Research Article – Vol. 02, Issue 04, No. 03

Advanced in Engineering and Intelligence Systems

Journal Web Page: https://aeis.bilijipub.com

Application of Fuzzy system on Settlement of shallow footing on granular soil

Soroush Lesani[1,*]

[1] *Department of Civil Engineering, Iran University of Science and Technology, Tehran, Iran*

Highlights

- Exploration of settlement simulation in cohesionless soils, addressing the complex characteristics of cohesionless soil texture.
- Development of integrated data intelligence algorithms, combining the reptile search algorithm (RSA) with support vector regression (SVR) and adaptive neuro-fuzzy inference system (ANFIS).
- Evaluation of the RSANF and RSSVR systems, showcasing adept capabilities in estimating shallow foundation settlement (S_m).
- Achievement of smaller performance index (PI) values for RSANF during the training and test stages, indicating superior precision compared to RSSVR.
- Introduction of a novel approach to enhance the accuracy of forecasting models, contributing to a better understanding of predicting settlement in cohesionless soils.

Article Info

Received: 20 August 2023
Received in revised: 21 October 2023
Accepted: 30 December 2023
Available online: 31 December 2023

Keywords

Granular soil;
Shallow footing;
Settlement;
Machine learning;
Fuzzy system;
Reptile search algorithm

Abstract

The complex characteristics of cohesionless soil texture necessitate an examination of settlement simulation in cohesionless materials, making it a fundamental area of inquiry. This article discusses the development of integrated data intelligence algorithms with the objective of improving the reliability and accuracy of estimate results for shallow foundations (S_m) on cohesionless soils. The proposed models integrate the reptile search algorithm (RSA) with support vector regression (SVR) analysis and adaptive neuro-fuzzy inference system ($ANFIS$). Based on the findings of the research, it can be seen that the $RSSVR$ and $RSANF$ systems have shown adept capabilities in the domain of estimate. During the training stage, the $RSANF$ simulation gained the smallest performance index (PI) value of 0.0668, which was smaller than the PI of 0.0993 for $RSSVR$. Similarly, during the test stage, the $RSANF$ simulation acquired a PI of 0.0904, which was fewer than the PI of 0.1038 related to $RSSVR$. The constituents mentioned above offer a new approach to enhance the precision of forecasting models and further our comprehension of predicting the S_m.

Nomenclature

Variables	*Description*	*Variables*	*Description*
ANFIS	*Adaptive neuro-fuzzy inference system*	*PI*	*Performance index*
ANN	*Artificial neural network*	*PSO*	*Particle swarm optimization*
B	*Footing width*	*q*	*Footing net applied pressure*
BPNN	*back-propagation neural network*	*q_u*	*Carrying capacity*
D_f/H	*Depth of footing embedment*	*R^2*	*Coefficient of determination*

* Corresponding Author: Soroush Lesani

Email: S.Le01614314@gmail.com

L/B	*The ratio of footing length to B*	*RAE*	*Relative absolute error*
MAE	*Mean absolute error*	*RBF*	*Radial basis function*
MARS	*Multivariate adaptive regression splines*	*RRSE*	*Root relative squared error*
MLP	*multi-layer perceptron*	*RSA*	*Reptile search algorithm*
MPMR	*Minimax Probability Machine Regression*	*RSM*	*Response surface method*
N	*Average SPT below count*	*SCA*	*Sine-cosine algorithm*
NID	*Neural Interpretation Diagram*	S_m	*shallow foundations*
NRMSE	*Normalized root mean squared error*	*SVR*	*Support vector regression*

1. Introduction

Due to the intrinsic stochasticity of the soil formation process, the soil parameters display a considerable level of inherent uncertainty that is dependent on the current site circumstances. Hence, it is essential to consider all variations in its properties before building any structure on the ground. The soil's carrying capacity (q_u) and foundation settlement (S_m) are the two main factors that influence the planning of shallow foundations. Furthermore, settlement is the main factor that primarily affects the design of shallow foundations rather than q_u [1]. In the conventional design of shallow foundations constructed on cohesive soil, empirical formulas that are based on settlement criteria and permissible q_u are commonly utilized. To ascertain the allowable q of a shallow foundation, it is imperative to divide the ultimate q_u by a factor of safety. Moreover, the factor of safety approach is frequently utilized to compute the S_m owing to its simplicity and its capacity to establish a direct correlation with settlement. However, the simulation does not incorporate the possible uncertainties related to the parameters that affect the characteristics of the soil [2–4]. Phoon [5] has established that the variability observed in test data can be attributed to two factors: errors in testing and variations in soil deposits. The research conducted a comprehensive reliability analysis considering the diverse soil properties range. The input parameters were considered stochastic factors, and diverse algorithms were employed to examine their influence on the targets. Several academics have utilized a probabilistic approach, utilizing either data from standard penetration tests or consolidation tests, to develop algorithms based on probabilistic principles with the aim of predicting total S_m [6]. Cherubini [7] introduced a probabilistic approach to ascertain the ultimate q_u of a shallow foundation footing on cohesionless soil. This method entails employing a proficient solution for the friction angle. Easa [8] utilized a probabilistic methodology that incorporated two variables, namely effective frictional angle and soil unit weight, to ascertain the ultimate bearing capacity q_u of shallow foundations that are located on cohesionless soil. The study's results suggest that alterations in unit weight have a substantial effect on the ultimate q_u of shallow foundations.

Various academics have used numerous methodologies to undertake reliability analysis. To create an approximated polynomial function, Babu and Srivastava [9] used the response surface method (RSM) and a valid set of soil characteristics as a parameter set for the estimation of S_m and q_u. In order to predict shallow S_m on granular soil, Shahin et al. [10] first defined two separate modeling approaches, namely multi-layer perceptron (MLPs) and B-spline Neuro-fuzzy networks. The largest surface S_m resulting from tunneling was predicted in the research of Hasanipanah et al. [11] using the particle swarm optimization artificial neural network ($PSO-ANN$) method. The results showed that the $PSO-ANN$ outperformed the standard ANN system in terms of reliability in forecasting the maximum surface S_m. According to Tarawneh's [12] research, the utilization of back-propagation neural network ($BPNN$) in predicting N_{60}-value through CPT data resulted in high accuracy ($R^2 = 0.95, MAE = 2.88$).

Recent literature may be used to reference more machine learning approaches and the reliability analysis conducted [13–16]. The S_m induced by liquefaction for buildings with weak foundations may be calculated simply because to fascinating research, which defined the process in detail. The multivariate adaptive regression splines ($MARS$) method is chosen [17]. To determine the S_m of a shallow strip footing resting on granular soils as a consequence of the interplay of static and cyclic loads, an ANN framework equation was created in study. To explore how input elements impact output, the Neural Interpretation Diagram (NID) is utilized [18]. The potential of the soft computing technique—more particularly, Gaussian process regression (GPR)—was investigated to anticipate q_u of cohesionless soils under shallow foundations. Based on the results obtained, it can be concluded that the generated GPR proved to be a dependable technique for the precise estimation of shallow foundations situated on soil with low cohesion [19]. ANNs have been employed in various studies to predict the settlement of a foundation on sandy soil. The demonstrated precision of the proposed formula highlights the significant potential of ANNs as a forecasting tool for predicting foundation settlement on sandy soils [20]. A recent study

has shown specific interest in utilizing recently created machine learning methods, such as the combined optimization-based *ANFIS* methods, as efficient means for predicting S_m over cohesion soil properties. The findings indicate that the $ANFIS-PSO$ exhibited superior predictive capacity compared to the other models employed, and demonstrated precise and reliable predictive data intelligence [21]. The study explores the reliability of shallow foundations in meeting settlement requirements through the implementation of three soft computing techniques. Based on the results of the analysis, it was determined that the Minimax Probability Machine Regression ($MPMR$) system outperformed other systems [21]. The study aimed to predict the settlement of a shallow foundation on granular soil by employing a mathematical model. The findings indicate that *ANN* models possess the ability to predict intricate relationships among non-linear variables, as observed in the present case [22]. Moreover, regarding to the considered methods and optimization algorithm in this study (i.e., SVR, $ANFIS$ and RSA), it is worth to assess the successful application of them in prediction processes. So limited usages were reported in the literature such as predicting COVID-19 [23], analyzing groundwater level [24], and monthly streamflow prediction [25].

The current article presents the establishment of combined data intelligence algorithms aimed at providing more dependable and precise S_m estimation outcomes. These models combine the reptile search algorithm (RSA) with support vector regression (SVR) analysis and adaptive neuro-fuzzy inference system ($ANFIS$). Optimization techniques facilitated the identification of the optimal primary variable value for both the SVR model and the $ANFIS$. The viability of the blended algorithms under investigation was established based on the consideration of S_m throughout the literature review.

The main benefit of the proposed networks is their ability to accurately model the complex relationship between the inputs (important qualities) and the outputs (goal parameters) without the need for a predetermined mathematical equation. Following this, the results are evaluated and contrasted with the findings documented in prior academic literature. This work is noteworthy for various reasons, including the incorporation of RSA, the comparison with pertinent literature, and the exploitation of a comprehensive dataset including several input variables. The aforementioned parts provide a novel methodology for improving the accuracy of forecasting models and deepening our understanding of predicting S_m.

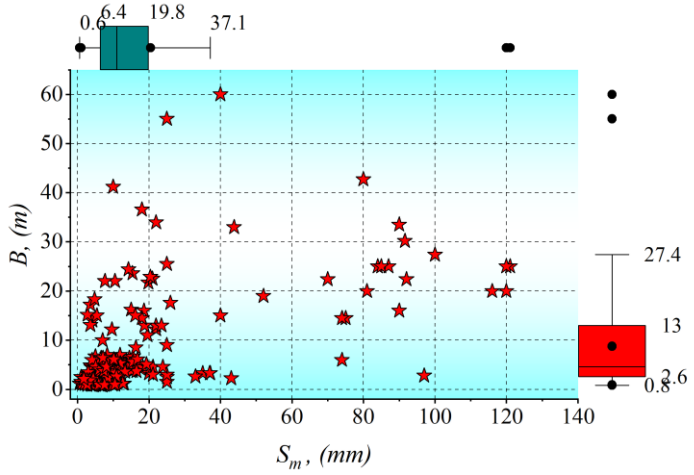
2. Dataset

The dataset employed in this study to develop prediction models was obtained from various previously published scholarly articles. The dataset comprises essential information regarding the in-situ measurement of S_m for the footing and soil under diverse conditions. The criteria include several modifications to the geometry requirements of the footing as well as changes to the soil's classification and characteristics. Interestingly, adding a wide variety of data to the dataset might improve the precision of the created algorithms. Important parameters were introduced as input, like footing width (B), footing net applied pressure (q), average SPT below count (N), the ratio of footing length to B (L/B), and depth of footing embedment (D_f/H), to predict S_m [10,17,18,20,22]. The dataset includes around 190 occurrences of in-situ analyses of elastic S_m [26]. Based on the data dividing literature [27–29], the collection experienced a partitioning procedure into two separate phases, namely the training and testing data phases, with a percentage of 75% and 25%, respectively, to accomplish the desired purpose. The statistical characteristics of the dependent and independent variables throughout the training and testing phases are shown in Table 1. Fig. 1 shows the distribution of the dependent and independent variables.

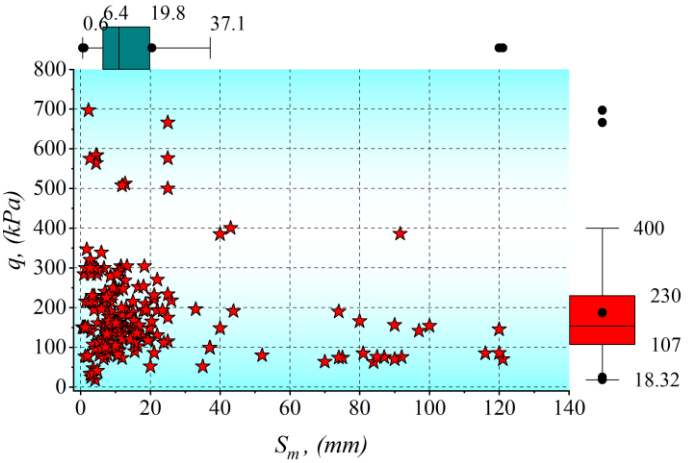
Table 1. The dependent and independent variables' characteristics

Data	*Index*	B (m)	q (kpa)	N	$\frac{D_f}{H}$	L/B	S_m (mm)
Train	*Max.*	*60*	*697*	*60*	*3.44*	*10.60*	*121*
	Min.	*0.80*	*18.32*	*4.0*	*0.0*	*1.00*	*0.6*
	St.d.	*9.0428*	*123.401*	*13.09*	*0.5286*	*1.731*	*25.4178*
	Avg.	*7.99*	*186.13*	*24.58*	*0.51*	*2.22*	*19.43*
	median	*4.55*	*151.60*	*20.00*	*0.47*	*1.60*	*11.00*
	Skew.	*2.4351*	*2.0062*	*0.847*	*2.5254*	*2.005*	*2.435*
	Kur.	*7.867*	*4.555*	*0.158*	*8.7215*	*4.623*	*5.0631*
Test	*Max.*	*55*	*584*	*60*	*3.0*	*9.9*	*120*
	Min.	*0.9*	*25*	*4.0*	*0.0*	*1.0*	*1.3*
	St.d.	*12.587*	*121.4229*	*14.66*	*0.70183*	*1.9676*	*29.456*

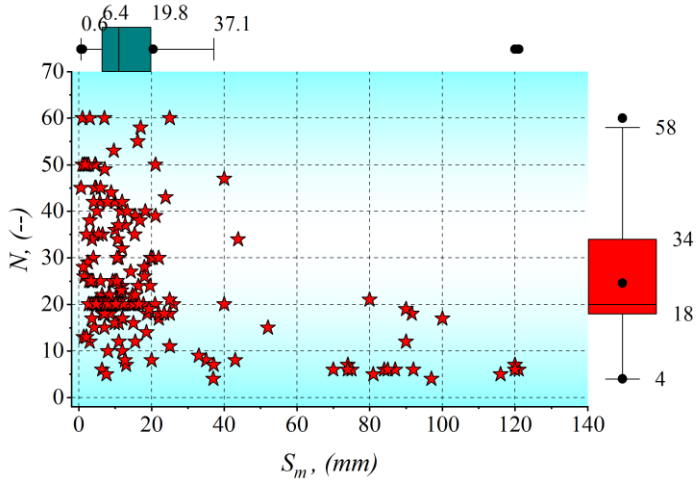
Avg.	*11.15*	*190.01*	*24.55*	*0.61*	*2.11*	*23.51*
median	*5.10*	*162.00*	*20.00*	*0.40*	*1.10*	*10.90*
Skew.	*1.647*	*1.41*	*0.672*	*1.9967*	*2.588*	*2.05446*
Kur.	*2.564*	*2.619*	*-0.304*	*4.184*	*6.755*	*3.681*



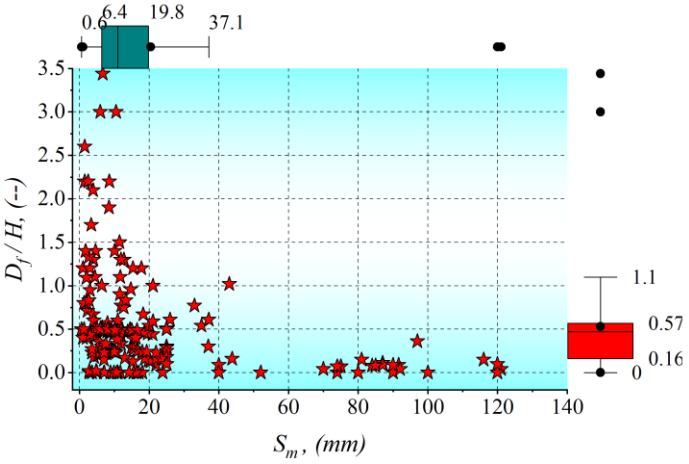
(i_1)



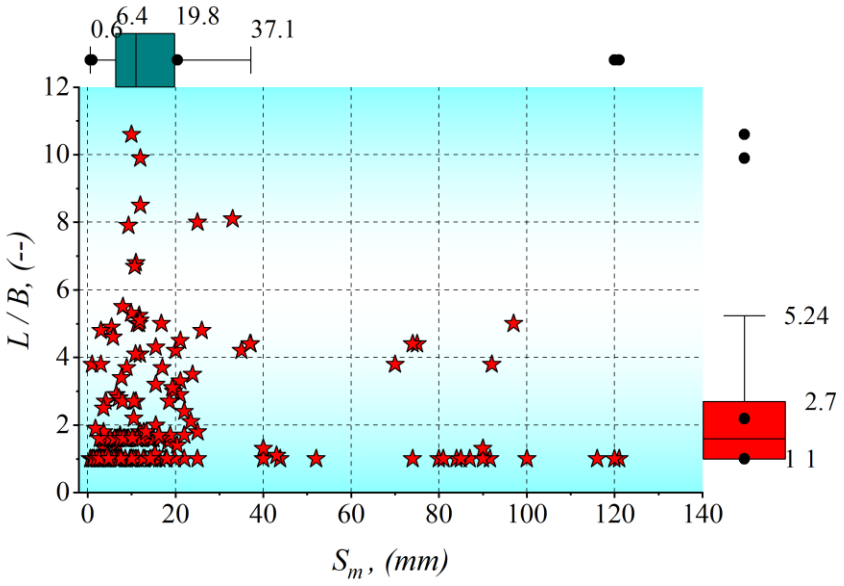
(i_2)



(i_3)



(i_4)



(i_5)

Fig. 1. Distribution of the input variables versus S_m

3. Used optimization and methods

3.1. Reptile search algorithm (RSA)

The reptile search method mostly emulates crocodiles' predation method and the social manner in nature [30]. The first occurs during the exploration stage as the encirclement technique, and the second occurs during the exploitation stage as the hunting technique. Before the iteration begins, a collection of potential answers is often subjected to a normal MH method. It is a created approach at random.

$$Z = \begin{bmatrix} Z_{1,1} & \dots & Z_{1,j} & Z_{1,d-1} & Z_{1,d} \\ Z_{2,1} & \dots & Z_{2,j} & \dots & Z_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Z_{N-1,1} & \dots & Z_{N-1.j} & \dots & Z_{N-1,d} \\ Z_{N.1} & \dots & Z_{N,j} & Z_{N.d-1} & Z_{N,d} \end{bmatrix} \quad (1)$$

Based on this equation, $Z_{i,j}$ shows the i_{th} crocodile's j_{th} dimension. N stands for whole crocodiles, and d shows dimension. Z stands for N potential answers accidentally produced by Eq. (2).

$$Z_{i,j} = rand \times (up - low) + low \quad (2)$$

Based on Eq. (2), low and up show the lowest limit and highest limits of the optimization issue, and $rand$ denotes an accidental number.

The distinctive feature of RSA's worldwide search is the encircling manner. Two manners—aerial and abdominal walk—make up the procedure. Crocodiles are often prevented from approaching food by these two manners. Nevertheless, since it is a worldwide search of the entire resolved spatial domain, the crocodile will accidentally discover the general region of the specific meal after multiple searches tries. Make sure the phase can be continually adjusted to the subsequent developmental phase in the meantime. The procedure is often restricted to the first stage of the whole cycle. Eq. (3), seen below, mimics the crocodile's encircling action.

$$Z_{i,j}(t+1) = \begin{cases} best_j(t) \times \left(-\eta_{i,j}(t)\right) \times \beta - R_{i,j}(t) \times r, & t \le \frac{T_{Max}}{4} \\ best_j(t) \times Z_{r_1,j} \times ES(t) \times r, & \frac{T_{Max}}{4} \le t < \frac{2 \times T_{Max}}{4} \end{cases} \quad (3)$$

According to this equation, $best_{i,j}(t)$ shows the finest-situated crocodile at t iterations, and r shows an accidental number in the range [0,1]. T_{Max} shows the highest iteration. $\eta_{i,j}$ stands for the i_{th} crocodile's operator at the j_{th} dimension and presented by Eq. (4). The sensitive variable β, which controls the search precision, is stated in the original text as 0.1 [30]. Eq. (5) is utilized to reduce the searched region and computes $R_{i,j}$, which is the reduction function. A random number between 1 and N is called r_1, and the crocodile's jth dimension is called $Z_{r_1,j}$. Eq. (6) states that $ES(t)$ is an optional lowering possibility ratio between 2 and -2.

$$\eta_{i,j} = best_{i,j}(t) \times P_{i,j} \quad (4)$$

$$R_{i,j} = \frac{best_j(t) - x_{r_2,j}}{best_j(t) + \varepsilon} \quad (5)$$

$$ES(t) = 2 \times r_3 \times \left(1 - \frac{1}{T}\right) \quad (6)$$

Based on these equations, r_2 stands for an accidental number in the range $[1, N]$, and r_3 stands for a number in the range [-1,1]. ε shows a low value. $P_{i,j}$, updated as in Eq.

(7), stands for the percentage gap between the crocodile in the ideal location and the present location.

$$P_{i,j} = \alpha + \frac{z_{i,j} - M(z_i)}{bset_j(t) \times (up_j - low_j) + \varepsilon'} \quad (7)$$

In this equation, $M(z_i)$ shows the z_i Crocodile's mean location is presented in Eq. (8). low_j and up_j show the lowest limit and highest limit of the j_{th} dimension. α stands for a sensitive variable to maintain the search precision the same as β. It was fixed to 0.1 in the main study [30].

$$M(z_i) = \frac{1}{n}\sum_{j=1}^{n} z_{i,j} \quad (8)$$

The hunting procedure of local symbolic exploitation, that has two tactics in this section: coordination and cooperation, is linked to the *RSA*'s search procedure. Following the encirclement technique's effect, the crocodiles nearly always know where to find their hunt, and their hunting tactics will make it simpler for them to come close to their hunt. Over multiple rounds, the development process often discovers the almost ideal potential answer. Eq. (9) presents its simulated crocodile hunting manner's mathematical model. In the iterations' second part, this development hunting technique is implemented.

$$z_{i,j}(t+1) = \begin{cases} best_j(t) \times P_{i,j}(t) \times r, & \frac{2 \times T_{Max}}{4} \le t < \frac{3 \times T_{max}}{4} \\ best_j(t) - \eta_{i,j}(t) \times \varepsilon - R_{i,j}(t) \times r, & \frac{3 \times T_{max}}{4} \le t < \frac{4 \times T_{Max}}{4} \end{cases} \quad (9)$$

In Eq. (9), $best_j$ stands for the crocodile's best location, $\eta_{i,j}$ show the i_{th} crocodile's operator at j_{th} dimension and is presented in Eq. (4). $R_{i,j}$ shows the decrease function, and this equation is utilized to reduce the searched region and calculated by Eq. (5). *RSA* is represented by the pseudocode of Algorithm 1.

Algorithm 1: the *RSA*'s framework [30]

input: The RSA's variables containing the sensitive variables α, β, crocodile size (N), and the highest generation T_{Max}
initializing n crocodile, z_i and compute f_i
specify the finest crocodile $best_j$
while ($t \le T_{max}$) **do**
Eq. (6) update the ES
for $i = 1$ to N **do**
for $i = 1$ to N **do**
using Eqs. (4), (5), and (7) compute the η, R, P
if $t \le T_{Max}/4$ **then**
$z_{i,j}(t+1) = best_j(t) \times \eta_{i,j} \times \beta - R_{i,j} \times r$
else if $\frac{T_{Max}}{4} \le t < 2 \times T_{Max}/4$
$z_{i,j}(t+1) = best_j(t) \times z_{i,j} \times ES(t) \times r$
else if $\frac{2 \times T_{Max}}{4} \le t < \frac{3 \times T_{max}}{4}$
$z_{i,j}(t+1) = best_j(t) \times P_{i,j}(t) \times r$
else
$z_{i,j}(t+1) = best_j(t) - \eta_{i,j}(t) \times \varepsilon - R_{i,j}(t) \times r$
end if
end for
end for
obtain the finest crocodile.
$t = t + 1$.
end while
output: the finest crocodile.

3.2. Support vector regression (SVR)

Support vector machines (*SVM*) were first developed to address categorization difficulties, and later they were expanded to address regression issues (*SVR*) [31]. This methodology's central concept is to find points near a hyperplane (support vectors) which optimize the distance among two-point categories (points higher and lower than the objective parameter) derived from the variance among the objective value and a threshold. By attempting to reduce the higher limit of the generalization fault rather than the training fault, this technique seeks to reduce the structural hazards in regression issues.

Since most practical issues exhibit nonlinear properties, the kernel notion may be added to the *SVR* technique. The intrinsic properties of the data may be mapped utilizing kernel functions. Numerous kernel functions, including radial basis, Gaussian, linear, and polynomial are employed in the literature. The linear kernel is employed in this paper's two case research that may be described as

$$k(x_i, x_j) = x_i' x_j \quad (10)$$

This is the same as the inner product of the observations $x_i'$and x_j in some characteristic areas. The variable cost ought to be established when this kernel is employed. This variable is acquired in this context using grid search during the cross-validation process.

On the one hand, *SVR*'s ability to take forecaster nonlinearities and then utilize them to enhance price forecasts is one of the key benefits of using it for the case research selected in this research, particularly for layer-1 in the *STACK* technique. Similarly, given the limited sample sizes in the chosen case research, it is appropriate to use this approach [31]. However, the data collection's fundamental drawback is connected to the selected kernel function. Making the wrong choice while using a function might result in inaccurate results. The main goal in the *SVR* process is to choose the optimal values for crucial parameters, notably C, $gamma$, and ε. The *RSA* method is used in conjunction with the *SVR* model to accomplish this goal. These parameters were examined in several combinations, covering a range of values. $Gamma$ ranged from 0.05 to 5, ε varied from 0.05 to 5, and C was investigated in the range of 1 to 700.

3.3. Adaptive neuro-fuzzy inference system (ANFIS)

According to fuzzy logic theory, a fuzzy interface system employs a set of statements, specifically fuzzy if-then rules, to translate the input to result data [32].

Before being completely de-fuzzified into crisp results, the fuzzy inputs are translated to fuzzy results. The *ANFIS* model is composed of five levels, every one of which is comprised of inputs and results from the one before it. The following describes the first level's adaptive nodes:

$$O_{1,i} = \mu_{A_{i-2}}(x) \quad (11)$$

$$O_{1,i} = \mu_{B_{i-2}}(y) \quad (12)$$

Based on these equations, y and x stand for the inputs, A and B show linguistic labels, $\mu_{A_{i-2}}(x)$ and $\mu_{B_{i-2}}(y)$ show the membership function's degrees, and $O_{1,i}$ stands for the first level's result. According to membership functions, the first level's inputs are fuzzified into membership values $(\mu_{A_{i-2}}(x))$.

In the *ANFIS* model, a bell-shaped function is often utilized:

$$\mu(x) = exp\left(-\left(\frac{x-c_i}{a_i}\right)^2\right) \quad (13)$$

In Eq. (13), a_i and c_i show the premise variables. The second level's results—known as firing strengths—are calculated using the multiplication of the incoming signals.

$$O_{2,i} = \omega_i = \mu_{A_i}(x).\mu_{B_i}(y) \quad (14)$$

According to Eq. (14), $O_{2,i}$ defines the second level's results. The firing power of the i_{th} node is compared to the firing power of entire rules in the third level to get the ratio. According to the inputs and fuzzy collections, the following rules are stated:

$$O_{3,i} = \overline{\omega_i} = \frac{\omega_i}{\sum \omega_i} = \frac{\omega_i}{\omega_1+\omega_2}$$
$$Rule(1) = if(x)is(A_1)and(y)is(B_1)then(f_1) = p_1x + q_1y + r_1$$
$$Rule(2) = if(x)is(A_2)and(y)is(B_2)then(f_2) = p_2x + q_2y + r_2 \quad (15)$$

Based on this equation, $O_{3,i}$ shows the third level's results, $p_1, q_1, r_1, p_2, q_2, r_2$ stand for consequent variables, A_1, A_2, B_1 and B_2 show the fuzzy collections and the fuzzy rules presented by f_1 and f_2. Using Eq. (14) and fuzzy if-then rules, the fourth level calculates every rule's result utilizing the weight assigned in the previous level.

$$O_{4,i} = \overline{\omega_i}.f_i, f = 1,2 \quad (16)$$

According to Eq. (16), $O_{4,i}$ show the fourth level's result and f_i shows the if-then rules in Eq. (6). In the fifth level, the total of every result is calculated.

$$O_{5,i} = f_{out} = \sum_{i=1}^{2} \overline{\omega_i}.f_i \quad (17)$$

In this equation, $O_{5,i}$ stands for the general result. The expertise of experts may be used to define the variables of an *ANFIS* model, like the premise and consequent variables. To determine its variables, a hybrid approach according to forward and backward passes may be utilized. The premise variables may be assessed using the backward pass's gradient descent approach, though the consequent variables can be calculated using the least squares technique in the forward pass. A hybrid approach, nevertheless, may not converge quickly enough or properly determine the premise and ensuing variable values. An *ANFIS* model's structure is shown in Fig. 2. In this task, the variables of membership functions of the produced *ANFIS* were trained (optimized) using the *RSA* algorithm.

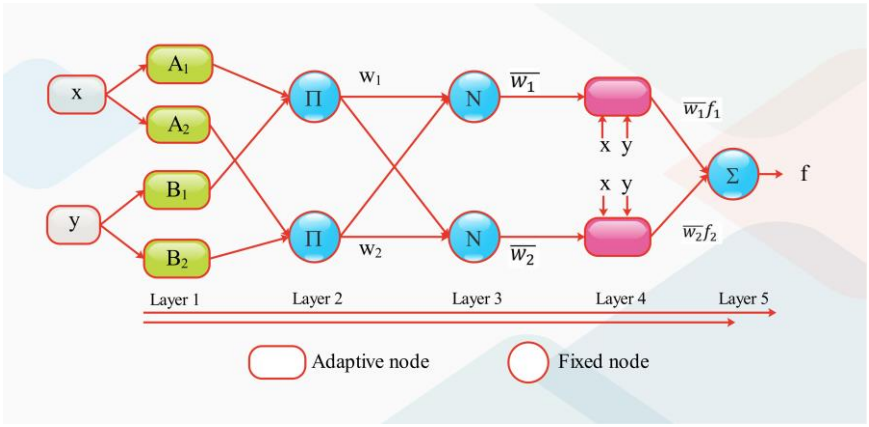


Fig. 2. The structure of the *ANFIS*

3.4. Indices

Different correlation and error measures may be used to assess the effectiveness of the constructed models. To achieve the intended objective, many indices were taken into account, including the coefficient of determination (R^2), normalized root mean squared error ($NRMSE$), relative absolute error (RAE), root relative squared error ($RRSE$), mean absolute error (MAE), and performance index (PI). The greater the correlation (R^2) and the lesser the error metrics ($NRMSE$, RAE, $RRSE$, and MAE), the more robust the models.

$$R^2 = \left(\frac{\sum_{d=1}^{D}(m_d - \bar{m})(z_d - \bar{z})}{\sqrt{[\sum_{d=1}^{D}(m_P - m)^2][\sum_{d=1}^{D}(z_d - \bar{z})^2]}}\right)^2 \quad (18)$$

$$RMSE = \sqrt{\frac{1}{D}\sum_{d=1}^{D}(z_d - m_d)^2} \quad (19)$$

$$NRMSE = RMSE/\bar{z} \quad (20)$$

$$RAE = \frac{\sum_{d=1}^{D}|m_d - z_d|}{\sum_{d=1}^{D}|m_d - \bar{m}|} \quad (21)$$

$$RRSE = \sqrt{\frac{\sum_{d=1}^{D}(m_d - z_d)^2}{\sum_{d=1}^{D}(m_d - \bar{m})^2}} \quad (22)$$

$$MAE = \frac{1}{D}\sum_{d=1}^{D}|z_d - m_d| \quad (23)$$

$$PI = \frac{1}{\bar{m}}\frac{RMSE}{\sqrt{R^2 + 1}} \quad (24)$$

where:

m_d : The observations
$\bar{m}$: The average of observations
z_d : The estimations from models
$\bar{z}$: The average of estimations from models
D : Number of data

4. Results and discussion

To calculate S_m, the output of the suggested hybrid SVR and $ANFIS$ models—abbreviated as $RSSVR$ and $RSANF$—was assembled and explained below. The considered data was split randomly into the training and examining sections by 75 and 25%, respectively. Fig. 3 presents the ratio of predicts to assessments for $RSSVR$ and $RSANF$ in a violin plot, along with the correlation between predicts and assessments S_m. Table 2 presents a comparison of various measures, such as R^2, $NRMSE$, RAE, $RRSE$, MAE, and PI, that were computed to attain the desired gain. In addition, it may be beneficial to compare the findings of a study with earlier presented articles to validate the predictions and improve their precision and thoroughness. The research attempted to establish a feasible comparison with the hybrid $ANFIS - PSO$ [21] and $SCA - RBFNN$ [33] models.

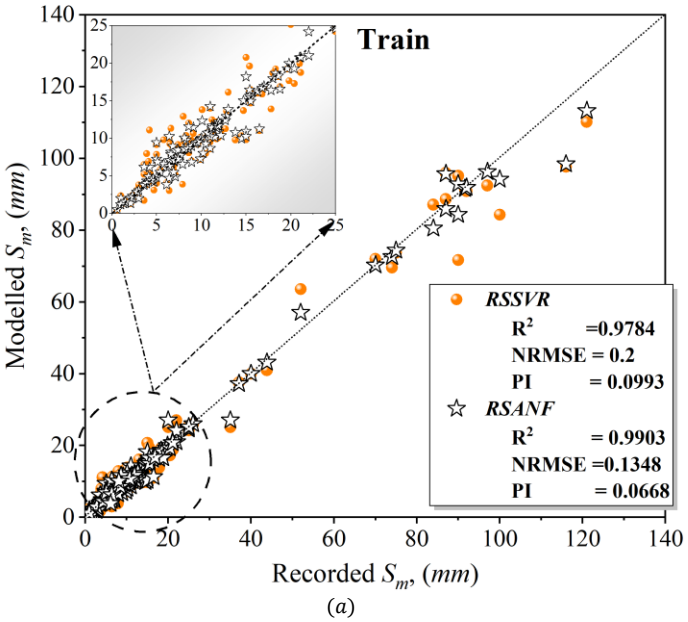
According to the study's results, it appears that the $RSSVR$ and $RSANF$ systems have demonstrated proficient abilities in estimation. The results indicate that the $RSANF$ and $RSSVR$ models have high R^2 values, with the $RSANF$ achieving 0.9903 and 0.9814 for its training and testing components, respectively, and the $RSSVR$ achieving 0.9784 and 0.9753 for its training and testing components. It is important to thoroughly analyze and evaluate the generated systems to determine the most optimal methodology. Throughout the training phase, it was noted that the $RSANF$ simulation demonstrated reduced values for all error metrics, including $NRMSE$, RAE, MAE, PI, and $RRSE$, when compared to the $RSSVR$ system. About the $RSANF$ framework, it was observed that the $NRMSE$ statistics underwent a reduction from 0.2 to 0.1348 during the training section and from 0.2003 to 0.175 during the testing phase. Additional error measurements, namely

RAE, *RRSE*, and *MAE*, also demonstrate the superior achievement of the *RSANF* simulation in comparison to the *RSSVR* approach. With regards to the *PI* indicator, it was observed that a smaller number indicates better effectiveness. Specifically, during the training stage, the *RSANF* simulation gained the smallest *PI* value of 0.0668, which was smaller than the *PI* of 0.0993 for *RSSVR*. Similarly, during the test stage, the *RSANF* simulation acquired a *PI* of 0.0904, which was fewer than the *PI* of 0.1038 related to *RSSVR*. The presented arguments and clarifications indicate that the *ANFIS* approach, when combined with *RSA*, can be classified as a superior model, despite the *RSSVR*'s efficacy in the prediction process.

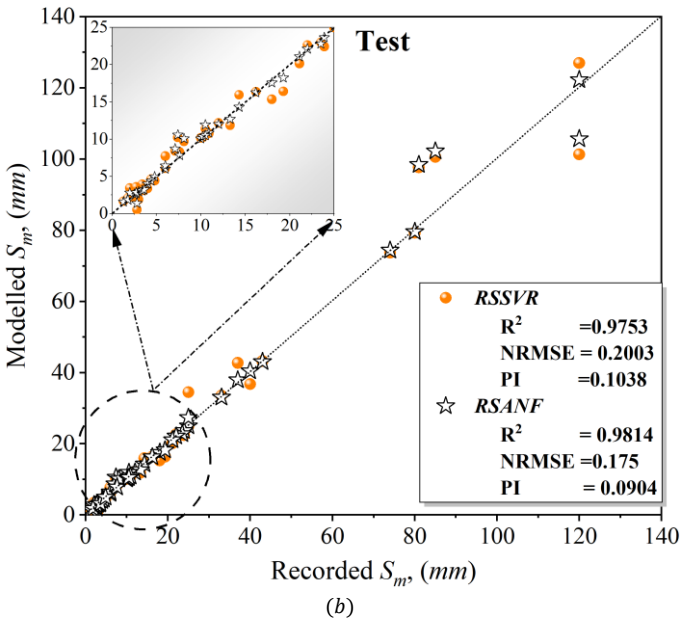
A thorough evaluation establishes the dependability of the models derived from pertinent scholarly sources [21] [33]. Table 2 illustrates the comparative advantage of our proposed *RSANF* over the prior research works referenced in the scholarly literature. The indices that were deemed valid were utilized for comparison purposes, such as R^2, *MAE*, and *PI*. The $ANFIS-PSO$ technique, as reported in reference [21], resulted in a significant increase in the R^2 value for the *RSANF* model. Specifically, the R^2 value improved from 0.9025 to 0.9903 during the learning stage and from 0.739 to 0.9814 during the examining stage. Furthermore, the *RSANF* model that was formulated exhibits superior performance in comparison to $SCA-RBFNN$ [33]. This is evidenced by the increase in the R^2 values and the decrease in both the *MAE* and *PI* values.

Table 2. The models' performance

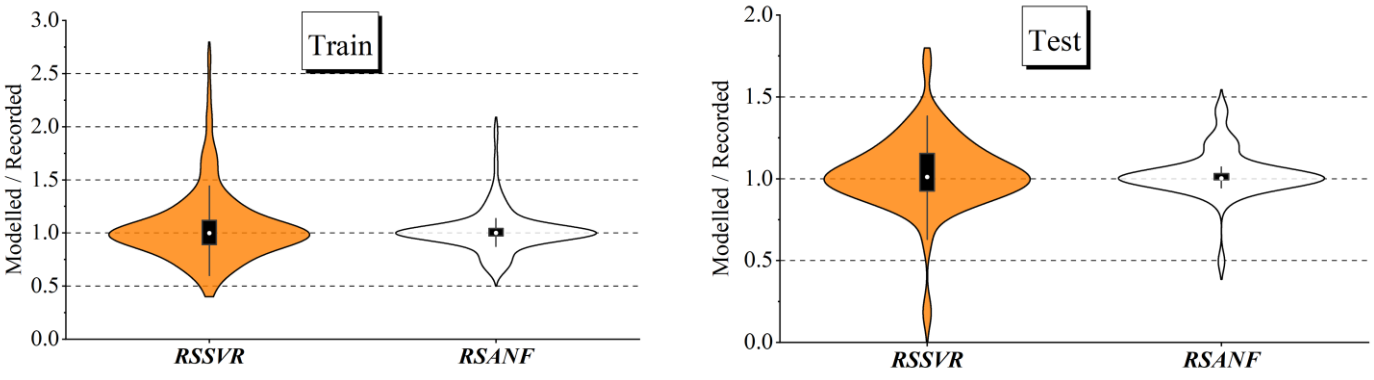
Model	*Index*					
	R^2	*NRMSE*	*RAE*	*RRSE*	*MAE*	*PI*
Train						
(This study)RSSVR	*0.9784*	*0.2*	*0.1414*	*0.151*	*2.2963*	*0.0993*
(This study)RSANF	*0.9903*	*0.1348*	*0.0831*	*0.1018*	*1.3497*	*0.0668*
[21]ANFIS − PSO	*0.9025*	–	–	–	–	–
[33]SCA − RBFNN	*0.9565*	–	–	–	*3.8642*	*0.1407*
Test						
(This study)RSSVR	*0.9753*	*0.2003*	*0.1132*	*0.1646*	*2.3246*	*0.1038*
(This study)RSANF	*0.9814*	*0.175*	*0.0745*	*0.1437*	*1.5304*	*0.0904*
[21]ANFIS − PSO	*0.739*	–	–	–	–	–
[33]SCA − RBFNN	*0.9422*	–	–	–	*5.1257*	*0.156*



(*a*)



(*b*)



(*c*) (*d*)

Fig. 3. The findings of the created analysis

The study's effectiveness heavily relies on the quality and quantity of the dataset used. Future studies should focus on obtaining more extensive and higher-quality data to enhance the reliability and generalizability of the proposed algorithms. While the study emphasizes the ability to capture complex relationships, the interpretability of the models may be limited. Researchers could consider methods for explaining model predictions, especially in applications where interpretability is crucial.

Future studies could provide a clear and detailed comparison with existing models, including various evaluation metrics. Future research could explore ensemble methods that combine multiple models (e.g., combining *RSA*, *SVR*, and *ANFIS*) to improve predictive accuracy and reduce model bias. Develop techniques to enhance the interpretability of the models, especially in applications where understanding the model's decision-making process is crucial. Applying the proposed algorithms to real-world scenarios and industries where estimating S_m is relevant, such as in environmental monitoring, materials science, or geophysics.

5. Conclusions

The current article presents the establishment of combined data intelligence algorithms aimed at providing more dependable and precise shallow foundations (S_m) on cohesionless soils' estimation outcomes. These models combine the reptile search algorithm (*RSA*) with support vector regression (*SVR*) analysis and adaptive neuro-fuzzy inference system (*ANFIS*). Optimization techniques facilitated the identification of the optimal primary variable value for both the *SVR* model and the *ANFIS*.

- According to the study's results, it appears that the *RSSVR* and *RSANF* systems have demonstrated proficient abilities in estimation. The results indicate that the *RSANF* and *RSSVR* models have high R^2 values, with the *RSANF* achieving 0.9903 and 0.9814 for its training and testing components, respectively, and the *RSSVR* achieving 0.9784 and 0.9753 for its training and testing components.
- Throughout the training phase, it was noted that the *RSANF* simulation demonstrated reduced values for all error metrics, including *NRMSE*, *RAE*, *MAE*, *PI*, and *RRSE*, when compared to the *RSSVR* system. During the training stage, the *RSANF* simulation gained the smallest *PI* value of 0.0668, which was smaller than the *PI* of 0.0993 for *RSSVR*. Similarly, during the test stage, the *RSANF* simulation acquired a *PI* of 0.0904, which was fewer than the *PI* of 0.1038 related to *RSSVR*.
- The $ANFIS-PSO$ technique, as reported in reference [21], resulted in a significant increase in the R^2 value for the *RSANF* model. Specifically, the R^2 value improved from 0.9025 to 0.9903 during the learning stage and from 0.739 to 0.9814 during the examining stage. Furthermore, the *RSANF* model that was formulated exhibits superior performance in comparison to $SCA-RBFNN$ [33]. This is evidenced by the increase in the R^2 values and the decrease in both the *MAE* and *PI* values.
- The study's application in practical efforts can be significant in the field of civil engineering and construction. For example, improving foundation design, cost savings, safer structures, environmental impact, risk mitigation. In summary, the practical applications of the research lie in its potential to enhance the efficiency, safety, and cost-effectiveness of construction projects on

cohesionless soils. These applications can benefit engineers, construction companies, regulators, and the broader construction industry by improving the predictability and management of settlement-related challenges.

REFERENCES

[1]J. H. Schmertmann, "Static cone to compute static settlement over sand," *Journal of the Soil Mechanics and Foundations Division*, vol. 96, no. 3, pp. 1011–1043, 1970.

[2]J. M. Duncan, "Factors of safety and reliability in geotechnical engineering," *Journal of geotechnical and geoenvironmental engineering*, vol. 126, no. 4, pp. 307–316, 2000.

[3]R. V Whitman, "Organizing and evaluating uncertainty in geotechnical engineering," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 126, no. 7, pp. 583–593, 2000.

[4]H. F. Schweiger, R. Thurner, and R. Pöttler, "Reliability analysis in geotechnics with deterministic finite elements—Theoretical concepts and practical application," *International Journal of Geomechanics*, vol. 1, no. 4, pp. 389–413, 2001.

[5]K. K. Phoon, "Potential application of reliability-based design to geotechnical engineering," in *4th Colombian Geotechnical Seminar*, 2002, pp. 1–24.

[6]R. J. Krizek, R. B. Corotis, and H. H. El-Moursi, "Probabilistic analysis of predicted and measured settlements," *Canadian Geotechnical Journal*, vol. 14, no. 1, pp. 17–33, 1977.

[7]C. Cherubini, "Reliability evaluation of shallow foundation bearing capacity on c'φ'soils," *Canadian Geotechnical Journal*, vol. 37, no. 1, pp. 264–269, 2000.

[8]R. Ray, D. Kumar, P. Samui, L. B. Roy, A. T. C. Goh, and W. Zhang, "Application of soft computing techniques for shallow foundation reliability in geotechnical engineering," *Geoscience Frontiers*, vol. 12, no. 1, pp. 375–383, 2021.

[9]G. L. S. Babu and A. Srivastava, "Reliability analysis of allowable pressure on shallow foundation using response surface method," *Comput Geotech*, vol. 34, no. 3, pp. 187–194, 2007.

[10]M. A. Shahin, H. R. Maier, and M. B. Jaksa, "Neural and neurofuzzy techniques applied to modelling settlement of shallow foundations on granular soils," in *Proc., Int. Congress on Modeling and Simulation, MODSIM2003*, Citeseer, 2003, pp. 1886–1891.

[11]M. Hasanipanah, M. Noorian-Bidgoli, D. Jahed Armaghani, and H. Khamesi, "Feasibility of PSO-ANN model for predicting surface settlement caused by tunneling," *Eng Comput*, vol. 32, pp. 705–715, 2016.

[12]B. Tarawneh, "Predicting standard penetration test N-value from cone penetration test data using artificial neural networks," *Geoscience Frontiers*, vol. 8, no. 1, pp. 199–204, 2017.

[13]R. Sarkhani Benemaran, M. Esmaeili-Falak, and A. Javadi, "Predicting resilient modulus of flexible pavement foundation using extreme gradient boosting based optimised models," *International Journal of Pavement Engineering*, pp. 1–20, 2022.

[14]R. S. Benemaran and M. Esmaeili-Falak, "Predicting the Young's modulus of frozen sand using machine learning approaches: State-of-the-art review," *Geomechanics and Engineering*, vol. 34, no. 5, pp. 507–527, 2023.

[15]M. Esmaeili-Falak and R. S. Benemaran, "Ensemble deep learning-based models to predict the resilient modulus of modified base materials subjected to wet-dry cycles," *Geomechanics and Engineering*, vol. 32, no. 6, pp. 583–600, 2023.

[16]R. S. Benemaran, "Application of extreme gradient boosting method for evaluating the properties of episodic failure of borehole breakout," *Geoenergy Science and Engineering*, vol. 226, p. 211837, 2023.

[17]G. Zheng, W. Zhang, H. Zhou, and P. Yang, "Multivariate adaptive regression splines model for prediction of the liquefaction-induced settlement of shallow foundations," *Soil Dynamics and Earthquake Engineering*, vol. 132, p. 106097, 2020.

[18]S. K. Sasmal and R. N. Behera, "Prediction of combined static and cyclic load-induced settlement of shallow strip footing on granular soil using artificial neural network," *International Journal of Geotechnical Engineering*, 2018.

[19]M. Ahmad *et al.*, "Prediction of ultimate bearing capacity of shallow foundations on cohesionless soils: A gaussian process regression approach," *Applied Sciences*, vol. 11, no. 21, p. 10317, 2021.

[20]N.-V. Luat, K. Lee, and D.-K. Thai, "Application of artificial neural networks in settlement prediction of shallow foundations on sandy soils," *Geomechanics and Engineering*, vol. 20, no. 5, p. 385, 2020.

[21]M. Mohammed, A. Sharafati, N. Al-Ansari, and Z. M. Yaseen, "Shallow foundation settlement quantification: application of hybridized adaptive neuro-fuzzy inference system model," *Advances in Civil Engineering*, vol. 2020, pp. 1–14, 2020.

[22]T. Gnananandarao, R. K. Dutta, and V. N. Khatri, "Application of artificial neural network to predict the settlement of shallow foundations on cohesionless soils," in *Geotechnical Applications: IGC 2016 Volume 4*, Springer, 2019, pp. 51–58.

[23]T. Jithendra and S. Sharief Basha, "A Hybridized Machine Learning Approach for Predicting COVID-19 Using Adaptive Neuro-Fuzzy Inference System and Reptile Search Algorithm," *Diagnostics*, vol. 13, no. 9, p. 1641, 2023.

[24]T. Jithendra and S. S. Basha, "Analyzing groundwater level with hybrid ANN and ANFIS using metaheuristic optimization," *Earth Sci Inform*, pp. 1–31, 2023.

[25]R. M. Adnan *et al.*, "Application of novel binary optimized machine learning models for monthly streamflow prediction," *Appl Water Sci*, vol. 13, no. 5, p. 110, 2023.

[26]M. A. Shahin, *Use of artificial neural networks for predicting settlement of shallow foundations on cohesionless soils*. University of Adelaide, Department of Civil and Environmental Engineering ..., 2003.

[27]M. Aghayari Hir, M. Zaheri, and N. Rahimzadeh, “Prediction of rural travel demand by spatial regression and artificial neural network methods (Tabriz County),” *Journal of Transportation Research*, vol. 20, no. 4, pp. 367–386, 2023.
[28]C. Yang, H. Feng, and M. Esmaeili-Falak, “Predicting the compressive strength of modified recycled aggregate concrete,” *Structural Concrete*, vol. 23, no. 6, pp. 3696–3717, 2022.
[29]X. Shi, X. Yu, and M. Esmaeili-Falak, “Improved arithmetic optimization algorithm and its application to carbon fiber reinforced polymer-steel bond strength estimation,” *Compos Struct*, vol. 306, p. 116599, 2023.
[30]L. Abualigah, M. Abd Elaziz, P. Sumari, Z. W. Geem, and A. H. Gandomi, “Reptile Search Algorithm (RSA): A nature-inspired meta-heuristic optimizer,” *Expert Syst Appl*, vol. 191, p. 116158, 2022.
[31]M. Awad, R. Khanna, M. Awad, and R. Khanna, “Support vector regression,” *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*, pp. 67–80, 2015.
[32]J. S. Tang, “ANFIS: Adaptive network based fuzzy inference systems,” *IEEE Trans. Syst. Cybern*, vol. 23, pp. 515–520, 1993.
[33]W. Wang, W. Zhang, and Z. Zhang, “A hybrid estimation procedure for modeling shallow foundation’s settlement: RBF-optimized neural network,” *Journal of Intelligent & Fuzzy Systems*, no. Preprint, pp. 1–10.