Research Article – Vol. 02, Issue 02, No. 01

Advanced in Engineering and Intelligence Systems

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

Appraising the Pile Settlement rates by Support Vector Regression Optimized Using the Novel Optimization Algorithms

Argyros Maris[1,*]

[1] *Univ Western Macedonia, Dept Commun & Digital Media, Fourka 52100, Kastoria, Greece*

Highlights

- Evaluating the effective parameters in predicting the pile settlement and to present a trained network, the support vector regression (SVR).
- Hybrid Flow Direction Algorithm (SVR-FDA) and Biogeography-Based Optimization (SVR-BBO) methods have been utilized to find the optimal conclusions.
- For predictions of pile settlement, all models have the coefficient of determination (R2) larger than 0.995 and 0.994, respectively.
- Furthermore, between four hybrid algorithms, SVR-FDA could be proposed as the best model to obtain the most accuracy in the prediction of pile settlement.

Article Info

Received: 22 January 2023
Received in revised: 03 April 2023
Accepted: 03 April 2023
Available online: 16 June 2023

Keywords

Pile Settlement;
Machine learning;
Support vector regression;
Flow Direction Algorithm;
Biogeography-Based Optimization;
RMSE.

Abstract

To ensure the safety of constructions such as bridge-owned structures, they must be immunized for the operational period. Considering the Pile settlement (PS) factor has to be an important project issue, much attention is paid to prevent damage before construction. Various items are considered to evaluate the movement of the piles that certainly help to understand a future picture of the project over the loading period. Most intelligent mathematical strategies in calculating the pile motion are operated. In this regard, the present research has used a machine learning technique: vector regression (SVR). That two optimizers were used to find the key variables of SVR accurately. Biogeography-Based Optimization (BBO) and Flow Direction Algorithm (FDA) were coupled with SVR to create the SVR-FDA and SVR-BBO frameworks. Moreover, several metrics have been used to assess the overall performance of models. The R^2 of the training phase for SVR-FDA was found 99.39 percent shows a great modeling process, while the *RMSE* of this model was calculated 0.4286 mm. The OBJ index as a comprehensive indicator including *MAE*, *RMSE,* and R^2 was obtained 0.2499 mm.

1. Introduction

A number of researches performed on computing the reaction of piles to the exposing to the viable axial masses and loads are mentioned in lectures with the assist of associated research [1], [2]. Existing expertise of how piles reply to loads has caused enhancements in many techniques that researchers can use to assess pile settlement. The techniques stated in various studies had been referred to through numerous researches on this field, especially research [3], [4]. Referred strategies vary from computation strategies from the complexity term view, utilizing analytical and realistic answers, however using the finite difference, and finite element ways are numerical. Practically, on the equal time, designing strategies for the pile is based completely on the sheets under the soil layers as the compressible ones.

* Corresponding Author: Argyros Maris
Email: Argyros.M01614314@gmail.com

Thus, soil compressive sheets beneath a pile were generally accepted because it has been a tangible issue of design and the risk that can outstandingly enhance the settlement rate of piles. An investigation [5] suggested additional settlement rates corresponding to the low-level soil sheets, which are too vital for pile geometry, physical features of the ground, and limited analysis. Studies of this matter apparently are barely done that analytically presented manmade computations and solutions which are often not practical for current individual soil layers.

Besides, a study proposed a way to check the displacement of the pile and introduced the theoretic functions for the investigations on the factor of earth pressure [6], [7]. Most research essentially assesses the pile movement, but all are used directly to lack a ground reflection model. Also, Artificial Neural Networks (ANNs) and well-branched machine learning solutions are used in several studies, Liu et al. [8] Lee and Lee [9], Che et al. [10], Shanbeh et al. [11], and Hanna et al. [12]. Training data, in several studies, were augmented, and motion tests were selected to create a model capable of predicting the ultimate bearing capacity of piles, for which training samples were selected from data collected for developed Artificial Intelligent (AI) models. One study using ANN attempted to calculate results from the settling properties of piles embedded in rock. This data set for the training phase was collected using the real data report of pile subsidence [13].

Regression methods have widely been operated, like Gaussian trend regression, multivariate spline adaptive regression, as well as mini-max stochastic regression machine [14]– [20]. Strategies to issues related to the field of geotechnical engineering are being explored using the method of gene expression programming [19], [21]– [24]. Many studies have examined the mentioned way for assigning vertical bearing capacity in piles [21]. The novel architecture was devised based on GEP [22]. Another article used algorithms to compute the UCS for rock, consisting of GEP, support vector machines, and MLP [24]. The capability of a support vector machine for appraising the displacement of piles socketed in rocks is an allowable fact [25]. In fact, support vector machines present high-accuracy and dependable results. In addition, another study evaluated the ultimate bearing capacity of piles using this method [26], [27]. In this regard, the input dataset contains practical and soil properties that are field-measured and pile samples and the foundation dimensions.

The predominant intention of the current paper is to compute the displacement of the pile in the rocky ground via means of SVR to show a realistic view of pile motion by two hybrid models. To this action, these models will be combined by the applicable optimization algorithms to shape the SVR efficiently to estimate pile settlement primarily based on an in-situ dataset. To attain those goals, SVR used algorithms including Flow Direction Algorithm (FDA) and Biogeography Based Optimization (BBO) to compute the optimized coefficient values to enhance simulation accuracy while reducing the complexity of the calculations' process. In-field datasets for current research in terms of pile movement rates and soil characteristics are given relevant to the Malaysia capital (Kuala Lumpur): the Klang Valley Mass Rapid Transit (KVMRT) network.

The proposed frameworks named SVR-FDA and SVR-BBO strive to feed data of the pile settlement (PS) rates by the ratio of 70 and 30 percent for, respectively, training and testing phases. For running models, the data of UCS of rocks, the column length to diameter ratio, the pile loads, the N_{SPT}, the column length beneath the soil to the length of the pile in the rock ratio, values of a penetration test have been used to appraise the PSs for the KVMRT [28]. Besides, reliability investigation of the developed frameworks demands the indicators that in current research, the indices of *R*, *OBJ, MAE,* and *RMSE* have been utilized to analyze the models' results.

2. Materials and Methodology

2.1.Study area: Klang Valley Rapid Transit System (KVMRT)

Kuala Lumpur (in Malaysia) is the most populated and rapid development city. This congested region and the buildings surrounding and infrastructures motivate building the Klang Valley Mass Rapid Transit (KVMRT). The KVMRT is built to cross the Federal Territory and Selangor State of Kuala Lumpur, especially joining areas within the Klang Valley region. The 51 km of lines in KVMRT includes 35 stations wherein involved both underground and surface constructions. That the total underground tunnel length is approximately 9.5 km. This megaproject constructed in Kuala Lumpur reduces the traffic jam problem, including many piles for bridge support, which is chosen for this study. Fig. 1 exhibits the KVMRT's location in Malaysia.

2.2. Introducing initial data set

A number of piles are to be socketed into various rocks like limestone, phyllite, sandstone, and granite. The records of 96 piles as granite-based in KVMRT were analyzed. The SanTrias class granites are utilized for the KVMRT project. Materials and subsoil datasets were collected to identify general geological features from the pile site. According to the architectural assessments of the ground basis, it is made up of residual rocks. The gathered dataset is the bedrock

depth between 70 cm underground and up to 1400 meters. Therefore, the process of getting samples and related datasets to excavate the piles' records in the study area is described as follows.

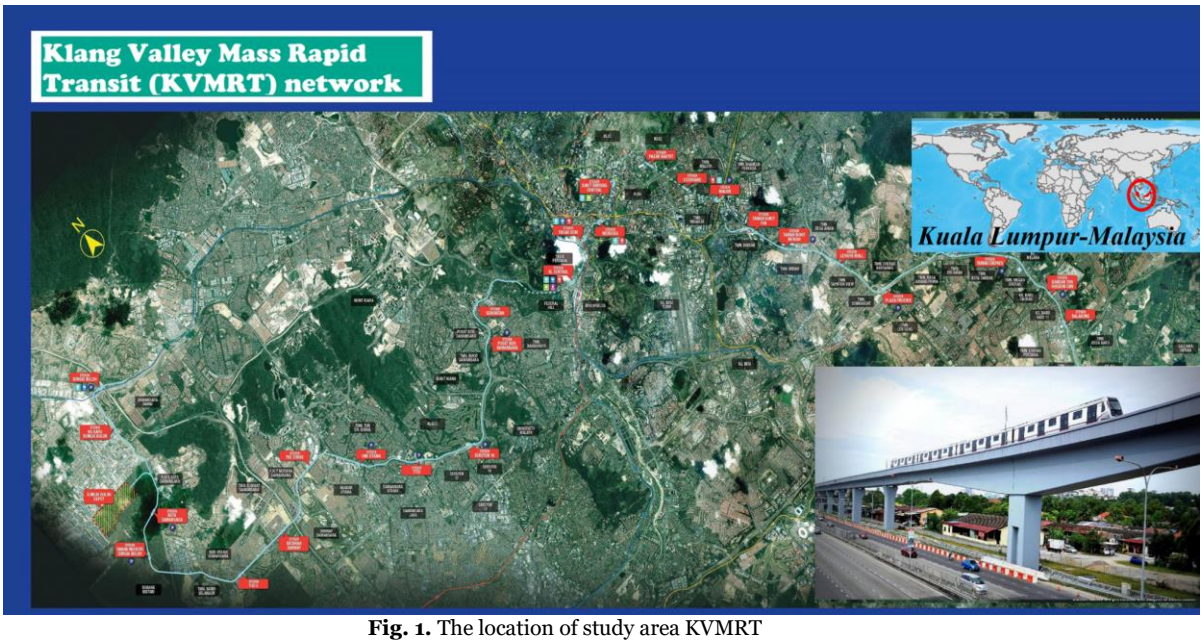


Fig. 1. The location of study area KVMRT

a) The recorded rocks' masses from moderately to highly weathered
b) The bottom and top value of rocks' UCS is based on the parameter of ISRM, alternatively, at the level of 25 to 68 MPa, [29].
c) Registering the bore log data under the 16.5m, that is extremely weathered soil, and the prevailing sort of soil is composed of high mud including sand accompanied by an at least of 4 to and at most of 167 of parameter N_{SPT} lower than 300 mm, alternatively.
d) A great deal of land under the surface level of the ranges 7.5 - 27.0 meters varies with the context of N_{SPT} rates further than 50 m deep for each 300 mm.

Creating the best dataset with effective dependencies was the first step in making the predictive structure. It is necessary to indicate the most important factors shifting the model's outcomes. The above tests were performed using pile analysis settings by Pile Dynamic, Inc. It was referred to earlier that the diameter and length of columns are variables that impact the pile settlement amount.

In this regard, several variables were opted to assess the ramification of the geometry of the pile:

a) the ratio of the column length under the soil to the that in the rock layer (Ls/Lr); b) the ratio of the column length to the diameter of the pile (Lp/D); c) UCS of rock d) N_{SPT} of rock e) pile load masses (directly affects the settlement, so the ultimate potential bearing Q_u of the pile is considered). So, several parameters were selected as data feeding models to appraise the PS values. The inputs' summery feeding the models of this study are shown in Table 1.

Table 1. The dataset feeding models of SVR

Item	Symbol	Unit	Max	Min	S. deviation	Average
Pile length to diameter Ratio	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	42701	12409	803	2454
Soil length to socket length ratio	L_s/L_r	-	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 sample of a pile based on relevant PS has got specific color.

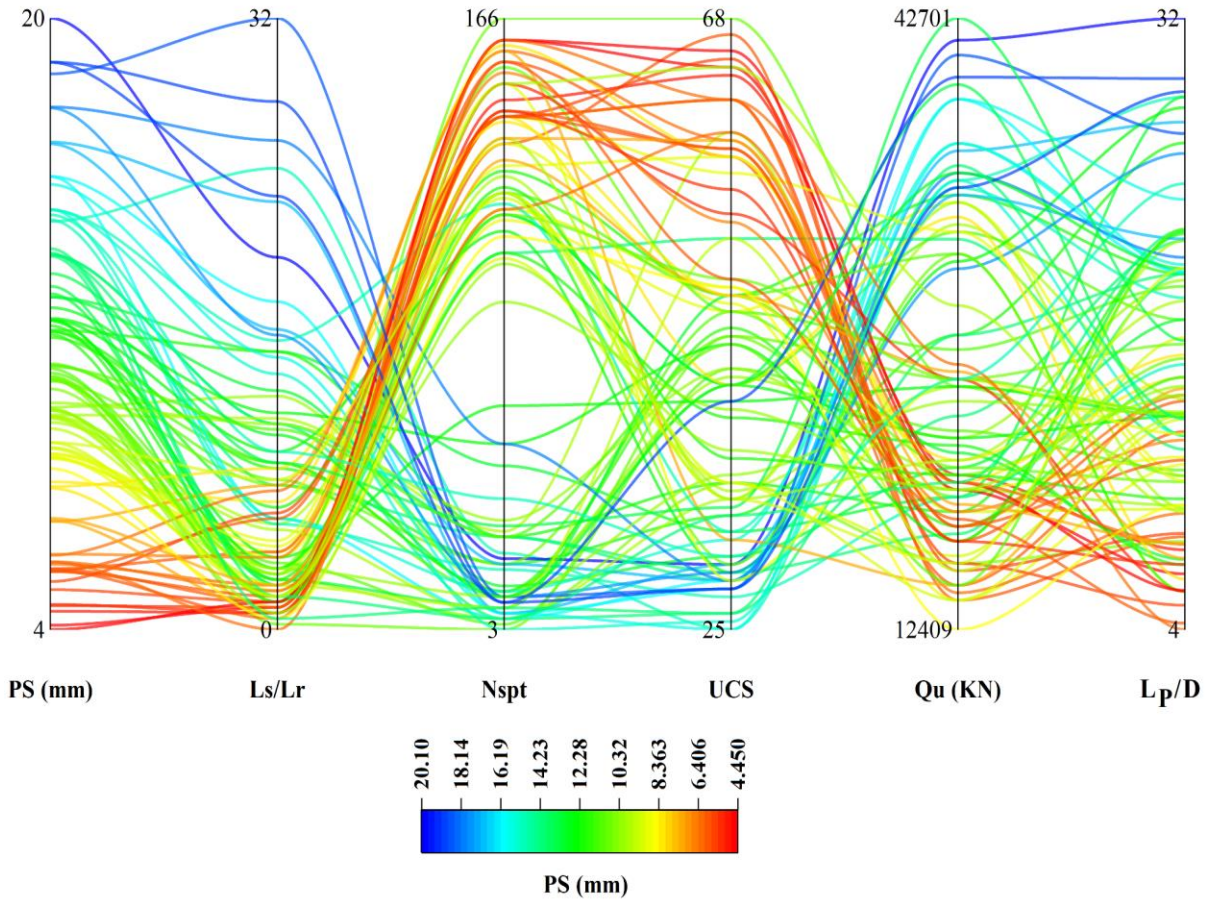


Fig. 2. The input and target values diagram

2.3. Support vector regression, SVR

The SVR machine learning method was suggested for calculating the regression matters [30]. Support vector regression is used for the regression sorting, wherein the tolerance range (ε) is considered for specifying regression. Class categorizing to regression for the SVR approach can create an optimized hyper-plane. This solution seems to be owned by the learning techniques (as the supervised type) for finding the answers for matters given the regression alongside having the following function [31].

$$min_{w,b} = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*)$$
$$s.t. \quad \begin{Bmatrix} y_i - (w^T x_i + b) \le \varepsilon + \xi_i \\ (w^T x_i + b) - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{Bmatrix} \qquad (1)$$

The Eq. (1), the parameter of ξ is the amount of boundary violation; C shows regularizing variables in a queue; w is the factor weight; b represents bias, and ε denotes the deviation rate from the hyper-plane.

Two terms of fitness function are brought up via Eq. (2) and (3):

$$\frac{1}{2}\|w\|^2 \qquad (2)$$

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

The former equation was proposed to enhance the gap between the samples and hyper-plane, then preserve the interval among the samples with the hyperplane; the latter equation acts as an adjusting tool. The variables of w, b were computed over the solvation of function as the target of hyperplane boundaries [32]. For the current research, the quadratic objective function is operated for desirable outcomes. The essential SVR duty is to solve determinative variables as optimal magnitudes containing $sigma$, C, and ε. Achieving mentioned key factors requires a smart algorithm, in which the optimization process via the two optimizers of FDA and BBO was coupled to the SVR for appraising them optimally. It should be noted that the above mentioned $sigma$, C, and ε can boost the SVR accuracy in modeling PS that are brought in Table 2.

Table 2. The key factors of variables magnitudes of each optimizer

		SVR-FDA	SVR-BBO
Training phase	C	0.437	0.577
	EPSILON	4154	958
	sigma	4.125	4.617
Testing Phase	C	0.138	0.259
	EPSILON	2747	302
	sigma	4.848	2.479

2.4. Biogeography-based optimization

The Biogeography Basis Optimization (BBO) algorithm is considered as the metaheuristic way designed with the paradigm of geographic distribution, immigration, and emigration of species in the ecosystem [33]. This optimization algorithm assumes that the ecosystem includes a small number of residencies. Various parameters called fitness index variables affect the quality of each residency for a specimen, such as climate, food, and resources of water. The Habitat Suitability Index (HSI) is deemed the criterion that reflects the residency condition. While the residencies are full or the HSI value is high, the specimens migrate from that residency to a low HSI. All habitats provide a valid solution, and suitability is the variable of decision. In the optimizing operation, responses with smaller targets have a higher value of the Habitat Suitability Index. This algorithm uses two operators: "migration" and "mutation." Migration operators are used to getting neighbors of existing responses, and mutation operators are used to exploring and navigating new responses.

Considering residencies accompanied by the size of HS, the residencies are registered from their function for cost. The desirability of the residencies in the categorized classes is indicated through Eq. (4).

$$HSI_i = -i + HS + 1 \tag{4}$$

Values of emigration and immigration values are calculated as follows:

$$\mu_i = \frac{HSI_i}{HS} \tag{5}$$

$$\lambda_i = 1 - \frac{HSI_i}{HS} \tag{6}$$

Where the variable of μ_i shows emigration and λ_i denotes immigration magnitudes.

The given Fig. 5 shows the migration process of the BBO. Here, the highest emigration and immigration speed value is supposed to be 1. Migration from the j^{th} decision variable of r^{th} habitat to the decision variable of i^{th} habitat is:

$$DV_j^k = \alpha DV_j^i + (1 - \alpha) DV_j^r \tag{7}$$

2.5. Flow Direction Algorithm

The Flow Direction Algorithm used in the present research is designed by the paradigm of runoff moving to different sides that happens in the watershed after the rainfall events. Firstly, such a solution creates a primitive population over the basin or the answer search area. The primitive parameters of the FDA algorithm consist of the number of neighbors defined with β, population number that is considered with α, and neighborhood radius shown with Δ. In this optimizer, the initial flows' positions are computed by the Eq. (8):

$$FlowX(i) = lb + rand \times (ub - lb) \tag{8}$$

Wherein the $randn$ is an accidental value in the range of [zero-one] with uniform distributing, $FlowX$ denotes the positions of flows, lb, and ub denote the low and up limits of the decision variables, alternatively. The parameter of β as the neighborhood close to the flowing cells is considered, and positions can be found by Eq. (9).

$$Neighbor\ X(j) = FlowX(i) + rand\ n \times \Delta \tag{9}$$

In the Eq. (9), the $Neighbor\ X$ parameter shows the neighbor position; $rand\ n$ represents a random value with a standard deviation of 1; The small numbers of neighborhood radius (Δ) searches in the ranges for the parameters in finding through the bigger range. In this regard, searching on a broader range will result in more diverse answers, increasing your chances of getting closer to the best answer. This is called global searching. The solution gets closer to an optimal global solution and finds

a small range of results to have the global solution optimally with the desired accuracy. This is named the local searching, and it helps to strike a balance between the aforementioned features. The Δ direction, moreover, represents an arbitrary position for diversity.

$$\Delta = [rand \times Xrand - rand \times FlowX(i)] \times \|BestX - FlowX(i)\| \times W \quad (10)$$

In which the parameter of $rand$ shows the accidental number that is uniformly distributed, $Xrand$ shows the random position that is generated by Eq. (10), W denotes a nonlinear weight with accidental magnitudes in the range of [0 to inf]. In Eq. (10), the first term is $FlowX$ that goes to a random position called $Xrand$. Moreover, with the second term with enhancing iteration, the variable of $FlowX$ is close to the parameter of $BestX$ that the distance of Euclidian between $FlowX$ and $BestX$ is bringing to zeros.

2.6. *Evaluative indicators for developed SVR-FDA and SVR-BBO*

For evaluating the accuracy of the SVR-BBO and SVR-FDA models to appraise the pile settlement rates for the train and test phases, several indicators are shown in Table 3.

Table 3. Evaluative criteria to analyze the models' performance

Evaluation criteria	Nomenclature	Relations	Assessment
Variance account factor	VAF	$(11)\left(1-\frac{\mathrm{var}(t_n-y_n)}{\mathrm{var}(t_n)}\right)*100$	Higher is desirable
Mean absolute error	MAE	$(12)\frac{1}{N}\sum_{n=1}^{N}\lvert p_n-t_n\rvert$	Lower is desirable
Root mean squared error	RMSE	$(13)\sqrt{\frac{1}{N}\sum_{n=1}^{N}(p_n-t_n)^2}$	Lower is desirable
Pearson's correlation coefficient	R2	$2\ (14)(\frac{\sum_{n=1}^{N}(t_n-\bar{t})(p_n-\bar{p})}{\sqrt{[\sum_{n=1}^{N}(t_n-\bar{p})^2][\sum_{n=1}^{N}(p_n-\bar{p})^2]}})$	Higher is desirable
Statistical parameters, including the various error indices	OBJ	$\left(\frac{n_{train}-n_{test}}{n_{train}+n_{test}}\right)\frac{RMSE_{train}+MAE_{test}}{R^2_{train}+1}+$ $(15)\left(\frac{2n_{train}}{n_{train}+n_{test}}\right)\frac{RMSE_{test}-MAE_{test}}{R^2_{test}+1}$	Lower is desirable [34]

For the variables in equations (11-15), the estimated subsidence rates of pile samples are shown via p_N; t_n denotes target value of measured PS; the average pile settlement measurements are considered by $\bar{t}$; the averaged calculated PSs are indicated using $\bar{p}$. Moreover, alternatively, the ntrain and ntest variables try to show the number of samples gathered for the train and test phase.

3. Results and discussions

The results of SVR are presented in this section by obtaining the SVR-FDA and SVR-BBO models, which are machine learning techniques developed to predict the pile settlement rate. Therefore, we consider the modeling complexity and the cost accompanying increased accuracy of PS estimation, and these issues should be addressed by considering the optimizer used in current research. Modeling is performed using the environment of MATLAB. A detailed chart of measured pile settlement extents for the KVMRT project used as the study area is indicated in Fig. 4, wherein data collected were enrolled in models as 70% of data for training and 30% for testing stages.

As shown from Fig. 3, the placement of points is well distributed for training and testing steps that in each phase, there are varying rates for pile settlements. In Fig. 4, the modeling results have been exhibited in front of measured ones.

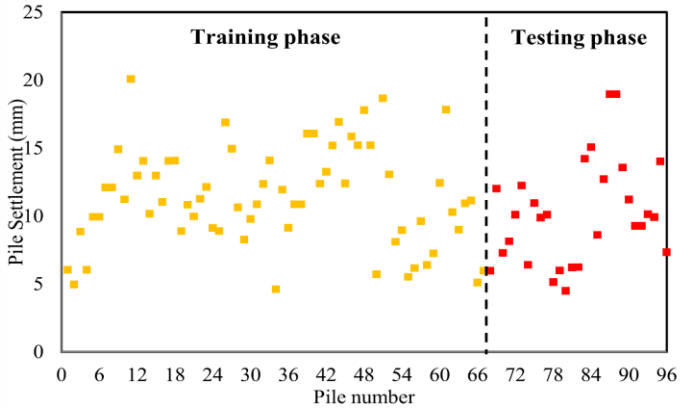


Fig. 3. The data feeding models to the simulation of PSs

SVR-FDA (b) modeling was performed to estimate the pile settlement, and the results are in Fig. 4 shows *RMSE* indicator, and R^2 was calculated, respectively, 0.484 and 0.992 mm. However, the best-fit trend line shows the relating accuracy of modeling located close to the dotted bisector, overestimation of the pile settlement near the pile number of 11, and underestimation of the piles with a higher number of this rate are obvious. The slope of the best-fit trendline of SVR-FDA is 0.92, which shows good modeling, while for SVR-BBO (a), this rate is 0.897, which has a lower correlation. Also, the *RMSE* and R^2 of SVR-BBO with the mentioned rates imply the desirable modeling quality but in comparison with another model (b) the BBO optimizer could do its job a bit lower with the difference of 0.29 32.53 percent, respectively, for R^2 and *RMSE*.

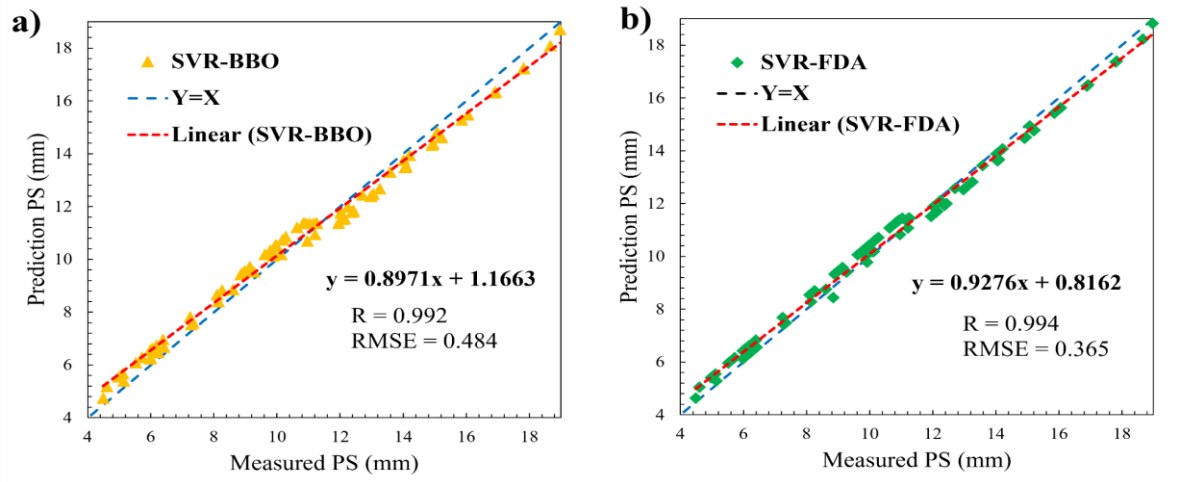


Fig. 4. The modeling results of PS computed by: (a) SVR-BBO and (b) SVR-FDA

Table 4 shows modeling functions for each framework with the *R^2, RMSE, MAE, OBJ*, and *VAF* criteria. The results of training and testing stages represent the same percentage as considering R^2. In the training phase, the FDA optimizer is defined better by having 0.994 and for BBO, 0.991. Also, *VAF* shows the same characteristics of the two models with the rates of 99.982 and 99.931 for SVR-FDA and SVR-BBO, respectively. *MAE*, *RMSE*, and *OBJ* represent a great mismatch to model the rates of pile settlement. For the *RMSE*, this index has shown a large value of 0.429 mm for SVR-FDA and 0.557 mm for SVR-BBO, with a difference of 29.87 percent. While for the testing stage, the FDA optimizer has obtained 0.134 mm, and BBO could get 0.247 mm for RMSE, respectively, with an 84.82 percent difference. The indicator of *MAE* also has a large difference of 82.59 percent in favor of SVR-FDA with an error rate of 0.132. On the other hand, the *OBJ* index shows that SRV-BBO estimates the PS value 0.367 mm as error and 0.250 for SVR-FDA, to which the difference is calculated to be 47 percent, and it is notable that this comprehensive assessment criterion has encompassed all of the error indexes for training and testing phases.

Table 4. Results of models' assessment

			SVR-FDA	SVR-BBO	Average
Criteria used	Training step	R2	0.994	0.991	0.992
		RMSE	0.429	0.557	0.493
		MAE	0.426	0.549	0.487
		VAF	99.982	99.931	99.957
	Testing step	R2	0.9995	0.9985	0.999
		RMSE	0.134	0.247	0.19
		MAE	0.132	0.241	0.187
		VAF	99.998	99.981	99.99
OBJ			0.25	0.367	0.309

For having a suitable view of modeling differences, Fig. 5 wants to show the differences in the simulation of the PS rate of each model.

Fig. 5 calculated the change percent of PS modeled by SVR-BBO from SVR-FDA. As shown from Fig. 5, except for the pile number of three that has been modeled with the difference near to 12%, the remaining 95 piles are modeled in the different range of ±4. However, the pile with the number of 94 has a little more exceeding rate rather than other more than four percent. With this figure, it is definite that appraising PS rate by both models is done in either area of positive and negative relative to each of models that in some piles, for example, 37-49 the PS magnitude calculated by FDA have been higher rates compared BBO. On the other hand, the piles with 53-59 have higher rates for the SVR-BBO than SVR-FDA.

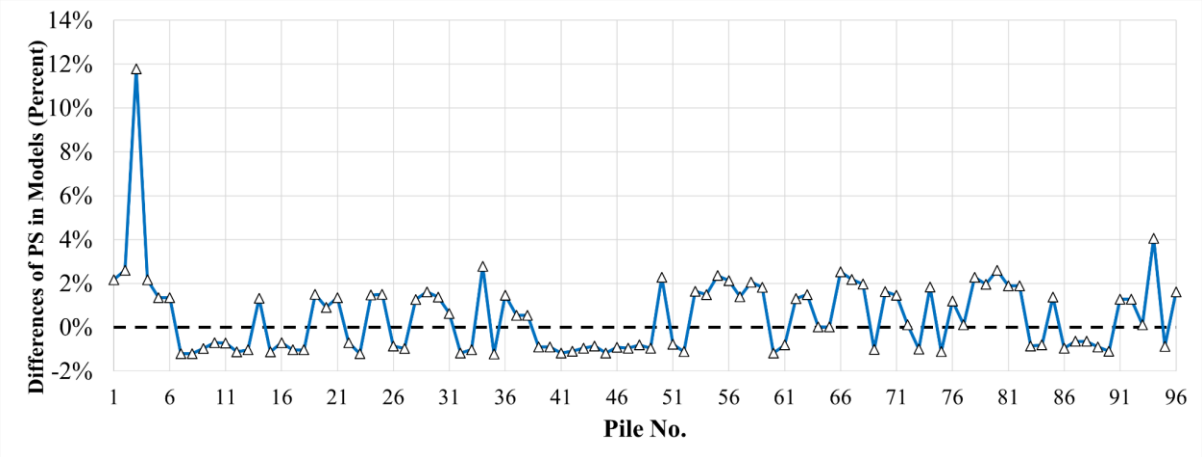


Fig. 5. Difference of modeled PS using SVR-BBO and SVR-FDA

To get a specific idea of the accuracy in modeling, the PS rate modeled is compared to the measured value to display modeling errors in each pile. In Fig. 6, there are several cases where the measured value and the proposed model's result do not match. Most simulations are performed correctly, as they can be implemented in both test and train stages. Figure 6 shows which pile and to what extent the deviation is among the model and the field data. SVR-FDA(a) was simulated accurately, similar to the measurements rates shown in the figure. However, for piles 7 and 55, the error rate between measurements and simulations is greater than other PS modeled. In the same way, this story also applies to SVR-BBO (b). Passing the dotted line as the boundary between the test and train phases further improves simulation accuracy.

In light of analyzing error rates in both modeling processes, Fig. 7 has indicated the trend of error conditions for each pile settlement case. At first look, it is clear that the error rates in the training phase for both models are greater than the testing phase with low error magnitudes. For SVR-FDA, the error domine is between +10% to -5%, that this range for SVR-BBO is +13% to -5%. Actually, with this proof, it can be perceptable that the former model has done its task at a better level than the later model. As shown from Fig. 7, pile number 2 with the error rate of 8.80%, number 4 with -7.23%, 34 with 9.48% are the cases with high error rates. While for SVR-BBO, these ranges have spread to greater error rates of 9.56% for pile one, 11.62% for pile two, 9.54% for pile four, 12.53% for pile 34. But the modeling PS in the test phase is acceptable with the error domain of +3%

to -1% for SVR-FDA and +6% to -2% for SVR-BBO that the former model has had the better result than a latter model.

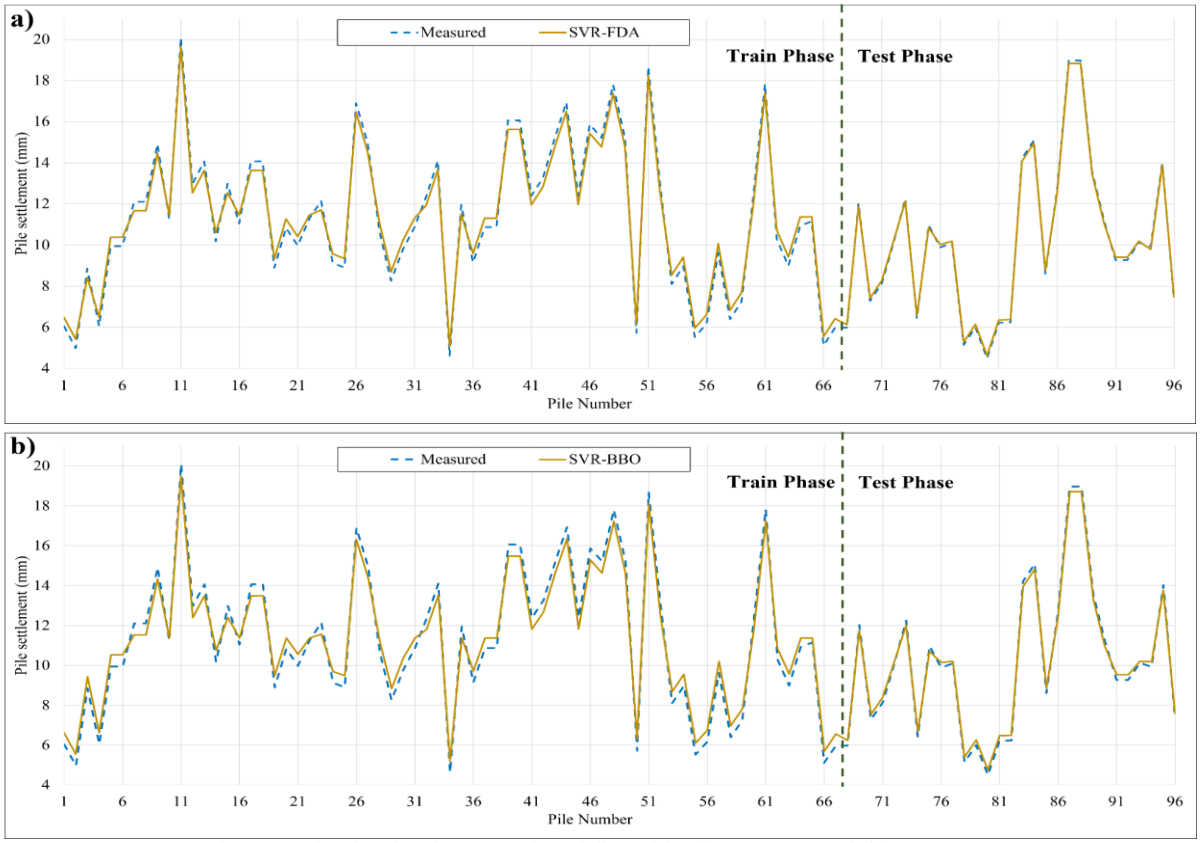
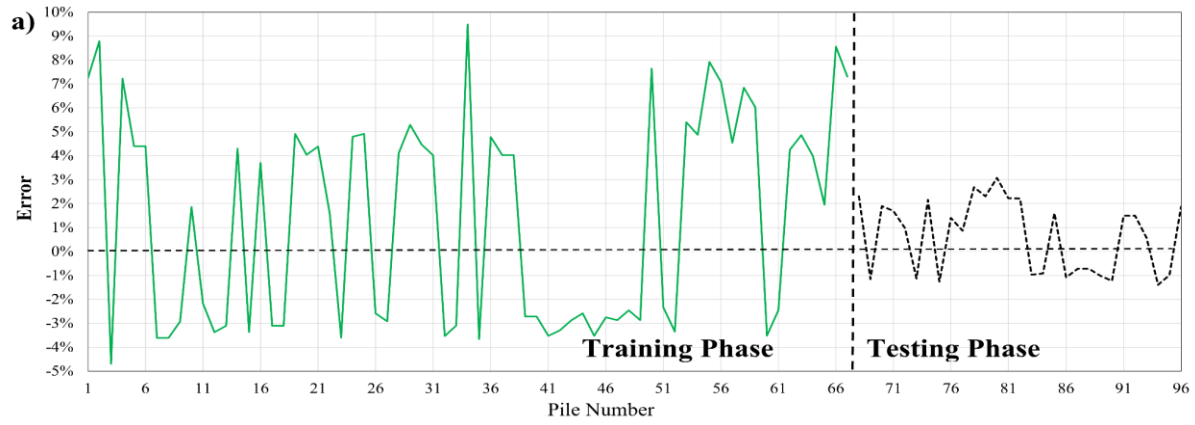


Fig. 6. PS simulated and measured modeling with: (a) SVR-FDA and (b) SVR-BBO



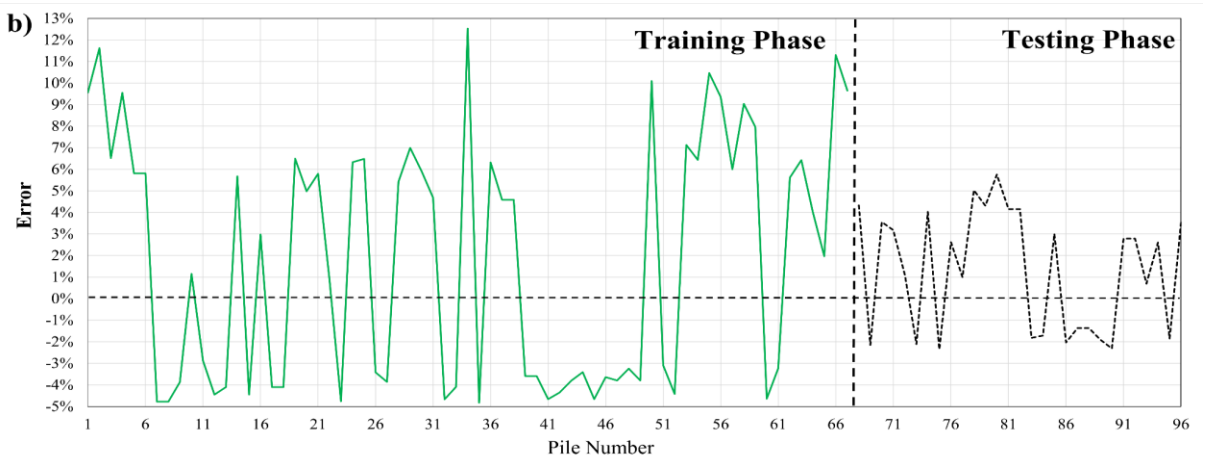


Fig. 7. The error percentage of PS modeled in, (a) SVR-FDA, and (b) SVR-BBO

In this section, the normal distribution of the error for the proposed model is given. In Fig. 8, the error distribution of both models shows the same pattern of non-uniform error spreading along with the horizontal axis of error. The normal distribution curves for SVR-BBO and SVR-FDA have a bold difference: the former model has the flattened curve while the latter has the bell-shaped one. In this regard, neither model has a harmonic pattern of errors. These various errors around zero resulted in a flat distributed error curve. Concentrating the errors of the SVR-FDA model is around -4%, while for SVR-BBO, this is around -2%.

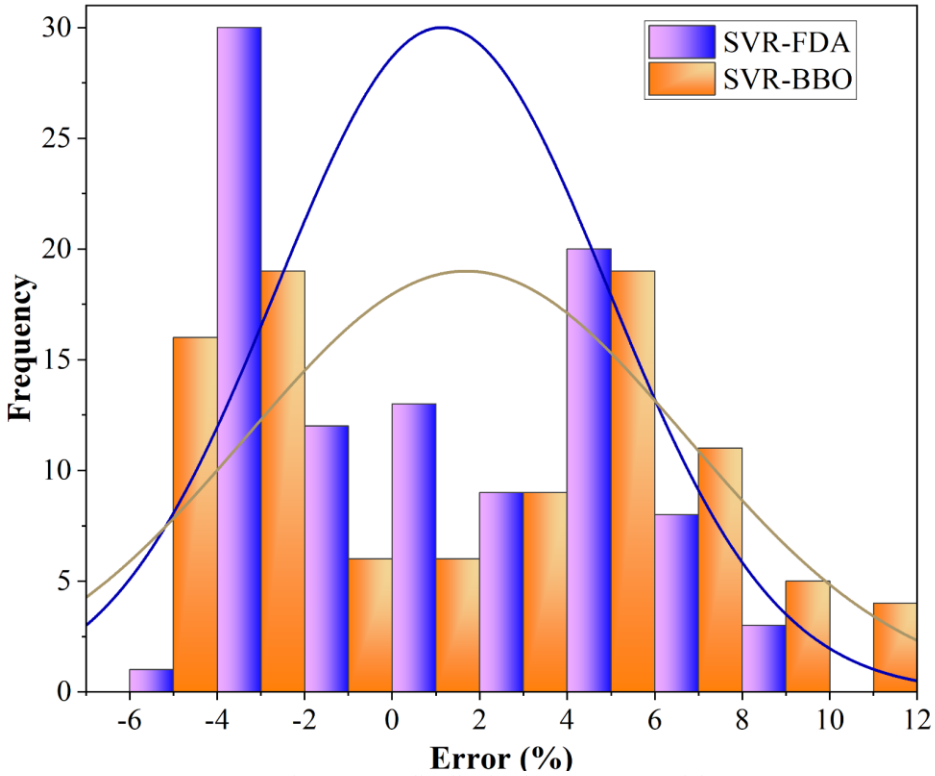


Fig. 8. Error distribution in SVR-FDA model

4. Conclusion

To immunize structures such as towers and bridged ones, it must be safe over the operating period. The factor of the pile settlement (PS), as an important issue of the project, must be considered before the operation that is often avoided. Various items are considered to evaluate the movement of the pile and will help you understand the future picture of the project during the using the project. Most intelligent mathematical strategies are managed when calculating a pile movement. In this regard, the

present study used the machine learning method, Support Vector Regression (SVR), to model the PS rates measured as the field data. Two optimizers were used to determine key SVR variables accurately. Biogeography-Based Optimization (BBO) and Flow Direction Algorithms (FDA) were combined with SVR to generate SVR-FDA and SVR-BBO frameworks. Five metrics were chosen to assess the capability of both proposed models apprising the PS magnitudes. The correlation index of R^2 for SVR-FDA in the training phase was at the desirable rate near 100 percent, 0.993 for SVR-FDA, and 0.990 for SVR-BBO. These rates for the testing phase were closer to each other, with a difference of 0.1 percent. For *RMSE* in the training stage, the SVR-FDA could obtain the mistake range of 0.428 mm in appraising PS, while SVR-BBO could get this value of 0.556 mm with the discrepancy rate of 29.87%. This index for the testing phase got the better rates that for SVR-FDA was about 0.133 mm and for SVR-BBO 0.247 mm with the difference of 84.82 percent. However, the scale of rates is millimeters. The error indicator of MAE also had a similar pattern to appraise the PS. In the training phase for the SVR-FDA, the MAE was obtained 0.426 mm, which was 28.81 percent higher than SVR-BBO. Moreover, the OBJ index, including the various training and testing phases criteria, gives us a better view of each model's performance. This indicator shows the error rate of modeling calculated the error rates for SVR-FDA and BBO as 0.249 and 0.367 mm, respectively. The difference of 46.99% is the sign of FDA capability to appraise the PS values. That this goodness is definite in the error rates for each pile brought up in Fig. (6) and (7). For SVR-FDA, the error domine is located between +10% to -5%, and this range for SVR-BBO is +13% to -5%. However, as reaching the testing phase, the errors of the modeling process reduce to the range of -1% to 3% for SVR-FDA and -2% to +6% for SVR-BBO; this event implies the role of the training phase and preparing for testing phase with lower error. Finally, the results of the FDA were better than BBO, with better values of indices.

REFERENCES

[1] H. G. Poulos, "Pile behaviour—theory and application," *Geotechnique*, vol. 39, no. 3, pp. 365–415, 1989.

[2] M. F. Randolph, "Science and empiricism in pile foundation design," *Géotechnique*, vol. 53, no. 10, pp. 847–875, 2003.

[3] H. G. Poulos, "Pile group settlement estimation—Research to practice," *Foundation Analysis and Design: Innovative Methods*, pp. 1–22, 2006.

[4] D. P. Stewart, R. J. Jewell, and M. F. Randolph, "Design of piled bridge abutments on soft clay for loading from lateral soil movements," *Geotechnique*, vol. 44, no. 2, pp. 277–296, 1994.

[5] H. G. Poulos, *Tall building foundation design*. CRC Press, 2017.

[6] Y. Zhang, X. Hu, D. D. Tannant, G. Zhang, and F. Tan, "Field monitoring and deformation characteristics of a landslide with piles in the Three Gorges Reservoir area," *Landslides*, vol. 15, no. 3, pp. 581–592, 2018.

[7] Y. Zhang, D. C. Richardson, O. S. Barnouin, P. Michel, S. R. Schwartz, and R.-L. Ballouz, "Rotational failure of rubble-pile bodies: influences of shear and cohesive strengths," *Astrophys J*, vol. 857, no. 1, p. 15, 2018.

[8] H. Liu, T. J. Li, and Y. F. Zhang, "The application of artificial neural networks in estimating the pile bearing capacity," 1997.

[9] I.-M. Lee and J.-H. Lee, "Prediction of pile bearing capacity using artificial neural networks," *Comput Geotech*, vol. 18, no. 3, pp. 189–200, 1996.

[10] W. F. Che, T. M. H. Lok, S. C. Tam, and H. Novais-Ferreira, "Axial capacity prediction for driven piles at Macao using artificial neural network." AA Balkema Publishers, Leiden, 2003.

[11] M. Shanbeh, D. Najafzadeh, and S. A. H. Ravandi, "Predicting pull-out force of loop pile of woven terry fabrics using artificial neural network algorithm," *Industria Textila*, vol. 63, no. 1, pp. 37–41, 2012.

[12] A. M. Hanna, G. Morcous, and M. Helmy, "Efficiency of pile groups installed in cohesionless soil using artificial neural networks," *Canadian Geotechnical Journal*, vol. 41, no. 6, pp. 1241–1249, 2004.

[13] A. T. C. Goh, "Pile Driving Records Reanalyzed Using Neural Networks," *Journal of Geotechnical Engineering*, vol. 122, no. 6, pp. 492–495, Jun. 1996, doi: 10.1061/(ASCE)0733-9410(1996)122:6(492).

[14] M. Pal and S. Deswal, "Modelling pile capacity using Gaussian process regression," *Computers and Geotechnics*, vol. 37, no. 7–8, pp. 942–947, 2010.

[15] P. Samui, "Determination of friction capacity of driven pile in clay using Gaussian process regression (GPR), and minimax probability machine regression (MPMR)," *Geotechnical and Geological Engineering*, vol. 37, no. 5, pp. 4643–4647, 2019.

[16] E. Momeni, M. B. Dowlatshahi, F. Omidinasab, H. Maizir, and D. J. Armaghani, "Gaussian process regression technique to estimate the pile bearing capacity," *Arabian Journal for Science and Engineering*, vol. 45, no. 10, pp. 8255–8267, 2020.

[17] W. G. Zhang and A. T. C. Goh, "Multivariate adaptive regression splines for analysis of geotechnical engineering systems," *Computers and Geotechnics*, vol. 48, pp. 82–95, 2013.

[18] R. S. Benemaran and M. Esmaeili-Falak, "Optimization of cost and mechanical properties of concrete with admixtures using MARS and PSO," *Computers and Concrete*, vol. 26, no. 4, pp. 309–

316, 2020.
[19] L. Teodorescu and D. Sherwood, “High energy physics event selection with gene expression programming,” *Computer Physics Communications*, vol. 178, no. 6, pp. 409–419, 2008.
[20] T.-T. Le and M. V. Le, “Development of user-friendly kernel-based Gaussian process regression model for prediction of load-bearing capacity of square concrete-filled steel tubular members,” *Materials and Structures*, vol. 54, no. 2, pp. 1–24, 2021.
[21] I. Alkroosh and H. Nikraz, “Correlation of pile axial capacity and CPT data using gene expression programming,” *Geotechnical and Geological Engineering*, vol. 29, no. 5, pp. 725–748, 2011.
[22] A. Mollahasani, A. H. Alavi, and A. H. Gandomi, “Empirical modeling of plate load test moduli of soil via gene expression programming,” *Computers and Geotechnics*, vol. 38, no. 2, pp. 281–286, 2011.
[23] A. Ozbek, M. Unsal, and A. Dikec, “Estimating uniaxial compressive strength of rocks using genetic expression programming,” *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 5, no. 4, pp. 325–329, 2013.
[24] S. R. Dindarloo, “Prediction of blast-induced ground vibrations via genetic programming,” *International Journal of Mining Science and Technology*, vol. 25, no. 6, pp. 1011–1015, 2015.
[25] S. Alemdag, Z. Gurocak, A. Cevik, A. F. Cabalar, and C. Gokceoglu, “Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming,” *Engineering Geology*, vol. 203, pp. 70–82, 2016.
[26] C. I. Teh, K. S. Wong, A. T. C. Goh, and S. Jaritngam, “Prediction of pile capacity using neural networks,” *Journal of computing in civil engineering*, vol. 11, no. 2, pp. 129–138, 1997.
[27] A. Soleimanbeigi and N. Hataf, “Prediction of settlement of shallow foundations on reinforced soils using neural networks,” *Geosynthetics International*, vol. 13, no. 4, pp. 161–170, 2006.
[28] M. A. Shahin, H. R. Maier, and M. B. Jaksa, “Predicting settlement of shallow foundations using neural networks,” *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 128, no. 9, pp. 785–793, 2002.
[29] A. W. Hatheway, “The complete ISRM suggested methods for rock characterization, testing and monitoring; 1974–2006.” Association of Environmental & Engineering Geologists, 2009.
[30] L. Wang, *Support vector machines: theory and applications*, vol. 177. Springer Science & Business Media, 2005.
[31] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.
[32] A. Al-Fugara, M. Ahmadlou, A. R. Al-Shabeeb, S. AlAyyash, H. Al-Amoush, and R. Al-Adamat, “Spatial mapping of groundwater springs potentiality using grid search-based and genetic algorithm-based support vector regression,” *Geocarto Int*, pp. 1–20, 2020.
[33] D. Simon, “Biogeography-based optimization,” *IEEE transactions on evolutionary computation*, vol. 12, no. 6, pp. 702–713, 2008.
[34] G. Pazouki, E. M. Golafshani, and A. Behnood, “Predicting the compressive strength of self-compacting concrete containing Class F fly ash using metaheuristic radial basis function neural network,” *Structural Concrete*, Feb. 2021, doi: 10.1002/suco.202000047.