Front. Struct. Civ. Eng.
https://doi.org/10.1007/s11709-021-0689-9

RESEARCH ARTICLE

Unconfined compressive strength prediction of soils stabilized using artificial neural networks and support vector machines

Alireza TABARSA[a*], Nima LATIFI[b], Abdolreza OSOULI[c], Younes BAGHERI[d]

[a] *Depatrment of Civil Engineering, Faculty of Engineering, Golestan University, Gorgan 49138-15759, Iran*
[b] *Terracon Consultants, Inc., Nashville, TN 37211, USA*
[c] *Civil Engineering Department, Southern Illinois University, Edwardsville, IL 62026-1800, USA*
[d] *Faculty of Engineering, Mirdamad Institute of Higher Education, Gorgan 49166-53989, Iran*
[*] *Corresponding author. E-mail: a.tabarsa@gu.ac.ir*

ABSTRACT This study aims to improve the unconfined compressive strength of soils using additives as well as by predicting the strength behavior of stabilized soils using two artificial-intelligence-based models. The soils used in this study are stabilized using various combinations of cement, lime, and rice husk ash. To predict the results of unconfined compressive strength tests conducted on soils, a comprehensive laboratory dataset comprising 137 soil specimens treated with different combinations of cement, lime, and rice husk ash is used. Two artificial-intelligence-based models including artificial neural networks and support vector machines are used comparatively to predict the strength characteristics of soils treated with cement, lime, and rice husk ash under different conditions. The suggested models predicted the unconfined compressive strength of soils accurately and can be introduced as reliable predictive models in geotechnical engineering. This study demonstrates the better performance of support vector machines in predicting the strength of the investigated soils compared with artificial neural networks. The type of kernel function used in support vector machine models contributed positively to the performance of the proposed models. Moreover, based on sensitivity analysis results, it is discovered that cement and lime contents impose more prominent effects on the unconfined compressive strength values of the investigated soils compared with the other parameters.

KEYWORDS unconfined compressive strength, artificial neural network, support vector machine, predictive models, regression

1 Introduction

The construction of soft soils requires soil improvement. One reliable method for ground modification is soil stabilization using cementitious materials such as cement and lime. In the presence of cement and lime, pozzolanic reactions can improve the mechanical properties of soils [1]. It is noteworthy that cement and lime are energy-intensive products that can increase the material cost for ground modification. Furthermore, the use of agricultural waste as a cementation material has been proposed to address environmental and economic concerns [2]. A vast proportion of agricultural lands constitutes paddy farms.

In terms of weight, 72% of the product is rice. Additionally, 20%–22% and 5%–8% of the cultivation are husk and bran, respectively. Therefore, a large volume of rice husks is available [3]. If rice husk is burned at an appropriate temperature, the achieved rice husk ash (RHA) will have a significant silica content. As a type of pozzolan, RHA can be used together with lime and cement mixtures to improve soil behavior. Pozzolanic stabilizers bind soil particles to each other, resulting in the reduced water absorption capacity of clay particles [4].

Generally, extensive and time-consuming experimental studies are required to evaluate the effect of different cementitious materials on soil behavior. Therefore, intelli-

Article history: Received Dec 10, 2019; Accepted Mar 18, 2020

gence-based techniques such as artificial neural networks (ANNs) or support vector machines (SVMs) have been utilized to determine soil stiffness and strength [5–7]. Other methods, including multivariate adaptive regression spline models, have also been used to estimate maximum dry density and unconfined compressive strength (UCS) values [8]. Regarding stiffness, He and Li [9] used an ANN to demonstrate that the shear modulus can be estimated reasonably using shear strain, confining pressure, curing time, and fiber and lime contents. With regard to strength, Gunaydin et al. [10] reported that the ANN performed better in estimating the UCS of soils compared with traditional regression or statistical models. They discovered that the flexibility and adaptability of the ANN model in generalizing data is its major advantage over multivariable regression models in predicting UCS values. Furthermore, ANNs have been used to predict the UCS values of a mixture comprising cement, clay, and peat [11,12]; soil mixed with air-foam and fishing net waste [13], bottom ash, jute, and steel fibers; as well as geopolymer stabilized soils [14,15]; micro silica-lime stabilized sand [16]; and liquefiable soil [17]. In general, the multilayer ANN architecture has demonstrated better performance in predicting the UCS values of two types of fat clay and one type of lean clay stabilized with various proportions of soil/water/cement in several curing times compared with the radial basis function or genetic algorithms. Furthermore, it has been demonstrated that strength predicted by models using the ANN is more representative than the stiffness predicted by such models [12]. As mentioned earlier, the SVM, which is dependent on a predetermined kernel function [18], has been used successfully to predict the UCS values of soils [19,20], jet grouting columns and volcanic rocks [21,22], slope reliability [23], bearing capacity of piles [24], soil permeability [25], and liquefaction [26–28]. A more detailed explanation will be provided in the following sections. It is noteworthy that these methods have not been used comparatively for treated soils. This study was conducted to predict the UCS values of silty sand and high plasticity silt treated with 5%–15% cement, lime, and rice husk ash (CLR) mixtures. To determine the UCS improvement rate, the curing time was varied from 7 to 60 d. Moreover, a set of treated samples without RHA was tested to identify its effect on soil strength. Two types of ANN architectures and an SVM method using two kernel functions were adopted to predict and compare the UCS values of the treated samples.

2 Materials and methods

2.1 Materials

The database used in this investigation was obtained from previous experimental studies [4], where two types of locally available soils were used. One type of soil was collected from Matang Kerat Telunjuk, Bandar Baharu in Kedah, and the other was collected from Nibong Tebal, USM Engineering campus, both in Malaysia. The gradation curves of the investigated soils are plotted in Fig. 1. The geotechnical properties of the soil samples determined based on the ASTM standard tests are summarized in Table 1.

The cement and lime used in this study were ordinary Portland cement and hydrated high-calcium lime ($Ca(OH)_2$), respectively. Moreover, the ash used in this study was procured by burning rice husks in a gas furnace. RHA demonstrates more prominent pozzolanic properties if incinerated under controlled conditions. The RHA used

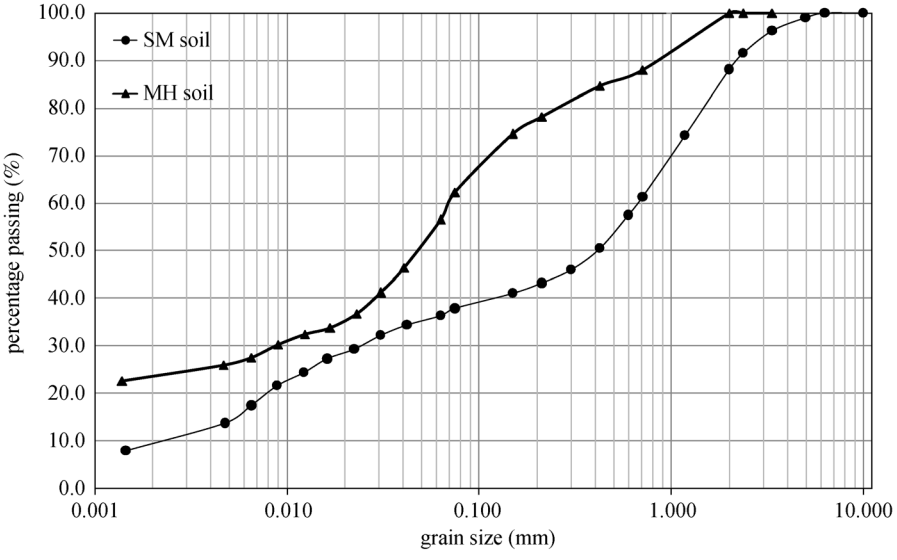


Fig. 1 Grain size distribution for used soils.

Table 1 Summary of basic properties of investigated soils

soil properties	quantity (SM soil)	quantity (MH soil)	test method
sand (%)	51.70	43.52	
silt (%)	28.60	33.87	
clay (%)	7.80	22.61	
organic content (%)	1.2	12.8	
specific gravity	2.55	2.38	(ASTM D854)
USCS classification	silty sand	high plasticity silt	(ASTM D2487)
liquid limit (%)	48.10	55.40	(ASTM D4318)
plastic limit (%)	31.50	36.40	(ASTM D4318)
plasticity index (%)	16.60	19.00	(ASTM D4318)
optimum moisture (%)	16.30	24.65	(ASTM D698)
maximum dry unit weight (kN/m^3)	17.52	14.67	(ASTM D698)
pH	4.78	4.21	(ASTM D4972)

in this study contained as much as 90% silica, indicating the appropriate pozzolanic properties of this type of additive. In this study, the effects of various proportions of CLR on the strength of two types of soils under different conditions, including the dry unit weights and cutting times, were investigated.

To evaluate the mechanical behavior of CLR-treated soil (Table 2), UCS tests were conducted according to ASTM D2166.

Table 2 Experimental procedure for UCS tests

variable	UCS test for soil
C (%)	1.25, 1.875, 2.5, 3.125
L (%)	2.5, 3.75, 5, 6.25
R (%)	1.25, 1.875, 2.5, 3.125
curing time (d)	7, 28, 60
dry unit weight, γ_d(kN/m^3)	14, 15, 16, 17
soil type	SM, MH

3 Predictive model development

3.1 Database

To determine the strength characteristics of CLR-treated soil specimens under different conditions, soft computing techniques such as ANNs and SVMs were adopted. The data used to estimate the UCS values of CLR-treated soils were obtained by conducting UCS tests on 137 specimens (Appendix A).

3.2 ANNs

To mathematically simulate the brain learning method using mesh interconnection among cells, ANNs were used. ANNs are mathematical tools that attempt to imitate the neural network of the human brain as well as the nervous system. These tools can define the interactive effects of different variables in a complicated process. Their design is based on the neural structure of the brain, in which living neurons and dendrites are represented by silicon and wires. The brain learns from experience [29] and stores information in a systematic pattern. The process of storing information as patterns and using them in solving problems are imitated in the field of computational mechanics. In a neural network, a neuron can be considered as the fundamental processing element. A biological neuron receives inputs from other sources. Generally, after combining these inputs, a nonlinear operation is conducted on the result, and the final output is provided. Similar to biological neurons of the human brain, ANNs comprise a large number of nodes. Neurons are connected by links and interact with each other. The nodes representing neurons receive data as input and perform simple operations on them. Subsequently, they pass the outcome to the other nodes. The output determined at each node is known as its activation or node value. Feedforward and feedback are two types of ANN topologies. The efficiency of an ANN in predicting new events based on previous history is affected by their architecture, learning algorithm, and robustness. Furthermore, ANNs have been used for function approximation, regression analysis, classification, data processing, robotics, and control [29]. In typical deep learning applications, neural networks are developed based on a number of inputs and outputs to minimize the difference between predicted and real values. This process, which is typically performed manually, requires a large number of datasets and is often susceptible to different types of errors [30]. Hence, a deep neural network was introduced for the bending analysis of Kirchoff plates exhibiting various shapes and subjected to different loads and boundary conditions. This technique is regarded as “mesh free” and can be employed to estimate continuous functions. There-

fore, it is considered an appropriate method to analyze thin-plate bending problems [31].

In this study, ANN models based on feedforward multilayer perceptron neural networks [12] were adopted. Among the abovementioned experimental data, 60% was randomly employed for training, 20% was randomly employed for validation, and the remainder was used to test the proposed models. To perform the training, validating, and testing process, a computer program was developed in MATLAB. Generally, two methods can be used to avoid overfitting a model. This problem can be avoided by training the model using more data or by altering the complexity of the network. In this study, a sufficient input sample was provided to evade the overfitting problem. Subsequently, the appropriate network structure was selected to avoid pruning. As a typical approach in artificial intelligence, the "early stopping" method is used to address the overfitting problem. In this regard, the number of epochs increases when the network error for the test data (validation set) decreases. By increasing the number of epochs beyond a certain level, the neural network error will no longer decrease, demonstrating the overfitting problem. In this study, this problem was avoided by investigating different numbers of epochs and their corresponding neural network errors. Moreover, it has been established that the underfitting problem will not occur owing to the number of used data and the proposed architecture of the neural network.

In these proposed models, soil type, curing time, dry unit weight (γ_d), cement, RHA, and lime contents were considered as input parameters, whereas the UCS value was used as the output. To achieve the best predictive model, an appropriate ANN architecture must be selected. Two ANN models, ANN1 and ANN2, were selected to predict the UCS values of CLR-treated soil specimens. The architectures of ANN1 and ANN2 are illustrated in Figs. 2 and 3, respectively. As shown, the ANN1 model was constructed using seven neurons in a single hidden layer, whereas the ANN2 model comprised two hidden layers with four neurons in each layer. Table 3 shows the characteristics of the ANN models used in this investigation. In these figures, W and b represent the weight and bias, respectively.

The performance of the developed ANN models was evaluated based on the average of absolute percentage error (*AAPE*) and correlation coefficient (*R*). These indices are provided in Eqs. (1) and (2), respectively.

$$AAPE = \frac{\sum_{i=1}^{n} \frac{|(x_t)_i - (x_p)_i|}{(x_t)_i}}{n} \times 100 \ , \qquad (1)$$

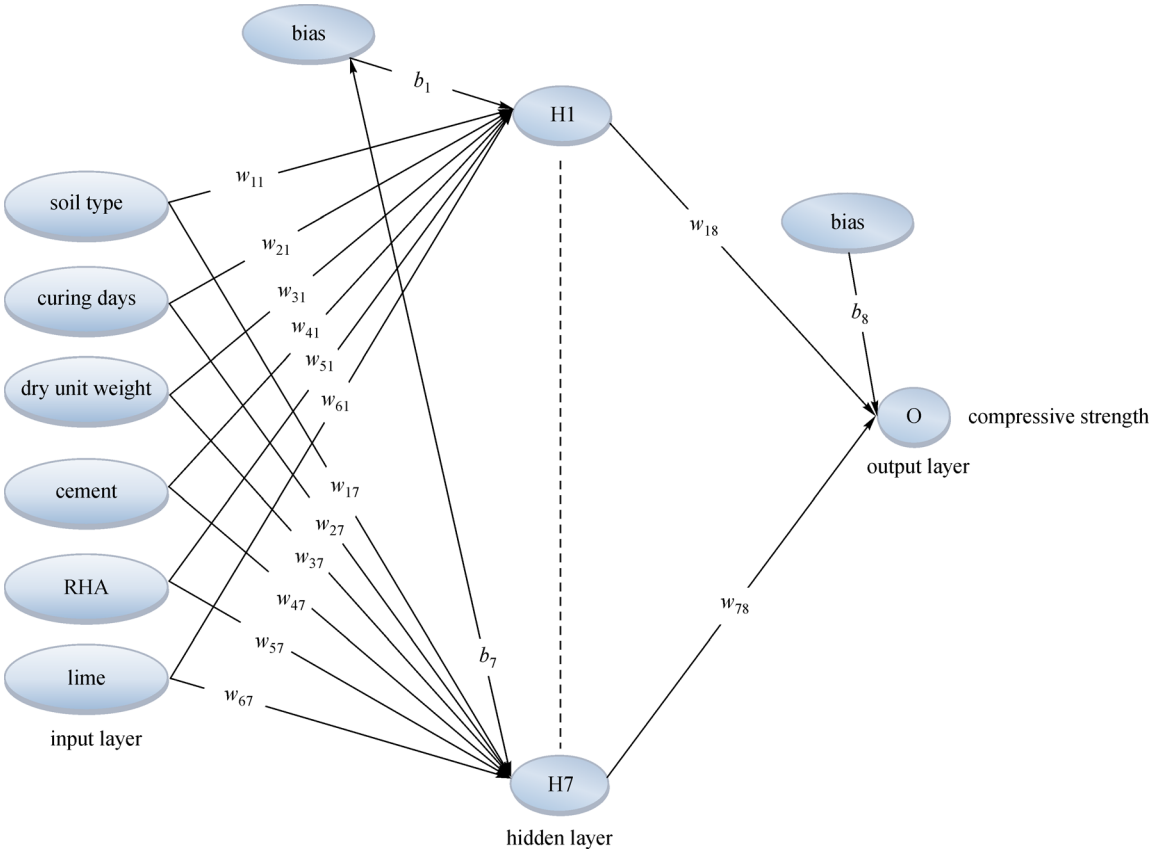


Fig. 2 Developed ANN1 model for UCS estimation.

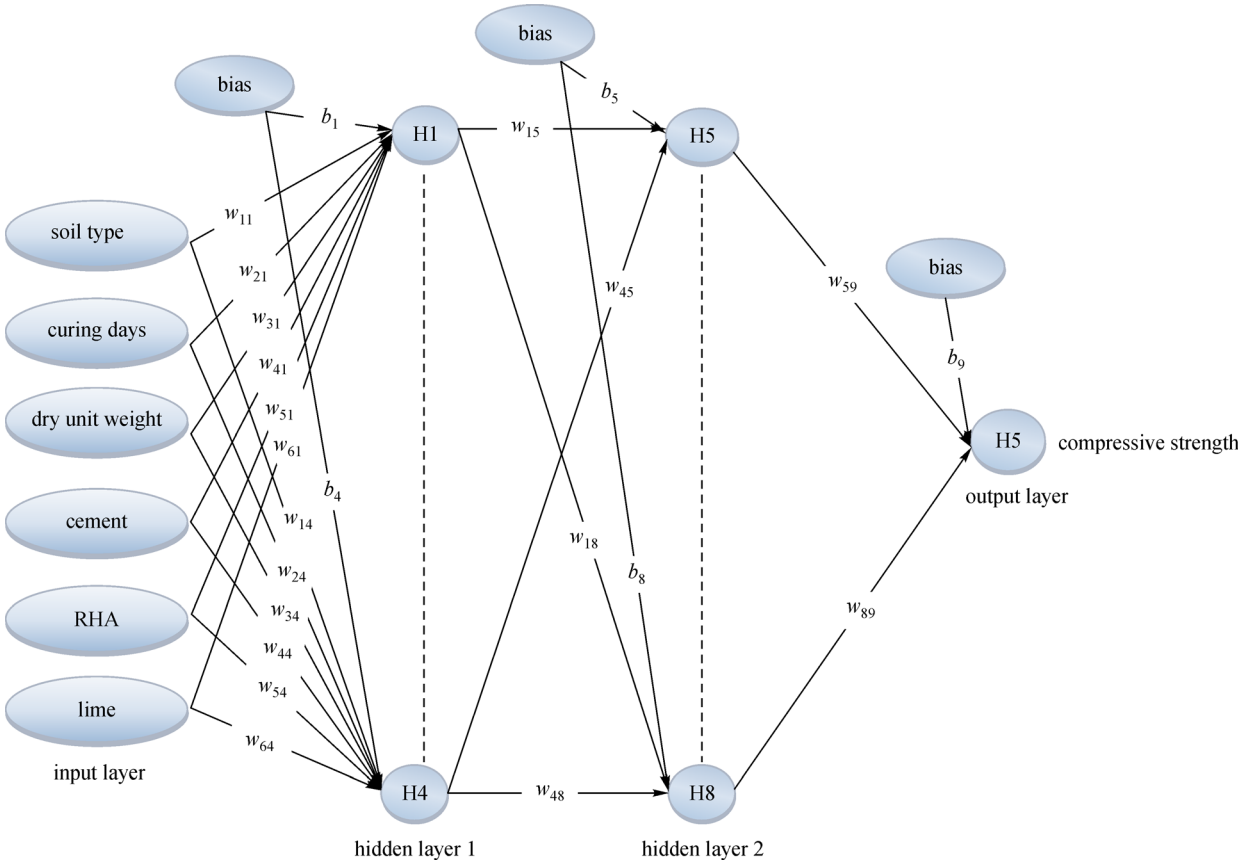


Fig. 3 Developed ANN2 model for UCS estimation.

Table 3 ANN characteristics

characteristic	description/value
number of hidden layers	one and two layers
number of optimum neurons in hidden layer	7 and 4
training algorithm	Levenberg–Marquardt (back-propagation)
activation function in hidden layer	sigmoid
activation function in output layer	linear

$$R = \frac{\sum_{i=1}^{n}[(x_t)_i - (\overline{x}_t)_i][(x_p)_i - (\overline{x}_p)_i]}{\sqrt{\sum_{i=1}^{n}[(x_t)_i - (\overline{x}_t)_i]^2[(x_p)_i - (\overline{x}_p)_i]^2}}, \quad (2)$$

where x_t and x_p are the target and ANN predicted results, respectively; $\overline{x}_t$ and $\overline{x}_p$ are the means of the target and predicted values, respectively; n is the number of data points.

To evaluate the generalizing capacity of the ANN, information criteria (such as the Akaike information criterion (AIC) and Bayesian information criterion) are used to determine the "insample" fit (i.e., fit to the training data) and penalize the complexity of the models. Nonetheless, the best criterion to determine the most appropriate ANN model is yet to be elucidated. According to the literature, the weighted information criterion is used more frequently than the root mean square error. Furthermore, the AIC is the most popular method used in linear and nonlinear models. In the proposed ANN models, the AIC was introduced as an evaluation index. The popular form of the AIC is as follows:

$$AIC = \log\left(\frac{\sum_{i=1}^{T}[(x_t)_i - (x_p)_i]^2}{T}\right) + \frac{2m}{T}, \quad (3)$$

where T and m indicate the amount of data and the weight, respectively [32].

3.3 Support vector machine

SVM analysis, as a new type of learning algorithm, was first introduced by Vapnik et al. [33] in 1992 and then continuously used by other researchers. The SVM is a widely used machine learning tool for classification and

regression analyses. It operates based on the structural risk minimization (SRM) principle. According to this principle, the complexity of the model is balanced with its success in fitting the training data [18,33]. In this method, the SRM principle is used to minimize errors, while empirical risk minimization is used in other methods such as AANs. The main objective of SRM is to minimize the empirical risk and maximize the generalizing ability of the model simultaneously [34]. The SVM can be easily analyzed mathematically because it corresponds to a linear method in a highly nonlinear input space. An SVM is a special class of algorithms. The use of kernel functions, absence of local minima, sparseness of the solution, and capacity control achieved from acting on the margin or on the number of support vectors are among its major specifications. The capacity of the system is dominated by parameters that are independent of the featured-space dimensions [35]. During classification, the data are segmented into two sets, i.e., training and testing sets. The training set comprises one target and several attributes. The SVM is primarily aimed at providing a model using the training data. Subsequently, it predicts the target value of the test data based on their attributes [36]. The support vector regression (SVR) technique is typically used in linear regression problems. Nonetheless, this method can be used in nonlinear regression problems if the training patterns are first mapped into a higher-dimensional feature space where linear regression is applicable.

SVR and the SVM use similar principles for classification. However, Vapnik et al. [33] introduced a new type of loss function. Using this function, known as the ε-insensitive loss function, the concept of margin is incorporated into SVM regression techniques. According to the basic assumption of the ε-insensitive loss function, if the error is within the range of ε, the model will not contain any prediction error. In other words, the error is acceptable if it is within the range of ε. The width of the ε-insensitive zone is determined by the parameter ε. This parameter is used to fit the training data, and its value is specified based on the number of support vectors used to establish the regression function [34]. To achieve larger values of ε, fewer support vectors are selected. In other words, larger values of ε will result in more flat approximations.

The main objective is to obtain a function $f(x)$ with a deviation of ε from the actual target, i.e., y, using the training data. This function must be as flat as possible. A loss function with an ε-insensitive zone is expressed as shown in Eq. (4).

$$L_\varepsilon(y) = |y - f(x)|_\varepsilon = \begin{cases} 0, & |y - f(x)| \leqslant \varepsilon, \\ |y - f(x)| - \varepsilon, & \text{othewise}, \end{cases} \tag{4}$$

where $L_\varepsilon(y)$ is the loss function.

The function must have the largest ε from the target achieved for all training data. Moreover, it must be as flat as possible. Therefore, the function

$$\phi(w,\xi^*,\xi) = \left(\frac{\|w\|^2}{2}\right) + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), \tag{5}$$

is minimized, subject to

$$\begin{cases} y_i - [(wx_i) + b] \leqslant \varepsilon + \xi_i^*, \\ [(wx_i) + b] - y_i \leqslant \varepsilon + \xi_i^*, \ i = 1,2,3,...,n. \\ \xi_i, \xi_i^* \geqslant 0. \end{cases} \tag{6}$$

In this function, $\|w\|^2$ is the Euclidian norm of the weight vector, and the constant $C \succ 0$ is the penalty parameter. In addition, ξ and ξ^* are slack variables. In the regularized form of the risk function (Eq. (5)), the first term $\frac{\|w\|^2}{2}$ is the structure risk and is used to control the smoothness or complexity of the function. The second term $C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$ indicates the empirical risk. Therefore, both terms must be minimized [34,37].

The constant ε used in the loss function and penalty parameter "C" is determined by the user. The parameter "C" controls the tradeoff between the complexity of the model and a specific tolerance that is acceptable and can be considered as a regulatory factor [37]. The use of kernel functions causes SVR to be regarded as a nonparametric method. Meanwhile, computational problems arising from the high dimensionality of the feature space can be avoided using such functions [35]. The kernel function is regarded as the dot product of the training patterns, and it is less expensive and much faster for solving computational problems. Hence, SVR can be performed in a higher-dimensional feature space without overcomputing the feature vectors in that particular space. Furthermore, using SVR in convex optimization problems results in a unique and global solution. It is noteworthy that a kernel function is required for the implicit mapping of training patterns into a feature space in nonlinear regression problems. Thus far, different types of kernel functions, including linear, polynomial, radial basis, and sigmoid ones, have been presented [37]. However, polynomial and radial basis functions demonstrated higher efficiency in geotechnical engineering [38,39].

In this investigation, SVM models using polynomial functions (SVM-poly) and Gaussian radial basis functions (SVM-RBF) were implemented in MATLAB. These two functions are expressed as shown in Eqs. (7) and (8), respectively.

$$k\langle x_k \cdot x\rangle = (\langle x_k \cdot x\rangle + 1)^d, \tag{7}$$

$$k\langle x_k \cdot x\rangle = \exp\left(\frac{-\|x_k - x\|^2}{2\sigma^2}\right). \tag{8}$$

In these equations, x is the input vector; d and σ indicate the degree of polynomial and width of RBF functions, respectively, where both are defined by the user.

Details regarding the SVM technique are available in [33,35]. To evaluate the efficacy of the model, two statistical parameters, including the *R*, *AAPE*, and *AIC* were used.

In this study, to achieve the optimum SVM models, the parameter *C* was determined using the training and test (validation) datasets. The value of ε is effective for the number of support vectors used to construct the regression function. In other words, less support vectors are selected to achieve larger values of ε. Moreover, the larger values of ε will result in more flat approximations. Therefore, both *C* and ε will be effective for the complexity of the model (but in different respects). Different combinations of kernel parameters, including *d*, σ, penalty parameter *C*, and loss function factor ε were adopted on a trial-and-error basis. Subsequently, the performance of the SVM models was determined for each combination of the abovementioned parameters. The best performance of the investigated models using different kernel functions corresponded to the optimal values of the kernel and loss function parameters, which are summarized in Table 4.

4 Results and discussion

The performances of the ANN1 and ANN2 models were compared, the results of which are shown in Table 5. In the ANN1 model, the values of *AAPE* for the training, validating, and testing datasets were determined to be 3.979, 11.598, and 12.349, respectively, whereas the corresponding values obtained by the ANN2 model were 4.394, 8.247, and 10.325, respectively. In the ANN1 model, the values of *R* corresponding to the training, validating, and testing datasets were 0.9998, 0.9976, and 0.9957, respectively. Meanwhile, in the ANN2 model, the values of *R* corresponding to the training, validating, and testing datasets were calculated to be 0.9997, 0.9952, and 0. 9979, respectively.

The AIC values for the testing datasets were determined to be 4.289 and 3.979 for the ANN1 and ANN2 models, respectively. According to the results, the values of AIC are typically lower than those of *AAPE* as an evaluation index in soft computing techniques. These results indicate that the ANN1 and ANN2 models can predict the UCS values both similarly and reasonably.

Figures 4 and 5 show scatter plots corresponding to the ANN1 and ANN2 models for the training and testing datasets, respectively. As shown, the points were distributed well around the line of equality, indicating that the UCS values estimated by the ANN1 and ANN2 models did not differ significantly from the experimental results.

The statistical performance of SVM models using different kernel functions was evaluated based on the values of *R* and *AAPE* (Table 6). The results indicate that the SVM-RBF model performed slightly better than the SVM-poly model. However, the difference between the *R* values of the training and testing datasets was not significant, demonstrating the high generalizing capacity of the developed models. Moreover, the *AAPE* value of the SVM-RBF, calculated for the testing datasets, was approximately 13% lower than that of the SVM-poly. Therefore, it can be concluded that the correlation coefficient and prediction error of the SVM model depended on the type of kernel function used in the model.

Figures 6 and 7 show the UCS values estimated by the SVM model using two different kernel functions as well as laboratory results used as training and testing datasets. As shown, both functions exhibited a reasonable level of precision in predicting the UCS values. However, the values predicted by the SVM-RBF model were much more consistent compared with those obtained by the SVM-poly model.

Using a large number of support vectors might signify the overfitting (low level of generalizing capacity) of the testing data [18]. In this study, the numbers of support vectors were determined to be 58 and 67 for the radial basis and polynomial kernel functions, respectively. This indicates the relative superiority of radial basis functions. Furthermore, it shows a tradeoff between the complexity of the model and its generalizing capacity. Figure 8 shows the scatter plots corresponding to the testing datasets used by the AANs and SVM. As shown, the developed models

Table 4 Optimal values of kernel, penalty, and loss functions used in SVM models

SVM models	kernel parameters	optimum value of *C*	optimum value of ε	number of support vectors
SVM-poly	$d = 3$	1	0.0100	67
SVM-RBF	$\sigma = 4$	30	0.0210	58

Table 5 Statistical performance of ANN models

dataset name	no. of data	ANN1			ANN2		
		R	*AAPE* (%)	*AIC*	*R*	*AAPE* (%)	*AIC*
training	83	0.9998	3.979	3.436	0.9997	4.394	3.511
validating	27	0.9976	11.598	4.103	0.9952	8.247	4.381
testing	27	0.9957	12.349	4.289	0.9979	10.325	3.979

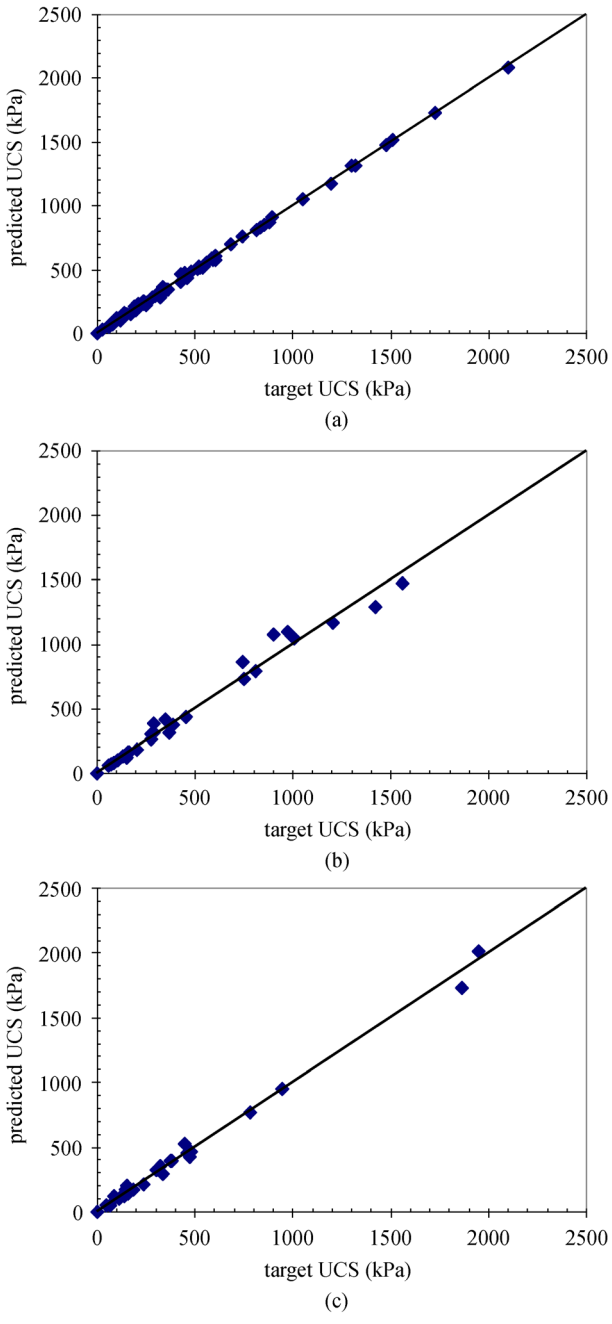


Fig. 4 Correlation of measured and predicted UCS values by ANN1: (a) training; (b) validating; (c) testing datasets.

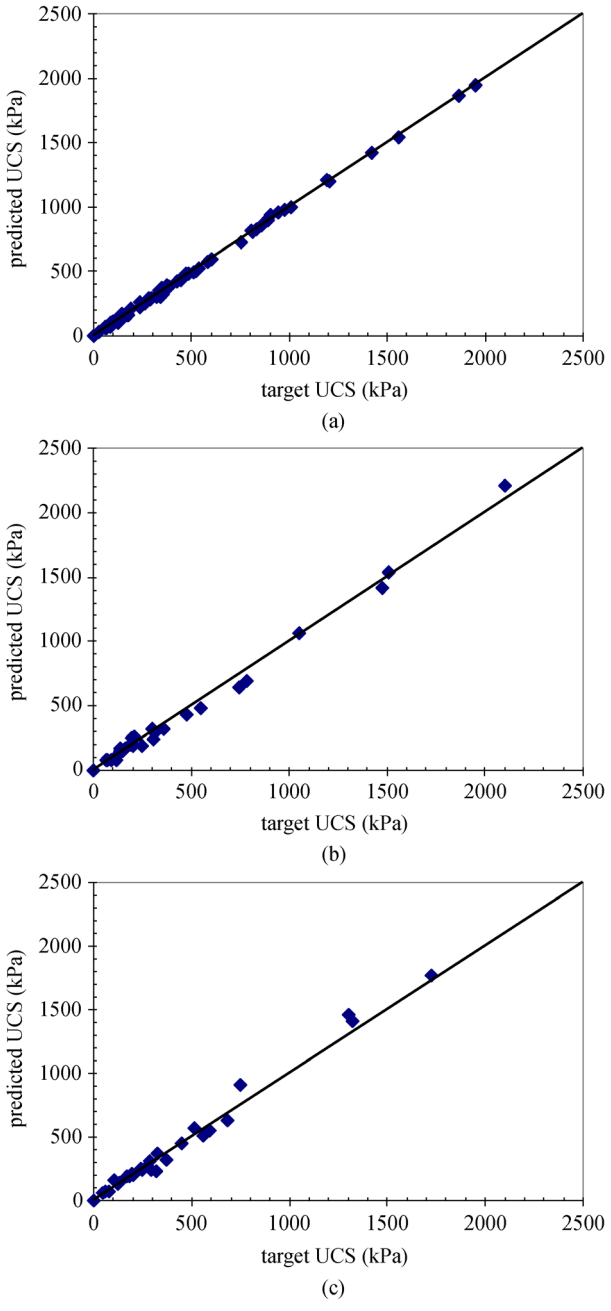
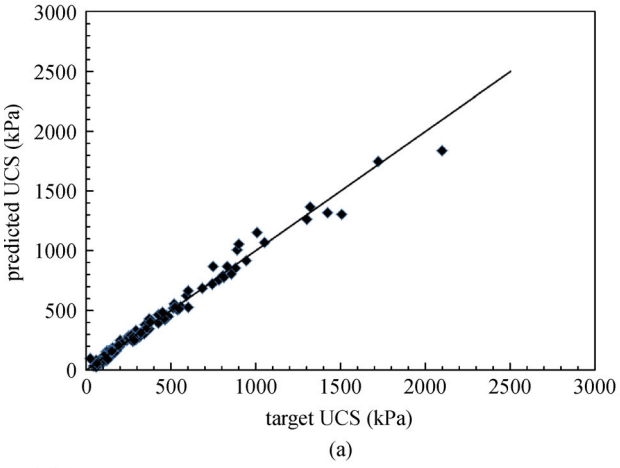


Fig. 5 Correlation of measured and predicted UCS values by ANN2: (a) training; (b) validating; (c) testing datasets.

Table 6 Statistical performances of SVM models

data set name	no. of data	SVM- Poly		SVM-RBF	
		R	*AAPE* (%)	*R*	*AAPE* (%)
training	110	0.9913	4.718	0.9965	3.880
testing	27	0.9918	5.286	0.9970	4.606

demonstrated reasonable accuracy in generalizing the testing datasets. Similar to models already developed, the use of these models is often limited and case dependent. The *AAPE* values of the testing datasets used by the SVM-RBF, SVM-poly, ANN2, and ANN1 models were determined to be 4.606, 5.286, 10.325, and 12.349,



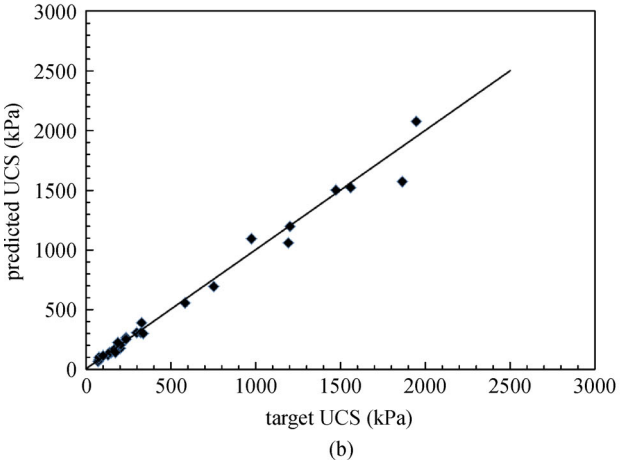
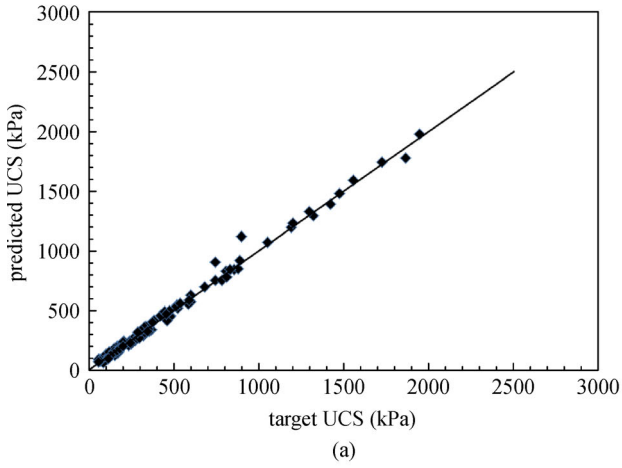


Fig. 6 Correlation between laboratory UCS values and those predicted by SVM-poly model: (a) training and (b) testing datasets.



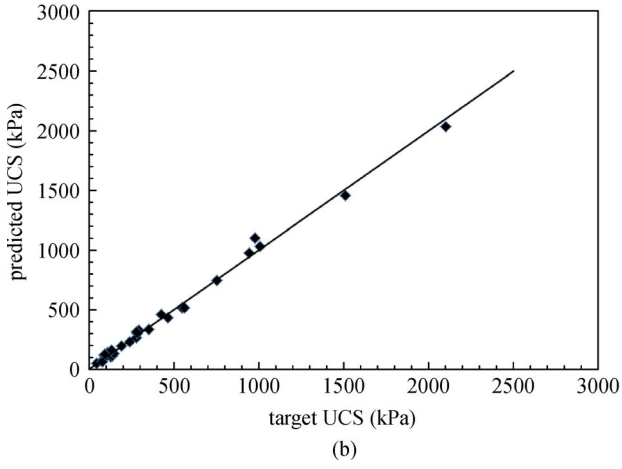


Fig. 7 Correlation between laboratory UCS values and those predicted by SVM- RBF: (a) training and (b) testing datasets.

respectively. By comparing the values of R and $AAPE$ obtained by the ANN and SVM models, it can be concluded that the SVM models performed better than the ANN models.

In this investigation, a whisker plot box was plotted to evaluate the error distribution of the predicted values by the ANN and SVM models corresponding to the testing datasets. Box plots, as exploratory data analysis tools, are generally used to provide the statistical summaries of the underlying prediction error distribution [40]. They are used to display the overall pattern of response in a group and can be considered as a suitable method to display characteristics such as the response range. Furthermore, the upper and lower boundaries of the whisker box represent the 0.25 and 0.75 absolute percentage errors, respectively. Moreover, the whiskers are lines extending from each end of the box to 1.5 times the interquartile range and also extending from the top and bottom of each box. In addition, the error distribution is determined by the length of the interquartile range, and the error median is displayed by a thick solid line in the box [18]. The cross symbols shown in the plot represent the values beyond the ends of the whiskers, i.e., the outlier data.

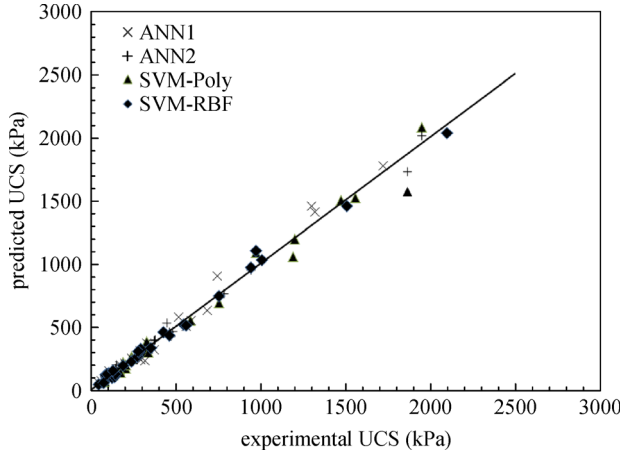


Fig. 8 ANN and SVM predicted UCS against experimental UCS for testing dataset.

As shown in Fig. 9, the smallest length of the interquartile range was related to the ANN model,

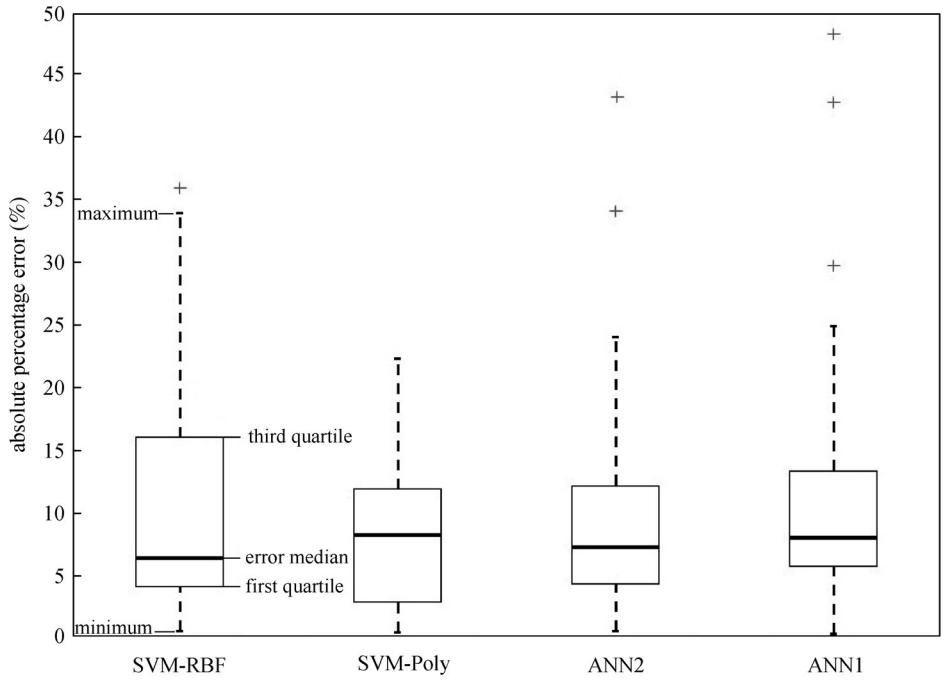


Fig. 9 Prediction error variability of ANN and SVM models corresponding to testing datasets.

suggesting less variability in the error distribution compared with other predictive models. Furthermore, it was discovered that the SVM-RBF exhibited more variabilities in its error distribution, while exhibiting the lowest error median. The variations in the interquartile range among the ANN1, ANN2, and SVM-poly models were similar. In the SVM-poly model, the error median was located almost at the middle of the interquartile range, and no outlier data were observed. It is noteworthy that the multilayer perceptron neural network contains more outlier data than the SVM. Because the error median is less than the mean obtained by the ANN1, ANN2, and SVM-RBF models, the error distribution magnitude was reduced. In addition, the error mean obtained by SVM-poly was approximately equal to the error median, demonstrating that the error distribution was symmetric and the distribution normal. In all models, the two ends of whiskers (minimum and maximum) were unequal to the median. In general, it can be concluded that the SVM-poly performed better, with a symmetric error distribution, smaller error, and lower error variability compared with the other models. The minor difference observed in the performance of the kernel functions might be attributed to their flexibility in mapping data as well as the formation of optimal regression functions that depended on the distribution of the data in the featured space [18].

5 Regression model and sensitivity analysis

Multiple regression analysis is extensively used to develop models for geotechnical characteristics, such as the strength parameters of soils. In this study, the efficiencies of two artificial intelligence–based models, namely, the ANN and SVM, were compared with that of regression analysis. Therefore, a linear regression model was assumed to predict the UCS values using 137 experimental datasets (Eq. (9)).

$$UCS = k_0 + k_1 D + k_2 T + k_3 C + k_4 L + k_5 R. \quad (9)$$

In the equation above, the dry unit weight (kN/m^3), curing time (d), as well as cement (%), lime (%), and rice husk ash (%) contents are independent variables denoted by D, T, C, L, and R, respectively. In addition, k_0 to k_5 are regression coefficients. Based on observations, the regression models used to estimate the UCS values (kPa) of the investigated soils were obtained using Eqs. (10)–(11).

For SM soil:

$$UCS = -4601.940 + 312.884D + 2.337T - 19.667\text{e}16C + 9.833\text{e}16L + 20.944R, \quad (10)$$

For MH soil:

$$UCS = -856.391 + 70.594D + 2.299T - 1.618\text{e}16C + 0.809\text{e}16L - 3.74R. \quad (11)$$

Based on the results, the R values obtained by the

regression models were 0.9080 and 0.8933 for SM and MH soils, respectively, which were lower than the values determined using the ANN and SVM. Therefore, the two artificial intelligence–based models are superior to regression analysis in predicting UCS values. Figures 10 and 11 show the relationship between the experimental data and the corresponding UCS values estimated using the regression model for the soils investigated.

Sensitivity analysis can be considered as a valuable method for quantitatively determining the effect of any change in input parameters on the output data. Following the determination of regression models for predicting the UCS values of soils treated with CLR, a sensitivity analysis was performed to estimate the effect of input variables on the estimated UCS value as the target parameter.

To determine the effective degree of the investigated parameters corresponding to SM and MH soil, Eqs. (12) and (13) were used. Using these equations, the sensitivity of the input variables was evaluated.

$$N_i = f_{xi}^{\max} - f_{xi}^{\min}, \tag{12}$$

$$S_i = \frac{N_i}{\sum_{i=1}^{n} N_i} \times 100, \tag{13}$$

where $f_{xi}^{\max}$ and $f_{xi}^{\min}$ are the highest and lowest estimated values corresponding to the ith input domain, respectively, if all other input parameters are assumed to be their mean values. Moreover, n and S_i are the number of input parameters and the effective degree of the ith parameter, respectively [41].

Table 7 shows the results of the sensitivity analysis for the suggested regression models. As shown, among the different parameters analyzed in this investigation, the cement and lime contents exerted the highest effect on the UCS. These findings agree well with those of previous investigations [4]. Although the RHA affected the UCS less compared with the cement and lime contents, it is still considered as an important parameter. This significance can be attributed to its considerable effect on problematic soils as well as its environmental benefits.

6 Conclusions

The aim of this study was to evaluate the strength behavior of SM and MH soils treated with CLR by conducting UCS tests. Hence, different soil samples were prepared with various CLR contents. The UCS test results indicated that adding RHA to both types of soil (SM and MH) resulted in a considerable increase in the strength of the investigated soils. However, the rate of strength increase of the CLR-treated soil specimens was higher than that of the CL-treated specimens. Using RHA as a waste material will reduce the consumption of cement and lime, thereby demonstrating its environmental and economic advantages. In addition, the feasibility of using ANN and SVM methods to estimate the UCS values of CLR-treated soils was investigated in this study. In the ANN models, different numbers of hidden layers and architectures were used. Furthermore, two types of kernel functions, i.e., polynomial and Gaussian, were used in the SVM models.

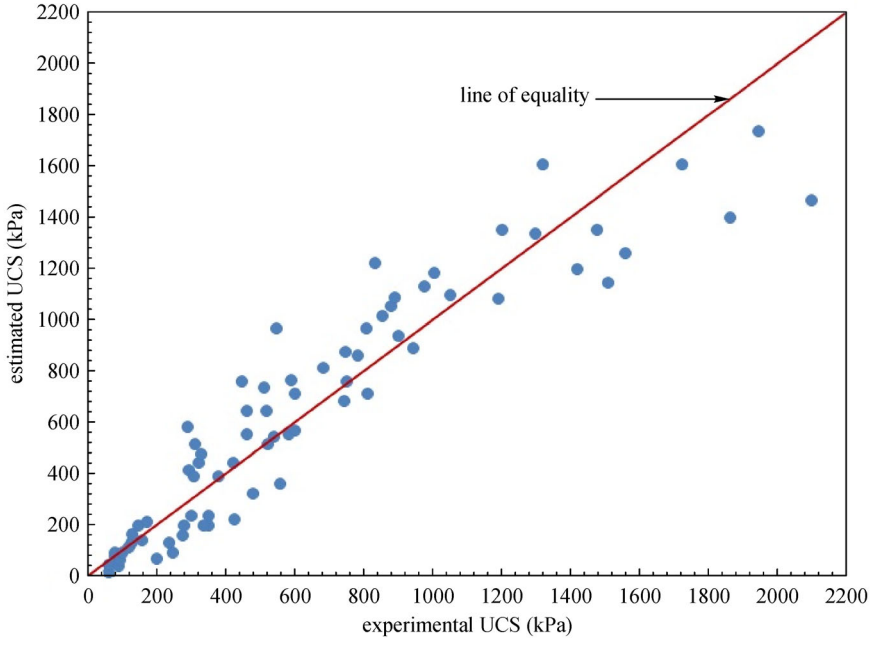


Fig. 10 Estimated values obtained via regression analysis vs. experimental data for SM soil.

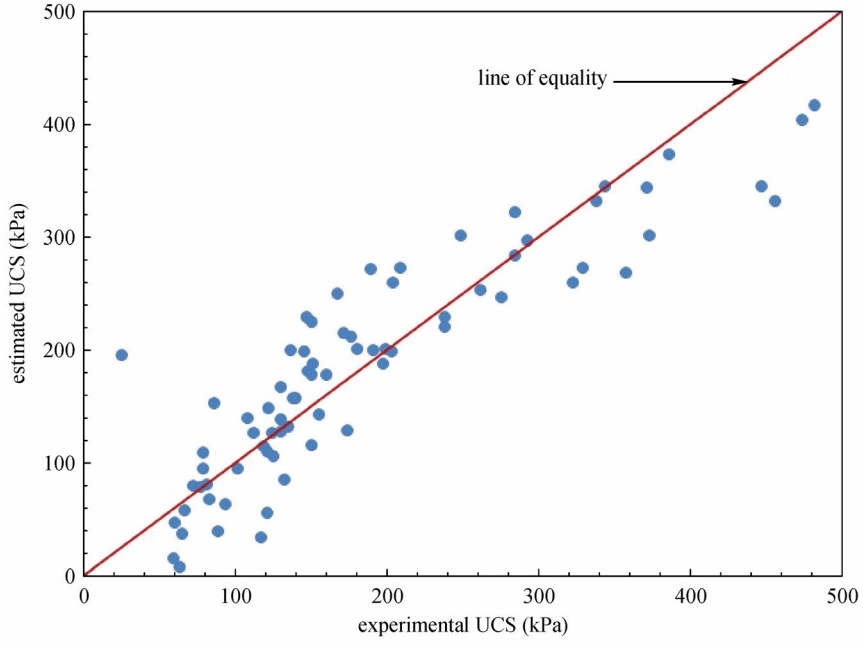


Fig. 11 Estimated values obtained by regression analysis vs. experimental data for MH soil.

Table 7 Sensitivity analysis results of input parameters in terms of UCS values

soil	effective degree (%)				
	D	*T*	*C*	*L*	*R*
SM	0.001	0.001	50.013	49.811	0.174
MH	0.001	0.002	71.627	28.212	0.158

In these models, soil type, dry unit weight, cement, lime, RHA content, and curing time were considered as input variables.

A database containing the results of UCS tests conducted on CLR-treated soil specimens was used to develop the models. Subsequently, the values of *R* and *AAPE* corresponding to the training and testing datasets were compared. By evaluating two popular shallow machine learning architectures, it was discovered that the SVM models performed better compared with the ANN models. Moreover, the type of kernel function used was effective in improving the accuracy of the developed models. By comparing the values of *R* and *AAPE* obtained by these models, it can be concluded that the SVM-RBF exhibited higher accuracy in predicting the UCS values. Additionally, these findings indicated that the two artificial-intelligence-based models (ANN and SVM) possessed higher capability compared with multiple regression analysis in predicting the UCS values. Moreover, based on the sensitivity analysis results obtained by statistical software, it was discovered that the cement and lime contents exerted a more prominent effect on the UCS values of the investigated soils compared with other parameters.

Acknowledgements The authors of this paper would like to acknowledge the support provided (No. 981861) by Golestan University.

Appendix A

Experimental variables and test results

test no.	soil type	γ_d (kN/m^3)	curing time (d)	*C* (%)	*L* (%)	*R* (%)	q_u (kPa)
1	SM	14	7	0.000	0.000	0.000	105
2	SM	14	7	1.250	2.500	1.250	247
3	SM	14	7	1.875	3.750	1.875	300
4	SM	14	7	2.500	5.000	2.500	320
5	SM	14	7	3.125	6.250	3.125	289
6	SM	15	7	1.250	2.500	1.250	294

(*Continued*)

test no.	soil type	γ_d (kN/m^3)	curing time (d)	C (%)	L (%)	R (%)	q_u (kPa)
7	SM	15	7	1.875	3.750	1.875	461
8	SM	15	7	2.500	5.000	2.500	447
9	SM	15	7	3.125	6.250	3.125	546
10	SM	16	7	1.250	2.500	1.250	512
11	SM	16	7	1.875	3.750	1.875	684
12	SM	16	7	2.500	5.000	2.500	854
13	SM	16	7	3.125	6.250	3.125	833
14	SM	17	7	1.250	2.500	1.250	879
15	SM	17	7	1.875	3.750	1.875	975
16	SM	17	7	2.500	5.000	2.500	1300
17	SM	17	7	3.125	6.250	3.125	1321
18	SM	14	28	1.250	2.500	1.250	274
19	SM	14	28	1.875	3.750	1.875	349
20	SM	14	28	2.500	5.000	2.500	423
21	SM	14	28	3.125	6.250	3.125	603
22	SM	15	28	1.250	2.500	1.250	329
23	SM	15	28	1.875	3.750	1.875	584
24	SM	15	28	2.500	5.000	2.500	752
25	SM	15	28	3.125	6.250	3.125	807
26	SM	16	28	1.250	2.500	1.250	589
27	SM	16	28	1.875	3.750	1.875	746
28	SM	16	28	2.500	5.000	2.500	1192
29	SM	16	28	3.125	6.250	3.125	1203
30	SM	17	28	1.250	2.500	1.250	890
31	SM	17	28	1.875	3.750	1.875	1421
32	SM	17	28	2.500	5.000	2.500	1865
33	SM	17	28	3.125	6.250	3.125	1724
34	SM	14	60	1.250	2.500	1.250	425
35	SM	14	60	1.875	3.750	1.875	559
36	SM	14	60	2.500	5.000	2.500	602
37	SM	14	60	3.125	6.250	3.125	812
38	SM	15	60	1.250	2.500	1.250	539
39	SM	15	60	1.875	3.750	1.875	744
40	SM	15	60	2.500	5.000	2.500	944
41	SM	15	60	3.125	6.250	3.125	1052
42	SM	16	60	1.250	2.500	1.250	783
43	SM	16	60	1.875	3.750	1.875	900
44	SM	16	60	2.500	5.000	2.500	1509
45	SM	16	60	3.125	6.250	3.125	1476
46	SM	17	60	1.250	2.500	1.250	1005
47	SM	17	60	1.875	3.750	1.875	1559
48	SM	17	60	2.500	5.000	2.500	2099
49	SM	17	60	3.125	6.250	3.125	1947
50	SM	14	7	1.250	2.500	0.000	202

(*Continued*)

test no.	soil type	γ_d (kN/m^3)	curing time (d)	C (%)	L (%)	R (%)	q_u (kPa)
51	SM	14	7	1.875	3.750	0.000	279
52	SM	14	7	2.500	5.000	0.000	306
53	SM	14	7	3.125	6.250	0.000	312
54	SM	14	28	1.250	2.500	0.000	237
55	SM	14	28	1.875	3.750	0.000	352
56	SM	14	28	2.500	5.000	0.000	377
57	SM	14	28	3.125	6.250	0.000	462
58	SM	14	60	1.250	2.500	0.000	336
59	SM	14	60	1.875	3.750	0.000	480
60	SM	14	60	2.500	5.000	0.000	522
61	SM	14	60	3.125	6.250	0.000	517
62	MH	11	0	0.000	0.000	0.000	25
63	MH	11	7	1.250	2.500	1.250	60
64	MH	11	7	1.875	3.750	1.875	63
65	MH	11	7	2.500	5.000	2.500	65
66	MH	11	7	3.125	6.250	3.125	83
67	MH	11	7	3.750	7.500	3.750	81
68	MH	12	7	1.250	2.500	1.250	67
69	MH	12	7	1.875	3.750	1.875	73
70	MH	12	7	2.500	5.000	2.500	79
71	MH	12	7	3.125	6.250	3.125	108
72	MH	12	7	3.750	7.500	3.750	86
73	MH	13	7	1.250	2.500	1.250	112
74	MH	13	7	1.875	3.750	1.875	122
75	MH	13	7	2.500	5.000	2.500	148
76	MH	13	7	3.125	6.250	3.125	176
77	MH	13	7	3.750	7.500	3.750	150
78	MH	14	7	1.250	2.500	1.250	204
79	MH	14	7	1.875	3.750	1.875	239
80	MH	14	7	2.500	5.000	2.500	262
81	MH	14	7	3.125	6.250	3.125	285
82	MH	14	7	3.750	7.500	3.750	293
83	MH	11	28	1.250	2.500	1.250	117
84	MH	11	28	1.875	3.750	1.875	121
85	MH	11	28	2.500	5.000	2.500	133
86	MH	11	28	3.125	6.250	3.125	150
87	MH	11	28	3.750	7.500	3.750	174
88	MH	12	28	1.250	2.500	1.250	125
89	MH	12	28	1.875	3.750	1.875	130
90	MH	12	28	2.500	5.000	2.500	140
91	MH	12	28	3.125	6.250	3.125	198
92	MH	12	28	3.750	7.500	3.750	199
93	MH	13	28	1.250	2.500	1.250	160
94	MH	13	28	1.875	3.750	1.875	191

(*Continued*)

test no.	soil type	γ_d (kN/m^3)	curing time (d)	C (%)	L (%)	R (%)	q_u (kPa)
95	MH	13	28	2.500	5.000	2.500	238
96	MH	13	28	3.125	6.250	3.125	323
97	MH	13	28	3.750	7.500	3.750	329
98	MH	14	28	1.250	2.500	1.250	276
99	MH	14	28	1.875	3.750	1.875	358
100	MH	14	28	2.500	5.000	2.500	373
101	MH	14	28	3.125	6.250	3.125	456
102	MH	14	28	3.750	7.500	3.750	447
103	MH	11	60	1.250	2.500	1.250	121
104	MH	11	60	1.875	3.750	1.875	135
105	MH	11	60	2.500	5.000	2.500	138
106	MH	11	60	3.125	6.250	3.125	151
107	MH	11	60	3.750	7.500	3.750	181
108	MH	12	60	1.250	2.500	1.250	151
109	MH	12	60	1.875	3.750	1.875	136
110	MH	12	60	2.500	5.000	2.500	147
111	MH	12	60	3.125	6.250	3.125	204
112	MH	12	60	3.750	7.500	3.750	209
113	MH	13	60	1.250	2.500	1.250	168
114	MH	13	60	1.875	3.750	1.875	189
115	MH	13	60	2.500	5.000	2.500	249
116	MH	13	60	3.125	6.250	3.125	338
117	MH	13	60	3.750	7.500	3.750	344
118	MH	14	60	1.250	2.500	1.250	285
119	MH	14	60	1.875	3.750	1.875	371
120	MH	14	60	2.500	5.000	2.500	386
121	MH	14	60	3.125	6.250	3.125	474
122	MH	14	60	3.750	7.500	3.750	482
123	MH	11	7	1.250	2.500	0.000	44
124	MH	11	7	1.875	3.750	0.000	60
125	MH	11	7	2.500	5.000	0.000	60
126	MH	11	7	3.125	6.250	0.000	77
127	MH	11	7	3.750	7.500	0.000	79
128	MH	11	28	1.250	2.500	0.000	88
129	MH	11	28	1.875	3.750	0.000	94
130	MH	11	28	2.500	5.000	0.000	101
131	MH	11	28	3.125	6.250	0.000	124
132	MH	11	28	3.750	7.500	0.000	156
133	MH	11	60	1.250	2.500	0.000	118
134	MH	11	60	1.875	3.750	0.000	130
135	MH	11	60	2.500	5.000	0.000	130
136	MH	11	60	3.125	6.250	0.000	145
137	MH	11	60	3.750	7.500	0.000	172

References

1. Cristelo N, Glendinning S, Fernandes L, Pinto T A. Effects of alkaline activated fly ash and Portland cement on soft soil stabilization. Acta Geotechnica, 2013, 8(4): 395–405
2. Anwar Hossain K M. Stabilized soils incorporating combinations of rice husk ash and cement kiln dust. Journal of Materials in Civil Engineering, 2011, 23(9): 1320–1327
3. Muthadhi A, Kothandaraman S. Optimum production conditions for reactive rice husk ash. Materials and Structures, 2010, 43(9): 1303–1315
4. Bagheri Y, Ahmad F, Ismail M A M. Strength and mechanical behavior of soil–cement–lime–rice husk ash (Soil–CLR) mixture. Materials and Structures, 2014, 47(1–2): 55–66
5. Shahin M A, Maier H R, Jaksa M B. Data division for developing neural networks applied to geotechnical engineering. Journal of Computing in Civil Engineering, 2004, 18(2): 105–114
6. Narendra B S, Sivapullaiah P V, Suresh S, Omkar S N. Prediction of unconfined compressive strength of soft grounds using computational intelligence techniques: A comparative study. Computers and Geotechnics, 2006, 33(3): 196–208
7. Kalkan E, Akbulut S, Tortum A, Celik A. Prediction of the unconfined compressive strength of compacted granular soils by using inference systems. Environmental Geology (Berlin), 2009, 58(7): 1429–1440
8. Suman S, Mahamaya M, Das S K. Prediction of maximum dry density and unconfined compressive strength of cement stabilized soil using artificial intelligence techniques. International Journal of Geosynthetics and Ground Engineering, 2016, 2: 1–11
9. He S, Li J. Modeling nonlinear elastic behavior of reinforced soil using artificial neural networks. Applied Soft Computing, 2009, 9(3): 954–961
10. Gunaydin O, Gokoglu A, Fener M. Prediction of artificial soil's unconfined compression strength test using statistical analyses and artificial neural networks. Advances in Engineering Software, 2010, 41(9): 1115–1123
11. Shrestha R. Deep soil mixing and predictive neural network models for strength prediction. Dissertation for the Doctoral Degree. Cambridge: University of Cambridge, 2012
12. Wang O, Al-Tabbaa A. Preliminary model development for predicting strength and stiffness of cement-stabilized soils using artificial neural networks. In: ASCE International Workshop on Computing in Civil Engineering. Los Angeles, CA, 2013, 299–306
13. Park H I, Kim Y T. Prediction of strength of reinforced lightweight soil using an artificial neural network. Engineering Computations, 2011, 28(5): 600–615
14. Mozumder R, Laskar A I. Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using artificial neural network. Computers and Geotechnics, 2015, 69: 291–300
15. Güllü H, Fedakar H I. On the prediction of unconfined compressive strength of silty soil stabilized with bottom ash, jute and steel fibers via artificial intelligence. Geomechanics and Engineering, 2017, 12(3): 441–464
16. Ghorbani A, Hasanzadehshooiili H. Prediction of UCS and CBR of microsilica-lime stabilized sulfate silty sand using ANN and EPR models: Application to the deep soil mixing. Soil and Foundation, 2018, 58(1): 34–49
17. Erzin Y, Ecemis N. The use of neural networks for CPT-based liquefaction screening. Bulletin of Engineering Geology and the Environment, 2015, 74(1): 103–116
18. Mozumder R A, Laskar A I, Hussain M. Empirical approach for strength prediction of geopolymer stabilized clayey soil using support vector machines. Construction & Building Materials, 2017, 132: 412–424
19. Wang J, Xing Y, Cheng L, Qin F, Ma T. The prediction of mechanical properties of cement soil based on PSO-SVM. In: International Conference on Computational Intelligence and Software Engineering. Wuhan: IEEE, 2010, 10: 1–4
20. Shi X, Liu Q, Lv X. Application of SVM In prediction the strength of cement stabilized soil. Applied Mechanics and Materials, 2012, 160: 313–317
21. Tinoco J, Gomes Correia A, Cortez P. Support vector machines applied to uniaxial compressive strength prediction of jet grouting columns. Computers and Geotechnics, 2014, 55: 132–140
22. Ceryan N. Application of support vector machines and relevance vector machines in predicting uniaxial compressive strength of volcanic rocks. Journal of African Earth Sciences, 2014, 100: 634–644
23. Zhao H. Slope reliability analysis using a support vector machine. Computers and Geotechnics, 2008, 35(3): 459–467
24. Kordjazi A, Pooya Nejad F, Jaksa M B. Prediction of ultimate axial load-carrying capacity of piles using a support vector machine based on CPT data. Computers and Geotechnics, 2014, 55: 91–102
25. Singh V K, Kumar D, Kashyap P S, Singh P K, Kumar A, Singh S K. Modelling of soil permeability using different data driven algorithms based on physical properties of soil. Journal of Hydrology (Amsterdam), 2019, 580: 124223
26. Xue X, Yang X. Seismic liquefaction potential assessed by support vector machines approaches. Bulletin of Engineering Geology and the Environment, 2016, 75(1): 153–162
27. Hoang N C, Bui D T. Predicting earthquake-induced soil liquefaction based on a hybridization of kernel Fisher discriminant analysis and a least squares support vector machine: A multi-dataset study. Bulletin of Engineering Geology and the Environment, 2018, 77(1): 191–204
28. Rahbarzare A, Azadi M. Improving prediction of soil liquefaction using hybrid optimization algorithms and a fuzzy support vector machine. Bulletin of Engineering Geology and the Environment, 2019, 78(7): 4977–4987
29. Haykin S S. Neural Networks: A Comprehensive Foundation. New York: Prentice Hall, ISBN 0132733501. OCLC 38908586. 1999
30. Anitescu C, Atroshchenko E, Alajlan N, Rabczuk T. Artificial Neural Network Methods for the solution of second order boundary value problems. Computers, Materials & Continua, 2019, 59(1): 345–359
31. Guo H, Zhuang X, Rabczuk T. A deep collocation method for the bending analysis of Kirchhoff plate. Computers, Materials & Continua, 2019, 59(2): 433–456
32. Aladag C H, Egrioglu E, Gunay S, Basaran M A. Improving weighted information criterion by using optimization. Journal of

Computational and Applied Mathematics, 2010, 233(10): 2683–2687

33. Vapnik V N, Golowich S E, Smola A. Support vector method for function approximation, regression estimation and signal processing. In: NIPS'96: Proceedings of the 9th International Conference on Neural Information Processing Systems. San Mateo, CA: Morgan Kaufmann, 1996
34. Kohestani V R, Hassanlourad M. Modeling the mechanical behavior of carbonate sands using artificial neural networks and support vector machines. International Journal of Geomechanics, 2016, 16(1): 04015038
35. Smola A J, Scholkopf B. A Tutorial on Support Vector Regression. Report No. NC2-TR-1998-030, Neuro COLT2 Technical Report Series. 1998
36. Hsu C W, Chang C C, Lin C J. A Practical Guide to Support Vector Classification. Technical report. Taipei, China: Taiwan University, 2016
37. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel Based Learning Methods. London: Cambridge University, 2000
38. Goh A T, Goh S. Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data. Computers and Geotechnics, 2007, 34(5): 410–421
39. Samui P. Support vector machine applied to settlement of shallow foundations on cohesionless soils. Computers and Geotechnics, 2008, 35(3): 419–427
40. Al-Anazi A F, Gates I D. Support vector regression to predict porosity and permeability: Effect of sample size. Computers & Geosciences, 2012, 39: 64–76
41. Tran K Q, Satomi T, Takahashi H. Tensile behaviors of natural fiber and cement reinforced soil subjected to direct tensile test. Journal of Building Engineering 2019, 24: 1–10