Journal of Soft Computing in Civil Engineering 5-3 (2021) 32-57

Contents lists available at SCCE

Journal of Soft Computing in Civil Engineering

Journal homepage: www.jsoftcivil.com

# Performance Evaluation of Machine Learning Algorithms for Seismic Retrofit Cost Estimation Using Structural Parameters

**N. Safaeian Hamzehkolaei[1*], M. Alizamir[2]**

1. Assistant Professor, Department of Civil Engineering, Bozorgmehr University of Qaenat, Qaen, Iran
2. Ph.D., Department of Civil Engineering, Hamedan Branch, Islamic Azad University, Hamedan, Iran
Corresponding author: *nsafaeian@buqaen.ac.ir*

https://doi.org/10.22115/SCCE.2021.284630.1312

## ARTICLE INFO

Article history:
Received: 05 May 2021
Revised: 21 July 2021
Accepted: 10 August 2021

Keywords:
Seismic retrofit cost;
Random forest (RF);
Extreme learning machine (ELM);
Classification and regression tree (CART);
Multivariate adaptive regression spline (MARS).

## ABSTRACT

Estimation of the seismic retrofit cost (SRC) is a complicated task in construction projects. In this study, the performance of four machine learning algorithms (MLAs), including Random Forest (RF), Extreme Learning Machine (ELM), Classification and Regression Tree (CART), and Multivariate Adaptive Regression Spline (MARS), was examined in estimating SRC values. The total floor area (TFA), number of stories (NS), seismic weight (SW), seismicity (S), soil type (ST), plan configuration (PC), and structural type (STT) were considered as structural input variables. To achieve the best performance of applied MLAs, twenty-two scenarios based on different combinations of input variables were considered. The correlation coefficient (r), Root Mean Squared Error (RMSE), Adjusted R-squared, and Nash-Sutcliffe efficiency (NSE) metrics together with the Taylor diagram were used to compare the accuracy of applied models. A sensitivity analysis using the RReliefF algorithm showed that TFA, SW, and PC are the most influential parameters, whereas the ST and STT have negative influences on SRC values. Comparison analysis results indicated that the ELM model with r of 0.896, RMSE of 0.081, and NSE of 0.758 had the best performance among other employed MLAs. Also, the RF regression achieved the second rank. In conclusion, the ELM model with single-layer feedforward neural network was superior to other data-driven models; therefore, it can be applied as an efficient tool for estimating SRC values using structural input parameters.

How to cite this article: Safaeian Hamzehkolaei N, Alizamir M. Performance evaluation of machine learning algorithms for seismic retrofit cost estimation using structural parameters. J Soft Comput Civ Eng 2021;5(3):32–57. https://doi.org/10.22115/scce.2021.284630.1312.

# 1. Introduction

Prediction of seismic retrofit cost (SRC) of structures is a complex task because of different effective parameters in each building. Finding a reliable tool for estimating SRC in construction projects is one of the concerns of project managers due to limited resources. Steel belts, Shotcrete, and fiber-reinforced polymer (FRP) are common retrofit actions that can be implemented for any masonry earthquake-prone buildings to improve structural performance, e.g., increasing the amount of the lateral strength for mitigating the corresponding risks [1]. Also, it can be mentioned that, before any decision-making regarding the establishment of a strategy for earthquake-prone buildings, the SRC value should be accurately predicted to decrease the risk level for unreinforced structures in the high seismic regions.

Previous studies developed SRC estimation models using different data-driven approaches [1–6]. All the reported methods that applied to estimate SRC may suffer from different limitations, e.g., uncertainties in the tuning of effective parameters and linear nature of the applied models that may result in overtraining and other problems that can take place in the training phase. To overcome the mentioned shortcoming, this study suggested a novel training model, a single hidden layer feedforward neural network (SHLFFNN)-based Extreme Learning Machine (ELM) that has several advantages compared with existing regression models. Results of previous works verify that the ELM model can be successfully applied in parameter estimation in different fields [7–10]. Alizamir et al. (2019) [7] applied several machine learning models, including ELM, multi-layer perceptron artificial neural network (MLPANN), and radial basis function (RBF), in modeling groundwater level fluctuations. They found that the ELM model provides better results than other compared models. Yaseen et al. (2019) [8] compared the performance of support vector regression (SVR) and ELM models for river flow forecasting. They concluded that the ELM as an intelligent expert system could be used effectively to forecast flow in rivers. Al-Shamiri et al. (2019) [9] compared the ability of ELM and MLPANN models for predicting high-strength concrete compressive strength and showed that ELM performed better than the MLPANN. More recently, Nayak et al. (2021) developed an ELM-based model for assessing the compressive strength of concrete [10].

In the last decade, data-driven approaches, including Random Forest (RF) regression [11–13], ELM [7–10,14–17], Multivariate Adaptive Regression Spline (MARS) [13,18–20], Classification and Regression Tree (CART) [21–24], and Extreme Gradient Boosting (XGBoost) method [11,13,18] have been successfully applied in different fields [25]. Regarding estimation of SRC, several studies have also been conducted using artificial intelligence models.

Chen and Huang [2] have investigated the performance of linear regression and ANN to estimate retrofit costs and duration of reconstruction projects for schools in Taiwan due to reconstruction following earthquake damage. The results showed that ANN yields better prediction results than the regression model, and the floor area provides a good basis for estimating the cost and duration of school reconstruction projects. Jafarzadeh et al. [3] provided a comprehensive dataset for SRC prediction from 158 public school buildings with a framed structure in Iran. Jafarzadeh et al. (2014) have suggested a series of nonparametric artificial neural network (ANN) models for

SRC prediction of earthquake-prone school buildings with a framed structure [4]. They performed a sensitivity analysis for finding the most effective structural parameters in SRC and introduced the total area of building as the key predictor of SRC. Jafarzadeh et al. [5] applied a multi-linear regression (MLR) model to estimate the SRC. In this study, fourteen independent influential variables were considered for training and testing MLR models. Based on the backward elimination (BE) regression analysis, they concluded that building age and compliance with seismic design code are insignificant predictors of SRC, whereas building total plan area, number of stories, structural type, seismicity, soil type, weight, and plan irregularity are the most statistically influential variables on SRC [5].

In another study, Jafarzadeh et al. (2015) [6] have investigated a retrofit cost predictive model for confined masonry structures using statistical regression analysis. A series of stepwise regression models have been developed using reliable data collected from 183 masonry school buildings in Iran. The mortar quality and concrete quality of confinement elements have been considered as input parameters for SRC. Similar to framed buildings, the total floor area was defined as the most important factor in estimating SRC for confined masonry structures [6]. Nasrazadani et al. (2017) [1] have presented a probabilistic cost model for SRC prediction as a continuous function of the desired retrofit level (or performance gain) using linear Bayesian regression based on their own collected database from 167 retrofits of masonry school buildings in Iran. They claimed that the proposed model by quantifying the significant uncertainties in SRC modeling using Bayesian regression could be employed for risk and reliability analysis. The pre-retrofit building value and the increase in lateral strength were also determined as the most important predictors of SRC [1].

Fung et al. (2017) have developed a standard linear regression-based model to estimate SRC by considering only the interaction between seismicity and performance objective and showed that a simple model with different combinations of predictors has better accuracy and lower error than the FEMA 156 model [26]. The training process in [26] was based on the “hold-out” method. In another study, Fung et al. (2018) [27] have developed a model to estimate structural SRC for typical federal buildings by considering the building construction type and square footage as essential factors affecting SRC. More recently, Fung et al. (2020) [28] employed a Generalized Linear Model (GLM) to predict SRC in terms of structural parameters based on the historical data provided by FEMA 156. The nested K-fold cross-validation was applied to not only use all of the data during the training phase but also to perform both model selection and model evaluation. The developed GLM-based framework is able to provide a fast approximation of the SRC especially for decision-makers with large building portfolios.

It can be concluded from the previously studies that the ELM model has not yet been applied in estimating SRC. This research is the first study that applies ELM to estimate SRC values in order to implement an efficient policy analysis for risk mitigation. The main objectives of this research are: (1) to investigate the influence of different sets of input parameters for predicting structural SRC, (2) to investigate the performance of ELM, CART, MARS, and RF algorithms for estimating SRC, and (3) to evaluate the models' uncertainty and sensitivity analysis.

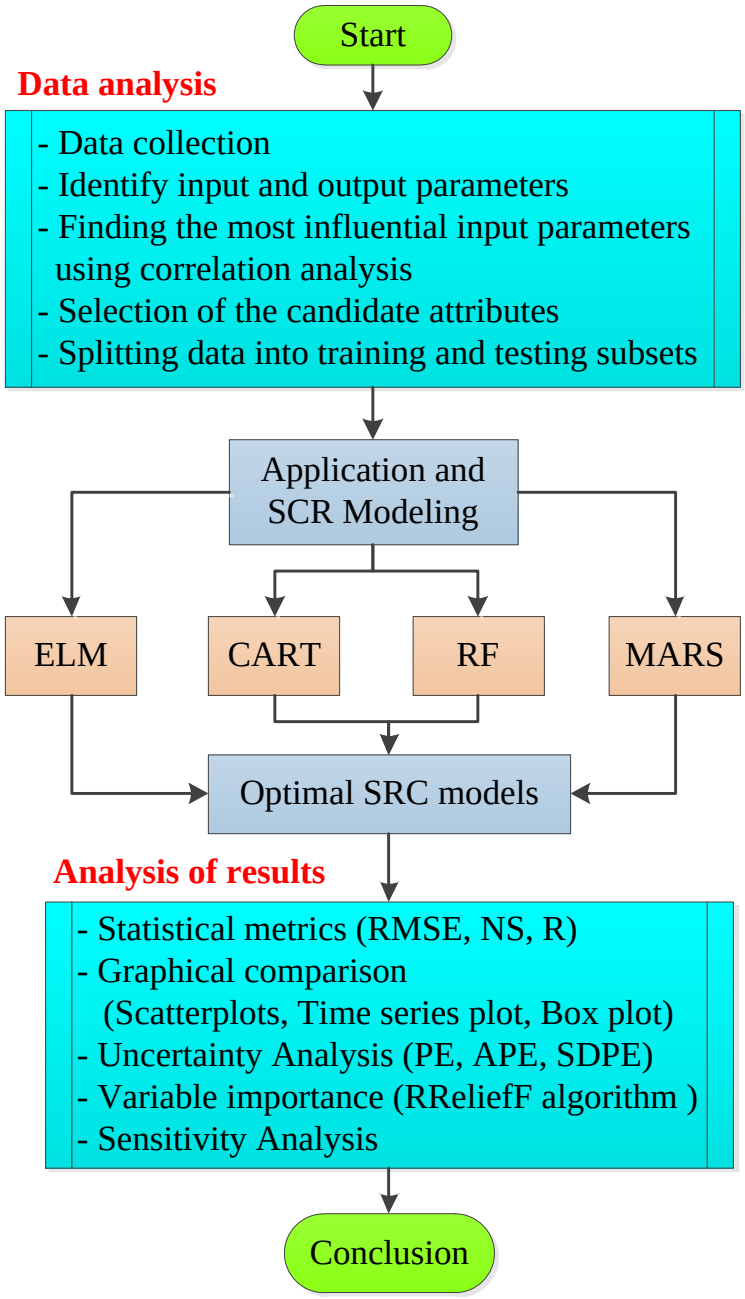



**Fig. 1.** Workflow of the proposed predictive models for SRC estimation.

A general workflow of the implementation of the proposed data-driven techniques is shown in Fig. 1.

The rest of this paper is structured as follows: An overview of different employed machine learning algorithms is presented in Section 2. The experimental dataset is presented in Section 3. Performance evaluation of different approaches and the uncertainty and sensitivity analysis results using the RReliefF algorithm is presented in Section 4. Finally, Section 5 discusses the concluding remarks.

## 2. Overview of different machine learning algorithms

### 2.1. Extreme learning machine (ELM)

The training process in traditional ANNs is based on an iterative approach using a gradient descent algorithm to tuning weights and biases that may result in slow training speed and/or local minima problems. ELM is one of the newly developed training algorithms for SHLFFNN. The main feature of the new training process in ELM is that it can randomly assign weights and biases. Also, the output weights in the ELM model can be analytically calculated using generalized inverse mathematical operation. The high learning speed is one of the crucial properties of ELM that leads to better generalization capability compared with traditional ANNs.

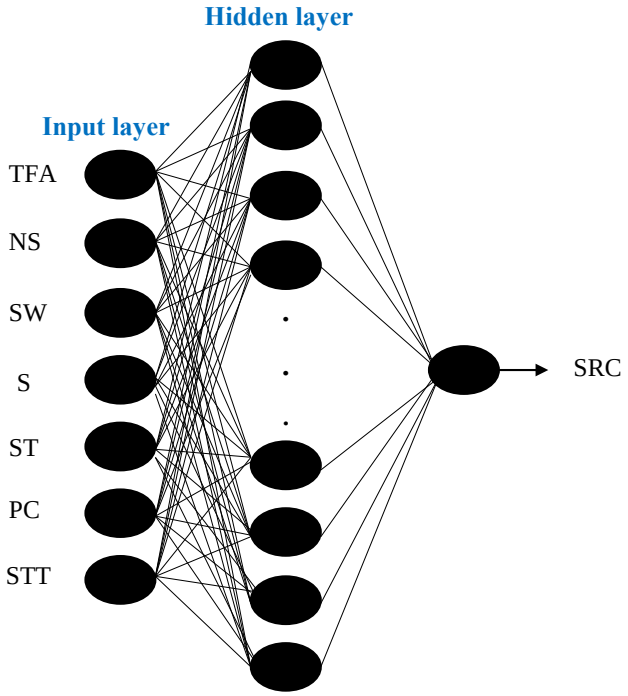



**Fig. 2.** A general structure of an ELM model used in this study for SRC estimation.

A typical SHLFFNN having $L$ hidden nodes, activation function $h(x)$, and $M$ samples can be expressed as [29]:

$$f_L(X_j) = \sum_{i=1}^{L} \beta_i h(W_i X_i + b_i) = t_i \qquad j = 1,2,\dots,M \tag{1}$$

where $W_i$ and $b_i$ denote input weights and biases, respectively.

By applying the least-square (LS) technique, ELM determines input weights ($W_i$) and biases ($b_i$) in order to compute the output weights $\beta$. In addition, activation function is defined as:

$$Minimize: ||H\beta - T||^2 \; and \; ||\beta|| \tag{2}$$

where $H$ is the hidden layer output matrix defined in Eq. (3).

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \dots & h_L(x_1) \\ \vdots & \vdots & \vdots \\ h_1(x_N) & \vdots & h_L(x_N) \end{bmatrix} \tag{3}$$

The main aim of ELM is minimizing $||\beta||$ that is equal to maximization of $\frac{2}{||\beta||}$. By applying the least square method, the relation between the Moore–Penrose generalized inverse ($H^+$) and the output weights ($\beta$) for SHLFFNN can be obtained as:

$$\beta = H^+ Y \tag{4}$$

The general architecture of the ELM model for SRC estimation is shown in Fig. 2. More details about ELM can be found in [30,31].

## 2.2. Classification and regression tree (CART)

CART is a type of the recursive data-driven approach that is applicable for both categorical and continuous variables [32]. For the first one, a classification tree for classifying such classes using some exogenous rules is employed. For the latter, a regression tree for prediction problems using predictors (input variables) and responses can be established [21]. Three main stages of CART are as follows [22,23]:

1. Establishing maximum tree via squared residuals minimization (SRM) approach.

2. Finding best parameters for tree size using cross-validation and optimization techniques. The complexity element (cp) can be considered to improve the procedure of selecting the best tree size.

3. Generating or classifying new data using established rules and trees. After completing this step, new outputs can be calculated for each of the new predictors.

CART algorithm divides the independent variables dataset from parent nodes using a binary-dividing process to generate child nodes based on their purity. For minimizing impurity of the samples, impurity measure can be defined as [24]:

$$\Delta i(s,t) = i(t) - p_L^i(t_L) - p_R^i(t_R) \tag{5}$$

where $i(t)$, $p_L^i(t_L)$, and $p_R^i(t_R)$ denote impurity before dividing process, left child node, and right child node, respectively. Also, in CART model the Gini index ($I_G$) is applied to choose the best split as follows:

$$I_G\left(t_{X(x_i)}\right) = 1 - \sum_{j=1}^{m} f\left(t_{X(x_i)}, j\right)^2 \tag{6}$$

where $f\left(t_{X(x_i)}, j\right)$ is the subset of the observed values by considering leave $j$ at node $t$.

In this study, a regression tree constructed to estimate SRC values using structural input parameters.

## 2.3. Multivariate adaptive regression spline (MARS)

MARS is one of the nonlinear approaches that can be applied to estimate numeric parameters by considering complex properties of inputs and output parameters [33]. This method is based on a divide-and-conquer strategy and divides the training data into separate splines of varying slopes [20]. MARS implements a mathematical relationship between effective parameters using basis functions (BFs), without any assumption regarding predictors and responses. BFs can be generated by stepwise searching through possible univariate candidate knots. The two main steps of MARS are the forward phase and the backward phase. By using the first one, appropriate input parameters can be identified and in the second phase, the unnecessary samples will be removed to enhance the model performance by using the Generalized Cross-Validation (GCV) approach. The GCV for N samples of the training set can be calculated as:

$$GCV = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - f(x_i)\right)^2}{\left(1 - \frac{M + d \times (M-1)/2}{N}\right)^2} \tag{7}$$

where $M$ is the number of BFs, $d$ denotes the penalty for each BF.

For a given target variable $y$ and the predictor variable $X = \left[X_1, X_2, \ldots, X_p\right]$, a general MARS model can be defined as:

$$y = f\left(X_1, X_2, \ldots, X_p\right) + \mathrm{e} = f(X) + e \tag{8}$$

where $f(x)$ denotes the predicted response, and $e$ is error of fitting. By considering a linear combination of BFs, Equation 8 is generally expressed as [20]:

$$f(X) = \beta_0 + \sum_{m=1}^{M} \beta_m \lambda_m(X) \tag{9}$$

where $\beta_m$ and $\lambda_m(m = 1,2, \ldots, M)$ denote constant coefficients that can be calculated using least square method. Two types of BFs are used the MARS model for mapping from parameter $X_p$ (input parameters) to response $Y$.

$$\begin{aligned} Y &= max(0, X - c), \\ Y &= max(0, c - X) \end{aligned} \tag{10}$$

where $c$ denotes the threshold value. More details about MARS can be found in [20,33–35].

## 2.4. Random forest (RF)

Breiman [36] suggested the RF technique as one of the types of decision trees using different subsets of data. RF by applying an ensemble learning approach enhances weak learners by using

a voting scheme. In the first step, RF generates tree samples using the original training elements by employing bootstrap sampling [37,38]. Then, the generated trees grow at each node, and parameters of the RF methodology are tuned to find the optimal ones. At the next step, ensemble averaging is applied to estimate responses. Finally, Out-of-bag (OOB) error estimation is calculated by using the data [39]. In this stage, two statistical parameters, coefficients of determination and mean square error (MSE) of the OOB, are calculated to investigate the established model.

$$MSE_{OOB} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n} \tag{11}$$

$$R_{RF}^2 = 1 - \frac{MSE_{OOB}}{\bar{\sigma}_y^2} \tag{12}$$

where $n$ is the number of samples, $\bar{\sigma}_y$ is the variance of OOB, and $y_i$ and $\bar{y}$ denote the observed and estimated values, respectively. Different steps of the applied RF are shown in Fig. 3.

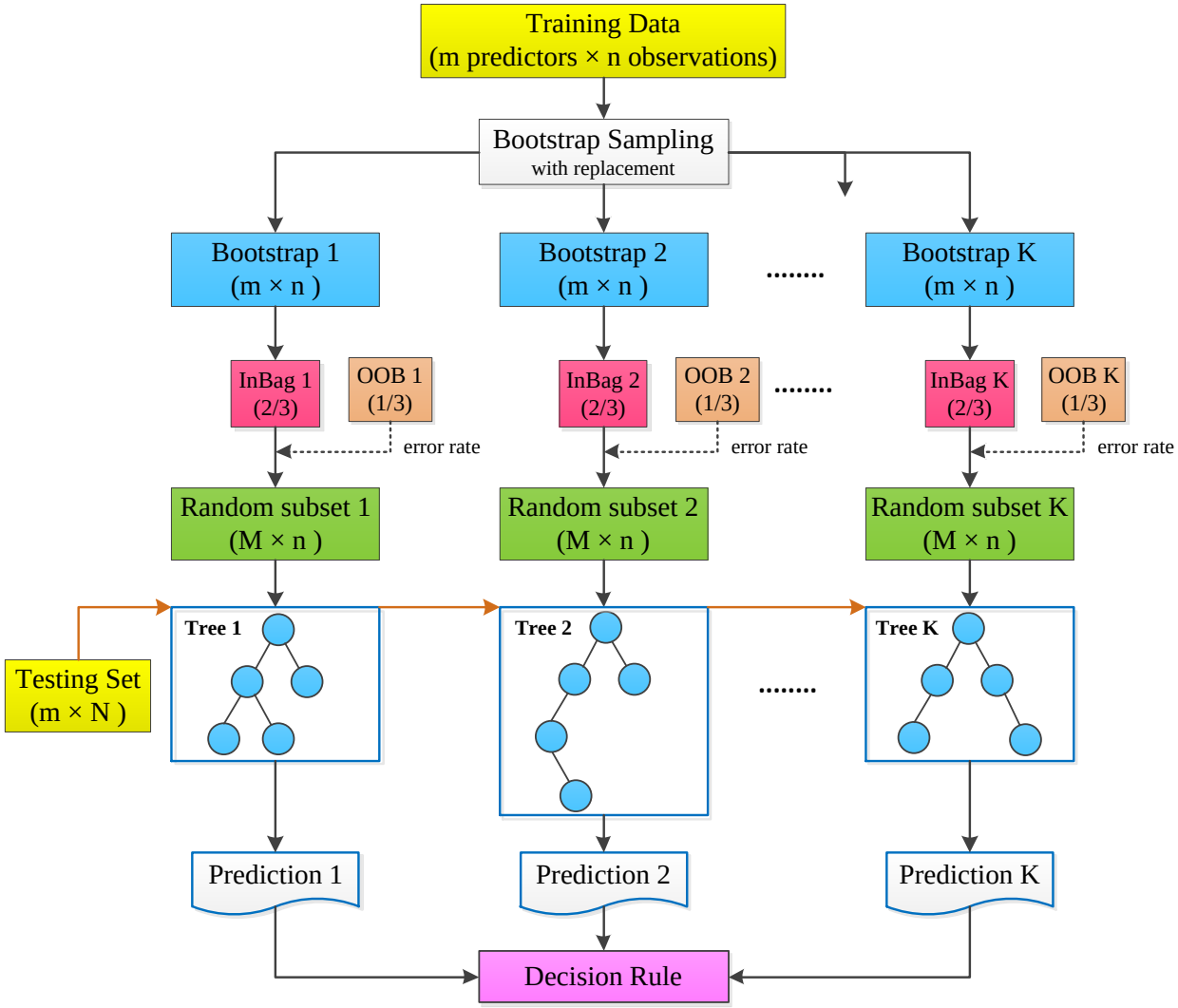



**Fig. 3.** Flowchart of the RF algorithm in estimating SRC values.

### 2.5. Model assessment criteria

In this study, the Nash–Sutcliffe model efficiency (NSE) coefficient, root mean square error (RMSE), the correlation coefficient ($r$), and the Adjusted R-squared $\left(R^2_{adj}\right)$ as standard statistical indices are utilized for performance comparison of the applied models. These indicators are expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left((SRC)_{io} - (SRC)_{ip}\right)^2};\ 0 \le RMSE < \infty \tag{13}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left((SRC)_{io} - (SRC)_{ip}\right)^2}{\sum_{i=1}^{n}((SRC)_{io} - (\overline{SRC})_{io})^2};\ -\infty < NSE \le 1 \tag{14}$$

$$r = \frac{\sum_{i=1}^{n}((SRC)_{io} - (\overline{SRC})_{io})\left((SRC)_{ip} - (\overline{SRC})_{ip}\right)}{\sqrt{\sum_{i=1}^{n}((SRC)_{io} - (\overline{SRC})_{io})^2 \sum_{i=1}^{n}\left((SRC)_{ip} - (\overline{SRC})_{ip}\right)^2}};\ 0 < r \le 1 \tag{15}$$

$$R^2_{adj} = 1 - \left[\frac{(1-R^2)(n-1)}{(n-k-1)}\right]; \tag{16}$$

where $(SRC)_{io}$ and $(SRC)_{ip}$ denote the observed and predicted SRC values, respectively. $(\overline{SRC})_{io}$ and $(\overline{SRC})_{ip}$ are the average of the observed and predicted SRC values, respectively. Also, $n$ and $k$ indicate the number of samples in the dataset and the number of independent variables used in the model, respectively. The RMSE metric ($0 \le RMSE < \infty$) with an optimum value of 0 is utilized for comparing the accuracy of the applied models. The $r$ index ($0 < r \le 1$) with an ideal value of 1 indicates the competence of the employed predictors for SRC prediction. The NSE index ($-\infty < NSE \le 1$) is used to evaluate the goodness of fit of the devolved models. For a perfect fit between observed and predicted SRC (i.e., in the situation with a zero error variance), the resulting NSE equals 1 ($NSE = 1$). Actually, the $NSE = 0$ denotes the model has the same predictive power as the mean of observed SRC, whereas negative values (NSE < 0) indicate that the observed mean performs better than the developed SRC model. It is also worth mentioning that, when applied for regression analysis, the NSE is equivalent to the coefficient of determination ($R^2$). The $R^2_{adj}$ indicator takes into account the number of independent variables used for predicting the target variable. If R-squared does not increase significantly on the addition of a new independent variable, the value of Adjusted R-squared will actually decrease. On the other hand, if on adding a new independent variable we see a significant increase in R-squared value, then the Adjusted R-squared value will also increase.

## 3. Experimental dataset

The dataset used in this study consists of 158 data points, provided by Jafarzadeh et al. [3], which were collected from earthquake-prone public school buildings with a framed structure in Iran. The general dataset includes information about fourteen variables influencing SRC that are

reflected in construction tender documents. Based on the previous studies by Jafarzadeh et al. [3–5], the total floor area (TFA), number of stories (NS), seismic weight (SW), seismicity (S), soil type (ST), plan configuration (PC), and structural type (STT) are considered as the influential input variables. The total SRC values comprising structural costs and the costs of architecture and finishes (also known as "restoration cost") are considered as output variables.

The statistical properties of the used dataset for estimating SRC are listed in Table 1. Because the parameters have different dimensions, converging the models may be difficult. Moreover, to show better generalization performance of the applied models, data were normalized between the range of 0 and 1. These data were randomly divided into training and testing subsets. It should be mentioned that 80% of the data (126 samples) was applied as a training subset and the remaining 20% (32 samples) was used for the testing subset. Besides, the "hold-out" method was used for both model evaluation and model selection, where a subset of the data is held out during the training procedure for all applied models. Moreover, the size of the training and testing subsets is also equal for all employed models. More details about this dataset can be found in [3].

**Table 1**
Basic statistical properties of the dataset.

| Data set | Variable | Average | Min. | Max. | St. Dev. |
|---|---|---|---|---|---|
| Training data (126 samples) | TFA ($m^2$) | 1832.8 | 260 | 6100 | 826.41 |
| | NS | 3.23 | 1 | 5 | 0.91 |
| | SW (ton) | 2080.7 | 15 | 7801 | 1052.9 |
| | PC | 2.76 | 1 | 3 | 0.45 |
| | S | 2.66 | 2 | 4 | 0.50 |
| | ST | 0.79 | 0 | 1 | 0.40 |
| | STT | 3.84 | 1 | 6 | 2.27 |
| | SRC ($10^3$ U.S.$) | 95.40 | 11.23 | 293.32 | 46.78 |
| Testing data (32 samples) | TFA ($m^2$) | 1982.5 | 187 | 4035 | 816.82 |
| | NS | 3.18 | 1 | 5 | 0.93 |
| | SW (ton) | 2091.8 | 170 | 6357 | 1041.7 |
| | PC | 2.87 | 2 | 3 | 0.33 |
| | S | 2.62 | 2 | 3 | 0.49 |
| | ST | 0.68 | 0 | 1 | 0.47 |
| | STT | 3.62 | 1 | 6 | 2.33 |
| | SRC ($10^3$ U.S.$) | 92.87 | 11.28 | 252.60 | 49.32 |

## 4. Results and discussion

In this section, the results of the SRC values predicted from ELM, CART, MARS, and RF were compared with the observed SRC values to investigate the accuracy of suggested models. The MATLAB 2014b is utilized to implement the applied data-driven techniques.

### 4.1. Performance evaluation of different approaches

Based on the several input parameters (TFA, NS, SW, S, ST, PC, and STT), seven different scenarios are investigated to estimate SRC using structural parameters with the minimum RMSE

value for the test dataset. In the first and second scenarios, only one and two variables are considered as input parameters, respectively. The number of the selected input parameters is subsequently increased such that all input variables are considered in the last scenario. Moreover, for each scenario, different combinations of the input variables are also investigated to find the most influential combination of the input variables on the target. Finally, twenty-two different cases based on the different combinations of input parameters have been investigated. The performance of the employed models has been assessed for each scenario in terms of the correlation coefficient $r$, RMSE, $R^2_{adj}$, and NSE indices. It is worth mentioning that, the most influential parameters for each scenario are defined based on the correlation analysis.

**Table 2**
Results of ELM models for estimation of retrofit cost.

| Input combination | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **NSE** | **r** | $R^2_{adj}$ | **RMSE** | **NSE** | **r** | $R^2_{adj}$ |
| **One input (scenario 1)** | | | | | | | | |
| TFA | **0.0985** | **0.615** | **0.784** | **0.6119** | **0.0988** | **0.643** | **0.821** | **0.6311** |
| NS | 0.1434 | 0.185 | 0.431 | 0.1784 | 0.1424 | 0.259 | 0.517 | 0.2343 |
| SW | 0.0986 | 0.615 | 0.784 | 0.6119 | 0.1348 | 0.336 | 0.625 | 0.3139 |
| S | 0.1514 | 0.091 | 0.302 | 0.0837 | 0.1646 | 0.011 | 0.168 | < 0 |
| ST | 0.1565 | 0.029 | 0.172 | 0.0212 | 0.1635 | 0.023 | 0.158 | < 0 |
| PC | 0.1483 | 0.128 | 0.359 | 0.1210 | 0.155 | 0.122 | 0.378 | 0.0927 |
| STT | 0.1582 | 0.008 | 0.094 | <0.001 | 0.1655 | <0.001 | 0.062 | < 0 |
| **Two inputs (scenario 2)** | | | | | | | | |
| TFA, NS | **0.0962** | **0.633** | **0.796** | **0.6270** | **0.0946** | **0.673** | **0.839** | **0.6504** |
| TFA, SW | 0.0968 | 0.629 | 0.793 | 0.6230 | 0.0956 | 0.666 | 0.84 | 0.6430 |
| NS, SW | 0.1101 | 0.519 | 0.72 | 0.5112 | 0.1373 | 0.311 | 0.569 | 0.2635 |
| TFA, PC | 0.0961 | 0.633 | 0.796 | 0.6270 | 0.0971 | 0.655 | 0.843 | 0.6312 |
| NS, PC | 0.1317 | 0.313 | 0.559 | 0.3018 | 0.1265 | 0.415 | 0.652 | 0.3747 |
| SW, PC | 0.1063 | 0.552 | 0.743 | 0.5447 | 0.1266 | 0.415 | 0.651 | 0.3747 |
| **Three inputs (scenario 3)** | | | | | | | | |
| TFA, NS, SW | 0.0974 | 0.624 | 0.79 | 0.6148 | 0.0985 | 0.645 | 0.818 | 0.6070 |
| TFA, NS, PC | 0.0916 | 0.667 | 0.817 | 0.6588 | 0.0962 | 0.662 | 0.837 | 0.6258 |
| NS, SW, PC | **0.094** | **0.649** | **0.806** | **0.6404** | **0.0904** | **0.701** | **0.868** | **0.6690** |
| **Four inputs (scenario 4)** | | | | | | | | |
| TFA, NS, SW, PC | **0.09** | **0.679** | **0.824** | **0.6684** | **0.0893** | **0.708** | **0.897** | **0.6647** |
| TFA, NS, SW, S | 0.093 | 0.65 | 0.808 | 0.6384 | 0.0905 | 0.699 | 0.859 | 0.6544 |
| **Five inputs (scenario 5)** | | | | | | | | |
| TFA, NS, SW, PC, S | **0.0902** | **0.677** | **0.823** | **0.6635** | **0.09** | **0.704** | **0.887** | **0.6471** |
| TFA, NS, SW, PC, ST | 0.0945 | 0.647 | 0.803 | 0.6323 | 0.973 | 0.654 | 0.839 | 0.5875 |
| **Six inputs (scenario 6)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST | **0.0895** | **0.682** | **0.826** | **0.6660** | **0.0843** | **0.74** | **0.893** | **0.6776** |
| **Seven inputs (scenario 7)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST, STT | **0.0904** | **0.676** | **0.822** | **0.6568** | **0.0814** | **0.758** | **0.896** | **0.6874** |

The results of the three performance evaluation indicators (RMSE, NSE, and r) for the employed data-driven techniques, namely the ELM, CART, MARS, and RF, are provided in Tables 2-5, respectively. According to Tables 2-5, it can be seen that, in the first scenario (only one input

variable), all employed models achieved the best performance (r, NSE, $R^2_{adj}$, and RMSE) for the TFA parameter for both training and testing data. However, the performance of the ELM model (RMSE=0.0988, NSE=0.643, r=0.821, and $R^2_{adj}$=0.6311) is better than other compared models for testing data (Table 2). Based on the obtained results in the in the first scenario (Tables 2-5), the TFA variable was chosen as the most effective parameter to estimate SRC values. Similar results were reported by Jafarzadeh et al. [4–6].

**Table 3**
Results of CART models for estimation of retrofit cost.

| Input combination | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **NSE** | **r** | $R^2_{adj}$ | **RMSE** | **NSE** | **r** | $R^2_{adj}$ |
| **One input (scenario 1)** | | | | | | | | |
| TFA | **0.0763** | **0.769** | **0.877** | **0.7671** | **0.1214** | **0.461** | **0.712** | **0.4430** |
| NS | 0.1415 | 0.207 | 0.454 | 0.2006 | 0.1442 | 0.241 | 0.497 | 0.2157 |
| SW | 0.0858 | 0.708 | 0.841 | 0.7056 | 0.1387 | 0.297 | 0.563 | 0.2736 |
| S | 0.1514 | 0.091 | 0.302 | 0.0837 | 0.1646 | 0.011 | 0.168 | < 0 |
| ST | 0.1565 | 0.029 | 0.172 | 0.0212 | 0.1635 | 0.023 | 0.158 | < 0 |
| PC | 0.1483 | 0.128 | 0.359 | 0.1210 | 0.155 | 0.122 | 0.378 | 0.0927 |
| STT | 0.1582 | 0.008 | 0.094 | < 0.001 | 0.1657 | -0.002 | 0.062 | < 0 |
| **Two inputs (scenario 2)** | | | | | | | | |
| TFA, NS | **0.0662** | **0.826** | **0.909** | **0.8232** | **0.1173** | **0.497** | **0.759** | **0.4623** |
| TFA, SW | 0.0518 | 0.893 | 0.945 | 0.8913 | 0.125 | 0.429 | 0.702 | 0.3896 |
| NS, SW | 0.0754 | 0.775 | 0.88 | 0.7713 | 0.1549 | 0.123 | 0.426 | 0.0625 |
| TFA, PC | 0.0762 | 0.77 | 0.877 | 0.7663 | 0.1216 | 0.46 | 0.711 | 0.4228 |
| NS, PC | 0.1269 | 0.361 | 0.601 | 0.3506 | 0.1287 | 0.395 | 0.642 | 0.3533 |
| SW, PC | 0.09 | 0.679 | 0.824 | 0.6738 | 0.1508 | 0.169 | 0.457 | 0.1117 |
| **Three inputs (scenario 3)** | | | | | | | | |
| TFA, NS, SW | **0.0491** | **0.904** | **0.951** | **0.9016** | **0.1081** | **0.573** | **0.773** | 0.5273 |
| TFA, NS, PC | 0.0661 | 0.827 | 0.909 | 0.8227 | 0.1175 | 0.495 | 0.758 | 0.4409 |
| NS, SW, PC | 0.0514 | 0.895 | 0.946 | 0.8924 | 0.12 | 0.474 | 0.739 | **0.4176** |
| **Four inputs (scenario 4)** | | | | | | | | |
| TFA, NS, SW, PC | **0.0487** | **0.906** | **0.951** | **0.9029** | **0.1027** | **0.614** | **0.804** | **0.5568** |
| TFA, NS, SW, S | 0.0511 | 0.896 | 0.947 | 0.8926 | 0.128 | 0.425 | 0.699 | 0.3398 |
| **Five inputs (scenario 5)** | | | | | | | | |
| TFA, NS, SW, PC, S | **0.0501** | **0.9** | **0.948** | **0.8958** | **0.1033** | **0.61** | **0.8** | **0.5350** |
| TFA, NS, SW, PC, ST | 0.6 | 0.835 | 0.912 | 0.8281 | 0.112 | 0.512 | 0.762 | 0.4182 |
| **Six inputs (scenario 6)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST | **0.0497** | **0.902** | **0.949** | **0.8971** | **0.0979** | **0.649** | **0.825** | **0.5648** |
| **Seven inputs (scenario 7)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST, STT | **0.0763** | **0.769** | **0.877** | **0.7553** | **0.1214** | **0.461** | **0.712** | **0.3038** |

In the second scenario with combinations of two input variables, the combination of the TFA and NS variables (TFA, NS) has provided better results than other input combinations in the ELM (Table 2), CART (Table 3), and RF (Table 5) models during the testing phase. However, it is obviously seen from the results of the second scenario in Tables 2-5 that the ELM model (RMSE=0.0946, NSE=0.673, r=0.839, and $R^2_{adj}$=0. 6504) provided better results than the other compared algorithms for the test data. In this scenario, MARS model (Table 4) achieved the best

performance during testing period in the case of combination of the TFA and PC as input variables. Also, it should be mentioned that the MARS model with RMSE=0.0992, NSE=0.64, and r=0.824, and $R^2_{adj}$=0. 6152 achieved the second rank in the second scenario (Table 4).

In the third scenario, the best combination of input variables in each model is different from the others. However, as can been from Table 2, the combination of the NS, SW, and PC variables yielded better output than other combinations in the ELM model. In this case, the ELM model (RMSE=0.0904, NSE=0.701, r=0.868, and $R^2_{adj}$=0.669) showed better performance than the CART, MARS, and RF models for the test subset. In this scenario, the RF model with a combination of the TFA, NS, and PC variables in Table 5 achieved the second rank (RMSE=0.092, NSE=0.688, and r=0.853) after the ELM model during the testing phase.

**Table 4**
Results of MARS models for estimation of retrofit cost.

| Input combination | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **NSE** | **r** | $R^2_{adj}$ | **RMSE** | **NSE** | **r** | $R^2_{adj}$ |
| **One input (scenario 1)** | | | | | | | | |
| TFA | **0.1004** | **0.6** | **0.775** | **0.5968** | **0.1053** | **0.595** | **0.796** | **0.5815** |
| NS | 0.1415 | 0.207 | 0.454 | 0.2006 | 0.1442 | 0.241 | 0.497 | 0.2157 |
| SW | 0.1086 | 0.532 | 0.729 | 0.5282 | 0.1421 | 0.262 | 0.538 | 0.2374 |
| S | 0.1515 | 0.091 | 0.301 | 0.0837 | 0.1646 | 0.011 | 0.168 | < 0 |
| ST | 0.1589 | 0 | <0.001 | < 0 | 0.1657 | -0.002 | <0.001 | < 0 |
| PC | 0.1483 | 0.128 | 0.359 | 0.1210 | 0.155 | 0.122 | 0.378 | 0.0927 |
| STT | 0.1589 | 0 | <0.001 | < 0 | 0.1657 | <0.001 | <0.001 | < 0 |
| **Two inputs (scenario 2)** | | | | | | | | |
| TFA, NS | 0.0887 | 0.688 | 0.829 | 0.6829 | 0.1057 | 0.592 | 0.792 | 0.5639 |
| TFA, SW | 0.1007 | 0.597 | 0.773 | 0.5904 | 0.1113 | 0.547 | 0.76 | 0.5158 |
| NS, SW | 0.1107 | 0.514 | 0.717 | 0.5061 | 0.1385 | 0.299 | 0.561 | 0.2507 |
| TFA, PC | **0.0976** | **0.622** | **0.789** | **0.6159** | **0.0992** | **0.64** | **0.824** | **0.6152** |
| NS, PC | 0.135 | 0.278 | 0.527 | 0.2663 | 0.133 | 0.354 | 0.609 | 0.3094 |
| SW, PC | 0.1069 | 0.547 | 0.74 | 0.5396 | 0.1267 | 0.413 | 0.651 | 0.3725 |
| **Three inputs (scenario 3)** | | | | | | | | |
| TFA, NS, SW | 0.096 | 0.634 | 0.796 | **0.6250** | 0.1188 | 0.484 | 0.722 | 0.4287 |
| TFA, NS, PC | **0.1008** | **0.597** | **0.722** | 0.5871 | **0.1075** | **0.578** | **0.785** | **0.5328** |
| NS, SW, PC | 0.1016 | 0.591 | 0.768 | 0.5809 | 0.1105 | 0.554 | 0.776 | 0.5062 |
| **Four inputs (scenario 4)** | | | | | | | | |
| TFA, NS, SW, PC | **0.1007** | **0.598** | **0.773** | **0.5847** | **0.1078** | **0.575** | **0.785** | **0.5120** |
| TFA, NS, SW, S | 0.112 | 0.517 | 0.718 | 0.5010 | 0.125 | 0.429 | 0.701 | 0.3444 |
| **Five inputs (scenario 5)** | | | | | | | | |
| TFA, NS, SW, PC, S | **0.0936** | **0.652** | **0.807** | **0.6375** | **0.0984** | **0.646** | **0.816** | **0.5779** |
| TFA, NS, SW, PC, ST | 0.1091 | 0.538 | 0.735 | 0.5188 | 0.1121 | 0.511 | 0.763 | 0.4170 |
| **Six inputs (scenario 6)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST | **0.0959** | **0.635** | **0.797** | **0.6166** | **0.1** | **0.63** | **0.815** | **0.5412** |
| **Seven inputs (scenario 7)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST, STT | **0.0995** | **0.608** | **0.779** | **0.5847** | **0.1079** | **0.574** | **0.797** | **0.4497** |

The results of the fourth scenario in Tables 2-5 indicate that all models have shown the best performance for SRC prediction in the case of the combination of TFA, NS, SW, and PC as input variables. Results of Table 3 show that the CART model has the best performance indices among all other compared algorithms during the training procedure. However, comparing the results of developed models during the testing phase indicate that the CART model achieved the third rank among other compared algorithms. In this scenario, the ELM model significantly enhanced the accuracy of the CART, MARS, and RF models during the testing period. Also, the RF algorithm (RMSE=0.092, NSE=0.689, r=0.852, and $R^2_{adj}$=0.6429) showed the second optimum model for this case.

**Table 5.**
Results of RF models for estimation of retrofit cost.

| Input combination | Training | | | | Testing | | | |
|---|---|---|---|---|---|---|---|---|
| | **RMSE** | **NSE** | **r** | $R^2_{adj}$ | **RMSE** | **NSE** | **r** | $R^2_{adj}$ |
| **One input (scenario 1)** | | | | | | | | |
| TFA | **0.065** | **0.830** | **0.915** | **0.8286** | **0.115** | **0.515** | **0.727** | **0.4988** |
| NS | 0.141 | 0.206 | 0.454 | 0.1996 | 0.144 | 0.240 | 0.499 | 0.2147 |
| SW | 0.067 | 0.819 | 0.909 | 0.8175 | 0.146 | 0.221 | 0.518 | 0.1950 |
| S | 0.151 | 0.091 | 0.302 | 0.0837 | 0.164 | 0.010 | 0.168 | < 0 |
| ST | 0.156 | 0.029 | 0.172 | 0.0212 | 0.163 | 0.023 | 0.158 | < 0 |
| PC | 0.148 | 0.128 | 0.359 | 0.1210 | 0.154 | 0.124 | 0.378 | 0.0948 |
| STT | 0.158 | 0.008 | 0.094 | <0.001 | 0.165 | <0.001 | 0.057 | < 0 |
| **Two inputs (scenario 2)** | | | | | | | | |
| TFA, NS | **0.058** | **0.865** | **0.933** | **0.8628** | **0.105** | **0.593** | **0.785** | **0.5649** |
| TFA, SW | 0.052 | 0.889 | 0.947 | 0.8872 | 0.110 | 0.554 | 0.751 | 0.5232 |
| NS, SW | 0.067 | 0.821 | 0.911 | 0.8181 | 0.144 | 0.238 | 0.544 | 0.1854 |
| TFA, PC | 0.063 | 0.838 | 0.917 | 0.8354 | 0.106 | 0.585 | 0.779 | 0.5564 |
| NS, PC | 0.127 | 0.360 | 0.600 | 0.3496 | 0.132 | 0.359 | 0.616 | 0.3148 |
| SW, PC | 0.064 | 0.835 | 0.918 | 0.8323 | 0.133 | 0.352 | 0.604 | 0.3073 |
| **Three inputs (scenario 3)** | | | | | | | | |
| TFA, NS, SW | 0.053 | 0.887 | 0.947 | 0.8842 | 0.102 | 0.616 | 0.791 | 0.5749 |
| TFA, NS, PC | **0.062** | **0.844** | **0.924** | **0.8402** | **0.092** | **0.688** | **0.853** | **0.6546** |
| NS, SW, PC | 0.068 | 0.812 | 0.909 | 0.8074 | 0.128 | 0.398 | 0.632 | 0.3335 |
| **Four inputs (scenario 4)** | | | | | | | | |
| TFA, NS, SW, PC | **0.055** | **0.878** | **0.943** | **0.8740** | **0.092** | **0.689** | **0.852** | **0.6429** |
| TFA, NS, SW, S | 0.065 | 0.83 | 0.914 | 0.8244 | 0.112 | 0.513 | 0.761 | 0.4409 |
| **Five inputs (scenario 5)** | | | | | | | | |
| TFA, NS, SW, PC, S | **0.061** | **0.852** | **0.931** | **0.8458** | **0.093** | **0.683** | **0.855** | **0.6220** |
| TFA, NS, SW, PC, ST | 0.067 | 0.817 | 0.911 | 0.8094 | 0.0988 | 0.641 | 0.833 | 0.5720 |
| **Six inputs (scenario 6)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST | **0.059** | **0.858** | **0.936** | **0.8508** | **0.096** | **0.662** | **0.840** | **0.5809** |
| **Seven inputs (scenario 7)** | | | | | | | | |
| TFA, NS, SW, PC, S, ST, STT | **0.061** | **0.851** | **0.936** | **0.8422** | **0.098** | **0.649** | **0.838** | **0.5466** |

Comparing the results of the fifth scenario reveal that all employed models have achieved the best performance in both the training and testing phases for the combination of the TFA, NS, SW, PC, and S variables. In this case, the CART model has shown superior performance

compared to other algorithms during the training phase (Table 3). However, according to Table 5, the RF algorithm with RMSE=0.093, NSE=0.683, and r=0.855 presented better results than both CART and MARS approaches. Again, the ELM model (RMSE=0.09, NSE=0.704, r=0.887, and $R^2_{adj}$=0.6471) had the best accuracy among other compared algorithms.

In the sixth scenario, the capability of employed models for SRC prediction has been investigated by considering the combination of the TFA, NS, SW, PC, S, and ST as input variables. For this case, the CART and RF algorithms showed better performance than the ELM model in the training phase. Nevertheless, it is quite apparent from the results of this scenario in Tables 2-5 that the ELM model (RMSE=0.0843, NSE=0.74, and r=0.893) has yielded better SRC values than the other models for the testing subset. The RF algorithm with RMSE=0.096, NSE=0.6662, r=0.840, and $R^2_{adj}$=0.5809 could also be ranked as the second-best model for this case.

Finally, results of Tables 2-5 in the last scenario with the TFA, NS, SW, PC, S, ST, and STT variables, confirm that the ELM approach (RMSE=0.0814, NSE=0.758, and r=0.896) has superiority to CART, MARS, and RF approaches. Again the RF algorithm achieved the second rank in terms of all performance indices for the test subset. It can be seen from Table 3 that the CART model with RMSE=0.1214, NSE=0.461, and r=0.712 presented the worst results for this case among all compared algorithms.

**Table 6**
Comparison the results of the best models for estimation of the retrofit cost.

| Model | Input combination | Training | | | Testing | | |
|---|---|---|---|---|---|---|---|
| | | RMSE | NSE | r | RMSE | NSE | r |
| ELM | TFA, NS, SW, PC, S, ST, STT | 0.0904 | 0.676 | 0.822 | **0.0814** | **0.758** | **0.896** |
| RF | TFA, NS, SW, PC | 0.055 | 0.878 | 0.943 | 0.092 | 0.689 | 0.852 |
| CART | TFA, NS, SW, PC, S, ST | 0.0497 | 0.902 | 0.949 | 0.0979 | 0.649 | 0.825 |
| MARS | TFA, NS, SW, PC, S | 0.0936 | 0.652 | 0.807 | 0.0984 | 0.646 | 0.816 |
| MLP [4]* | TFA, NS, SW, PC, S, ST, STT | 0.2 | 0.831 | - | 0.247 | 0.734 | - |

* MLP: $R^2$ = 0.831, and MSE = 0.040 for train data, and $R^2$ = 0.734, and MSE = 0.061 for test data.

In Table 6, the best performance achieved by each model among different investigated scenarios and the corresponding input combinations is presented. For comparison purposes, results of the multilayer perceptron (MLP) neural network for SRC prediction, reported by Jafarzadeh et al. [4], are also presented in Table 6. As can be seen, in accordance with the ELM model, the best results of the MLP neural network have also been obtained in the last scenario [4]. Results of Table 6 verify that the ELM model had the best accuracy compared with the MLP and other proposed models for estimating the SRC values. As can be seen, while the CART model has the best training performance, the ELM model has better accuracy than other compared algorithms with regard to all performance indices (RMSE, NSE, and r) during the testing phase.

Fig. 4 illustrates the scatterplots of the predicted and the observed SRC values for the training dataset. These results are presented for the best scenario corresponding to each of the employed

models (defined in Table 6). It can be concluded from the reported regression coefficients in Fig. 4 that the CART prediction yielded better results than other developed models the training data. Also, it should be mentioned that the MARS model estimated SRC values with the highest error.

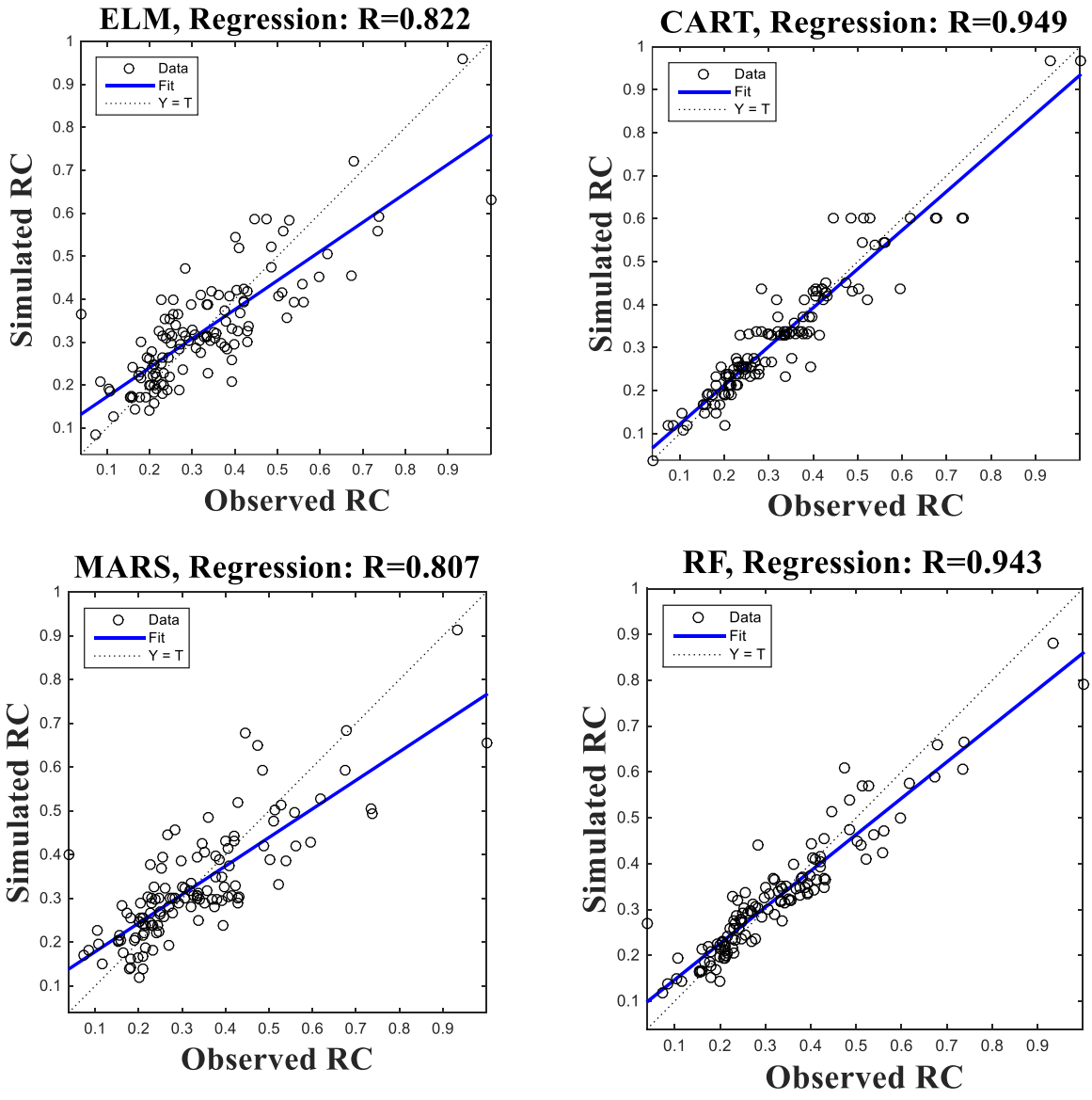



**Fig. 4.** Observed vs. predicted scatter plot for best SRC prediction models during the training phase.

The scatterplots between the observed and estimated SRC values for the testing subset are presented in Fig. 5. Comparison of regression coefficients of different models indicates that the ELM model achieved better performance than the CART, MARS, and RF models. Again, the MARS algorithm yielded the worst results in the estimation of the SRC values.

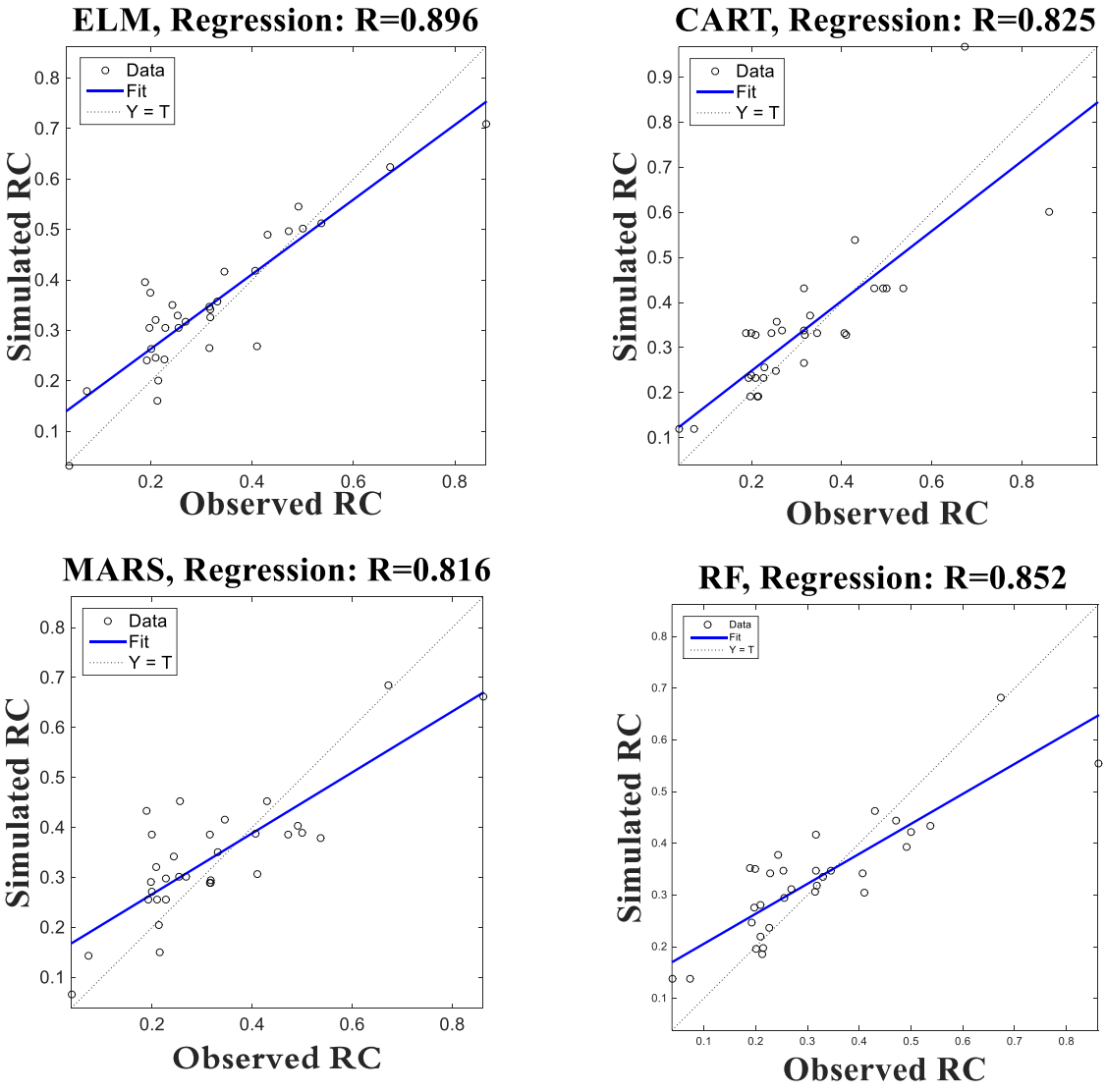
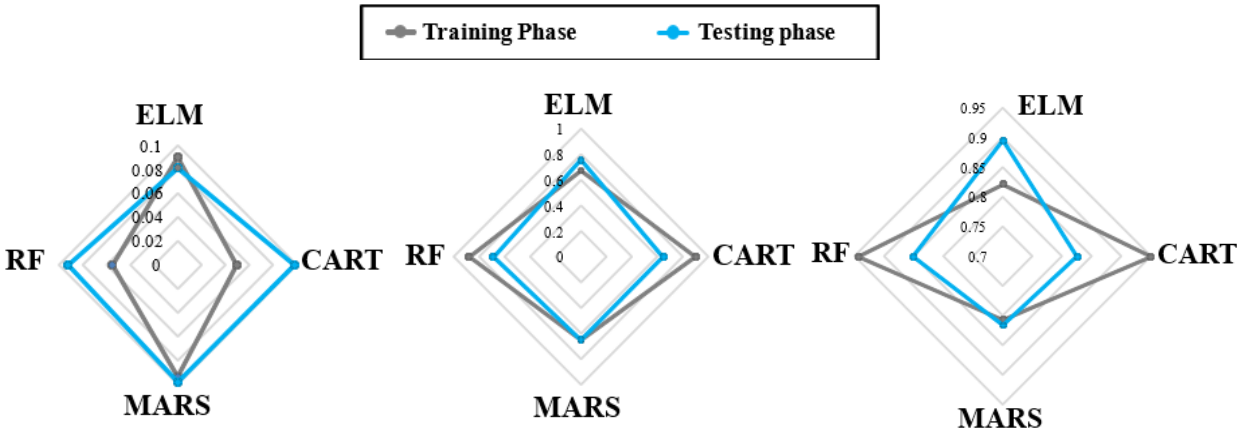



**Fig. 5.** Observed vs. predicted scatter plot for best SRC prediction models during the testing phase.




a) RMSE b) NSE c) r

**Fig. 6.** The Radar charts for the RMSE, NSE and r of the best developed SRC prediction models.

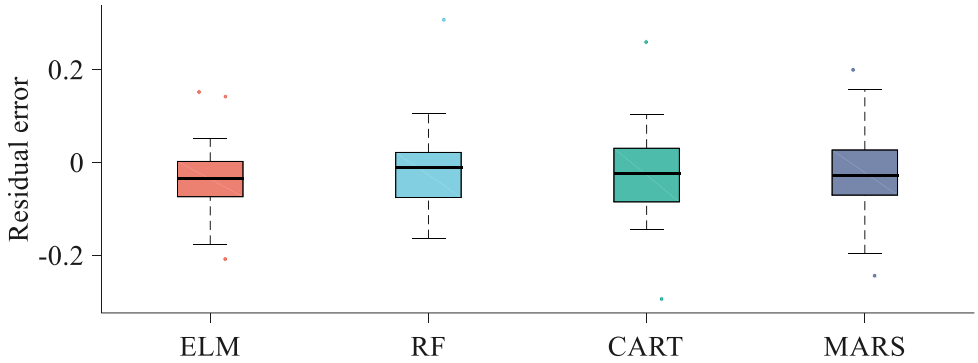



**Fig. 7.** Box plots of residual error of the best models for testing dataset.

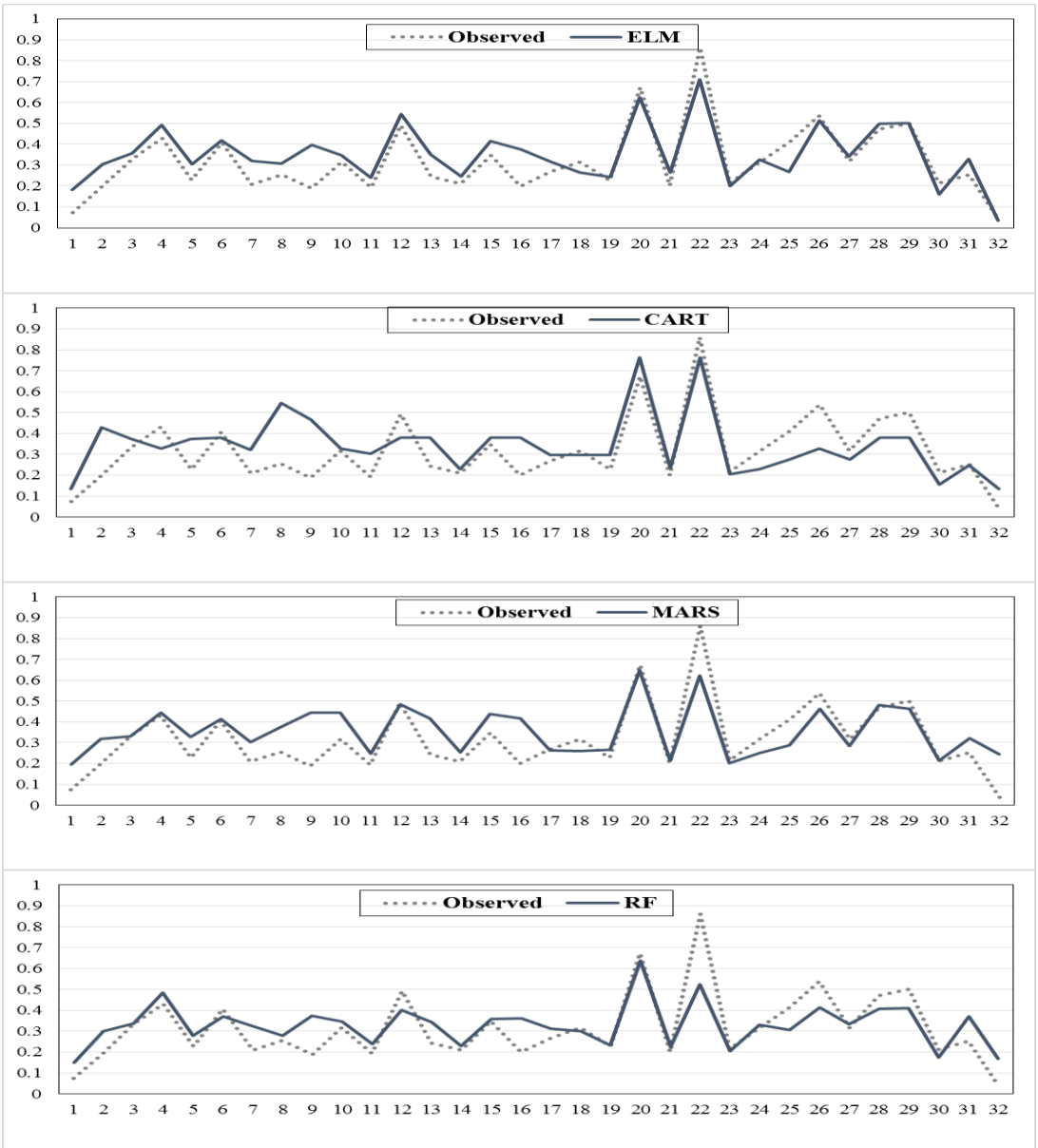



**Fig. 8.** Variation of the observed and estimated SRC values for the testing dataset.

Fig. 6 shows the radar chart for assessing three employed performance indices, including the RMSE, correlation coefficient r, and NSE values corresponding to the best developed SRC prediction models. In this figure, the results are depicted for both the training and testing data. According to Fig. 6, the ELM model had the best NSE and r values and the lowest RMSE during the testing phase. The RF model can also be ranked as the second optimal model for SRC prediction.

Fig. 7 shows the boxplots of standardized residual error values of the best models during the testing phase. The variations of the observed and predicted SRC values for the test dataset are also presented in Fig. 8. Comparing the residual error of different models in Fig. 7 indicate that the proposed ELM-based model for SRC prediction showed the lowest length compared with other employed algorithms. It is also obvious from Fig. 8 that the estimated SRC values of the ELM model thoroughly follow the corresponding observed ones. These results verify the superior performance of the proposed ELM model for prediction of the SRC values.

The histogram of the residual error of the best models is also presented in Fig. 9. These results are provided by considering the mean (μ) and standard deviation (σ) (SD) of the residual error corresponding to the testing dataset. As it is quite evident from Fig. 9, the ELM and the MARS models have yielded the lowest and highest value for the SD measure, respectively. It is in harmony with the trend of the RMSE values reported in Table 6.

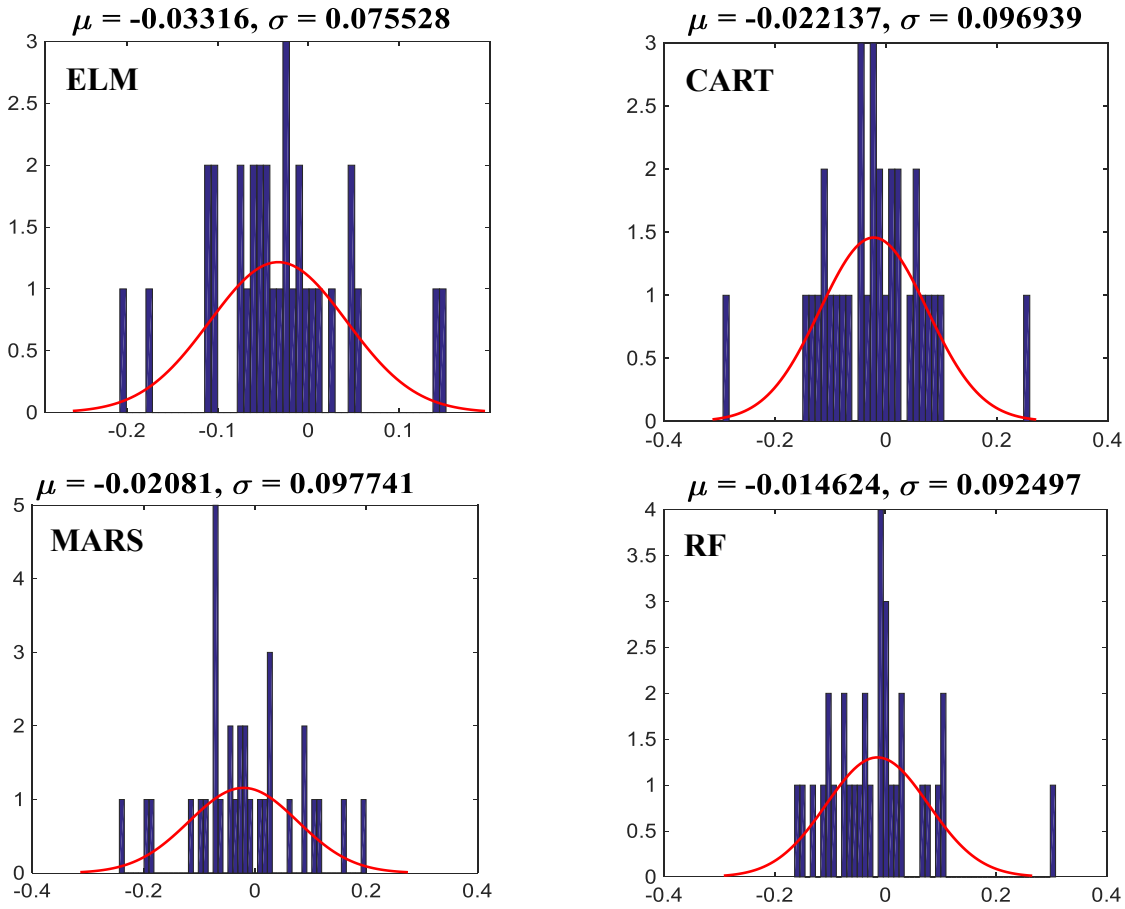



**Fig. 9.** Histogram of residual error for best models in estimation of retrofit cost.

Fig. 10 depicts the Taylor diagram for identifying the performance of different approaches in the last scenario, whereby the distance from the observed SRC values is a quantity of the centered RMS error in the simulated field. In this figure, the azimuth angle denotes to the correlation (r) between the predicted and observed SRC values while the radial distance from the origin signifies the ratio of the normalized SD of the simulation to that of the observation [40]. It is evident from Fig. 10 that the ELM model agrees best with observations while having the least RMS error (less than 0.08) and the highest correlation with observations among all other algorithms. The normalized SD of all models (radial distance from the origin) is clearly lower than that of the observed SRC values. However, the SD of the ELM and CART models (indicated by the dashed contour at approximately radial distance 0.14) are closer to the observed values. Moreover, the ELM model has a lower distance from the observed SRC values compared with the CART, MARS, and RF models. Therefore, it can be concluded that the proposed ELM model with TFA, NS, SW, PC, S, ST, and STT variables provided more accurate results compared with other employed machine learning algorithms. Fig. 10 also indicates that the RF model has a higher correlation with observations as well as a lower centered RMS error than both MARS and CART algorithms.

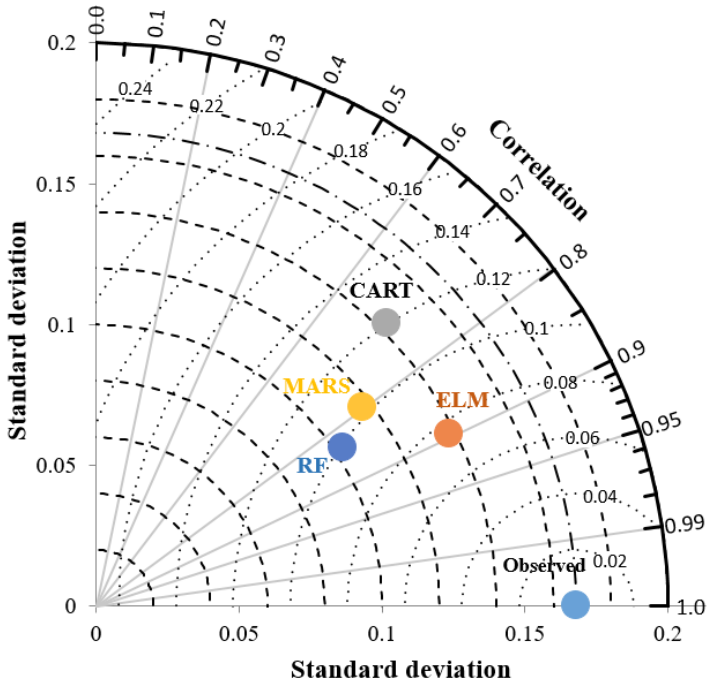



**Fig. 10.** Taylor diagram for the estimated SRC values in the last scenario.

## 4.2. Uncertainty analysis

Since the influential variables may have stochastic nature, uncertainty is inevitable in the applied models. Therefore, an uncertainty analysis is conducted to determine which model is more efficient during the testing period. For this purpose, the prediction error (PE), an average of the PE (APE), the SD of the PE (SDPE), the width of uncertainty band (WUB), and 95% PE interval (95% PEI) have been calculated in this study to quantify uncertainty. These parameters can be expressed as follows:

$$PE_i = (SRC)_{io} - (SRC)_{ip} \tag{17}$$

$$APE = \frac{1}{n}\sum_{I=1}^{n} PE_i \tag{18}$$

$$SDPE = \sqrt{\sum_{I=1}^{n} \frac{1}{n-1}(PE_i - APE)^2} \tag{19}$$

The results of the uncertainty analysis for the best model of each algorithm are provided in Table 7. It should be mentioned that the WUB and 95% PEI can be computed using $\pm 1.96\, SDPE$ and $PE \pm (APE + WUB)$, respectively. As can be seen from Table 7, the values of the APE (0.013) and SDPE (0.075) for the ELM are the lowest value compared with other models. In addition, the WUB for the ELM (± 0.148) is the lowest value compared with the RF (± 0.181), CART (± 0.190), and MARS (± 0.191) models. These results verify the efficiency of the ELM as an accurate data-driven technique to enhance estimation accuracy of the SHLFFNN.

**Table 7**
Uncertainty analysis results for the applied models for prediction of the SRC in the best scenario.

| Method | APE | SDPE | WUB | 95% PEI |
|---|---|---|---|---|
| ELM | 0.013 | 0.075 | ± 0.148 | (-0.037, 0.349) |
| RF | 0.014 | 0.092 | ± 0.181 | (-0.140, 0.360) |
| CART | 0.022 | 0.096 | ± 0.190 | (-0.091, 0.506) |
| MARS | 0.020 | 0.097 | ± 0.191 | (-0.028, 0.456 ) |

## 4.3. Sensitivity analysis

In this section, sensitivity analysis is performed using the RReliefF algorithm to identify the rank importance of different predictors affecting the SRC value. RReliefF algorithm uses intermediate weights to determine final weight of each predictor by penalizing predictors that provide different values for samples with the same outputs. The predictors that give different values to neighbors with different outputs are also rewarded. RReliefF algorithm works by analyzing the attribute ($A$) parameters and identifies related random ($R_i$) samples by finding the two nearest neighbors from two different classes (nearest hit H and nearest miss M). Based on the mentioned elements, the quality estimation ($W[A]$) is then calculated. A considerable difference between the two samples can result in lower quality estimation which is not acceptable. By considering the mentioned procedure for all samples, the quality estimation $W[A]$ can be computed as follows [41,42]:

$$W[A] = \frac{P_{diffC\backslash diffA}\, P_{diffA}}{P_{diffC}} - \frac{\left(1 - P_{diffC\backslash diffA}\right) P_{diffA}}{1 - P_{diffC}} \tag{20}$$

where $P_{diffA}$ denotes the difference of attribute A and the nearest instances, and $P_{diffC}$ indicates the difference of the estimated and nearest instances.

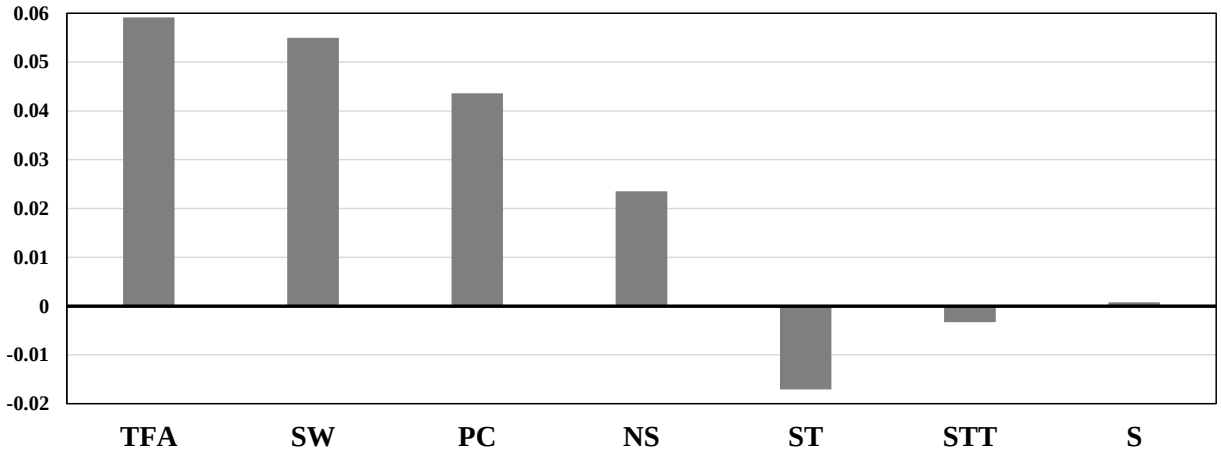



**Fig. 11.** The relative importance of predictors on the target variable using the RReliefF algorithm.

Results of the RReliefF algorithm-based sensitivity analysis, including the relative importance of predictors on the target SRC values, are presented in Fig. 11. Similar to the results of previous studies [2–6], it can be concluded from the weights presented in Fig. 11 that the TFA is the most effective parameter in SRC prediction. The SW parameter achived the secound rank of importance after TFA. An increase in the TFA and SW values considerably increases the SRC values for a given building. Therefore, several policies should be implemented to mitigating the risk by applying reliability analysis. The rank of input variables based on the RReliefF importance analysis is also presented in Table 8. It can be mentioned that the SW and the PC found also to be significant (positive effect) in process of SRC estimation. Moreover, the results of Table 8 demonstrate that the parameter S has the least influence in retrofit actions.

**Table 8**
Sensitivity analysis of influence of input parameters using the RReliefF algorithm and Gamma test.

| Input parameter | Rank |
|---|---|
| | RReliefF algorithm |
| TFA | 1 |
| NS | 4 |
| SW | 2 |
| S | 7 |
| ST | 5 |
| PC | 3 |
| STT | 6 |

## 5. Conclusion

One of the main goals of this research was to estimate the SRC using the relevant structural parameters and to identify the most influential variables in the retrofit cost. For this purpose, four different machine learning algorithms (MLAs), including the extreme learning machine (ELM), classification and regression tree (CART), multivariate adaptive regression spline (MARS), and

random forest (RF) regression, were employed to estimate the SRC values in the construction projects by considering the total floor area (TFA), number of stories (NS), seismic weight (SW), seismicity (S), soil type (ST), plan configuration (PC), and structural type (STT) as input variables affecting SRC values. The best prediction models for employed MLAs were determined by investigating several scenarios based on the different combinations of the input variables. The performances of the employed MLAs were compared in terms of four statistical indices including the correlation coefficient r, root mean square error (RMSE), Nash–Sutcliffe efficiency (NSE), Adjusted R-squared $\left(R_{adj}^2\right)$ indicator, and also the Taylor diagram. Data of the employed structural parameters were prepared and 80% of them was selected randomly and applied as a training subset using the hold-out method and the remaining 20% of the data was used as a testing subset.

Performance comparisons of different employed MLAs shown that the best ELM model with a combination set of all input variables provided the most accurate prediction for the SRC values among other compared algorithms. Furthermore, the RF regression with a combination set of the TFA, NS, SW, and PC as input variables achieved the second-rank best model. The results of the present study can be summarized as follows:

- This study suggested a reliable and efficient method based on the extreme learning machine to estimate SRC values using related structural parameters.
- The uncertainty analysis results for the best models of the applied MLAs indicated that the proposed ELM-based model for prediction of the SRC values has the lowest amount of uncertainty compared with other employed algorithms. Moreover, the RF regression was achieved the second rank in terms of the average, standard deviation, and width of the uncertainty band of the prediction error.
- A sensitivity analysis was also conducted using the RReliefF algorithm to investigate the importance of different effective variables on the SRC estimation. The RReliefF algorithm-based sensitivity results proved that the TFA, SW, and PC are the most influential input parameters, whereas the seismicity parameter (S) has the least influence in the retrofit actions. Moreover, both the soil type ST and structural type STT have shown a negative influence on the SRC value, while the effect of ST is also relatively higher than the effect of STT.

Results of this study verified that the developed ELM model with random weights outperforms traditional ANN and other compared algorithms in terms of error measures for estimating SRC. However, in general, designers should consider that there are different factors influencing the accuracy/efficiency of the model, including data quality and algorithm parameters. In addition, despite the high capability of MLAs, they might be subject to some inherent shortcomings due to different sources of uncertainties and external disturbances when applying in complex practical applications. The reliability-based analysis is a common approach that could be employed to deal with external disturbances and uncertainties in the input/output data. Finally, the boosting machine learning algorithms such as the least-square boosting could also be implemented to test their performance in order to enhance the accuracy of SRC prediction in future works.

## Funding

This research received no external funding.

## Conflicts of interest

The authors declare no conflict of interest.

## Authors contribution statement

N. Safaeian Hamzehkolaei, M. Alizamir: Conceptualization, Methodology, Data creation, Software, Writing–original draft. N. Safaeian Hamzehkolaei: Supervision, Validation, Writing – review & editing.

## References

[1] Nasrazadani H, Mahsuli M, Talebiyan H, Kashani H. Probabilistic Modeling Framework for Prediction of Seismic Retrofit Cost of Buildings. J Constr Eng Manag 2017;143:04017055. doi:10.1061/(asce)co.1943-7862.0001354.

[2] Chen WT, Huang Y. Approximately predicting the cost and duration of school reconstruction projects in Taiwan. Constr Manag Econ 2006;24:1231–9. doi:10.1080/01446190600953805.

[3] Jafarzadeh R, Ingham JM, Wilkinson S. A seismic retrofit cost database for buildings with a framed structure. Earthq Spectra 2014;30:625–37. doi:10.1193/080713EQS226.

[4] Jafarzadeh R, Ingham JM, Wilkinson S, González V, Aghakouchak AA. Application of Artificial Neural Network Methodology for Predicting Seismic Retrofit Construction Costs. J Constr Eng Manag 2014;140:04013044. doi:10.1061/(asce)co.1943-7862.0000725.

[5] Jafarzadeh R, Wilkinson S, González V, Ingham JM, Amiri GG. Predicting Seismic Retrofit Construction Cost for Buildings with Framed Structures Using Multilinear Regression Analysis. J Constr Eng Manag 2014;140:04013062. doi:10.1061/(ASCE)CO.1943-7862.0000750.

[6] Jafarzadeh R, Ingham JM, Walsh KQ, Hassani N, Ghodrati Amiri GR. Using Statistical Regression Analysis to Establish Construction Cost Models for Seismic Retrofit of Confined Masonry Buildings. J Constr Eng Manag 2015;141:04014098. doi:10.1061/(ASCE)CO.1943-7862.0000968.

[7] Alizamir M, Kisi O, Zounemat-Kermani M. Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data. Hydrol Sci J 2018;63:63–73. doi:10.1080/02626667.2017.1410891.

[8] Yaseen ZM, Deo RC, Hilal A, Abd AM, Bueno LC, Salcedo-Sanz S, et al. Predicting compressive strength of lightweight foamed concrete using extreme learning machine model. Adv Eng Softw 2018;115:112–25. doi:10.1016/j.advengsoft.2017.09.004.

[9] Al-Shamiri AK, Kim JH, Yuan TF, Yoon YS. Modeling the compressive strength of high-strength concrete: An extreme learning approach. Constr Build Mater 2019;208:204–19. doi:10.1016/j.conbuildmat.2019.02.165.

[10] Nayak, Sarat and Nayak, Sanjib and Panda S. Assessing Compressive Strength of Concrete with Extreme Learning Machine. J Soft Comput Civ Eng 2021;5:68–85. doi:10.22115/SCCE.2021.286525.1320.

[11] Zhang W, Zhang R, Wu C, Goh ATC, Wang L. Assessment of basal heave stability for braced excavations in anisotropic clay using extreme gradient boosting and random forest regression. Undergr Sp 2020:1–9. doi:10.1016/j.undsp.2020.03.001.

[12] Zhang W, Wu C, Li Y, Wang L, Samui P. Assessment of pile drivability using random forest regression and multivariate adaptive regression splines. Georisk 2021;15:27–40. doi:10.1080/17499518.2019.1674340.

[13] Zhang R, Wu C, Goh ATC, Böhlke T, Zhang W. Estimation of diaphragm wall deflections for deep braced excavation in anisotropic clays using ensemble learning. Geosci Front 2021;12:365–73. doi:10.1016/j.gsf.2020.03.003.

[14] Zhang Y, Wang Y, Zhou G, Jin J, Wang B, Wang X, et al. Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces. Expert Syst Appl 2018;96:302–10. doi:10.1016/j.eswa.2017.12.015.

[15] Yaseen ZM, Sulaiman SO, Deo RC, Chau K-W. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. J Hydrol 2019;569:387–408. doi:10.1016/j.jhydrol.2018.11.069.

[16] Sattar AMA, Ertuğrul ÖF, Gharabaghi B, McBean EA, Cao J. Extreme learning machine model for water network management. Neural Comput Appl 2019;31:157–69. doi:10.1007/s00521-017-2987-7.

[17] Wang X, Yang K, Kalivas JH. Comparison of extreme learning machine models for gasoline octane number forecasting by near-infrared spectra analysis. Optik (Stuttg) 2020;200:163325. doi:10.1016/j.ijleo.2019.163325.

[18] Zhang W, Zhang R, Wu C, Goh ATC, Lacasse S, Liu Z, et al. State-of-the-art review of soft computing applications in underground excavations. Geosci Front 2020;11:1095–106. doi:10.1016/j.gsf.2019.12.003.

[19] Zhang WG, Li HR, Wu CZ, Li YQ, Liu ZQ, Liu HL. Soft computing approach for prediction of surface settlement induced by earth pressure balance shield tunneling. Undergr Sp 2020. doi:10.1016/j.undsp.2019.12.003.

[20] Siahkali MZ, Ghaderi A, Bahrpeyma A, Rashki M. Estimating Pier Scour Depth : Comparison of Empirical Formulations. J AI Data Min 2021;9:109–28. doi:10.22044/jadm.2020.10085.2147.

[21] Singh R, Wagener T, Crane R, Mann ME, Ning L. A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA. Water Resour Res 2014;50:3409–27. doi:10.1002/2013WR014988.

[22] Arifuzzaman M, Gazder U, Alam MS, Sirin O, Mamun A Al. Modelling of Asphalt's Adhesive Behaviour Using Classification and Regression Tree (CART) Analysis. Comput Intell Neurosci 2019;2019. doi:10.1155/2019/3183050.

[23] Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. Sci Total Environ 2019;651:2087–96. doi:10.1016/j.scitotenv.2018.10.064.

[24] Salimi A, Faradonbeh RS, Monjezi M, Moormann C. TBM performance estimation using a classification and regression tree (CART) technique. Bull Eng Geol Environ 2018;77:429–40. doi:10.1007/s10064-016-0969-0.

[25] Pandey S, Kumar V, Kumar P. Application and Analysis of Machine Learning Algorithms for Design of Concrete Mix with Plasticizer and without Plasticizer. J Soft Comput Civ Eng 2021;5:19–37. doi:10.22115/scce.2021.248779.1257.

[26] Fung JF, Butry DT, Sattar S, McCabe SL. A methodology for estimating seismic retroft costs. Gaithersburg, MD: 2017. doi:10.6028/NIST.TN.1973.

[27] Fung JF, Butry DT, Sattar S, McCabe SL. Estimating structural seismic retrofit costs for federal buildings. Gaithersburg, MD: 2018. doi:10.6028/NIST.TN.1996.

[28] Fung JF, Sattar S, Butry DT, McCabe SL. A predictive modeling approach to estimating seismic retrofit costs. Earthq Spectra 2020;36:579–98. doi:10.1177/8755293019891716.

[29] Li L-L, Sun J, Tseng M-L, Li Z-G. Extreme learning machine optimized by whale optimization algorithm using insulated gate bipolar transistor module aging degree evaluation. Expert Syst Appl 2019;127:58–67. doi:10.1016/j.eswa.2019.03.002.

[30] Shariati M, Mafipour MS, Ghahremani B, Azarhomayun F, Ahmadi M, Trung NT, et al. A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement. Eng Comput 2020:1–23. doi:10.1007/s00366-020-01081-0.

[31] Dai B, Gu C, Zhao E, Zhu K, Cao W, Qin X. Improved online sequential extreme learning machine for identifying crack behavior in concrete dam. Adv Struct Eng 2019;22:402–12. doi:10.1177/1369433218788635.

[32] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Routledge; 2017. doi:10.1201/9781315139470.

[33] Friedman JH. Multivariate Adaptive Regression Splines. Ann Stat 2007;19:1–67. doi:10.1214/aos/1176347963.

[34] Ghanizadeh AR, Rahrovan M. Modeling of unconfined compressive strength of soil-RAP blend stabilized with Portland cement using multivariate adaptive regression spline. Front Struct Civ Eng 2019;13:787–99. doi:10.1007/s11709-019-0516-8.

[35] Shirzad A, Safari MJS. Pipe failure rate prediction in water distribution networks using multivariate adaptive regression splines and random forest techniques. Urban Water J 2019;16:653–61. doi:10.1080/1573062X.2020.1713384.

[36] Breiman L. Random forests. Mach Learn 2001;45:5–32. doi:10.1023/A:1010933404324.

[37] Benali L, Notton G, Fouilloy A, Voyant C, Dizene R. Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components. Renew Energy 2019;132:871–84. doi:10.1016/j.renene.2018.08.044.

[38] Baez-Villanueva OM, Zambrano-Bigiarini M, Beck HE, McNamara I, Ribbe L, Nauditt A, et al. RF-MEP: A novel Random Forest method for merging gridded precipitation products and ground-based measurements. Remote Sens Environ 2020;239:111606. doi:10.1016/j.rse.2019.111606.

[39] Tan K, Wang H, Chen L, Du Q, Du P, Pan C. Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. J Hazard Mater 2020;382:120987. doi:10.1016/j.jhazmat.2019.120987.

[40] Samadianfard S, Ghorbani MA, Mohammadi B. Forecasting soil temperature at multiple-depth with a hybrid artificial neural network model coupled-hybrid firefly optimizer algorithm. Inf Process Agric 2018;5:465–76. doi:10.1016/j.inpa.2018.06.005.

[41] Behbahani H, Amiri AM, Imaninasab R, Alizamir M. Forecasting accident frequency of an urban road network: A comparison of four artificial neural network techniques. J Forecast 2018;37:767–80. doi:10.1002/for.2542.

[42] Robnik-Šikonja M, Kononenko I. Theoretical and Empirical Analysis of ReliefF and RReliefF. Mach Learn 2003;53:23–69.