam-column joints using multivariate adaptive regression splines and symbolic regression. Vu and Hoang (2016) established a model for estimating punching shear capacity of fiber-reinforced polymer (FRP) reinforced concrete slabs using SVM. Santos et al. (2017) used ANN to detect structural anomalies in bridge structures. A more detailed review of ML applications in structural engineering has been given by Salehi and Burgueño (2018).

Various ML techniques exist for regression problems. The most widely used ML algorithms are the ANN family (Schalkoff 1997), SVM family (Hearst et al. 1998), and classification and regression trees (CART) family (Safavian and Landgrebe 1991) methods. In principle, these algorithms prescribe individual learners pursuing different objectives (i.e., each method has a specific way of adjusting its internal parameters to increase prediction performance) to establish a single prediction model. Each procedure and the models it generates have their strengths and weaknesses. By combining outputs of multiple ML techniques to benefit from the distinct capabilities of individual algorithms, another set of meta-algorithms, named ensemble learning methods, are formed. In ensemble learning, individual ML algorithms serve as the base learners, and their predictions are combined to create a single predictive model.

Numerous studies have demonstrated that models established using a committee of learners very often attain higher prediction accuracy and better robustness over individual learning algorithms (Zhou 2012; Erdal et al. 2013; Chou et al. 2014). According to Thomas (1997), the ability of ensemble methods to better generalize processes is due to three primary reasons. First, using a committee of decision makers, ensemble methods can compensate for the weaknesses of individual learners. Second, when the training data are small in size, obtaining a unique best learner might not be possible with individual learning algorithms. Multiple models may produce results of comparable accuracy. Combining various models yielding similar prediction performance may lead to more robust results, and hence the best learner. Lastly, a base learning method can be merely incapable of modeling a target process, yet an aggregation of multiple predictions from the same method may provide better approximations.

In this paper, a PHL prediction model for RC columns that outperforms the existing PHL estimation relationships is developed using the adaptive boosting (AdaBoost) ensemble learning technique. Ensemble methods are briefly introduced, and the AdaBoost algorithm is discussed in full detail. The training data set based on 133 reinforced concrete column tests conducted by various researchers in the period from 1984 to 2013 and compiled by Ning and Li (2015) is presented. An objective training procedure based on 10-fold cross-validation (CV) is defined and utilized to train the model based on the set of hyperparameters that maximize its prediction performance. Results from existing empirical relationships and models obtained using several well-known base learning methods are established to validate the accuracy of the AdaBoost model. Subsequently, a finite-element implementation is performed using distributed plasticity beam-column elements based on the force-based fiber section formulation and PHL estimations obtained from the developed AdaBoost model. The monotonic and cyclic responses of RC columns modeled based on AdaBoost predictions are examined, and the ability of these models to capture experimental results is displayed.

Novel Approach for Plastic Hinge Length Prediction

Ensemble Learning Methods

In general, constructing an ensemble learner consists of two steps. First, a set of base learners are generated using individual learning algorithms. Depending on the ensemble method, the learning algorithms used for base learners can be of the same (homogeneous) or different (heterogeneous) kind. In the second step, to establish the predictive model, outputs of the individual learners are either combined according to a set of rules or passed on to another group of learning algorithms as training features.

Ensemble learning techniques for regression can be categorized into three groups based on how the weak learners are trained and combined: bootstrap aggregation (bagging) (Breiman 1996a), boosting (Drucker et al. 1994; Freund and Schapire 1997), and stacked generalization (stacking) (Breiman 1996b). Bagging uses a collection of bootstrap samples for training the base learners. In bagging, bootstrap samples are obtained by randomly subsampling with replacement from the training data set where the likelihood of selection (weight) for each example is uniform. The base learners are trained in parallel, and their outputs are combined through simple averaging.

Boosting, on the other hand, fits reweighted versions of bootstrap samples from the training data to a sequence of weak learners. On the first round of training, a weak learner is trained using a bootstrap sample drawn with equal weighting coefficients for each example. After each boosting iteration, subsampling weights for data that are incorrectly predicted is increased, a new bootstrap sample is generated, and a new weak learner is trained. Once all the weak learners are trained, learners are aggregated through a weighted sum or median to produce the final model. In conjoining weak learners, the ones that better predict the data set are assigned a higher weight and vice versa. Unlike bagging and boosting, in

general, stacking utilizes heterogeneous learners. Stacking aggregates multiple regression models via a metaregressor. In stacking, base-level models are trained on a complete training set. The outputs of the base-level models then serve as the training features for the metaregressor. Stacking regressions can consist of multiple parallel or sequential processes at both the base level and metaregressor levels.

Evidently, in bagging, training of base learners is an independent process. In contrast, in boosting and stacking, base learners are strongly interdependent. The training process for each regressor is affected by the performance of the regressors from earlier rounds of training. From a computational efficiency viewpoint, the bagging method results in highly parallelized training. However, from the standpoint of prediction quality, bagging ensembles mainly result in reduced model variance, whereas boosting and stacking ensembles yield improvements in both model variance and bias (Zhou 2009).

In addition, several studies (Drucker 1997; Zenko et al. 2001; Solomatine and Shrestha 2004; Shrestha and Solomatine 2006) demonstrated that boosting and stacking methods consistently achieve better prediction accuracies over bagging ensembles. One downside of stacking ensembles is that, as a result of their deep and more complex architecture, they generally require a large training set (Ren et al. 2016). This study aims to create a predictive model for PHL of RC columns that improves over the existing relationships in terms of prediction accuracy, model variance, and bias. Given that data limitations mostly rule out stacking techniques and boosting algorithms bear a proven consistency in satisfying the modeling objectives of this study, boosting, in particular the AdaBoost algorithm, is determined as the most suitable tool for this application.

Boosting Algorithm: AdaBoost

Adaboost, first proposed by Freund and Schapire (1997), is a robust, flexible, and easy-to-interpret algorithm that has a demonstrated ability to satisfy the prediction requirements of various applications (Wu et al. 2008). The fundamental idea behind AdaBoost is simple. A strong learner can be attained by generating and combining a series of weak learners trained using reweighted bootstrap samples from the training set. The adaptive aspect of AdaBoost is reflected in the adjustments performed at the end of each training round to the subsampling weights used for generating the training set for the next weak learner. The weights for the incorrectly predicted examples in a training round are increased in later rounds such that more emphasis is placed on those challenging examples in the subsequent rounds of training and vice versa.

Define the training data set Θ_{Tr} as

$$\Theta_{Tr} = \{(\boldsymbol{X}_1, Y_1), (\boldsymbol{X}_2, Y_2), \ldots, (\boldsymbol{X}_k, Y_k)\} \tag{1}$$

where $\boldsymbol{X}_i$ = input (independent) variables; Y_i = output (dependent) variable; and k = total number of examples in the data set.

The goal is to compute a strong learner $F(\boldsymbol{X})$ that can predict Y. AdaBoost defines $F(\boldsymbol{X})$ as a combination of T weak learners as follows:

$$F(\boldsymbol{X}) = \sum_{t=1}^{T} G\{\alpha_t f_t(\boldsymbol{X})\} \tag{2}$$

where for $t \in [1, T]$, $f_t(\boldsymbol{X})$ and α_t = tth weak learner and corresponding learner weight, respectively; and $G\{\cdot\}$ = combination rule.

Given that each weak learner is trained on bootstrap samples drawn from the complete training set Θ_{Tr} using a set of subsampling weights $\boldsymbol{D}_t$ at a boosting iteration t, the weak learner $f_t(\boldsymbol{X})$ can be obtained by

$$f_t(\boldsymbol{X}) = \mathcal{L}(\Theta_{Tr}, \boldsymbol{D}_t) \tag{3}$$

where $\mathcal{L}(\cdot)$ = individual learning algorithm utilized by AdaBoost to train each weak learner; and $\boldsymbol{D}_t = \{w_{t,1}, w_{t,2}, \ldots, w_{t,k}\}$ = weight distribution of the training examples, in which $w_{t,i}$ denotes the subsampling weight of the ith example in the training data set Θ_{Tr}. For the first boosting iteration (i.e., $t = 1$), a uniform subsampling weight distribution of $w_{1,i} = 1/k$ is specified. Theoretically, any type of individual learning algorithm (e.g., ANN, SVM, or CART) can be used for $\mathcal{L}(\cdot)$, and $\boldsymbol{D}_t$ depends on the performance of the learners from the previous training rounds. In this paper, CART is adopted as the individual learning algorithm.

The following linear loss function is employed to evaluate the weak learner performance at each boosting iteration:

$$e_t = \sum_{i=1}^{k} w_{t,i} e_{t,i} = \sum_{i=1}^{k} w_{t,i} \frac{|Y_i - f_t(\boldsymbol{X}_i)|}{E_t} \tag{4}$$

where e_t = average loss; $e_{t,i}$ = relative loss of ith sample; and $E_t = \max(|Y_i - f_t(\boldsymbol{X}_i)|)$ is the maximum loss at iteration step t.

At iteration step t, the weight updating parameter β_t and learner weight α_t are calculated as follows (Solomatine and Shrestha 2004; Shrestha and Solomatine 2006):

$$\beta_t = \frac{e_t}{1 - e_t}, \qquad \alpha_t = \log \frac{1}{\beta_t} \tag{5}$$

At the end of each boosting iteration, updated subsampling weights $\boldsymbol{D}_{t+1}$ are attained using the following rule:

$$\boldsymbol{D}_{t+1} = \{w_{t+1,1}, w_{t+1,2}, \ldots, w_{t+1,k}\}$$
$$\text{where } w_{t+1,i} = \frac{w_{t,i} \beta_t^{1-e_{t,i}}}{\sum_{i=1}^{k} w_{t,i} \beta_t^{1-e_{t,i}}} \tag{6}$$

Once the iterations are complete, weak learners are combined using the following weighted median rule to generate a strong learner (Solomatine and Shrestha 2004; Shrestha and Solomatine 2006):

$$F(\boldsymbol{X}) = \inf\left[y \in \mathbb{R}: \sum_{t:f(\boldsymbol{X}) \le y} v\alpha_t \ge \frac{1}{2} \sum_{t=1}^{k} v\alpha_t\right] \tag{7}$$

where $\mathbb{R}$ = range for the output variable; and $v \in (0, 1]$ = regularization factor used to avoid overfitting.

Algorithm 1 shows the pseudocode for the aforementioned procedure. Adaptive learning scheme of AdaBoost is built on two types of weighting. One set of weights (α_t) is used for aggregating the weak learners to obtain the final model. The other set of weights that consists of $w_{t,i}$ is used for setting the subsampling probability of each example in the training data set. It can be seen from Eq. (5) that if the training accuracy of a weak learner is high, i.e., if e_t is low, β_t will be small. Consequently, the learner weight α_t will be high and that weak learner will have more so-called discourse power in the final combination. Similarly, as indicated in Eq. (6), if a sample is inaccurately predicted at a round of training, its subsampling weight is increased such that it is considered to a greater extent in the subsequent boosting iterations.

Algorithm 1. Pseudocode for AdaBoost

begin

Inputs: training data set Θ_{Tr}, individual learning algorithm (CART) $\mathcal{L}(\cdot)$, and number of boosting iterations T;

Initialize: subsampling weight distribution $\boldsymbol{D}_1(i) = w_{1,i} = 1/k$ for $i = 1, 2, \ldots, k$;

for $t = 1, 2, \ldots, T$ **do**

- generate the weak learner using the weighted training data

$$f_t(\boldsymbol{X}) \leftarrow \mathcal{L}(\Theta_{\mathrm{Tr}}, \boldsymbol{D}_t)$$

- calculate the sample loss and average loss

$$e_{t,i} \leftarrow \frac{|Y_i - f_t(\boldsymbol{X}_i)|}{E_t}, \qquad e_t \leftarrow \sum_{i=1}^{k} w_{t,i} e_{t,i}$$

- calculate the weight updating parameter

$$\beta_t \leftarrow \frac{e_t}{1 - e_t}$$

- update the weak learner weight

$$\alpha_t \leftarrow \log \frac{1}{\beta_t}$$

- update the subsampling weights for next round

$$\boldsymbol{D}_{t+1}(i) = w_{t+1,i} \leftarrow \frac{w_{t,i} \beta_t^{1-e_{t,i}}}{\sum_{i=1}^{n} w_{t,i} \beta_t^{1-e_{t,i}}}$$

end

Output: strong learner

$$F(\boldsymbol{X}) \leftarrow \inf\left[y \in \mathbb{R} : \sum_{t:f(\boldsymbol{X}) \le y} v\alpha_t \ge \frac{1}{2} \sum_{t=1}^{k} v\alpha_t \right]$$

end

Implementation Workflow

Using AdaBoost, the regression model for PHL can be established in four steps, as shown in Fig. 1. The first step is the collection of data to create a database containing the number of samples sufficient for adequate generalization. The model inputs and output are determined at this step. Next, the database is randomly split into training and test sets. The training set is used to train a predictive model using AdaBoost, as described in Algorithm 1. The test set is used to validate the performance of the trained predictive model. The 10-fold CV method is employed in the process of model training to identify optimal model parameters. A large number of hyperparameter combinations are drawn using the grid search method, and for each combination, several performance measures are calculated. The hyperparameters that yield the highest predictive performance are deemed suitable, and the resulting model is passed onto validation. If at the validation stage, model performance is satisfactory, the model is decided as the final predictive model. Finally, the validated model is used to predict the PHL for real seismic analysis purposes of RC structures.

Experimental Database for Plastic Hinge Length Prediction

Creating a predictive model for PHL via AdaBoost requires an experimental database for the training and test sets. The database needs to be comprehensive enough to cover the whole range of parameters believed to affect PHL formation. The mechanisms involved in plastic hinge development (e.g., moment gradient, tension shift, and strain penetration) shall all be adequately represented in the database such that the final predictive model can achieve

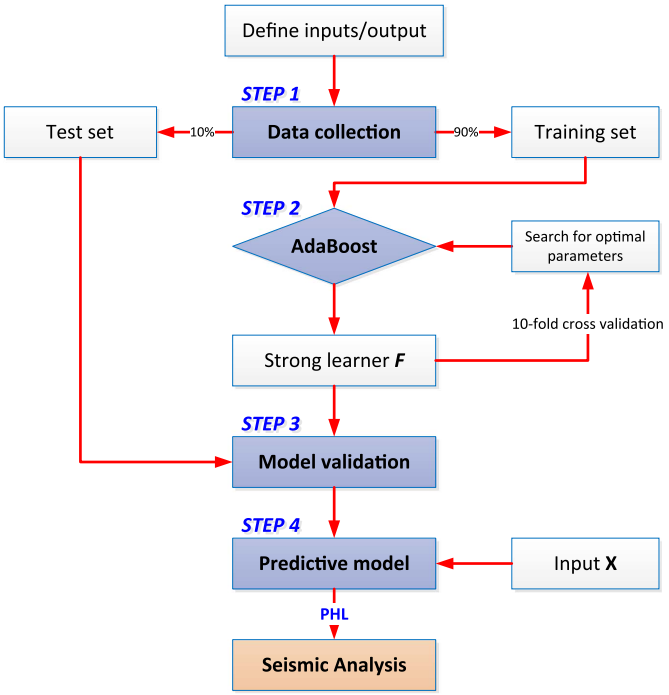

Fig. 1. Implementation workflow for AdaBoost training and seismic analysis of RC columns.

strong generalization and applicability. According to Friedman et al. (2001), for a database to be considered sufficient for training and testing, the number of examples in the database needs to be at least 10 times the number of independent variables. Based on Ning and Li (2015), data sets consisting of a minimum of 92 examples are sufficient for establishing a model satisfying generalization and applicability requirements for PHL prediction. In consideration of the stipulations discussed previously, in this study, the database of 133 PHL tests performed on RC columns collected by Ning and Li (2015) is employed for developing the AdaBoost model.

Each example in the data set consists of 12 specimen features as well as the PHL ratio (i.e., PHL divided by the cross-section depth). These 12 features, defined as the input vector $\boldsymbol{X}$ for the AdaBoost model, are composed of parameters such as geometric dimensions, material properties, reinforcement details, and the loading type for each test specimen. Specifically, specimen features included are the column height L, section width B and height H, concrete strength f_c, longitudinal reinforcement diameter d_b and yielding strength f_y, longitudinal reinforcing ratio ρ_s, transverse reinforcement spacing s and yielding strength f_{vy}, transverse reinforcing ratio ρ_v, axial load ratio n, and lateral loading type (LT), i.e., cyclic or monotonic.

Except for the lateral loading type, which is a categorical variable, all specimen features are numerical variables. The model output Y is a single variable: PHL ratio. Mean, lower, and upper bounds, and the standard deviation (SD) of the input and output variables are listed in Table 1. Fig. 2 shows the relative frequency distribution for each variable except for lateral loading type. Lateral loading type contains two kinds of loading. Monotonic loading makes up 36% of the data set and is assigned the integer label 1, and cyclic loading makes up 64% of the data set and is assigned the integer label 2.

PHL ratios listed in the experimental database are formed on measurements performed two different ways. The first method, direct measurement, defines PHL ratio based on physical observations

Table 1. Statistical properties of parameters included in PHL database

Parameter	Unit	Min	Max	Mean	SD	Type
L	mm	630.00	3,050.00	1,712.62	407.88	Input
B	mm	150.00	610.00	273.87	107.15	Input
H	mm	140.00	610.00	288.31	107.48	Input
f_c	MPa	24.80	175.00	70.27	30.89	Input
d_b	mm	8.00	32.00	16.39	5.27	Input
f_y	MPa	339.00	572.00	498.07	44.65	Input
ρ_s	%	0.82	6.10	2.31	0.91	Input
s	mm	50.00	220.00	93.59	28.92	Input
f_{vy}	MPa	325.00	952.00	530.92	158.69	Input
ρ_v	%	0.32	6.72	2.11	1.26	Input
n	—	0.00	0.77	0.35	0.19	Input
LT	—	1.00	2.00	1.64	0.48	Input
PHL ratio	—	0.35	2.43	1.11	0.39	Output

Note: Min = minimum; Max = maximum; and SD = standard derivation.

(e.g., the extent of cover concrete spalling, longitudinal reinforcement buckling, and transverse reinforcement yielding) or the curvature profile recorded along the column height. However, the second method, indirect measurement, calculates the PHL ratio by coupling a predefined curvature profile with top displacement and curvature measurements. Accordingly, the two approaches may lead to different PHL values. The differences caused by the use of these two distinct procedures are hard to reconcile. Hence, the dissimilarity in PHL measurements is attributed to the measuring error inherent to experimental measurements (Ning and Li 2015). In assigning PHL ratios to database examples, if only one of the two methods is used for a test, PHL ratio for that test is set equal to the available measurement. If the PHL ratio for an experiment is reported through both procedures, PHL ratio for that example is defined as the average of the two measurements.

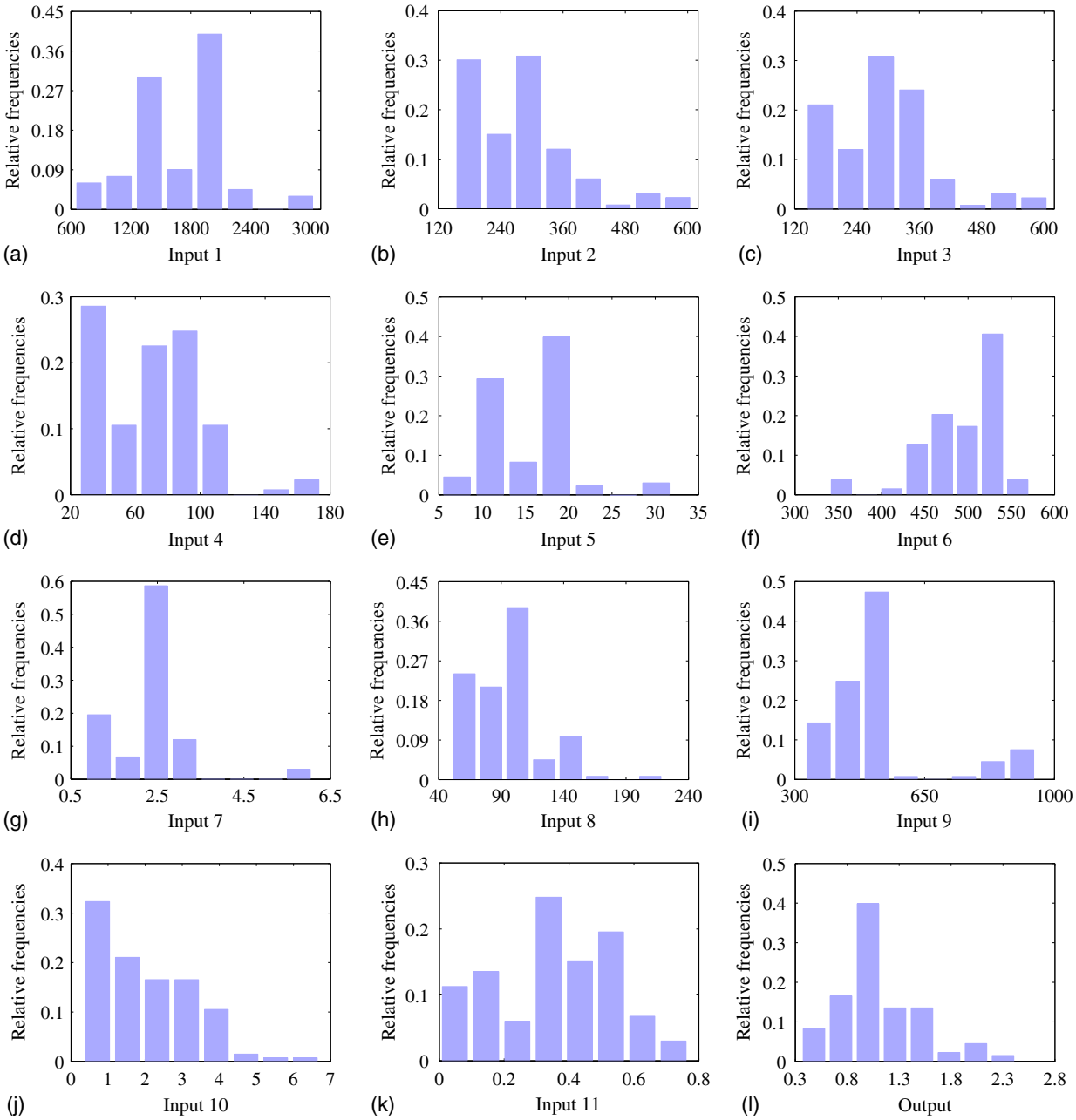

Fig. 2. Relative frequency distributions of the input and output variables.

J. Struct. Eng., 2021, 147(2): 04020332

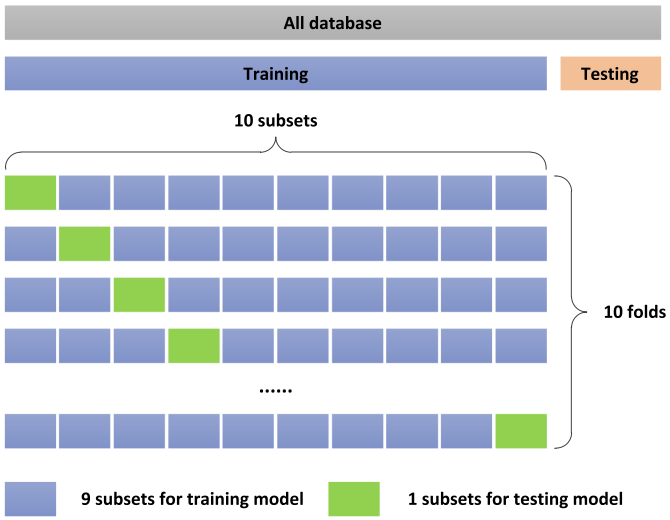

Fig. 3. Data partitioning used for 10-fold cross-validation and testing of the AdaBoost algorithm.

AdaBoost Model and Its Implications for Identifying Features Significant to Plastic Hinge Formation

Model Training and Validation

The AdaBoost algorithm is trained on the experimental database to create a predictive model for PHL of RC columns. The database of 133 experiments is divided into training and test sets for model training and validation, respectively. In model training, to determine the set of hyperparameters that result in the highest generalization capability, the performance of each parameter set is evaluated through the 10-fold CV method. For this purpose, the training set is randomly partitioned into 10 subsets with an equal number of samples. In each fold, nine sets are utilized for training, and the remaining set is held out for testing. This procedure is repeated 10 times using each of the 10 sets for testing exactly once. Fig. 3 provides a summary of the described process. After finding the optimal model hyperparameters using the 10-fold CV procedure, the prediction performance of the model trained on the optimal parameters is validated based on its performance on the test set.

The following performance measures are calculated to assess the effectiveness of the developed predictive model:

- Root-mean squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{m} (P_i - \text{GT}_i)^2}{m}} \tag{8}$$

- Mean absolute error (MAE)

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m} |P_i - \text{GT}_i| \tag{9}$$

- Mean absolute percentage error (MAPE)

$$\text{MAPE} = \frac{100\%}{m}\sum_{i=1}^{m} \left| \frac{P_i - \text{GT}_i}{\text{GT}_i} \right| \tag{10}$$

where P_i and GT_i = predicted and ground-truth values, respectively; and m = number of observations. RMSE represents the deviation of the predicted results from the ground truth, whereas MAE and MAPE depict the predictive error of the model. The smaller these measures are, the better performance the model offers.

In identifying the model hyperparameters, the effect of two framework-level parameters, namely the number of weak learners T and the regularization factor v, are investigated in detail using the grid search method. Fig. 4(a) displays the influence of these two parameters on the model performance in terms of the RMSE value averaged across all 10 folds. It is observed that RMSE is at a minimum, i.e., model performance is the highest, when $T = 100$ and $v = 0.01$, and the model performance decreases (in other words, RMSE increases) when the regularization factor and the number of weak learners are increased. Hence, the number of boosting iterations T is set as 100, and the regularization factor v is defined as 0.01. At the CART level, the objective is to set the hyperparameters so that each individual learner is a weak learner. Thus, at the individual learner level, an upper bound is not specified for the tree depth and the number of features to consider to make each split decision. The minimum number of examples required to split an internal node is set as two, and the minimum examples needed at a leaf node is fixed at one.

Fig. 4(b) shows a comparison of the model predictions against the ground-truth PHL ratios. The $y = x$ line marks the state where the predicted and ground-truth values are the same. Evidently, the

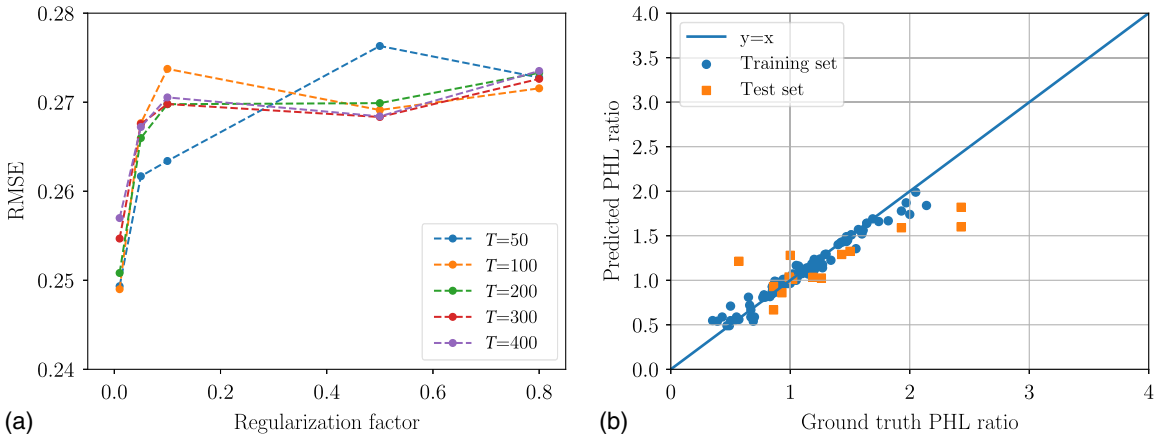

Fig. 4. (a) Sensitivity of model performance to framework-level hyperparameters; and (b) comparison of the PHL ratio predictions of the established model against ground-truth data.

Table 2. Predictive performance of the AdaBoost model at various development stages

Development stage	Type	Performance measures		
		RMSE	MAE	MAPE (%)
10-fold CV	Mean value	0.249	0.175	17.6
	SD	0.087	0.050	3.52
Training	—	0.073	0.043	4.89
Testing	—	0.363	0.271	22.48

predicted values almost coincide with the ground-truth training data. PHL ratios predicted for the test set exhibit a somewhat higher variation but still are very close to the ground-truth values.

Table 2 presents the model performance recorded in 10-fold CV, training, and testing stages in terms of RMSE, MAE, and MAPE. The average RMSE, MAE, and MAPE obtained for the developed model in the cross-validation stage are 0.249%, 0.175%, and 17.6%, respectively. SD corresponding to these values are 0.087%, 0.050%, and 3.52%, respectively. The results indicate that the selected hyperparameters result in an AdaBoost model that is stable, and the model is capable of predicting the PHL ratio of RC columns with exceptional accuracy as well as minimal variance. The training error is rather small; RMSE is 0.073, MAE is 0.043, and MAPE is 4.89%. When evaluated on the test set, the established model results in RMSE = 0.363, MAE = 0.271, and MAPE = 22.48%, respectively. Even though the model performance decreases to some extent for the test set, as discussed subsequently, the generalization capability of the model still far exceeds the existing models.

The authors previously studied the influence of using ANN, SVM, and CART as weak learners in AdaBoost for predicting the concrete compressive strength, and found that these three weak learner generation methods can all attain high levels of performance. However, the best prediction capability is attained using CART (Feng et al. 2020). Yet, to further confirm the validity of the decision to use CART at the individual level in this study, model performance using ANN and SVM is also investigated. Optimal hyperparameters for ANN and SVM are obtained using grid search and 10-fold CV. Fig. 5 demonstrates the variation of the averaged RMSE resulting from the 10-fold CV for ANN- and SVM-based weak learner types. ANN-based AdaBoost reaches the minimum RMSE when $T = 200$ and $v = 0.05$, whereas SVM-based AdaBoost reaches the minimum when $T = 50$ and $v = 0.01$. Generating weak learners using ANN and SVM yield satisfactory model performance with minimum RMSE values at 0.269 and 0.266, respectively. Nonetheless, using CART-based AdaBoost with its optimal hyperparameter combination results in higher predictive performance with a minimum RMSE of 0.249, as shown in Fig. 4(a). In short, using CART at the weak learner level results in better PHL prediction models.

Feature Importance Analysis

Machine learning models offer a notable performance increase over traditional statistical models. Yet, what machine learning models typically lack over conventional methods is the explicit description of the features selected by the learning algorithm to resemble the modeled process. Unlike a traditional statistical model, for a machine learning model, interpreting the relationship between model parameters and the output requires further examination. One procedure that is commonly utilized for this purpose is feature importance analysis. Through feature importance analysis, the significance of each model input toward predicting the designated outputs can be determined. Thus, an understanding of the mechanics of the modeled process can be attained. In this section, a feature importance analysis is performed to break down the relative contributions of model features for the developed AdaBoost model to shed light on the parameters of most statistical significance for plastic hinge development.

The significance of each feature to the AdaBoost model is calculated by averaging the importance measures over all the weak learners, where the importance of a feature is determined at the CART level by quantifying the total decrease in node impurity during the splitting process (Friedman et al. 2001). Fig. 6 shows the relative feature importance values determined for the 12 parameters used in developing the AdaBoost model. Results indicate that concrete strength f_c, axial load ratio n, longitudinal bar ratio ρ_s and strength f_y, and transverse bar spacing s are the parameters that are most critical to correct identification of PHL. Section height H and transverse bar ratio ρ_v have an importance of around 30% relative to the concrete strength. The contributions of the transverse bar strength f_{vy}, column length L, longitudinal bar diameter d_b, section width B, and loading type to PHL predictions of the AdaBoost model are rather small, i.e., around 10%–15% of concrete strength.

The findings of the performed feature importance analysis are in remarkable agreement with the recognized physical mechanisms behind plastic hinge formation. As discussed by Ning and Li (2015)

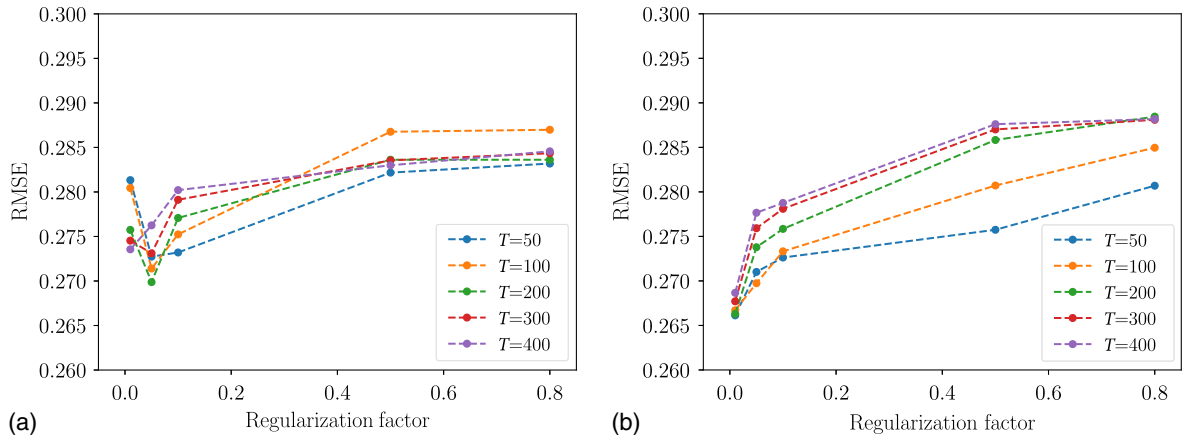

Fig. 5. Influence of weak learner type and framework-level hyperparameters on AdaBoost model performance for (a) ANN-based; and (b) SVM-based learners.

J. Struct. Eng., 2021, 147(2): 04020332

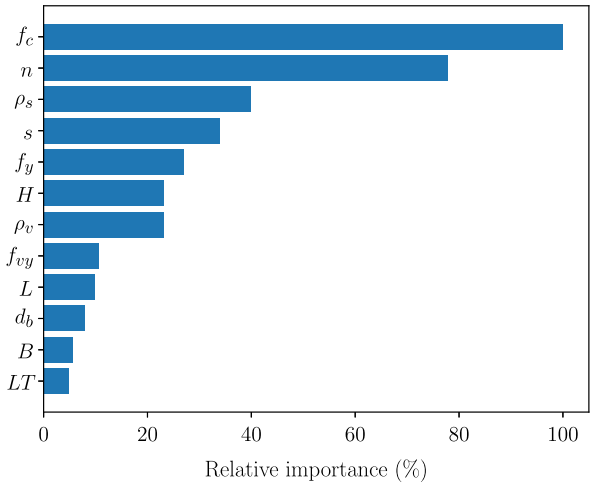

Fig. 6. Relative importance of the data set features for PHL predictions of the AdaBoost model.

and Bae and Bayrak (2008), plastic hinge development is governed by a combination of the flexural, shear, and bond-slip behavior of an RC section. The three variables determined to have very high feature importance, i.e., f_c, ρ_s, and f_y, are all directly related to these three principal mechanisms. Another parameter of great importance, axial load ratio n, is known to affect the flexure ductility level of RC elements—a property perceived important in plastic hinge formation. Section confinement levels are also crucial to plastic hinge development and are satisfactorily captured by the high feature importances determined for transverse bar properties s, ρ_v, and f_{yv}. Longitudinal bar behavior is described by the longitudinal reinforcement ratio ρ_s and bar diameter d_b, with the former being a better descriptor of PHL by the AdaBoost model, leaving the latter with lower feature importance. This observation does not necessarily mean d_b is not crucial in plastic hinge development; it could be an artifact of the dependence between ρ_s and d_b. Lastly, the mechanisms governing plastic hinge behavior are not known to be sensitive to section width B and loading type, explaining the small relative importance of these two features.

Comparison with Individual Learning and Linear Regression Methods

For additional validation, the generalization capability of the created AdaBoost model is tested against models generated via three widely used individual learning algorithms, namely ANN, SVM, and CART, and a 12-parameter linear regression model. The ANN implementations consist of six distinct fully connected, feedforward, three-layer regressors with different neuron densities in the hidden layer. The rectified linear unit (ReLU) (Nair and Hinton 2010) activation function is used for the hidden layer of all six implementations. Furthermore, weight optimization is performed using the Adam solver (Kingma and Lei Ba 2015) for a constant learning rate with a step size of 0.1. In developing the SVM models, the Gaussian kernel is adopted, and a total of 15 models are generated for different kernel coefficients γ and penalty parameters C. Multiple models are also trained using CART by performing a grid

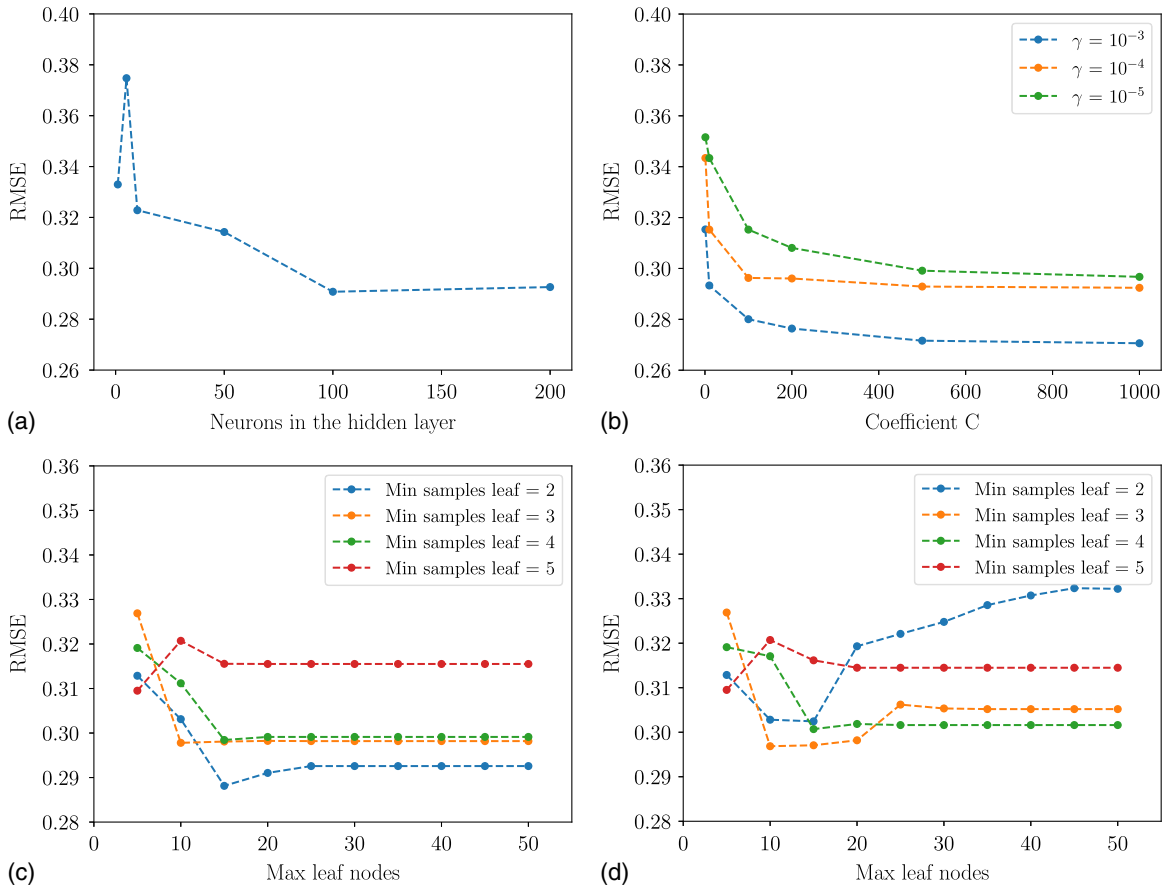

Fig. 7. Variations in the performance of the examined individual learning models for different hyperparameter configurations: (a) ANN; (b) SVM; (c) CART: Max tree depth = 5; and (d) CART: Max tree depth = 10.

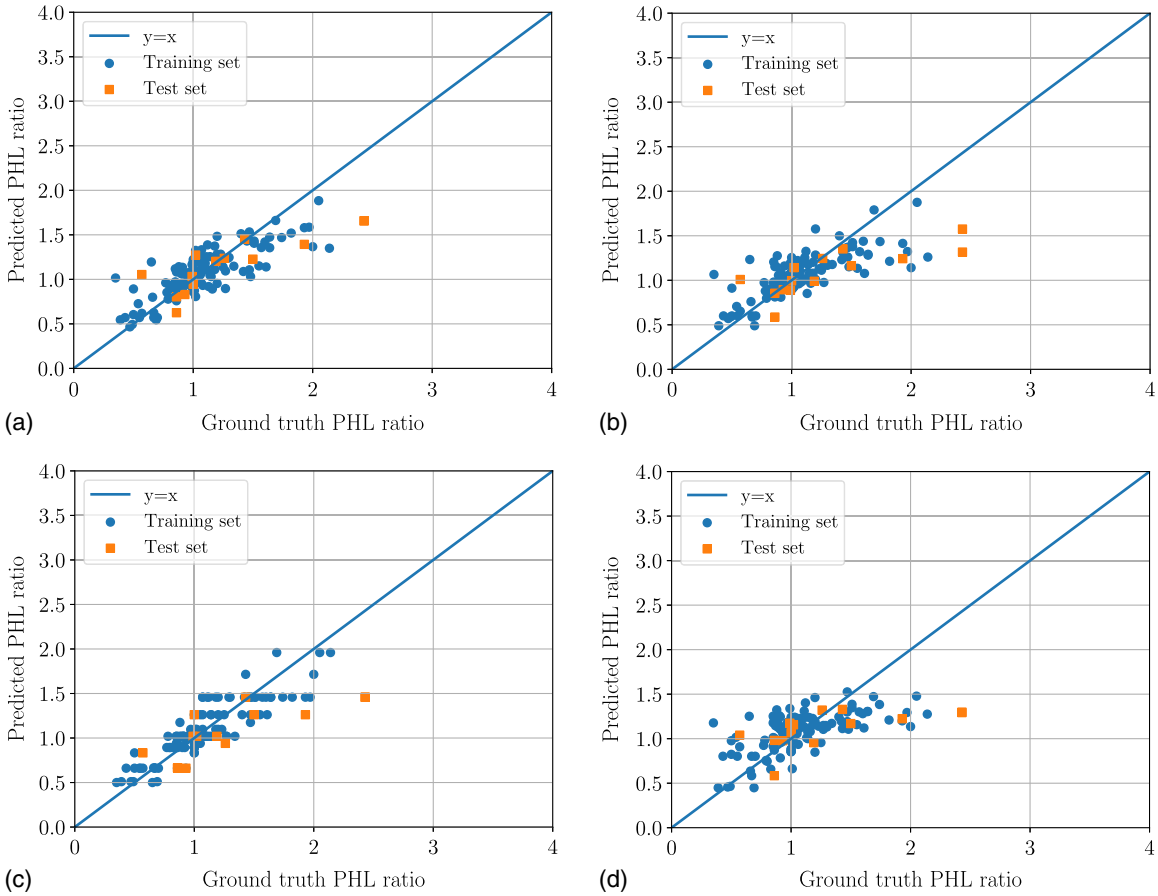

Fig. 8. Prediction results of the final (a) ANN; (b) SVM; (c) CART; and (d) linear regression models.

search over the hyperparameters maximum tree depth, maximum number of leaf nodes, and minimum number of samples to split a leaf node. For each learning algorithm, the parameters that result in best averaged 10-fold CV performance are determined as the optimal hyperparameters for that algorithm.

Fig. 7 shows the effect of the choice of hyperparameters on the model performance of the considered base learning algorithms in terms of RMSE averaged across all 10 cross-validation folds. ANN model with the best performance is achieved when the number of neurons is set to 100. The SVM model trained using the hyperparameters $C = 500$ and $\gamma = 10^{-3}$ attains the highest performance among the generated SVM models. Finally, the CART model with the minimum RMSE is achieved when the maximum tree depth, maximum number of leaf nodes, and minimum number of samples to split a leaf node are set to 5, 15, and 2, respectively.

Fig. 8 demonstrates the prediction results of the three individual learning models trained using the determined optimal set of hyperparameters, i.e., final models, as well as the 12-parameter statistical model created using linear regression. Table 3 provides a comparison between the prediction performance observed for these models and the developed AdaBoost model. Based on the calculated RMSE, MAE, and MAPE values, AdaBoost outperforms all four procedures, most likely as a result of the known advantages of ensemble learning methods over individual learning and traditional statistical methods (Erdal et al. 2013; Chou et al. 2014).

Table 3. Prediction performance of the final ANN, SVM, CART, linear regression, and AdaBoost models

		Performance measures		
Sets	Methods	RMSE	MAE	MAPE (%)
10-fold CV results	ANN	0.291	0.224	22.8
	SVM	0.270	0.210	20.9
	CART	0.288	0.213	21.2
	LR	0.326	0.244	25.1
	AdaBoost	0.249	0.175	17.6
Training results	ANN	0.212	0.157	15.8
	SVM	0.236	0.169	16.9
	CART	0.176	0.137	13.4
	LR	0.282	0.219	22.5
	AdaBoost	0.073	0.043	4.89
Testing results	ANN	0.371	0.260	19.8
	SVM	0.455	0.305	21.3
	CART	0.448	0.328	23.0
	LR	0.508	0.357	25.4
	AdaBoost	0.363	0.271	22.4

Comparison with the Existing Empirical Models

Existing Empirical PHL Models

In order to further examine the prediction performance of the proposed model, several empirical PHL models that are widely used in the literature are evaluated for comparison. The mathematical relationships for the models considered are summarized in Table 4. In the defined formulas, L is the column height, d_b is the longitudinal bar diameter, f_y is the yield stress of longitudinal reinforcement, H

Table 4. Plastic hinge length relationships for the empirical models considered in evaluating the performance of the final AdaBoost model

References	Model abbreviation	Relationship for plastic hinge length
Priestley and Park (1987)	EM1	$0.08L + 6d_b$
Paulay and Priestley (1992)	EM2	$0.08L + 0.022f_y d_b$
Panagiotakos and Fardis (2001)	EM3	$0.12\frac{L}{H} + 0.014\alpha_{sl} f_y d_b$
Lu et al. (2005)	EM4	$0.077L + 8.16d_b$
Berry et al. (2008)	EM5	$0.0375L + 0.12\frac{f_y d}{\sqrt{f_c}}$
Bae and Bayrak (2008)	EM6	$\left[0.3\frac{P}{P_0} + 3\frac{A_s}{A_g} - 0.1\right]L + 0.25H \geq 0.25H$
Ning and Li (2015)	EM7	$\left[0.042 + 0.072\frac{P}{P_0}\right]L + 0.298H + 6.407d_b$

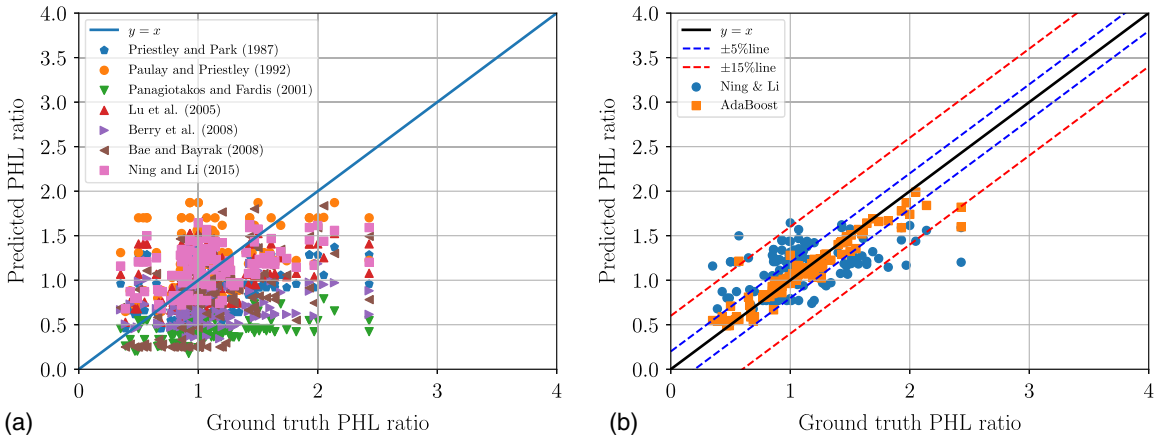

Fig. 9. (a) Prediction results of the considered existing empirical models; and (b) comparison of the AdaBoost and Ning and Li (2015) model predictions.

is the section height, α_{sl} is the coefficient for slip (which equals 1 if considering bar pullout effect or 0 if not), f_c is the concrete compressive strength, L is the distance from the critical section to point of contra-flexure, P is the applied axial load, A_g is the gross area of the section, A_s is the area of the longitudinal reinforcements, and P_0 is the nominal axial load capacity of the considered element. Most of the existing empirical models feature three to five independent variables, with the most complex relationship—namely that by Bae and Bayrak (2008)—consisting of seven parameters.

AdaBoost Model versus the Existing PHL Equations: Performance Comparison

In this section, the prediction performance of the AdaBoost model is compared against the empirical models summarized in Table 4. Fig. 9(a) shows a comparison of the PHL ratios predicted by the existing empirical models and the corresponding ground-truth PHL ratios. The results in the figure exhibit considerable variability, especially at PHL ratios less than 0.8 and greater than 1.5. For some specimens, PHL ratios predicted by the existing relationships vary from 0.5 to 2. The RMSE, MAE, and MAPE for empirical models EM1–EM7 and the AdaBoost model are given in Table 6. With RMSE = 0.3588, MAE = 0.288, and MAPE = 30.39%, the model by Ning and Li (2015) attains a better accuracy on the experimental database than all the other empirical models, yet it is considerably outperformed by the AdaBoost model.

Fig. 9(b) shows the predictions of the AdaBoost and Ning and Li (2015) models, as well as the lines marking the bounds for 5% and 15% deviations from the ground-truth line $y = x$. Nearly all of the PHL ratios predicted by the AdaBoost model are located in the high-confidence region [in other words, almost all AdaBoost predictions are within the $y = (1 \pm 0.05)x$ lines]. RMSE, MAE, and MAPE of the AdaBoost model are remarkably better than that of the existing empirical models at 0.137%, 0.067%, and 6.74%, respectively. The performance improvements of the AdaBoost model over the empirical model with the best performance, i.e., the model by Ning and Li (2015), are 86.25% in RMSE, 76.74% in MAE, and 77.82% in MAPE.

The mean and coefficient of variance (COV) of the ground-truth and predicted PHL values produced by each model are listed in Table 6. It can be seen that the proposed AdaBoost model performs the best with a mean of 1.00 and COV 11.5%. Additionally, the ground-truth/predicted PHL ratios obtained using the AdaBoost and Ning and Li (2015) models are shown in detail in Fig. 10. Even though the model by Ning and Li (2015) performs better than EM1–EM6, as indicated in Table 5, its prediction performance is significantly exceeded by the AdaBoost model, which displays a higher prediction accuracy and minimal prediction variance.

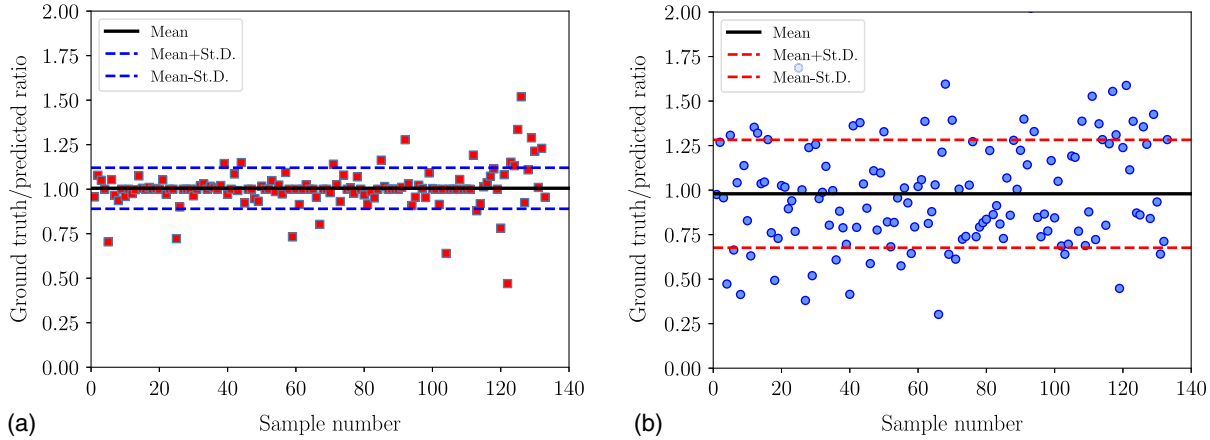

Fig. 10. Ground-truth/predicted PHL ratio results for (a) AdaBoost; and (b) Ning and Li (2015) models.

Table 5. Prediction performance of the AdaBoost model versus existing empirical models

Models		Performance measures				
		RMSE	MAE	MAPE (%)	Mean	COV (%)
Empirical model	EM1	0.4407	0.349	32.22	1.28	34.2
	EM2	0.4333	0.332	36.67	0.95	33.5
	EM3	0.7632	0.675	57.53	2.63	33.0
	EM4	0.4005	0.312	30.77	1.13	32.8
	EM5	0.5729	0.462	39.97	1.64	40.8
	EM6	0.5181	0.429	40.00	1.79	58.1
	EM7	0.3588	0.288	30.39	0.98	30.9
AdaBoost model	—	0.1375	0.067	6.74	1.00	11.5
Improvement over EM7 (%)	—	86.25	76.74	77.82	2.0	62.78

In order to provide additional insight into the performance of EM1–EM7, particularly in comparison with the AdaBoost model, Fig. 11 shows the relative frequency distribution of the ratio between the predicted and ground-truth PHL determined for each model. For the AdaBoost model, the figure displays a frequency distribution with a pronounced peak around unity showing little spread around that peak. This observation serves as a good indicator for the prediction accuracy of the AdaBoost model because, on average, its predictions for PHL ratios closely resemble the ground-truth values. It can also be seen in Fig. 11 that the models

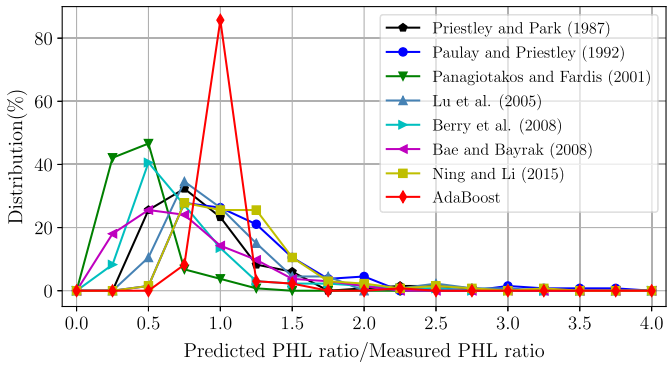

Fig. 11. Relative frequency distribution of the ratio between ground-truth and predicted PHL values. A predicted/measured ratio of 1 signifies perfect match between predicted and measured values.

by Priestley and Park (1987), Lu et al. (2005), Berry et al. (2008), and Bae and Bayrak (2008) underestimate the PHL ratios for 25%–37% of the examples, whereas the model by Panagiotakos and Fardis (2001) underestimates the PHL ratios for nearly 80% of the examples.

Furthermore, to demonstrate the consistency in the model performance for various section and loading configurations, the variation of the ground-truth/predicted PHL ratio is evaluated for four column properties at a wide range of values. These select properties, namely, concrete compressive strength f_c, axial load ratio n, longitudinal bar ratio ρ_s, and longitudinal bar strength f_y are features that are highly important to the predictions of the AdaBoost model. Fig. 12 displays the ground-truth/predicted PHL ratios obtained from the AdaBoost model for a broad selection of f_c, n, ρ_s, and f_y values, where the shaded area in each plot denotes the interval containing predictions within Mean $\pm$ 1SD. As shown in the figure, the ground-truth/predicted PHL ratio of the AdaBoost model is very close to 1.0 for all four column properties irrespective of the choice of parameter values, and the prediction variation is around 10% on average. These results imply that the AdaBoost model is stable, and its prediction performance exhibits minimal sensitivity to different sets of column properties.

Application to Finite-Element Analysis

PHL is crucial in simulating the cyclic behavior of RC columns using finite-element analysis. With displacement-based formulation, plastic hinge response is localized over the lengths of the

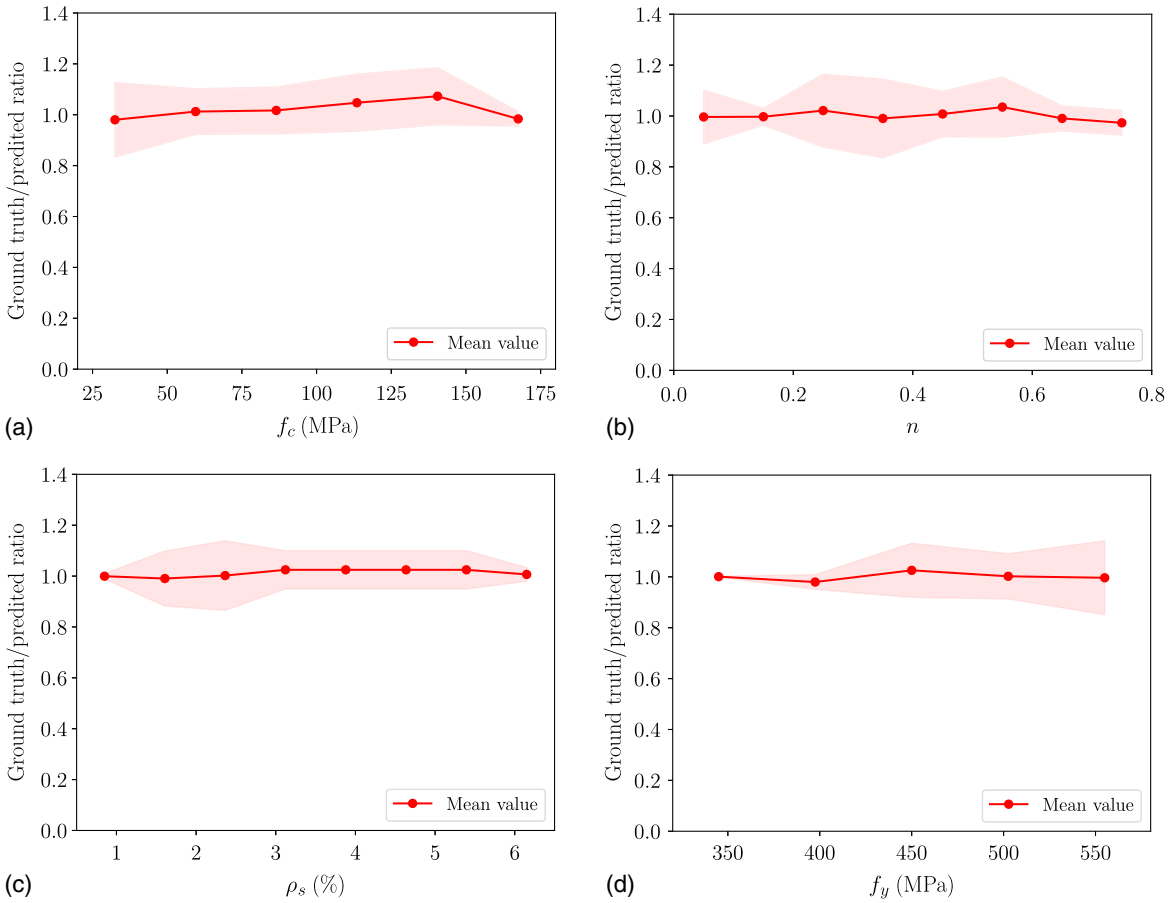

Fig. 12. Change in prediction accuracy of the AdaBoost model for different values of (a) concrete strength; (b) axial load ratio; (c) longitudinal bar ratio; and (d) longitudinal bar strength. Shaded areas denote the interval containing predictions within Mean ± 1SD.

elements undergoing nonlinear behavior. In force-based elements, plastic deformations are confined to the integration points within the regions where the yielding capacity of the member is exceeded. To achieve an objective response, in the former, element lengths can be determined based on PHL, whereas in the latter, integration weights at plastic hinge locations can be set equal to the PHL (Scott and Hamutçuoğlu 2008).

Typically, PHLs of RC columns are specified using empirical relationships. However, as demonstrated in the preceding section, the PHL predictions of the existing models are highly variable, making the process of identifying a realistic PHL value challenging. PHL values predicted by the developed AdaBoost model more closely resemble experimental results. As a result, the established AdaBoost model is expected to improve the veracity of finite-element simulations. This section will demonstrate how the PHL predictions of the developed model can be used for accurately simulating the response of RC columns to cyclic loading. Here, the discussion is limited to beam-column elements based on force-based fiber section formulation. This decision is made mainly due to the popularity of this type of element in modeling column elements, possibly due to its computational efficiency and mesh-independent response.

Force-Based Fiber Element

According to Spacone et al. (1996), for a force-based beam-column element, the relationship between the section force field $\boldsymbol{D}(x)$ and the element forces $\boldsymbol{Q}$ is defined

$$\boldsymbol{D}(x) = \boldsymbol{B}(x)\boldsymbol{Q} \tag{11}$$

where $\boldsymbol{B}(x)$ = matrix containing force interpolation functions.

The section deformations $\boldsymbol{d}(x)$ can be obtained through the fiber section model as follows:

$$\boldsymbol{d}(x) = \boldsymbol{f}_s(x)\boldsymbol{D}(x) \tag{12}$$

where $\boldsymbol{f}_s = [\boldsymbol{k}_s]^{-1}$ is the section flexibility which is the inverse of the section stiffness $\boldsymbol{k}_s$.

Using the virtual work principal, the element deformations $\boldsymbol{q}$ can be derived as follows:

$$\boldsymbol{q} = \int_L \boldsymbol{B}^T \boldsymbol{d}\mathrm{d}x \tag{13}$$

where L = element length.

By linearizing Eq. (13) with respect to the element forces, the element flexibility matrix $\boldsymbol{F}$ is obtained as follows:

$$\boldsymbol{F} = \frac{\partial \boldsymbol{q}}{\partial \boldsymbol{Q}} = \int_L \boldsymbol{B}^T \boldsymbol{f}_s \boldsymbol{B}\mathrm{d}x \tag{14}$$

The integral in Eq. (14) can be solved through numerical quadrature. In performing the numerical integration, the weights for the plastic hinge locations are typically set as the PHL to avoid a nonobjective nonlinear column response under strain-softening. Assuming plastic hinges are formed at both ends of the element, the integral can be written as follows:

Table 6. Specimens modeled using force-based fiber section beam-column elements

References	Specimen	L (mm)	B (mm)	H (mm)	f_c (MPa)	d_b (mm)	f_y (MPa)	ρ_s (%)	s (mm)	f_{vy} (MPa)	ρ_v (%)	n	PHL ratio
Barrera et al. (2011)	N30-10.5-C0-200	1,500	150	140	32.2	10	537	2.20	100	530	1.00	0.00	0.50
Barrera et al. (2011)	H60-10.5-C0-230	1,500	150	140	60.5	10	537	2.20	100	530	1.00	0.34	1.61
Tanaka (1990)	Unit 7	1,650	550	550	32.1	20	511	1.25	90	325	2.08	0.30	0.67
Paultre et al. (2001)	1006052	2,000	305	305	109.4	20	481	2.15	60	492	4.26	0.51	1.51
Bayrak (1998)	RS-9HT	2,010	250	350	71.2	20	454	2.74	80	542	3.44	0.34	1.02
Bayrak (1998)	WRS-22HT	2,010	350	250	91.3	20	521	2.74	70	465	3.92	0.31	1.11

$$\boldsymbol{F} = \boldsymbol{B}^T \boldsymbol{f}_s \boldsymbol{B}|_{x=\xi_1} L_{pI} + \int_{L_{pI}}^{L-L_{pJ}} \boldsymbol{B}^T \boldsymbol{f}_s \boldsymbol{B} \mathrm{d}x + \boldsymbol{B}^T \boldsymbol{f}_s \boldsymbol{B}|_{x=\xi_N} L_{pJ} \quad (15)$$

where ξ_1 and ξ_N = positions of the first and last integration points (IPs); N = number of IPs; and L_{pI} and L_{pJ} = PHLs at the two end nodes I and J. Generally, modified Gauss-Lobatto or modified Gauss-Radau integration is preferred for solving Eq. (15) (Scott and Hamutçuoğlu 2008).

Nonlinear Analysis of RC Columns

Six specimens with different geometric configurations are selected from the experimental database for modeling, among which two are subjected to monotonic loading, and four are subjected to cyclic loading. The geometric properties, reinforcement details, and material properties of the specimens are given in Table 6. One force-based element with five IPs is used in modeling all six specimens. The PHL values predicted by the AdaBoost model, as well as the upper- and lower-bound PHL predictions obtained from the empirical models EM1–EM7 are used in simulating the response of the column elements (Table 7 gives specifics). It can be observed that the PHLs predicted by the AdaBoost model are very close to the measured (i.e., ground truth) values. At the same time, there exists a large discrepancy between the lower- and upper-bound empirical model results.

The simulations are performed using the finite-element analysis framework OpenSees version 2.3.0 (McKenna 2011). Member cross sections are discretized using 36 concrete fibers for all six cases, where the number of steel fibers is determined according to the particular reinforcement layout for each specimen. For modeling the concrete and steel fibers, the Kent-Park (Kent and Park 1971) and Menegotto-Pinto (Menegotto and Pinto 1973) models are used, respectively. The parameters of the reinforced concrete fibers are calculated in terms of the concrete compressive strength f_c [ACI Committee 318 (ACI 2019)], i.e., peak strain $\epsilon_c = (700 + 172\sqrt{f_c}) \times 10^{-6}$, residual strength $f_{cu} = 0.2 f_c$, residual strain $\epsilon_{cu} = 0.005$, and tensile strength $f_t = 0.1 f_c$. The effect of rebar confinement is considered through the modified Kent-Park model (Scott et al. 1982). For the reinforcing steel, the elastic modulus is set as $E_s = 2.1 \times 10^5$ MPa, and the hardening ratio is specified as 1%. Both the monotonic and cyclic loading is applied in two consecutive steps. First, the axial load is imposed at the top node of the column through force control, and then the lateral loading is applied through displacement control.

Table 7. PHL values for the specimens selected for modeling (in mm)

Loading type	Specimen	Measured	AdaBoost	Empirical models: Lower bound	Empirical models: Upper bound
Monotonic	N30-10.5-C0-200	70.00	99.82	35.00	238.14
	H60-10.5-C0-230	225.40	175.00	262.50	355.25
Cyclic	Unit 7	368.50	423.50	143.44	396.98
	1006052	460.55	460.55	135.47	376.47
	RS-9HT	357.00	355.25	127.75	366.10
	WRS-22HT	277.50	262.50	146.75	390.00

Fig. 13 shows the load-displacement curves resulting from the nonlinear analyses performed for the six specimens. The simulation results based on PHL values predicted by the AdaBoost model, and lower- and upper-bound PHLs predicted by the empirical models are compared. Because the AdaBoost model results in PHL predictions very close to the ground-truth PHL values, both the monotonic loading and hysteresis curves computed for the columns modeled based on the AdaBoost predictions match experimental results consistently well. On the other hand, the simulated responses based on upper- and lower-bound PHL values calculated from the existing empirical relationships considerably differ from the experimental results. In general, the models based on upper-bound PHLs perform relatively better, given that the empirical relationship appears to have a higher likelihood of underestimating the PHL, as evidenced in Fig. 11. The models generated from lower-bound PHLs lead to limited plastic hinge development, resulting in larger primary stiffness and unrealistic stiffness degradation. The models using upper-bound PHLs overestimate the postpeak responses and cumulative energy dissipation.

Moreover, the yield load and displacement, peak load and displacement, and—for column models subjected to cyclic loading—cumulative energy dissipation values are determined from the responses displayed in Fig. 13. The results of the performed calculations are summarized in Table 8. For all five response parameters computed, the AdaBoost model gives results that are in close agreement with the experimental findings (denoted as measured in Table 8). In contrast, the results for models based on PHL predictions of empirical relationships span a wide range of values. For example, the nonlinear model based on the AdaBoost predictions can capture peak column displacement with an average relative error of 3.98%, whereas peak column displacements calculated based on existing empirical models range from 3.90% (upper PHL) to 11.16% (lower PHL). The average relative error in simulating the energy dissipation is 10.68% for the models based on the predictions of the AdaBoost model, whereas for the column models constructed using the PHL outputs of empirical models, the average relative error in the energy dissipation range from 12.23% (upper PHL) to 23.16% (lower PHL). In summary, PHL estimations produced by empirical models result in simulations of the structural response that are less dependable due to the clear connection between the localized behavior of a plastic hinge on the column stiffness, postpeak strength, and energy dissipation behavior.

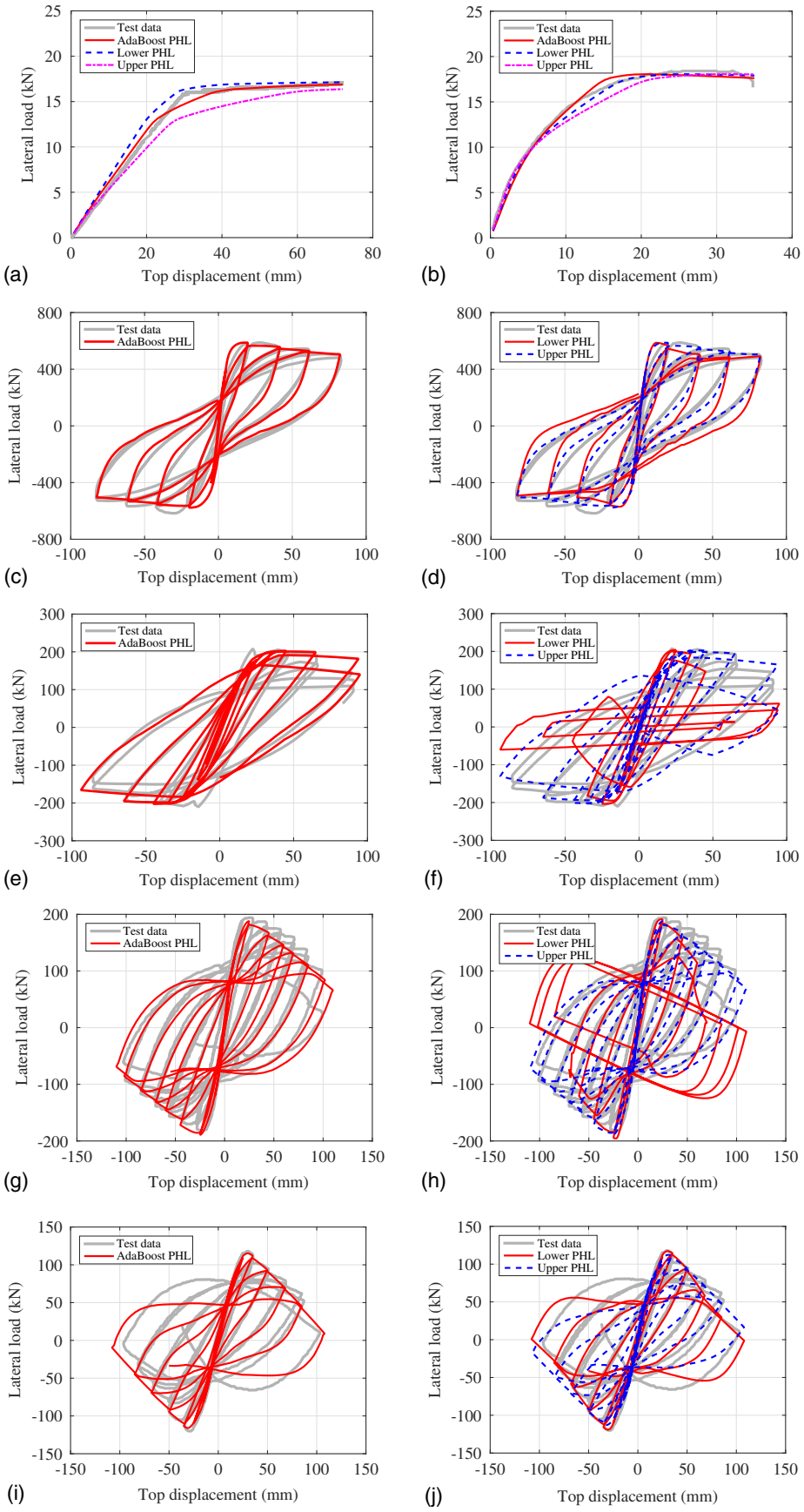

Fig. 13. Column responses simulated using PHL predictions of the AdaBoost model and the upper- and lower-bound PHL estimates of existing empirical models: (a) N30-10.5-C0-200; (b) H60-10.5-C0-230; (c) Unit 7 modeled using AdaBoost PHL; (d) Unit 7 modeled using empirical PHLs; (e) 1006052 modeled using AdaBoost PHL; (f) 1006052 modeled using empirical PHLs; (g) RS-9HT modeled using AdaBoost PHL; (h) RS-9HT modeled using empirical PHLs; (i) WRS-22HT modeled using AdaBoost PHL; and (j) WRS-22HT modeled using empirical PHLs.

Table 8. Response characteristics for the specimens

Specimen	Response characteristic	Measured	AdaBoost	Empirical models	
				Lower PHL	Upper PHL
N30-10.5-C0-200	Yield load (kN)	16.10	15.16	16.04	14.38
	Yield displacement (mm)	33.66	32.47	28.18	39.04
	Peak load (kN)	17.00	16.89	17.13	16.36
	Peak displacement (mm)	71.60	72.00	72.00	72.00
	Dissipated energy (J)	—	—	—	—
H60-10.5-C0-230	Yield load (kN)	15.77	15.72	15.40	15.19
	Yield displacement (mm)	13.14	12.41	13.58	15.03
	Peak load (kN)	18.40	18.05	18.05	18.05
	Peak displacement (mm)	34.80	35.00	35.00	35.00
	Dissipated energy (J)	—	—	—	—
Unit 7	Yield load (kN)	461.24	479.22	466.34	478.99
	Yield displacement (mm)	14.69	11.39	10.66	11.36
	Peak load (kN)	552.80	587.08	566.47	586.62
	Peak displacement (mm)	20.53	19.90	19.90	19.90
	Dissipated energy (J)	133,409	152,735	178,701	153,568
1006052	Yield load (kN)	162.53	181.91	182.20	183.13
	Yield displacement (mm)	33.05	26.52	18.66	24.43
	Peak load (kN)	198.87	202.72	204.01	202.32
	Peak displacement (mm)	46.39	45.00	25.00	45.00
	Dissipated energy (J)	92,996	79,255	51,578	81,427
RS-9HT	Yield load (kN)	194.00	149.92	156.24	149.59
	Yield displacement (mm)	25.00	16.43	19.07	16.31
	Peak load (kN)	194.00	181.27	183.23	181.06
	Peak displacement (mm)	25.00	25.00	25.00	25.00
	Dissipated energy (J)	88,631	85,028	101,153	84,351
WRS-22HT	Yield load (kN)	118.00	94.66	99.11	91.51
	Yield displacement (mm)	30.00	20.89	22.27	19.18
	Peak load (kN)	118.00	107.09	108.70	105.70
	Peak displacement (mm)	30.00	35.00	35.00	35.00
	Dissipated energy (J)	49,820	45,145	49,800	41,575

Conclusions

Improving on the current ability to simulate the inelastic behavior of RC columns requires (1) a better understanding of the factors that affect the plastic deformations causing the nonlinear response, and (2) the capability to determine the extent of these deformations accurately. Assuming these deformations are localized in small regions and can be lumped into plastic hinges, establishing models that are able to correctly predict the PHL of RC columns is crucial. The traditional approach for creating models to estimate PHL of RC columns is to develop empirical relationships in terms of a small number of parameters deemed to affect plastic hinge behavior. These models often exhibit limited prediction accuracy and relatively large variance; therefore, the applicability of their predictions to simulating RC column behavior is inadequate. Models developed using machine learning ensemble methods, in specific AdaBoost, can address the drawbacks of the existing models by delivering significant performance improvements.

An essential component of training an AdaBoost model is defining an objective framework to assess its performance. Especially for small data sets, such as the one utilized for this study, it is determined that using 10-fold cross-validation in connection with metrics that appropriately represent prediction performance is vital. By averaging the performance attained in each fold of the training set, the potential overfitting effects are canceled, and the performance of AdaBoost for a set of hyperparameters can be reliably determined. Subsequently, for a given data set, hyperparameters of AdaBoost can be tuned to maximize the predictive capability of the generated model.

The AdaBoost PHL prediction model is trained based on the experimental PHL data set compiled by Ning and Li (2015) using the optimal hyperparameters identified via 10-fold CV. The AdaBoost model is observed to predict the PHL ratio for the 133 experiments contained in the data set exceptionally well with most predictions within 5% of the ground-truth values. A comparison of the model's prediction performance shows that it outperforms the models produced by the popular base learning algorithms, artificial neural networks, support vector machines, and classification and regression trees, as well as linear regression. A similar conclusion can be made based on the AdaBoost model's performance against seven of the existing empirical relationships. PHL ratios estimated by the empirical models display a large scatter; the AdaBoost model's prediction performance far exceeds even the highest performing empirical model. All in all, the AdaBoost model demonstrates high generalization capability with minimal prediction variance, and its performance is consistent for a wide range of parameter values.

In order to develop an understanding of the parameters that are most statistically significant for plastic hinge development, the significance of each training feature to the AdaBoost model has been investigated. It is determined that 10 of the 12 features utilized in model training are highly essential for model predictions, with the parameters of most importance being concrete strength, axial load

ratio, longitudinal bar ratio, longitudinal strength, and transverse bar spacing.

The AdaBoost model is also utilized to calculate the weights of plastic hinge locations for force-based fiber elements given that using PHL at plastic hinge locations eliminates nonobjective element behavior. Monotonic and cyclic responses of six RC column specimens were selected as representative examples. The numerical results show that models based on AdaBoost PHL predictions successfully simulate the response of RC columns. On the contrary, simulation results of PHLs estimated using the existing relationships considerably digress from experimental results. This contrast in performance is believed to be due to empirical equations' limited ability to generate nonobjective PHL values. Models generated using PHL predictions of empirical equations often fail to realistically represent the localized damage regions of the columns; thus, their column responses are inconsistent with the test results.

The PHL prediction model developed in this study is made available on DesignSafe (Cetiner et al. 2020) for the effective dissemination of these research outcomes. This deployment of the model is designed to generate PHL predictions and column response curves based on a set of user entries. It allows interaction through an easy-to-use graphical user interface as well as at the source-code level.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available in a repository online (Cetiner et al. 2020) in accordance with funder data retention policies.

Acknowledgments

The authors would like to acknowledge financial support from the Natural Science Foundation of Jiangsu Province (Grant No. BK20170680) and the National Natural Science Foundation of China (Grant Nos. 51708106, 52078119) that enabled the first author to spend a term as a Visiting Scholar at the University of California, Los Angeles.

References

ACI (American Concrete Institute). 2019. *Building code requirements for structural concrete and commentary.* ACI 318-19. Farmington Hills, MI: ACI.

Bae, S., and O. Bayrak. 2008. "Plastic hinge length of reinforced concrete columns." *ACI Struct. J.* 105 (3): 290–300.

Barrera, A., J. Bonet, M. Romero, and P. Miguel. 2011. "Experimental tests of slender reinforced concrete columns under combined axial load and lateral force." *Eng. Struct.* 33 (12): 3676–3689. https://doi.org/10.1016/j.engstruct.2011.08.003.

Bayrak, O. 1998. "Seismic performance of rectilinearly confined high strength concrete columns." Ph.D. thesis, Dept. of Civil and Mineral Engineering, Univ. of Toronto.

Berry, M. P., D. E. Lehman, and L. N. Lowes. 2008. "Lumped-plasticity models for performance simulation of bridge columns." *ACI Struct. J.* 105 (3): 270–279.

Breiman, L. 1996a. "Bagging predictors." *Mach. Learn.* 24 (2): 123–140.

Breiman, L. 1996b. "Stacked regressions." *Mach. Learn.* 24 (1): 49–64.

Cetiner, B., D.-C. Feng, M. R. Azadi Kakavand, and E. Taciroglu. 2020. *A data-driven plastic hinge length prediction model for rectangular reinforced concrete columns.* DesignSafe-CI. Alexandria, VA: NHERI Network. https://doi.org/10.17603/ds2-cy48-d115.

Chou, J.-S., C.-F. Tsai, A.-D. Pham, and Y.-H. Lu. 2014. "Machine learning in concrete strength simulations: Multi-nation data analytics." *Constr. Build. Mater.* 73 (Dec): 771–780. https://doi.org/10.1016/j.conbuildmat.2014.09.054.

Drucker, H. 1997. "Improving regressors using boosting techniques." In Vol. 97 of *Proc., Fourteenth Int. Conf. on Machine Learning*, edited by D. Fisher, Jr, 107–115. Burlington, MA: Morgan Kaufmann.

Drucker, H., C. Cortes, L. D. Jackel, Y. LeCun, and V. Vapnik. 1994. "Boosting and other ensemble methods." *Neural Comput.* 6 (6): 1289–1301. https://doi.org/10.1162/neco.1994.6.6.1289.

Elmenshawi, A., T. Brown, and S. El-Metwally. 2012. "Plastic hinge length considering shear reversal in reinforced concrete elements." *J. Earthquake Eng.* 16 (2): 188–210. https://doi.org/10.1080/13632469.2011.597485.

Erdal, H. I., O. Karakurt, and E. Namli. 2013. "High performance concrete compressive strength forecasting using ensemble models based on discrete wavelet transform." *Eng. Appl. Artif. Intell.* 26 (4): 1246–1254. https://doi.org/10.1016/j.engappai.2012.10.014.

Feng, D. C., X. D. Ren, and J. Li. 2016. "Implicit gradient delocalization method for force-based frame element." *J. Struct. Eng.* 142 (2): 04015122. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001397.

Feng, D.-C., Z.-T. Liu, X.-D. Wang, Y. Chen, J.-Q. Chang, D.-F. Wei, and Z.-M. Jiang. 2020. "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach." *Constr. Build. Mater.* 230 (Jan): 117000. https://doi.org/10.1016/j.conbuildmat.2019.117000.

Feng, D.-C., and X.-D. Ren. 2017. "Enriched force-based frame element with evolutionary plastic hinge." *J. Struct. Eng.* 143 (10): 06017005. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001871.

Freund, Y., and R. E. Schapire. 1997. "A decision-theoretic generalization of on-line learning and an application to boosting." *J. Comput. Syst. Sci.* 55 (1): 119–139. https://doi.org/10.1006/jcss.1997.1504.

Friedman, J., T. Hastie, and R. Tibshirani. 2001. *The elements of statistical learning*. New York: Springer.

Hearst, M. A., S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf. 1998. "Support vector machines." *IEEE Intell. Syst. Appl.* 13 (4): 18–28. https://doi.org/10.1109/5254.708428.

Hjelmstad, K. D., and E. Taciroglu. 2003. "Mixed variational methods for finite element analysis of geometrically non-linear, inelastic Bernoulli-Euler beams." *Commun. Numer. Methods Eng.* 19 (10): 809–832. https://doi.org/10.1002/cnm.622.

Jeon, J.-S., A. Shafieezadeh, and R. DesRoches. 2014. "Statistical models for shear strength of RC beam-column joints using machine-learning techniques." *Earthquake Eng. Struct. Dyn.* 43 (14): 2075–2095. https://doi.org/10.1002/eqe.2437.

Kent, D. C., and R. Park. 1971. "Flexural members with confined concrete." *J. Struct. Div.* 97 (7): 1969–1990.

Kingma, D. P., and J. Lei Ba. 2015. "Adam: A method for stochastic optimization." In *Proc., 3rd Int. Conf. on Learning Representations*. Ithaca, NY: Cornell Univ.

Lu, Y., X. Gu, and J. Guan. 2005. "Probabilistic drift limits and performance evaluation of reinforced concrete columns." *J. Struct. Eng.* 131 (6): 966–978. https://doi.org/10.1061/(ASCE)0733-9445(2005)131:6(966).

McKenna, F. 2011. "OpenSees: A framework for earthquake engineering simulation." *Comput. Sci. Eng.* 13 (4): 58–66. https://doi.org/10.1109/MCSE.2011.66.

Mendis, P. 2002. "Plastic hinge lengths of normal and high-strength concrete in flexure." *Adv. Struct. Eng.* 4 (4): 189–195. https://doi.org/10.1260/136943301320896651.

Menegotto, M., and P. E. Pinto. 1973. "Method of analysis for cyclically loaded RC plane frames including changes in geometry and non-elastic behavior of elements under combined normal force and bending." In Vol. 11 of *Proc., IABSE Symp. on Resistance and Ultimate Deformability of Structures Acted on by Well Defined Repeated Loads*, 15–22. Zürich, Switzerland: International Association for Bridge and Structural Engineering.

Nair, V., and G. E. Hinton. 2010. "Rectified linear units improve restricted Boltzmann machines." In *Proc., 27th Int. Conf. on Machine Learning*, 807–814. Madison, WI: Omni Press.

Neuenhofer, A., and F. C. Filippou. 1997. "Evaluation of nonlinear frame finite-element models." *J. Struct. Eng.* 123 (7): 958–966. https://doi.org/10.1061/(ASCE)0733-9445(1997)123:7(958).

Ning, C.-L., and B. Li. 2015. "Probabilistic approach for estimating plastic hinge length of reinforced concrete columns." *J. Struct. Eng.* 142 (3): 04015164. https://doi.org/10.1061/(ASCE)ST.1943-541X.0001436.

Omran, B. A., Q. Chen, and R. Jin. 2016. "Comparison of data mining techniques for predicting compressive strength of environmentally friendly concrete." *J. Comput. Civ. Eng.* 30 (6): 04016029. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000596.

Pam, H., and J. Ho. 2009. "Length of critical region for confinement steel in limited ductility high-strength reinforced concrete columns." *Eng. Struct.* 31 (12): 2896–2908. https://doi.org/10.1016/j.engstruct.2009.07.015.

Panagiotakos, T. B., and M. N. Fardis. 2001. "Deformations of reinforced concrete members at yielding and ultimate." *ACI Struct. J.* 98 (2): 135–148.

Park, R., M. Priestley, and W. D. Gill. 1982. "Ductility of square-confined concrete columns." *J. Struct. Div.* 108 (4): 929–950.

Paulay, T., and M. N. Priestley. 1992. *Seismic design of reinforced concrete and masonry buildings*. New York: Wiley.

Paultre, P., F. Légeron, and D. Mongeau. 2001. "Influence of concrete strength and transverse reinforcement yield strength on behavior of high-strength concrete columns." *ACI Struct. J.* 98 (4): 490–501.

Pham, A.-D., N.-D. Hoang, and Q.-T. Nguyen. 2015. "Predicting compressive strength of high-performance concrete using metaheuristic-optimized least squares support vector regression." *J. Comput. Civ. Eng.* 30 (3): 06015002. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000506.

Priestley, M., and R. Park. 1987. "Strength and ductility of concrete bridge columns under seismic loading." *ACI Struct. J.* 84 (1): 61–76.

Ren, Y., L. Zhang, and P. N. Suganthan. 2016. "Ensemble classification and regression—Recent developments, applications and future directions." *IEEE Comput. Intell. Mag.* 11 (1): 41–53. https://doi.org/10.1109/MCI.2015.2471235.

Safavian, S. R., and D. Landgrebe. 1991. "A survey of decision tree classifier methodology." *IEEE Trans. Syst. Man Cybern.* 21 (3): 660–674. https://doi.org/10.1109/21.97458.

Salehi, H., and R. Burgueño. 2018. "Emerging artificial intelligence methods in structural engineering." *Eng. Struct.* 171 (Sep): 170–189. https://doi.org/10.1016/j.engstruct.2018.05.084.

Santos, A., E. Figueiredo, M. Silva, R. Santos, C. Sales, and J. C. Costa. 2017. "Genetic-based EM algorithm to improve the robustness of Gaussian mixture models for damage detection in bridges." *Struct. Control Health Monit.* 24 (3): e1886. https://doi.org/10.1002/stc.1886.

Schalkoff, R. J. 1997. Vol. 1 of *Artificial neural networks*. New York: McGraw-Hill.

Scott, B., R. Park, and M. Priestley. 1982. "Stress-strain behavior of concrete confined by overlapping hoops at low and high strain rates." *ACI Struct. J.* 79 (1): 13–27.

Scott, M., and O. Hamutçuoğlu. 2008. "Numerically consistent regularization of force-based frame elements." *Int. J. Numer. Methods Eng.* 76 (10): 1612–1631. https://doi.org/10.1002/nme.2386.

Sheikh, S. A., and S. S. Khoury. 1993. "Confined concrete columns with stubs." *ACI Struct. J.* 90 (4): 414–431.

Sheikh, S. A., D. V. Shah, and S. S. Khoury. 1994. "Confinement of high-strength concrete columns." *ACI Struct. J.* 91 (1): 100–111.

Shrestha, D. L., and D. P. Solomatine. 2006. "Experiments with AdaBoost.RT: An improved boosting scheme for regression." *Neural Comput.* 18 (7): 1678–1710. https://doi.org/10.1162/neco.2006.18.7.1678.

Siddique, R., P. Aggarwal, and Y. Aggarwal. 2011. "Prediction of compressive strength of self-compacting concrete containing bottom ash using artificial neural networks." *Adv. Eng. Software* 42 (10): 780–786. https://doi.org/10.1016/j.advengsoft.2011.05.016.

Solomatine, D. P., and D. L. Shrestha. 2004. "Adaboost.RT: A boosting algorithm for regression problems." In Vol. 2 of *Proc., 2004 IEEE Int. Joint Conf. on Neural Networks*, 1163–1168. New York: IEEE.

Spacone, E., F. C. Filippou, and F. F. Taucer. 1996. "Fiber beam-column model for non-linear analysis of RC frames. Part I: Formulation." *Earthquake Eng. Struct. Dyn.* 25 (7): 711–726. https://doi.org/10.1002/(SICI)1096-9845(199607)25:7<711::AID-EQE576>3.0.CO;2-9.

Tanaka, H. 1990. "Effect of lateral confining reinforcement on the ductile behaviour of reinforced concrete columns." Ph.D. thesis, College of Engineering, Univ. of Canterbury.

Thomas, G. D. 1997. "Machine learning research: Four current directions." *AI Mag.* 18 (4): 97–136.

Vu, D.-T., and N.-D. Hoang. 2016. "Punching shear capacity estimation of FRP-reinforced concrete slabs using a hybrid machine learning approach." *Struct. Infrastruct. Eng.* 12 (9): 1153–1161. https://doi.org/10.1080/15732479.2015.1086386.

Wu, X., et al. 2008. "Top 10 algorithms in data mining." *Knowl. Inf. Syst.* 14 (1): 1–37. https://doi.org/10.1007/s10115-007-0114-2.

Zenko, B., L. Todorovski, and S. Dzeroski. 2001. "A comparison of stacking with meta decision trees to bagging, boosting, and stacking with other methods." In *Proc., 2001 IEEE Int. Conf. on Data Mining*, 669–670. New York: IEEE.

Zhou, Z.-H. 2009. "Ensemble learning." In *Encyclopedia of biometrics*, edited by S. Z. Li, 270–273. Berlin: Springer.

Zhou, Z.-H. 2012. *Ensemble methods: Foundations and algorithms*. Boca Raton, FL: Chapman and Hall/CRC Press.