



Using Novel Optimization Algorithms with Support Vector Regression to Estimate Pile Settlement Rates

Lu Sun^{1,2} · Tinghui Li³

Received: 21 December 2022 / Accepted: 4 February 2024
© The Author(s), under exclusive licence to Indian Geotechnical Society 2024

Abstract To ensure structure safety, such as bridge structures, artificial intelligence ways are the most helpful way to predict failure factors such as pile settlement estimation. Various methods are taken into account for evaluating the movement of the piles, which helps to realize the future view of the project during the loading period. The most intelligent mathematical strategy to calculate the movement of the pile is applied. In this regard, support vector regression (SVR), a machine learning method, was used in this study, accompanying two optimizers to accurately determine the key SVR variables. The marine predator algorithm (MPA) and the grasshopper optimization algorithm (GOA) were combined with SVR to create the SVR-MPA and SVR-GOA frameworks. Additionally, several metrics were used to evaluate the model's overall performance. The R^2 for SVR-MPA in the training phase was found to be 0.997, showing a desirable modelling operation. At the same time, the RMSE was calculated 0.2843 mm and compared to the SVR-GOA, the differences are 1.16 and 106.38%, respectively, in favour of the former model. The comprehensive index of OBJ, including RMSE, MAE, and R^2 , was calculated at 0.2791 and 0.581 mm for models optimized by MPA and GOA, alternatively.

Keywords Pile · Structure settlement · Grasshopper optimization algorithm · Machine learning · Support vector regression · Marine predator algorithm

Introduction

Many methods evaluate the pile settlement (PS) from empirical methods through experiments in an efficient way to generate the perspectives before operating projects, especially the bridged ones. Paying attention to assessing sensitive constructions dealing with people's lives and considering intelligent solutions can help us find their actual conditions and calculate with low energy, time, and cost. Using computer-aid methods that are underlaid by the experimental data can give us a freer domain in terms of cost and time to realize the functioning of structures over the utilization period [1]. Several researchers have studied pile motion using the popular finite difference and finite element methods [2–6]. The key role of estimating the PS is played by the inputs of a model that wants to find the logical relationships between initial entering variables [7, 8].

However, testing on large models or full-scale experiments is required to adequately cover all aspects of pile-moving behaviour. Because these large-scale experimental researches are difficult to perform and expensive, many studies in the literature contain various simplified and numerical methods for predicting pile behaviour. Many experts have suggested a simplified method for predicting the initial pile motion responding to load settlement, given the elastic behaviour of the soil. In some simplified methods, piles are considered a means of reducing settlement [9], where the compacting ground soil layers support the pile structure's bearing capacity, which is why the piles contribute to controlling the settlement rates. In another simplified

✉ Lu Sun
Seliang_si@126.com

¹ School of Architecture and Engineering, Huzhou Vocational and Technical College, Huzhou 313000, Zhejiang, China

² Huzhou Key Laboratory of Green Building Technology, Huzhou 313000, Zhejiang, China

³ Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, Heilongjiang, China

method, the load is distributed between the piles and foundation regarding interactions between soil layers and pile [10]. Most simplified solutions do not adequately account for soil–structural interactions and cannot accurately predict the pile response to load settling.

In this regard, one article suggested a way to analyse the displacement of piles and presented the theoretical function for the examinations on earth pressure [11, 12]. Most experts assessed the PS, but all used the model that lacks the reflection of ground. However, artificial neural networks (ANNs) and techniques of machine learning have been operated in many types of research Lee and Lee [13], Hanna et al. [14], Liu et al. [15], Che et al. [16], and Shanbeh et al. [16]. Calibrating data for some research were used, and PS tests were done to generate the framework predicting the final capacity of the pile's bearing. Samples from the in situ piles are selected to calibrate artificial intelligent (AI) developed models. The research used the technique of ANN to appraise the PS rates from some useful traits of piles socketed in rocks [17].

Several studies have broadly used methods related to functioning with regressions [18–26] regarding using regression methods such as stochastic regression machine of min-max, adaptive regression of multivariate spline, and Gaussian trend regression [18, 23, 27, 28]. Also, many researchers use the powerful solution of support vector machines to compute the experimental parameters [29, 30]. The ability of this machine learning technique to estimate pile movements can be found in many studies with acceptable correlation results of more than 90 per cent [31, 32]. Further, one article evaluated the ultimate bearing capacity of piles using this method with desirable results [33, 34]. With this respect, the input dataset has properties of soil observed in a real field of projects, as the pile specimens should be provided.

Therefore, this study aims at estimating the PS factor as the key dependent variable in relation to many parameters of pile physical properties and the ground characteristics for the rocky region. For this reason, the support vector regression (SVR) as the means of model tries to indicate a perspective of pile movement over the operation period. As the novelty of the research and, secondly, estimating PS accurately, SVR is linked with the optimization algorithms to develop the SVR in the form of optimized frameworks with high accuracy besides the low-complex calculation net to compute PS in desirable conditions based on in situ measurements. SVR was coupled with marine predator algorithm (MPA) and the grasshopper optimization algorithm (GOA) to reach the goals. This study's pile profiles and soil features data were collected from the Klang Valley Mass Rapid Transit (KVMRT) network in Kuala Lumpur, Malaysia.

Modifying the SVR-MPA and SVR-GOA models feeding data of KVMRT piles are used by the ratio of 70 for calibration and 30 per cent for validation, respectively. In this

regard, the ratio of pile length to diameter, the rock hardness level showing with UCS, load masses effect over the pile, the standard penetration test (NSPT) results for rocks, the pile length beneath the soil, and the length of the pile in the rock was fed to models to estimate the PS for the megaproject of KVMRT [35]. Also, the indices of *R*, OBJ, VAF, MAE, and RMSE were opted to evaluate the impact of each proposed optimizer in fining pile settlement rates.

Materials and Methodology

Case Study: Klang Valley Rapid Transit System (KVMRT)

Kuala Lumpur (Malaysia) has confronted many congestion issues as a populated city. Several projects have been done to solve these and are on the table to check the feasibility of executions. The Klang Valley Mass Rapid Transit (KVMRT) project was designed and constructed to reduce the traffic masses. The project of KVMRT was built to pass Kuala Lumpur's Federal Territory and Selangor State, especially joining areas within the region of Klang Valley. The 51 km of KVMRT lines contain 35 stations in which the underground and surface constructions are involved. The total length of the underground tunnel is about 9.5 km. This mega project of KVMRT in Kuala Lumpur to decrease the traffic jam issues, including many piles to supporting bridges, has been chosen as the case study. Figure 1 shows the KVMRT's position in Kuala Lumpur.

Initial Data Set of Case Study

The 96 piles' information in the KVMRT project were gathered into rocks like limestone, phyllite, sandstone, and granite. The San-trias class granite rocks were used for the KVMRT project piles. Materials and subsoil datasets were collected to reach the information on geological profiles. Based on the assessments done, the ground comprises residual rock materials. According to the dataset collected, the bedrock is placed between the depth of 70 cm underground and more than 1400 m.

The methodology for collecting field data sets for materials and subsoil databases in infrastructure projects typically involves a combination of geotechnical investigation techniques, such as drilling, sampling, and testing. The collected data includes information on the type and properties of soil, rock, and other subsurface materials, as well as their depth and location. However, the following lines provide information on the field sampling and bore log data collected:

- The observed rock masses range from moderately to largely weathered.



Fig. 1 The location of the study area KVMRT

- The highest and lowest values for UCS, as per ISRM, are 68 and 25 MPa, respectively [36].
- Bore log data taken to a maximum depth of 16.5 m shows the presence of highly weathered soil. The primary soil type is a hard sandy mud with a minimum and maximum N_{SPT} of 4 and 167 blows per 300 mm, respectively.
- Subsoil materials with N_{SPT} values greater than 50 blows per 300 mm are observed in most areas at ground depth levels between 7.5 and 27.0 m.

Based on the gathered data, the recorded rocks' masses are weathered as moderate to extreme. However, the UCS parameter for the lowest and highest values of rocks was registered, respectively, from 25 to 68 MPa based on the parameter of ISRM [37]. Further, the parameter of N_{SPT} was reported for at least 4 to a high value of 167, respectively, per 30 cm. That the rock types vary, going from 750 to 2700 cm. Moreover, according to the registered data of bores under the depth of 16.5 m, the soil type is extremely weathered, as well as the soil type in, which is highly muddy with sand.

Making the leading dataset with effective impacts was the primary step to creating the predictive frameworks. It is fundamental to show the foremost vital variables processing the initial data of the model. The mentioned

tests were performed utilizing analysis settings by Pile Dynamic, Inc. For this respect, several variables opted to investigate the outcomes of pile geometry: (1) Pile length to the diameter fraction (L_p/D); (2) pile length in the soil to the pile length in the rock layer ratio (L_s/L_r); (3) N_{SPT} of rocks; (4) Q_u , that is ultimate potential bearing (directly affects the pile motion rate); (5) UCS of rock. These parameters feeding models were used to simulate the PS values in this regard. Table 1 indicates the summary of inputs used for the developed models.

To estimate the settlement rate of a pile, the parameters that influence it are the length and diameter of the pile. Two specific parameters, the ratio of the pile length in the soil layer to the pile length in the rock layer (L_s/L_r) and the ratio of the total pile length to the pile diameter (L_p/D), are analysed to determine the impact of pile geometry. Additionally, the model input for predicting pile subsidence is the UCS, which has a significant effect. The value of N_{SPT} is also considered as input to indicate the condition of the soil layer. The pile load is another important factor that directly affects the settlement, so the final pile-bearing capacity (Q_u) is included as input as well. In summary, five variables are selected as inputs to evaluate the pile settlement (SP).

Table 1 The feeding dataset for developed models

Parameters	Symbols	Unit	Max	Min	S. deviation	Average
Fraction of pile length to diameter	L_p/D	–	31.56	4.33	6.55	15.37
Settlement of pile	PS	mm	20.095	4.494	3.690	10.99
Uniaxial compressive strength	UCS	MPa	68.489	25.324	12.442	43.411
Standard penetration test	N	–	166.42	2.92	59.08	80.03
Ultimate potential bearing	Q_u	KN	42,701	12,409	803	2454
Soil length to socket length ratio	L_s/L_t	–	31.714	0.286	6.562	7.063

Also, Fig. 2 indicates measured inputs and target values (PSs) with a diagram in which each string shows one pile sample that, based on relevant PS, has a specific colour.

Support Vector Regression, SVR

The machine learning way of support vector regression (SVR) was designed to compute the regression boundaries [30]. SVR is operated for sorting regressions, in which the error range (ϵ) is considered to determine regressions. Categorizing classes of regressions can be operated for creating hyperplane boundaries. The SVR learning technique worked as the supervised and found the regressions explained in Eq. (1) [38].

$$\min_{w,b} = \frac{1}{2}w^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) \tag{1}$$

$$\text{s.t.} \left\{ \begin{array}{l} y_i - (w^T x_i + b) \leq \epsilon + \xi_i \\ (w^T x_i + b) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{array} \right\}$$

where the variable of ξ shows the boundary violations; w represents the weight of inputs; C denotes modifying a variable in order; b is the bias rate and denotes the rate of deviation from the hyperplane boundaries is shown with ϵ .

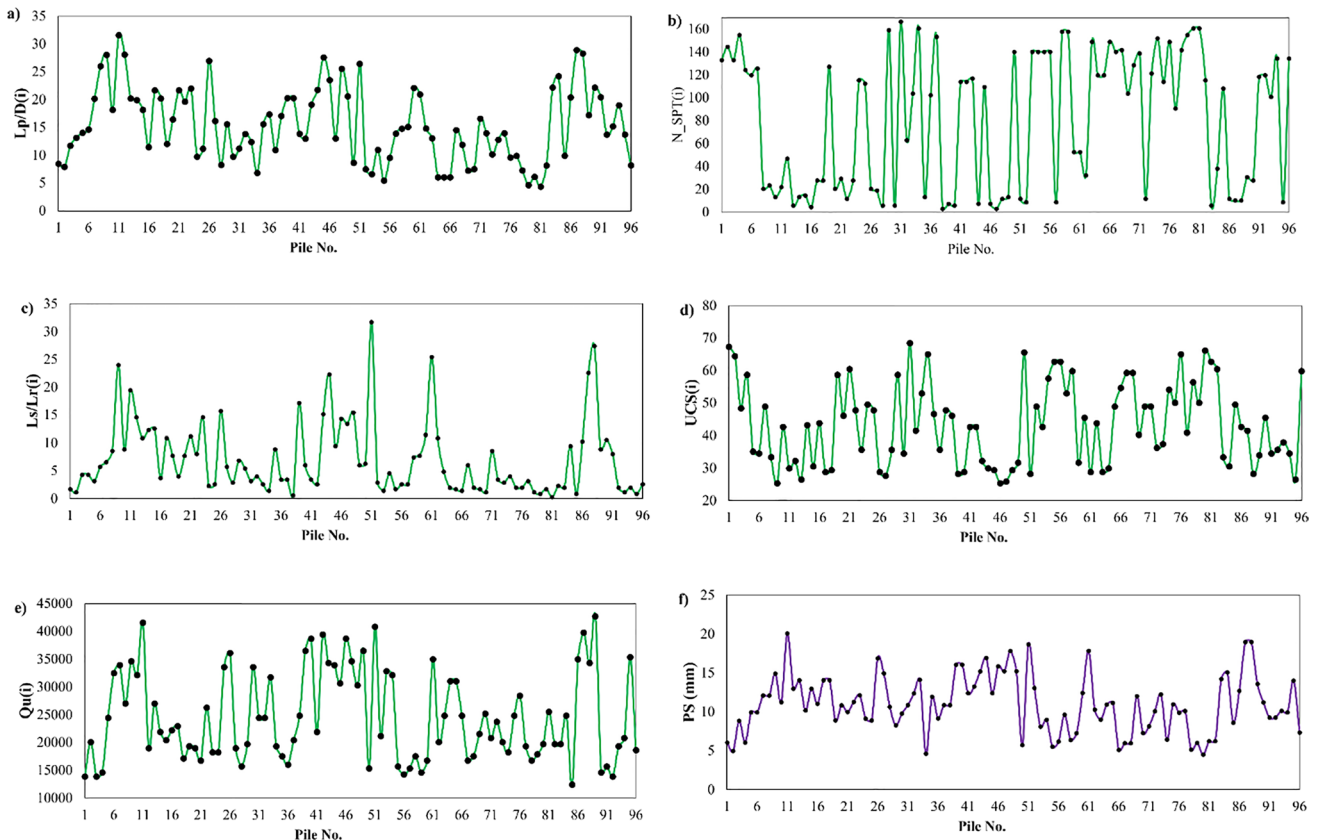


Fig. 2 The input and target value diagram: **a** L_p/D , **b** L_s/L_r , **c** N_{SPT} , **d** UCS, and **e** Q_u , **f** PS

The fitness function has 2 terms that are explained as follows:

$$\frac{1}{2} \|w\|^2 \tag{2}$$

$$C \sum_{i=1}^m (\xi_i + \xi_i^*) \tag{3}$$

Equation (2) was proposed for enhancing the area among the samples and boundaries of the hyperplane, afterwards preserving the intervals between the inputs with the boundaries. Equation (3) works like the modifying tool. The weight and bias were generated over solving the objective function as the target of the hyperplane’s boundaries [24]. This article uses the quadratic objective function to do the dedicated tasks accurately. The crucial SVR duty has been to figure out the key components as optimal magnitudes. To achieve the

Fig. 3 The flowchart of the marine predator algorithm (MPA)

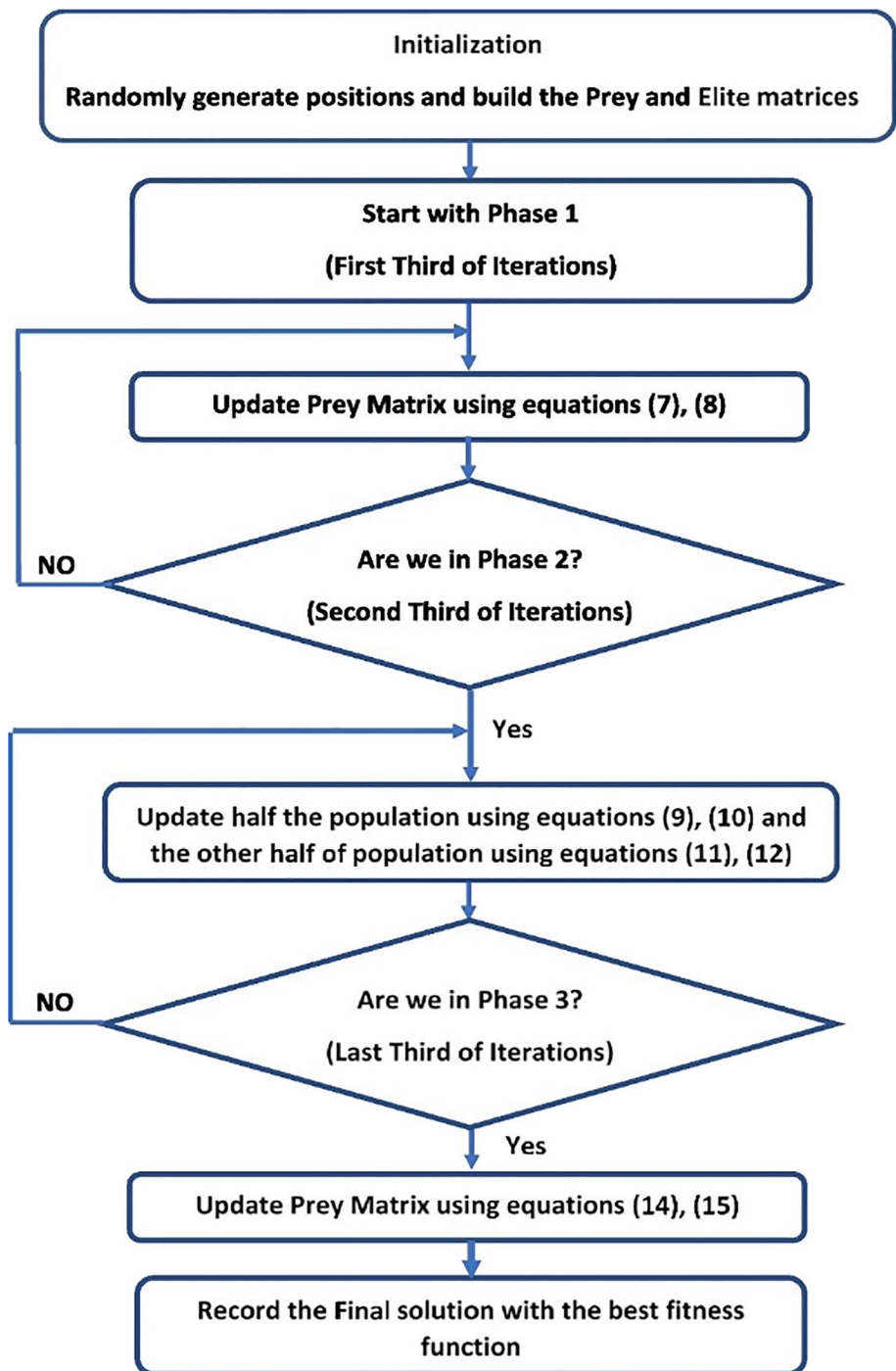


Table 2 Evaluating indices for analysing the proposed frameworks

Criteria	Nomenclatures	Equation	Assessment
Variance account factor	VAF	$\left(1 - \frac{\text{var}(t_n - y_n)}{\text{var}(t_n)}\right) * 100$ (18)	Higher is desirable
Mean absolute error	MAE	$\frac{1}{N} \sum_{n=1}^N p_n - t_n $ (19)	Lower is desirable
Root-mean-squared error	RMSE	$\sqrt{\frac{1}{N} \sum_{n=1}^N (p_n - t_n)^2}$ (20)	Lower is desirable
Pearson’s correlation coefficient	R^2	$\left(\frac{\sum_{n=1}^N (t_n - \bar{t})(p_n - \bar{p})}{\sqrt{[\sum_{n=1}^N (t_n - \bar{t})^2][\sum_{n=1}^N (p_n - \bar{p})^2]}}\right)^2$ (21)	Higher is desirable
Statistical parameters, including the various error indices	OBJ	$\left(\frac{n_{\text{train}} - n_{\text{test}}}{n_{\text{train}} + n_{\text{test}}}\right) \frac{\text{RMSE}_{\text{train}} + \text{MAE}_{\text{test}}}{R^2_{\text{train}} + 1} + \left(\frac{2n_{\text{train}}}{n_{\text{train}} + n_{\text{test}}}\right) \frac{\text{RMSE}_{\text{test}} - \text{MAE}_{\text{test}}}{R^2_{\text{test}} + 1}$ (22)	Lower is desirable [42]

key factors mentioned, such as σ , C , and ϵ needs a smart algorithm, the optimizing strategies of MPA and GOA were used to join with the SVR to appraise the above-mentioned factors at optimal rates.

Grasshopper Optimization Algorithm (GOA)

The grasshopper optimization algorithm (GOA), known as the colony base algorithm, can simulate the grasshoppers’ behaviour as insects to locate the best solution [39]. These creatures in adulthood reveal their colony base habits in going a long distance having traits of longer ranges and sudden motions. Equation (4) reveals the mathematical form of the habits.

$$x_i = S_i + G_i + A_i \tag{4}$$

In Eq. (4), the location of the grasshopper i is defined by x_i and the social interaction of grasshoppers is shown by S_i .

$$S_i = \sum_{\substack{j=1 \\ i \neq j}}^N s(d_{ij}) \hat{d}_{ij}, \quad d_{ij} = |x_i - x_j|, \quad \hat{d}_{ij} = \frac{x_i - x_j}{d_{ij}} \tag{5}$$

In Eq. (5), d_{ij} represents the distance among grasshoppers of i and j ; also \hat{d}_{ij} reveals the vector as being a unit among i th and j th grasshoppers. Equation (6) considers the strength of the social force (s).

$$s(x) = fe^{\frac{-x}{l}} - e^{-x} \tag{6}$$

where the parameters of f and l refer to the intensity of attraction and length, alternatively. Notably, the Nymph types of grasshoppers do not have wings, which leads to the wind being an important factor in moving direction.

$$A_i = u\hat{e}_w, \quad G_i = -g\hat{e}_g \tag{7}$$

Fig. 4 Pile settlement rates measured in project KVMRT

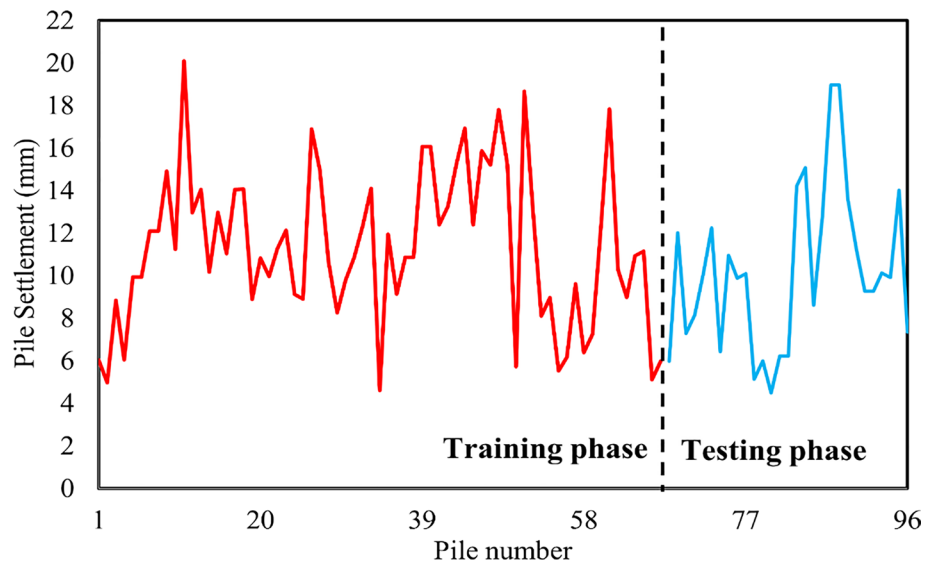


Table 3 The key factors of variables magnitudes of each optimizer

		SVR-MPA	SVR-GOA
Training phase	C	0.300	0.650
	EPSILON	751.634	653.142
	sigma	66.896	64.236
Testing phase	C	0.300	0.650
	EPSILON	1146	389.640
	sigma	9.278	9.487

where Eq. (7), the parameters of gravity force and wind advection of grasshoppers are mentioned with G and A . The variables of u and g are known as, respectively, the wind drifts and the gravity constants. Moreover, the unit vector specifying the wind advection directions and the gravity force directions are represented via \hat{e}_w and \hat{e}_g . Consequently, by contributing Eqs. (5)–(7), grasshoppers’ behaviours can be formed with the below relation.

$$X_i = \sum_{\substack{j=1 \\ i \neq j}}^N s(|x_j - x_i|) \frac{x_j - x_i}{d_{ij}} - g\hat{e}_d + u\hat{e}_w \tag{8}$$

Nonetheless, utilizing the mathematical model to tackle optimization problems is unfeasible as the grasshoppers promptly attain their comfort zone, and the swarm fails to converge towards a predetermined point. To address optimization issues, a revised version of Eq. (8) is suggested as follows according to [39]:

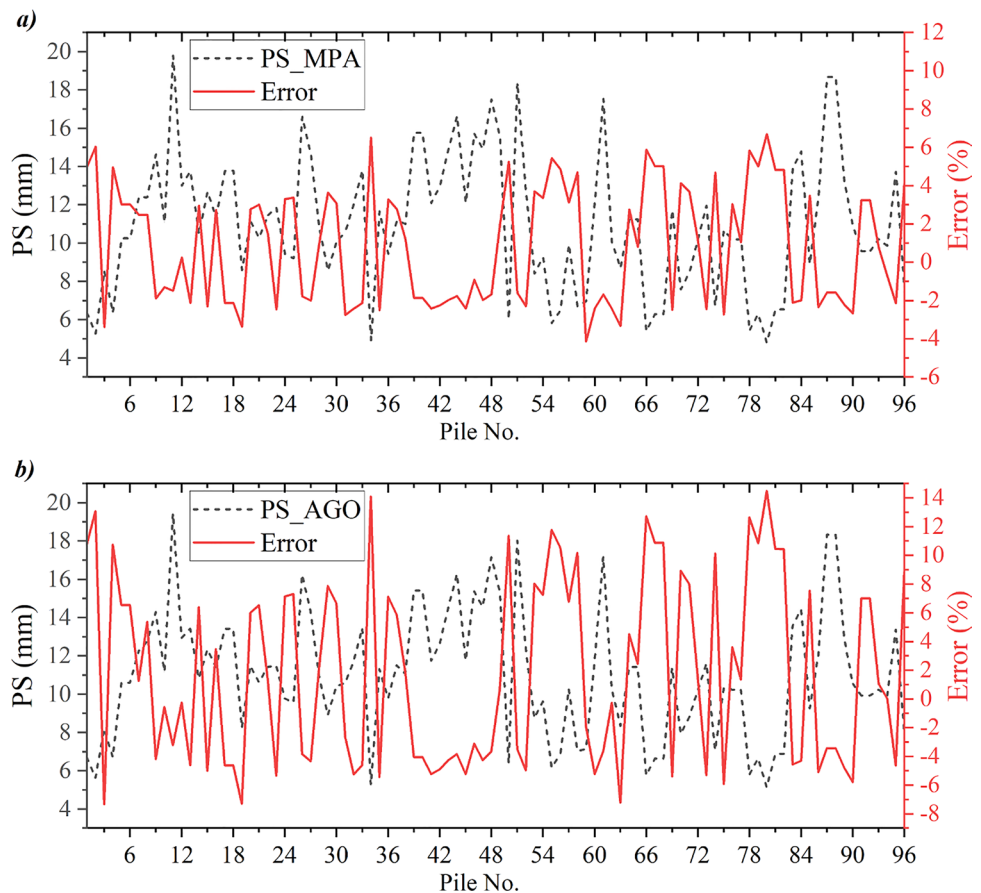
$$X_j^d = c \left[\sum_{\substack{j=1 \\ i \neq j}}^N c \frac{ub_d - lb_d}{2} s(|x_i^d - x_j^d|) \frac{x_i - x_j}{d_{ij}} \right] + \hat{D}_d \tag{9}$$

where the parameters of the upper and lower boundary are indicated via ub_d and lb_d ; the d th dimensions values are represented by \hat{D}_d ; the population number is indicated via N ; c is a reducing coefficient that improves the exploitation as the iteration number goes up, and increasing the iteration leads to a balance of exploration and exploitation.

$$c = c_{max} - Iter \times \frac{c_{max} - c_{min}}{M.Iter} \tag{10}$$

where for the current article, the values of c_{max} and c_{min} , respectively, are determined to be 1 and 0.0001.

Fig. 5 The errors and PS rates modelled by: **a** SVR-MPA and **b** SVR-GOA



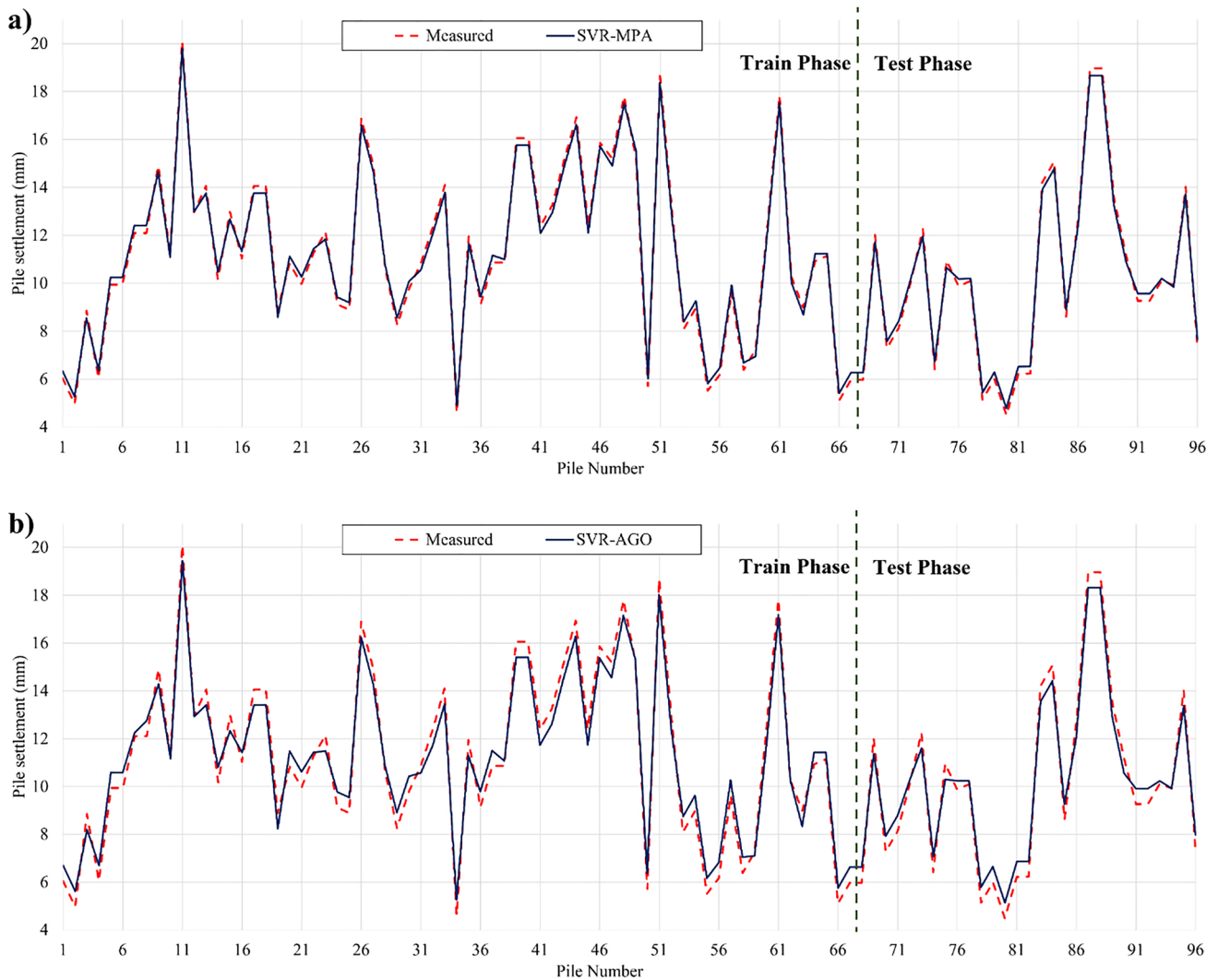


Fig. 6 The PS target values and modelled by: **a** SVR-MPA and **b** SVR-GOA

Marine Predator Algorithm (MPA)

The novel metaheuristic algorithm of the natural marine predators algorithm (MPA) was developed by Faramarzi et al. [40]. In the marine predators’ interactions with prey, predators employ a wide range of foraging strategies called MPA-inspired Brown and Levy accidental migration. In hunting grounds, if the focus of the prey is higher, the predators will use the Brown way, and if the prey is little, the predator will use the Levy method. Levi’s movements include jumps and fast motions that enhance the exploration speed. Brown motion involves fixed steps in the same task to optimize the work process. On the other hand, environmental issues such as fish aggregating devices (FADs) and eddy formations are among the factors changing the

behaviour of predators. Figure 3 shows the process of optimization by MPA [41].

The MPA main stages are explained as follows [6]:

When prey moves with the Brown movement, upgrades the matrices of prey with the following relations:

$$\overline{step}_j = \overline{R}_L \otimes \left[\overline{elite}_j - \left(\overline{R}_L \otimes \overline{prey}_j \right) \right] \tag{11}$$

$$\overline{prey}_j = \overline{prey}_j + \left(P \cdot \overline{R} \otimes \overline{step}_j \right) \tag{12}$$

The variable of R_L denotes the vectors containing changing numbers based on the movement of levy type. Another population society can be upgraded as:

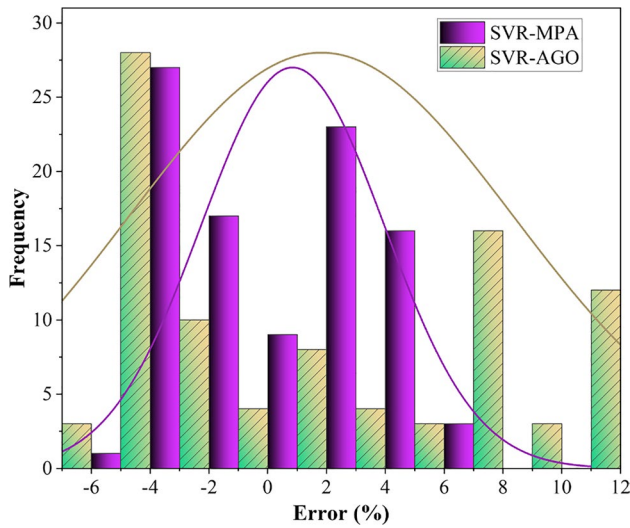


Fig. 7 The error distribution of models with the normal distribution curve

$$\overline{\text{step}}_j = \overline{R}_B \otimes (\overline{R}_B \otimes \overline{\text{elite}}_j) - \overline{\text{prey}}_j \tag{13}$$

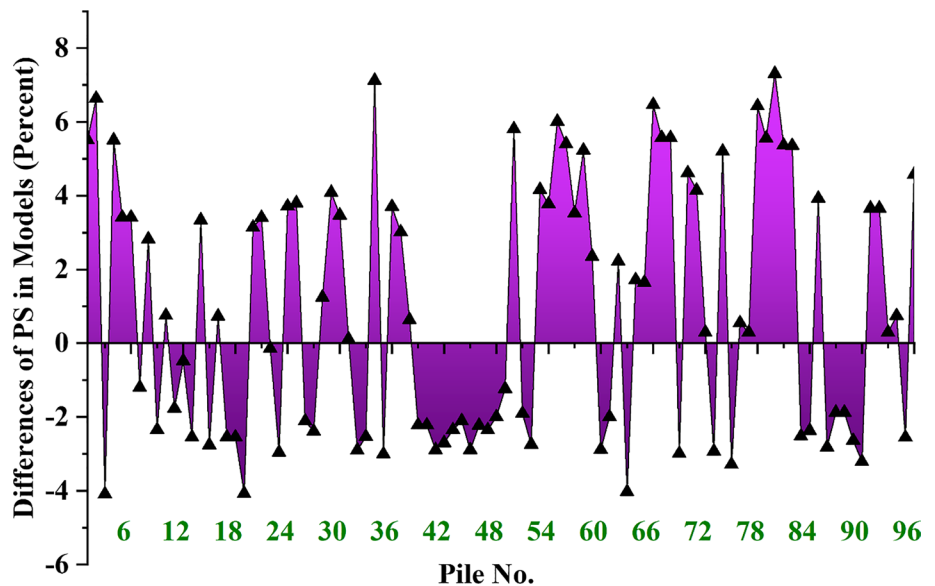
$$\overline{\text{prey}}_j = \overline{\text{elite}}_j + (P \cdot cf \otimes \overline{\text{step}}_j) \tag{14}$$

Afterwards, the elite matrix is equal to multiplying by *cf*. That the *cf* is defined as follows:

$$cf = [1 - (\text{iter}/\text{max} \times \text{iter})]^{(2 \cdot \text{iter}/\text{max} \times \text{iter})} \tag{15}$$

The predator moves using the movement type of Levy, and then the matrix of Preys can be upgraded via the relations (16) and (17):

Fig. 8 The difference in PS modelled using SVR-MPA and SVR-GOA



$$\overline{\text{step}}_j = \overline{R}_L \otimes (\overline{R}_L \otimes \overline{\text{elite}}_j) - \overline{\text{prey}}_j \tag{16}$$

$$\overline{\text{prey}}_j = \overline{\text{elite}}_j + (P \cdot cf \otimes \overline{\text{step}}_j) \tag{17}$$

After either of the iterations, the elite matrix would be upgraded, accompanied by the best answers, and the ultimate answers are presented after the last iteration.

Evaluation of Models: SVR-MPA and SVR-GOA

Five indicators were used to evaluate the accuracy of models SVR-GOA and SVR-MPA to estimate the pile settlement (PS) ranges in the calibration and validation phases, as revealed in Table 2.

In relations (18)–(22), the predicted subsidence of piles is indicated via p_N ; target value of measurements is indicated by t_n ; \bar{t} is showing the averaged pile settlement that is observed; the calculated PSs are indicated using \bar{p} as the average value. Afterwards, alternatively, the number of samples gathered for the train and test phase is indicated by the *ntrain* and *nstest* variables.

Results and Discussion

SVR-MPA and SVR-GOA are hybrid algorithms combining SVR with a population-based metaheuristic optimization algorithm. While SVR-MPA mimics the hunting behaviour of marine predators, SVR-GOA simulates the swarming behaviour of grasshoppers. These algorithms have shown promising results in improving the performance of SVR in estimating pile settlement.

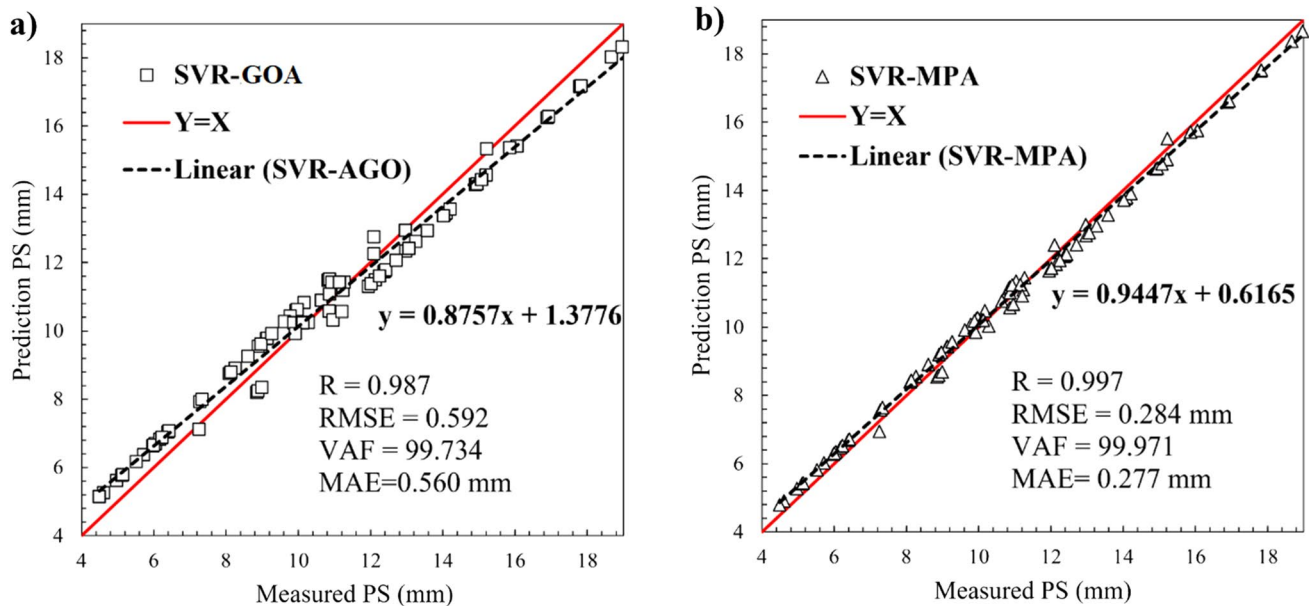


Fig. 9 Modelled PS with the measured values for **a** SVR-GOA and **b** SVR-MPA

The results of the PS model were generated via the SVR-MPA and SVR-GOA frameworks, which have used the machine learning technique to estimate the rates of pile settlements. Thus, we considered the modelling process to have the cost and complexity at the lowest level accompanied by the highest accuracy of PS estimation; this event is practical when using the optimization algorithms to modify SVR in solving kernels close to target values. Mathematics operations and models were performed in the MATLAB environment. Measured pile settlement rates for the project KVMRT as the case study are indicated using Fig. 4, in which 70% of data are considered for the training phase and 30% for the testing phase. As shown in Fig. 4, the placement of the PS rate for both testing and training phases ranges in various scales of values.

Table 3 shows the key variables of SVR optimized with MPA and GOA.

The error rates entered in the modelling process are shown in Fig. 5 for both SVR-MPA (a) and SVR-GOA (b). In this regard, the SVR-MPA (a) modelled pile settlement rates with lower errors than SVR-GOA (b). The former model has an error range of -4 to $+7\%$ of the suitable condition, while on the other hand, SVR-GOA (b) has an error range of -7 to $+15\%$. Extending the error range of 75% and more than 100% for mentioned lowest and highest bounds can signify the MPA optimization algorithm's capability to remove the error rates rather than the GOA method. However, the overestimated and underestimated PS values can be found in both the validation and calibration stages. For

SVR-GOA (b), the 80th pile is calculated with the highest error rate and for another model.

Focusing on the error, matters are discussed, and showing the target values with the modelling PS values can help us understand the models' performance to appraise the PS magnitudes. For this respect, Fig. 6 indicates the modelled PS magnitudes with the measured ones as the target values. Regarding the objective rates of PS based on Fig. 6, SVR-MPA (a) shows a better coincidence for a continuous blue solid line with a red dashed line representing the measured PSs. However, this status is conducted for two phases of training and testing. On the other side, SVR-GOA (b) could not do modelling operations as well as the others. According to Fig. 6 (b), overlapping the two lines mentioned above is not fitted as happened for SVR-MPA. Actually, in the calibrating phase, the GOA optimizer made the highest mistake computing the PS rates, especially for the piles of 78–82, with approximately 12% error rates. In comparison, this rate for the model with MPA optimizer was 5%. Therefore, the difference of 140% between the two models shows the goodness of the MPA algorithm.

As seen in the presented figures, the modelling process is so important that in this part, Fig. 7 tries to show the error distribution based on the relevant frequency. The condition of error distribution for both models does not obey a typical rule to create a tall bar showing the concentration of errors around a certain number, especially zero. In this regard, both optimization algorithms have

the most error accumulation, around -4% , with about 27 counts as the frequency. This event has led to creating the short-height bars for either proposed framework.

For comparing the differences of models to estimate the pile settlement rates, Fig. 8 is attempted to show the discrepancy percentage in simulations. Regarding Fig. 8, most of the piles experienced high rates of differences in modelling PS that reached 7 per cent for the pile number of 34. However, in some cases in which the PS estimated by two models are close to each other, reducing the differences near the zero line, such as 10, 12, 16, and 22, totally the average amount of differences reaches 3 per cent.

Figure 9 shows the best-fit line between the measured PS (horizontal axis) against the modelled one (vertical axis). In the light of the lower error rate for SVR-MPA (b), the best-fit line for the PS points is closer to the bisector line, rather than SVR-GOA with more error rates. Moreover, the correlation index of R_2 shows that the MPA algorithm has done its job better than the GOA optimizer. This fact results from the error rates shown via MAE and RMSE that are high for SVR-GOA, respectively, with 102 and 108% differences compared to SVR-MPA.

From the OBJ indicator viewpoint, which encompasses the error indexes of RMSE, MAE, and correlation factor of R^2 , the SVR-GOA framework (a) was rated for 0.581, about 108 per cent higher than the SVR-MPA with a value of 0.279.

Conclusion

Assessing the immunization of constructions such as bridges and buildings over the operation stage is paramount for experts and engineers. Considering essential factors such as the rate of piles' settlement (PS) as a vital issue in projects should be investigated before operating them. Various items have to be used to examine the pile movement, and these measures can help us realize the perspective of projects over the operation of the project. In this regard, several intelligent solutions are designed to calculate the rates of pile motions. Thus, the current research used machine learning SVR to model the magnitude of PSs for the constructed project of KVMRT in Malaysia. Two optimization algorithms, the marine predator algorithm (MPA) and the grasshopper optimization algorithm (GOA), were utilized to accurately determine SVR's main variables. The optimizers were combined with SVR to create SVR-MPA and SVR-GOA frameworks. To examine the capability of developed models, five metrics were considered apprising the PS

magnitudes. In this regard, the R^2 correlation index of SVR-MPA in the training phase was calculated at 0.996 at the desired level and 0.985 for SVR-GOA. These values for the testing phase were calculated with a 0.72% difference in favour of SVR-MPA, obtained at 0.998, and for SVR-GOA, 0.991. In the training stage, the RMSE index was calculated for SVR-MPA as 0.286 mm, while SVR-GOA was 0.590 mm with 106.38%. This index for validation for SVR-MPA was calculated at 0.281 mm and for SVR-GOA at 0.596 mm with a 112.5% difference. In addition, using the OBJ that includes the error criteria in both training and testing phases, the SVR-MPA was rated at 0.279 and the SVR-GOA 0.581. The condition of error distribution for both models did not obey a typical rule showing the concentration of errors around a certain number, especially zero. In this regard, both optimization algorithms had the most error accumulation, around the -4% , with about 27 counts as the frequency. In summary, the SVR-MPA model could remove the most errors seen in SVR-GOA through a better simulation process. Existing the error range of -7.34 to 14.48% entered by GOA compared to the MPA with the errors of -4.14 to 6.68% implies the latter model's accuracy and higher capability to reduce error rates.

Considering the project directly related to people's lives earns high attention. However, covering all aspects of pile settlement behaviour on a large and full scale is difficult and requires a large cost range regarding energy, time, and finance. To cease these costs and reach an adequate estimation of pile settlement, introduced models in the present study proved an acceptable convergence rate of estimation which defines their workability and performance in estimating pile settlement behaviour without mentioned costs and difficulty. Therefore, it is suggestible that these models and specially SVR-MPA, in practical situations and projects.

Funding Special Project for High-level Talents of Huzhou Vocational and Technical College (2020GY22). Higher Education Teaching Reform Project of Heilongjiang Province (SJGZ20200101). Basic Public Welfare Research Project of Zhejiang Province (LGF22D020002).

Declaration

Conflicts of interest The authors declare no competing interests.

References

1. Poulos HG (2006) Pile group settlement estimation—research to practice. In: Foundation analysis and design: innovative

- methods, American Society of Civil Engineers, Reston, VA, pp 1–22. [https://doi.org/10.1061/40865\(197\)1](https://doi.org/10.1061/40865(197)1)
2. Samui P, Kim D (2013) Least square support vector machine and multivariate adaptive regression spline for modeling lateral load capacity of piles. *Neural Comput Appl* 23:1123–1127. <https://doi.org/10.1007/s00521-012-1043-x>
 3. Loganathan N, Poulos HG, Xu KJ (2001) Ground and pile-group responses due to tunnelling. *Soils Found* 41:57–67. <https://doi.org/10.3208/sandf.41.57>
 4. Dias TGS, Bezuijen A (2015) Data analysis of pile tunnel interaction. *J Geotech Geoenviron Eng*. [https://doi.org/10.1061/\(ASCE\)GT.1943-5606.0001350](https://doi.org/10.1061/(ASCE)GT.1943-5606.0001350)
 5. Momeni E, Dowlatshahi MB, Omidinasab F, Maizir H, Armaghani DJ (2020) Gaussian process regression technique to estimate the pile bearing capacity. *Arab J Sci Eng* 45:8255–8267. <https://doi.org/10.1007/s13369-020-04683-4>
 6. Yakout AH, Attia MA, Kotb H (2021) Marine predator algorithm based cascaded PIDA load frequency controller for electric power systems with wave energy conversion systems. *Alexandria Eng J* 60:4213–4222
 7. Poulos HG (1989) Pile behaviour—theory and application. *Geotechnique* 39:365–415
 8. Randolph MF (2003) Science and empiricism in pile foundation design. *Géotechnique* 53:847–875
 9. Burland JB, Hancock RJR, Burland J (1977) Underground car park at the house of commons, Geotechnical aspects, Building research establishment, London
 10. Leong EC, Randolph MF (1994) Finite element modelling of rock-socketed piles. *Int J Numer Anal Methods Geomech* 18:25–47. <https://doi.org/10.1002/nag.1610180103>
 11. Zhang Y, Hu X, Tannant DD, Zhang G, Tan F (2018) Field monitoring and deformation characteristics of a landslide with piles in the Three Gorges Reservoir area. *Landslides* 15:581–592. <https://doi.org/10.1007/s10346-018-0945-9>
 12. Zhang Y, Richardson DC, Barnouin OS, Michel P, Schwartz SR, Ballouz R-L (2018) Rotational failure of rubble-pile bodies: influences of shear and cohesive strengths. *Astrophys J* 857:15
 13. Lee I-M, Lee J-H (1996) Prediction of pile bearing capacity using artificial neural networks. *Comput Geotech* 18:189–200. [https://doi.org/10.1016/0266-352X\(95\)00027-8](https://doi.org/10.1016/0266-352X(95)00027-8)
 14. Hanna AM, Morcous G, Helmy M (2004) Efficiency of pile groups installed in cohesionless soil using artificial neural networks. *Can Geotech J* 41:1241–1249. <https://doi.org/10.1139/t04-050>
 15. Liu H, Li TJ, Zhang YF (1997) The application of artificial neural networks in estimating the pile bearing capacity
 16. Che WF, Lok TMH, Tam SC, Novais-Ferreira H (2003) Axial capacity prediction for driven piles at Macao using artificial neural network
 17. Goh ATC (1996) Pile Driving Records Reanalyzed Using Neural Networks. *J Geotech Eng* 122:492–495. [https://doi.org/10.1061/\(ASCE\)0733-9410\(1996\)122:6\(492\)](https://doi.org/10.1061/(ASCE)0733-9410(1996)122:6(492))
 18. Le TT, Le MV (2021) Development of user-friendly kernel-based Gaussian process regression model for prediction of load-bearing capacity of square concrete-filled steel tubular members. *Mater Struct* 54:59. <https://doi.org/10.1617/s11527-021-01646-5>
 19. Samui P (2011) Prediction of pile bearing capacity using support vector machine. *Int J Geotech Eng* 5:95–102. <https://doi.org/10.3328/IJGE.2011.05.01.95-102>
 20. Samui P (2012) Application of relevance vector machine for prediction of ultimate capacity of driven piles in cohesionless soils. *Geotech Geol Eng* 30:1261–1270. <https://doi.org/10.1007/s10706-012-9539-9>
 21. Samui P (2012) Determination of ultimate capacity of driven piles in cohesionless soil: a multivariate adaptive regression spline approach. *Int J Numer Anal Methods Geomech* 36:1434–1439. <https://doi.org/10.1002/nag.1076>
 22. Kumar M, Kumar V, Rajagopal BG, Samui P, Burman A (2022) State of art soft computing based simulation models for bearing capacity of pile foundation: a comparative study of hybrid ANNs and conventional models. *Model Earth Syst Environ*. <https://doi.org/10.1007/s40808-022-01637-7>
 23. Zhang WG, Goh ATC (2013) Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Comput Geotech* 48:82–95. <https://doi.org/10.1016/j.compgeo.2012.09.016>
 24. Al-Fugara A, Ahmadlou M, Al-Shabeeb AR, AlAyyash S, Al-Amoush H, Al-Adamat R (2022) Spatial mapping of groundwater springs potentiality using grid search-based and genetic algorithm-based support vector regression. *Geocarto Int* 37:284–303. <https://doi.org/10.1080/10106049.2020.1716396>
 25. Javed MF, Amin MN, Shah MI, Khan K, Iftikhar B, Farooq F, Aslam F, Alyousef R, Alabduljabbar H (2020) Applications of gene expression programming and regression techniques for estimating compressive strength of Bagasse ash based concrete. *Crystals* 10:737. <https://doi.org/10.3390/cryst10090737>
 26. Masoumi F, Najjar-Ghabel S, Safarzadeh A, Sadaghat B (2020) Automatic calibration of the groundwater simulation model with high parameter dimensionality using sequential uncertainty fitting approach. *Water Supply* 20:3487–3501. <https://doi.org/10.2166/ws.2020.241>
 27. Samui P (2019) Determination of friction capacity of driven pile in clay using Gaussian process regression (GPR), and minimax probability machine regression (MPMR). *Geotech Geol Eng* 37:4643–4647. <https://doi.org/10.1007/s10706-019-00928-8>
 28. Teodorescu L, Sherwood D (2008) High energy physics event selection with gene expression programming. *Comput Phys Commun* 178:409–419. <https://doi.org/10.1016/j.cpc.2007.10.003>
 29. Chou J-S, Pham A-D (2015) Smart artificial firefly colony algorithm-based support vector regression for enhanced forecasting in civil engineering. *Comput Civ Infrastruct Eng* 30:715–732. <https://doi.org/10.1111/mice.12121>
 30. Wang L (2005) Support vector machines: theory and applications. Springer Science & Business Media
 31. Alemdag S, Gurocak Z, Cevik A, Cabalar AF, Gokceoglu C (2016) Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming. *Eng Geol* 203:70–82. <https://doi.org/10.1016/j.enggeo.2015.12.002>
 32. Shariati M, Mafipour MS, Ghahremani B, Azarhomayun F, Ahmadi M, Trung NT, Shariati A (2022) A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement. *Eng Comput* 38:757–779. <https://doi.org/10.1007/s00366-020-01081-0>
 33. Soleimanbeigi A, Hataf N (2006) Prediction of settlement of shallow foundations on reinforced soils using neural networks. *Geosynth Int* 13:161–170. <https://doi.org/10.1680/gein.2006.13.4.161>
 34. Teh CI, Wong KS, Goh ATC, Jaritngam S (1997) Prediction of pile capacity using neural networks. *J Comput Civ Eng* 11:129–138. [https://doi.org/10.1061/\(ASCE\)0887-3801\(1997\)11:2\(129\)](https://doi.org/10.1061/(ASCE)0887-3801(1997)11:2(129))
 35. Shahin MA, Maier HR, Jaksa MB (2002) Predicting settlement of shallow foundations using neural networks. *J Geotech Geoenvironmental Eng* 128:785–793. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:9\(785\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(785))
 36. Hatheway AW (2009) The complete ISRM suggested methods for rock characterization, testing and monitoring; 1974–2006. *Environ Eng Geosci* 15:47–48. <https://doi.org/10.2113/gsegeosci.15.1.47>
 37. Hatheway AW (2009) The complete ISRM suggested methods for rock characterization, testing and monitoring; 1974–2006

38. Vapnik V (2013) *The nature of statistical learning theory*. Springer science & business media
39. Saremi S, Mirjalili S, Lewis A (2017) Grasshopper optimisation algorithm: theory and application. *Adv Eng Softw* 105:30–47. <https://doi.org/10.1016/j.advengsoft.2017.01.004>
40. Faramarzi A, Heidarinejad M, Mirjalili S, Gandomi AH (2020) Marine predators algorithm: a nature-inspired metaheuristic. *Expert Syst Appl* 152:113377. <https://doi.org/10.1016/j.eswa.2020.113377>
41. Ramezani M, Bahmanyar D, Razmjoooy N (2021) A new improved model of marine predator algorithm for optimization problems. *Arab J Sci Eng* 46:8803–8826. <https://doi.org/10.1007/s13369-021-05688-3>
42. Pazouki G, Golafshani EM, Behnood A (2021) Predicting the compressive strength of self-compacting concrete containing Class F fly ash using metaheuristic radial basis function neural network. *Struct Concr*. <https://doi.org/10.1002/suco.202000047>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.