



Modelling of lateral effective stress using the particle swarm optimization with machine learning models

Erdal Uncuoğlu¹ · Levent Latifoğlu¹ · Abdullah Tolga Özer²

Received: 6 August 2021 / Accepted: 23 October 2021 / Published online: 13 November 2021
© Saudi Society for Geosciences 2021

Abstract

Predicting the lateral effective stress and the coefficient of lateral earth pressure at rest has a major importance in the design and analysis of many geotechnical problems. The purpose of this study is to predict the lateral effective stress without needing any experimental study or in-situ testing effort, by using the physical properties of sand which can be easily quantified in the laboratory. Therefore, the lateral effective stress values, σ'_h , were estimated by using particle swarm optimization-artificial neural network (PSO-ANN), particle swarm optimization-support vector regression (PSO-SVR) and particle swarm optimization-random forest (PSO-RF) approaches. The internal friction angles were back-calculated using the Jaky's formula utilizing the output of the PSO-ANN model were compared to that of measured experimentally in the laboratory. Thus, both the reliability of the model and the potential of Jaky's formula in predicting the K_0 coefficient were evaluated. The PSO-ANN model found out to be an effective tool to estimate accurately σ'_h in cohesionless soils. It is clearly seen that the predictive performance of the PSO-ANN model was better than that of the both PSO-SVR and PSO-RF models.

Keywords Sand · Lateral effective stress · PSO-ANN · PSO-SVR · PSO-RF · Internal friction angle

Introduction

The prediction of the initial stress state due to the soil's own weight has a great importance for realistic analysis of geotechnical design problems. The vertical effective stress (σ'_v) at any depth in a soil profile can easily be calculated by multiplying the unit weight of soil with the depth and extracting pore water pressure; however, estimating the lateral effective stress at the same depth is often a complex issue. Lateral effective stress (σ'_h) is affected from various parameters such as soil type, void ratio, grain size distribution, grain shape,

stress history, grain sphericity and mineralogy (Hayat 1992; Landva et al. 2000; Mayne and Kulhawy 2003; Chu and Gan 2004; Hanna and Al-Romhein 2008; Tian et al. 2009; Zhao et al. 2010; Hayashi et al. 2012; Talesnick 2012; Lee et al. 2013; Levenberg and Garg 2014; Yun et al. 2015; Gronbech et al. 2016; Wang et al. 2018).

The coefficient of lateral earth pressure at rest (K_0) is defined as the ratio of the lateral effective stress to vertical effective stress in a soil mass which is in elastic equilibrium under the condition of no lateral deformation.

$$K_0 = \sigma'_h / \sigma'_v \quad (1)$$

There are number of studies performed both in the laboratory and in situ to develop a reliable method to obtain the lateral effective stress (Sağlamer 1973; Sağlamer 1975; Abdelhamid and Krizek 1976; Massarch and Broms 1976; Krizek and Abdelhamid 1977; Edil and Dhowian 1981; Fukagawa and Ohta 1988; Ting et al. 1994; Hatanaka and Uchida 1996; Fioravante et al. 1998; Özer 2001; Teerachaikulpanich et al. 2007; Tong et al. 2013; Lee et al. 2013). There is a potential of disturbance when obtaining lateral effective stress values by using the in situ test methods from the drilling of the borehole and insertion of the test device. Laboratory test methods used

Responsible Editor: Zeynal Abiddin Erguler

✉ Erdal Uncuoğlu
erdalu@erciyes.edu.tr

Levent Latifoğlu
latifoglu@erciyes.edu.tr

Abdullah Tolga Özer
tolgaozer@gtu.edu.tr

¹ Department of Civil Engineering, Erciyes University, Kayseri, Turkey

² Department of Civil Engineering, Gebze Technical University, Kocaeli, Turkey

to predict the lateral effective stress in soils have disadvantages such as they require high quality undisturbed samples. In addition, they also require sophisticated test procedure which is costly and time-consuming. However, there are also studies performed using non-destructive field test methods such as seismic method and electrical resistivity method.

Laboratory test methods to define the K_0 coefficient are divided into two groups. These are K_0 -consolidation tests (horizontal strain is restricted, $\epsilon_3=0$) performed in the oedometer test cell and anisotropic consolidation tests ($\sigma'_3/\sigma'_1=\text{constant}$, minor and major principle stresses, respectively) performed in triaxial test systems.

In triaxial test systems, a flexible lateral boundary with a feedback system to maintain the position of vertical boundary of the specimen is used. One of the advantages of the triaxial set up is that the wall friction does not occur. The control of the test specimen under zero lateral deformation conditions as well as ensuring the uniformity of the effective stresses in the specimen can be considered as the disadvantages of this test. Rigid lateral boundaries are used in the oedometer tests, and the required zero lateral deformation condition is achieved. However, friction effect between the oedometer side wall and the test specimen cannot be accurately defined. The side friction in the oedometer cell may induce a variation in vertical stresses along the height of the test specimen. This problem can be solved by measuring the vertical stress at the mid-height of the sample or by averaging the vertical stress values measured on the top and bottom of the sample (Fukagawa and Ohta 1988; Teerachaikulpanich et al. 2007; Wang et al. 2018). Lirer et al. (2011) and Lee et al. (2014) stated that the error in measurement of K_0 value due to the deformation in the oedometer ring is quite small.

In experimental studies performed in the oedometer cell, it is confirmed that the range of lateral deformations occurred during vertical loading are smaller than the limit value ensuring the K_0 conditions. The lateral effective stress value can be measured directly with the pressure cells installed on the side walls of the oedometer test mold. Alternatively, lateral deformations in the thin wall oedometer cell can be directly related to the lateral effective stresses by using the strain gauges attached to thin wall oedometer cell.

Although there are many equations to estimate the value of K_0 in the geotechnical literature (Brooker and Ireland 1965; Fioravante et al. 1998; Federico and Elia 2009; Tong et al. 2013), the most widely accepted one is the Jaky's equation which calculates the value of K_0 as a function of the internal friction angle. Jaky (1944) proposed the following equation to calculate the K_0 values in normally consolidated soils.

$$K_{0-nc} = 1 - \sin\phi' \quad (2)$$

where ϕ' is the effective angle of internal friction.

Due to the listed disadvantages of both in situ and laboratory test methods, there is still no method that can estimate reliably the lateral effective stress and K_0 coefficient. Therefore, alternative approaches are needed for the prediction of lateral effective stress and K_0 coefficient which are easy to apply and can produce economical, fast and reliable solutions.

Uncuoğlu et al. (2008) developed an artificial neural network (ANN) model to predict the lateral effective stress in cohesionless soils using the results of the experimental program performed by Sağlamer (1973). Multilayer feedforward network models have been trained using Levenberg–Marquardt (LM) learning algorithm. Data set has been arranged as three different data groups including data subsets (e.g. training, testing and validation) different from each other to investigate the effect of data selection on the model performance. The relative importance of the selected input parameters on the output parameter were also evaluated by performing sensitivity analysis on the trained network model.

The purpose of this study is to predict the lateral effective stress due to the vertical pressure for a given relative density without needing any experimental effort, by using the physical properties of sand soil which can be easily quantified in the conventional soil laboratory. A data set consisting total of 445 data was used in this paper. The 371 data have been obtained from the 43 normal loading tests performed by Sağlamer (1973) on Kilyos, Ayvalık and Yalıköy sand samples. The remaining 74 data have been obtained from the 12 normal loading tests carried out by Özer (2001) on Şile sand samples. The data set includes relative density, D_r , unit weight, γ_s , particle size at percent finer 10%, D_{10} , particle size at percent finer 60%, D_{60} , mineralogical composition of sand, M , vertical effective stress values applied in oedometer tests, σ'_v and lateral effective stresses, σ'_h , corresponding to the vertical effective stresses. Only the quartz mineral percentage was taken into account for the mineralogical composition of sand samples to ensure unity between the data obtained from both experimental studies.

In the present study, the lateral effective stress values, σ'_h , were estimated by using particle swarm optimization-artificial neural network (PSO-ANN), particle swarm optimization-support vector regression (PSO-SVR) and particle swarm optimization-random forest (PSO-RF) approaches using an extended data set from different experimental studies (Sağlamer 1973; Özer 2001). Effects of the various input parameters on the values of the lateral effective stress have been extensively evaluated performing PSO analyses with different data sets consisting of various numbers of input parameters. The results of the PSO analyses have been compared with each other considering the values of the model performance parameters such as mean square error (MSE), mean absolute error (MAE), the correlation coefficient (R) and the coefficient of determination (R^2). Then, the input

parameters which produce close match between the measured and predicted values of lateral effective stresses have been determined. Performance results obtained from ANN models were compared with results obtained from SVR and RF models.

The lateral effective stress values of the sand soils used in the laboratory model tests available in the literature have been predicted by selected ANN model for different vertical effective stress values. The physical and strength properties of the sands have been obtained from the literature. Then, K_0 coefficient was calculated as the ratio between the lateral effective stress value predicted by ANN and the vertical effective stress value used as an input parameter. The internal friction angle, ϕ' , corresponding to the calculated K_0 coefficient was obtained by back calculation using the Jaky's formula. The internal friction angles obtained with back calculation were compared to that of quantified experimentally by triaxial compression tests in the laboratory. Thus, both the reliability of the model and the potential of Jaky's formula in predicting the K_0 coefficient were evaluated.

Even though there are various laboratory and in situ test methods to define lateral stress in soil medium, there is no consensus on which method can be used in the design. Authors tried to develop a model to estimate lateral stresses based on the physical properties of sands which can be easily quantified in any conventional soil laboratory. The K_0 coefficients were estimated using the lateral stress values obtained from the proposed model, and the internal friction angle values of sands were back-calculated from the K_0 coefficients. Therefore, the study is attempted to provide a basis for estimating the internal friction angle of sands by using index properties without the need for performing triaxial strength testing which requires high-quality undisturbed samples. This can be considered as a novel attempt to overcome the difficulties of collecting undisturbed samples for laboratory testing. The results obtained from the study is also considered to be a reference for future studies on increasing the predictive performance of lateral effective stress using the PSO feature selection with machine learning models.

Materials and methods

Experimental studies

Sağlamer (1973) performed oedometer tests on air-dried, uniform Kilyos, Ayvalık and Yalıköy sand samples to investigate the effects of grain size, grain shape, relative density and stress history on the coefficient of lateral earth pressure at rest. The oedometer cell used in the experimental study was 47 mm high and 17 mm thick and had an inner diameter of 76 mm. The maximum vertical stress applied during the

tests was 1960 kPa, and the measured radial deformation corresponding to this pressure value was 1.5×10^{-5} . During the loading, the lateral and vertical effective stresses in the sand samples were measured directly by piezoelectric measurement method using quartz pressure crystals installed at mid-height and bottom of the oedometer cell. Tests were carried out on sand samples prepared in loose, medium-dense and dense sand conditions. The sand samples used in the experiments were prepared by air pluviation method, tamping method and compaction with vibratory compactor for the loose, medium-dense and dense sand conditions, respectively. The relative density values of the prepared test samples were monitored by checking the weight of the sand in the oedometer cell. During the experimental studies, a total of 61 tests were carried out, 43 of which were under normal loading conditions and 18 were under unloading–reloading conditions.

Özer (2001) was investigated the determination of the lateral soil pressures and the coefficient of earth pressure at rest in cohesionless soils by thin wall oedometer technique performing consolidation tests on the air-dried, uniform, Şile sand samples. A thin wall oedometer cell used in the experimental studies was 62.5 mm high and 0.50 mm thick and had an inner diameter of 63.5 mm. The maximum vertical pressure applied in the experimental studies was 600 kPa, and the maximum lateral deformation developed at this pressure was measured as 11.79×10^{-5} . The lateral displacements on the side wall of the thin wall oedometer ring due to the vertical pressures applied during the loading were measured by strain gauges attached to the side of the thin wall oedometer ring. Then, the lateral stress for a given vertical pressure was computed multiplying the displacement value by the calibration coefficient. Tests were conducted on sand samples prepared in loose, medium-dense and dense sand conditions. The sand samples used in the tests were achieved compacting the sand with a weight corresponding to a certain relative density into the thin wall oedometer cell by hand tamping to create a homogeneous sample. The relative density values of the prepared test samples were monitored by checking the weight of the sand in the oedometer cell. During the experimental studies, a total of 12 normal loading tests were carried out.

The physical and mineralogical properties of sands used in the experimental studies are presented in Table 1. The relative density values of the sand samples used in the tests are summarized in Table 2.

Artificial intelligence studies

In recent years, machine learning techniques have been more widely applied to geotechnical problems (Wang and Akeju 2016; Armaghani et al. 2017; Sharma et al. 2017; Puri et al.

Table 1 The physical and mineralogical properties of sands

Reference	Location of the sand	Type	γ (kN/m ³)	U (D ₆₀ /D ₁₀)	e _{max}	e _{min}	Mineral composition (%)*
Sağlamer (1973)	Kilyos	Fine	26.67	1.25	0.81	0.45	Q(45), C + A (50), M (5)
	Ayvalık	Medium	25.89	1.30	0.91	0.59	Q(80), C + A (19), M (1)
	Yalıköy	Coarse	26.09	1.00	0.67	0.44	Q(99.9), M (0.1)
Özer (2001)	Şile	Fine	26.28	1.00	1.07	0.78	Q(62.12)
		Medium	25.79	1.07	1.00	0.74	C(10.35), MM(3.73)
		Coarse	25.99	1.16	0.93	0.71	F(18.8)

*Following abbreviations are used for the names of the minerals; Q quartz, C calcite, A aragonite, M magnetite, F feldspar, MM magmatic minerals.

**Maximum and minimum void ratios of the sand are denoted by e_{max} and e_{min}.

***U is the uniformity coefficient of sands.

Table 2 The relative density values of the sand samples

Location of the sand		Relative density, D _r
Kilyos		0.36, 0.39, 0.43, 0.47, 0.89, 0.90
Ayvalık		0.33, 0.64, 0.86
Yalıköy		0.33, 0.83
Şile	Fine	0.754, 0.386
	Medium	0.33, 0.61, 0.89
	Coarse	0.33, 0.61

2018; Pham et al. 2019; Ly and Pham 2020; Nguyen et al. 2020).

In this study, artificial intelligence techniques were used to obtain the lateral effective stress values depending on the data selection and prediction. The effect of feature selection using PSO on the modelling performance was analysed, and the modelling performance of the ANN model was compared with the SVR and the RF models. The flow of the proposed study is seen in Fig. 1.

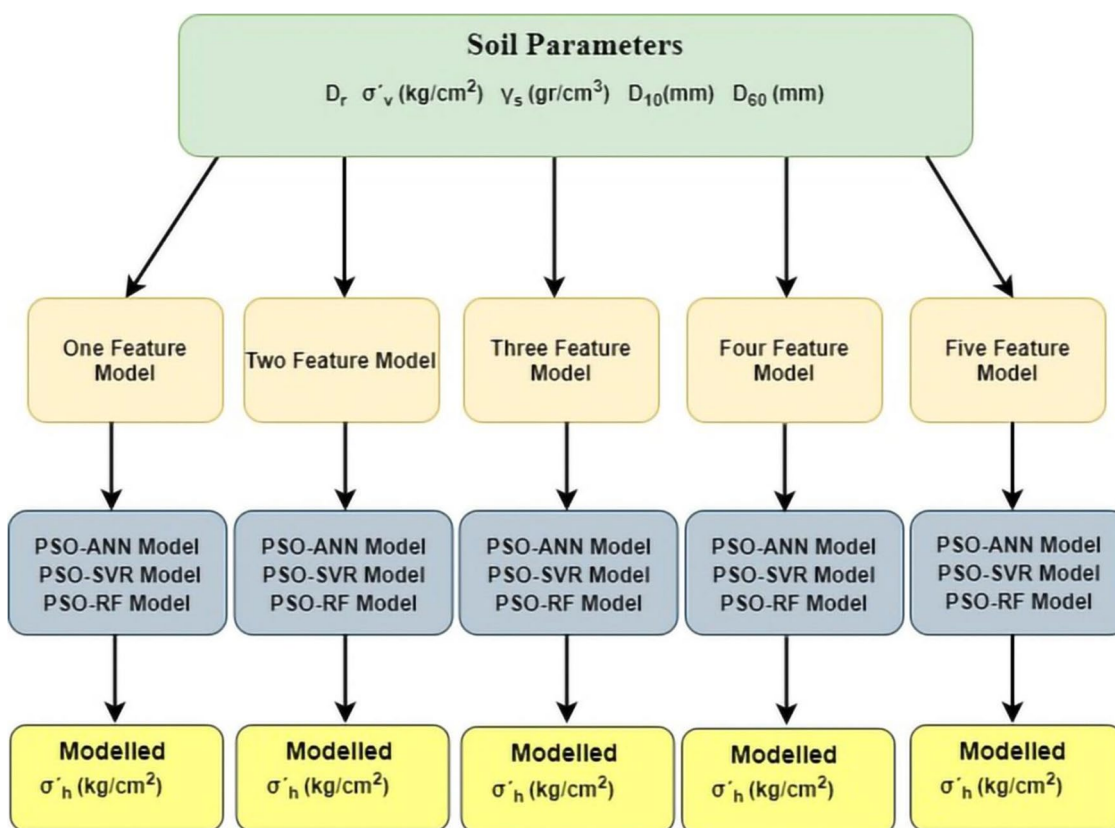


Fig. 1 The flowchart performed for modelling of lateral effective stress, σ'_h (kg/cm²)

The one feature model is used to obtain lateral effective stress by utilizing the most important one feature from the six features listed above. Similarly, for 2–5 feature models, the same approach is applied.

Particle swarm optimization (PSO) algorithm

Particle swarm optimization is a swarm intelligence-based optimization algorithm proposed by J. Kennedy and R. Eberhart in 1995. This algorithm simulates animal’s social behaviour of insects, herds, birds and fishes (Kennedy and Eberhart 1995). The literature shows that PSO has high potential for use in different optimization applications (Ghazvinian et al. 2019). These swarms follow a cooperative food-finding pattern, with each member of the swarm modifying the search pattern based on its own and other members’ learning experiences. PSO algorithm is focused on comparing the positions of individuals in the flock to the flock’s best-positioned individual. This rate of approach is a random condition, and much of the time, individuals in the flock get better in their new movements than they were before, and this process continues until the target is reached.

In particle swarm optimization, the displacement of individuals is done according to the below equations:

$$x_i(t + 1) = x_i(t) + v_i(t) \tag{3}$$

where $x_i(t)$ is position and $v_i(t)$ is velocity vector at t time.

The velocity vector is calculated in the particle swarm optimization as follows:

$$v_{ij}(t + 1) = w * v_{ij}(t) + c_1 r_{1j}(t) * (y_{ij}(t) - x_{ij}(t)) + c_2 r_{2j}(t) * (\hat{y}_j(t) - x_{ij}(t)) \tag{4}$$

where w is inertia weight constant, v_{ij} represents the velocity of i^{th} particle at the range of $j = 1 \dots n$, x_{ij} is the position of i^{th} particle at the range of $j = 1 \dots n$, y_{ij} shows the optimal position (pbest) of its own of i^{th} particle at the j^{th} range, and \hat{y}_j shows the optimal position (gbest) of the swarm at the j^{th} range. Also, c_1 and c_2 are positive acceleration constants, respectively. r_{1j} and r_{2j} are a random number generated between 0 and 1.

To determine their next locations in the search space, the PSO algorithm is led by personal experience (p_{best}), overall experience (g_{best}) and the current movement of the particles.

The general steps of the particle swarm optimization algorithm are as follows.

1. The first step is the creation of the population. The initial value and velocity of each particle are randomly assigned.
2. The second step is calculation of the fitness value. The fitness value of each particle is calculated according to the given objective function.

3. Third step is determination of the particle which has the best value. The fitness value calculated in the previous step is compared with the best personal value (p_{best}) found in the particle’s memory. If the result found in the previous step is better than the current “ p_{best} ” result, the new result is replaced with “ p_{best} .”
4. The fourth step is finding the global best particle. In the second step, the fitness value calculated for each particle is compared with the global best solution (g_{best}) kept in the memory of the program. If there is a better result, this result is replaced by “ g_{best} .” The comparison is performed for all particles.
5. The fifth step is the setting of the speed and position of each particle. The velocity variable of the particle is set according to the formula in Eq. (4), and the position of the particle is adjusted according to the formula in Eq. (3). This process is done separately for each particle.
6. Steps 2 to 5 are repeated until the stopping criteria or conditions are met.

Two important points should be considered when choosing the stopping criteria. First, the stopping condition should not cause early convergence of the algorithm, as this will only result in finding the regional best point. Second, if the stopping condition causes the fitness function to be calculated too high, the search computation cost increases, in which case it should be avoided.

Algorithm can be stopped when,

1. A predetermined maximum number of cycles is reached.
2. A desired result is found.
3. There is no improvement over a period of time.

In this study, PSO algorithm was used to determine the most important one to five features to be used in the modelling of lateral effective stress with ANN, SVR and RF models. As usually to find the best scores, all features in each other with large numbers of combinations have to be tried. The PSO helps the artificial algorithms by choosing the best feature or features in how many feature is requested by these algorithms. Therefore, PSO also can be called as pre-processor or feature selector algorithm.

During the feature selection process, PSO algorithm parameters such as inertia weight w , acceleration constants c_1 and c_2 were set to 0.2, 2 and 2, respectively, by trial and error (He et al. 2016). The number of populations was defined as 20, and stopping criteria were defined as 10 experimentally. It was determined by experimentally values of PSO algorithm parameters during the optimization process. Performance parameters between the modelled and observed data using ANN, SVR and RF models were used during construction of objective function of PSO algorithm. The error term of the objective function was obtained to

minimize the mean square error and maximize the determination coefficient between the modelled and observed data.

A plethora of optimization algorithms have been developed. Due to its numerous advantageous, such as fewer parameters, quicker speed and a simpler flow diagram, PSO is a widely desired form of heuristic algorithm (Hu et al. 2004). Therefore, PSO was employed for feature selection.

Artificial neural networks (ANN)

Artificial neural networks (ANN) are computational algorithms inspired by the information processing technique of the human brain. The organization of biological neural networks in the brain, as well as their ability to learn, recall and generalize, was mimicked by ANN. In accordance with the brain’s information processing method, ANN is a parallel distributed processor capable of storing and generalizing information after a learning process. ANN has the ability to produce solutions to many problems today. Similar to the functional features of the human brain, it has been successfully applied in subjects such as learning, association, classification, generalization, feature determination, optimization and prediction. ANN generates its own experiences based on the data from the samples and generates findings that allow similar decisions on similar issues.

ANN consists of artificial cells that are hierarchically connected to each other and can work in parallel. ANN is composed of processing elements that are connected to each other through weighted connections and each having its own memory. The information processing capabilities of the process elements that make up the network and their connections with each other create different ANN structures. Just as there are nerve cells in biological neural networks, there are artificial nerve cells in ANN. In engineering science, artificial nerve cells are referred to as process element as seen in Fig. 2. The input data is added to the sum function by multiplying it by the weight coefficients. These sum functions are then passed through a transfer function, and the output value of the neuron is defined as follow equation:

$$y_j = f(\sum_i w_{ij}x_i + \theta_j) \tag{5}$$

where j is the number of neurons, i is the number of inputs, x_i is the input signal, w_{ij} is the weight coefficient, and is the bias expression (or threshold).

The nerve cell receives information from the environment through the inputs (x_1, x_2, \dots, x_n). Inputs to the neural network may come from previous nerve cells or from the outside. Weights (w_1, w_2, \dots, w_i) are suitable coefficients that determine the effect of inputs received by ANN on the nerve cell. Each input has its own weight. The large value of a weight means that it is strongly connected to the artificial

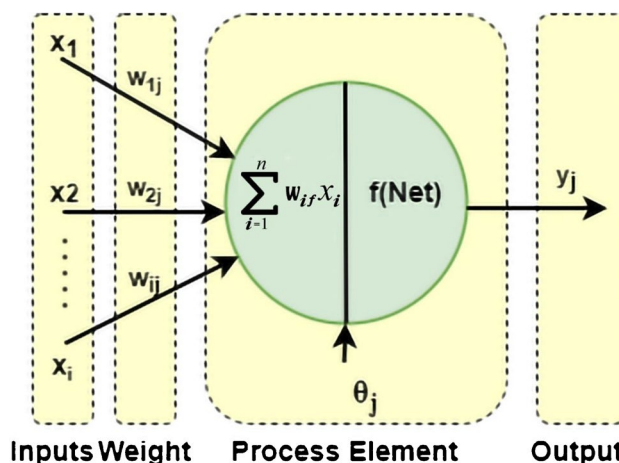


Fig. 2 Structure of process element

nerve cell of that input or important, and a small one means that it is weakly connected or not important (Haykin 1994; Braspenning et al. 1995, Citakoglu 2017).

The result of the addition function is passed to the result by passing f (Net) through the activation function. This function determines the output that the cell will produce in response to this input by determining the net input to the cell. Different formulas are used to measure the output in the activation function, just as they are in the summation function. It is the output value determined by the activation function. The output produced is sent to the outside world or to another cell (process). Generally, cells form a network of three layers, and they are positioned in parallel in each layer.

The multi-layer perceptron (MLP) model, which consists of an input layer, one or more hidden layers and an output layer, is the most commonly used version of ANN. The input layer’s processor elements function as a buffer, distributing input signals to the hidden layer’s processing elements. Artificial nerve cells come together to form the ANN. Nerve cells do not shape in a random order. To form a network, cells are usually arranged in three layers which are input layer, hidden layer and output layer, each layer parallel to the next.

Input layer: In this layer, the process element is responsible for receiving the information coming from the outside world and transferring it to the hidden layer. In some networks, there is no information processing at the input layer.

Hidden layers: Information from the input layer is processed and sent to the output layer. The processing of this information occurs in the hidden layer. There can be more than one hidden layer in a network.

Output layer: The process element in this layer processes the information from the hidden layer and produces the output that the network needs to produce for the information

from the input layer. The output produced is sent to the outside world.

The process elements in each of these three layers and the relationships between the layers are shown schematically in Fig. 3. In these structures, every nerve cell in one layer is connected to all nerve cells of the next layer. There are no connections between the nerve cells in the same layer or in the form of feedback, and it is *feed-forward* ANN (Hornik et al. 1989; Haykin 1994).

The sum of squared differences between the desired and actual values of the output neurons E can be calculated using the below equation:

$$E(w) = \sum (y_{dj} - y_i)^2 \tag{6}$$

where y_{dj} is the desired output value and y_i is the calculated output value.

The number of input parameters and hidden neurons in an ANN model has a major effect on the modelling efficiency. In this study, the number of neurons in the hidden layer was varied from one to twelve in the design of the ANN using the MATLAB program's *for* loop. The model's optimal architecture was obtained by achieving the lowest mean square error between actual and modelled data during training. In order to prevent overfitting, it is also important to determine the number of hidden layers. As a result, the number of hidden layers in this study is set to one to prevent overfitting. Back propagation neural network is the name given to the MLP model when it is supervised by a learning algorithm (BPNN). The feed-forward network, or BPNN, is the most widely used ANN model in modelling. The BPNN

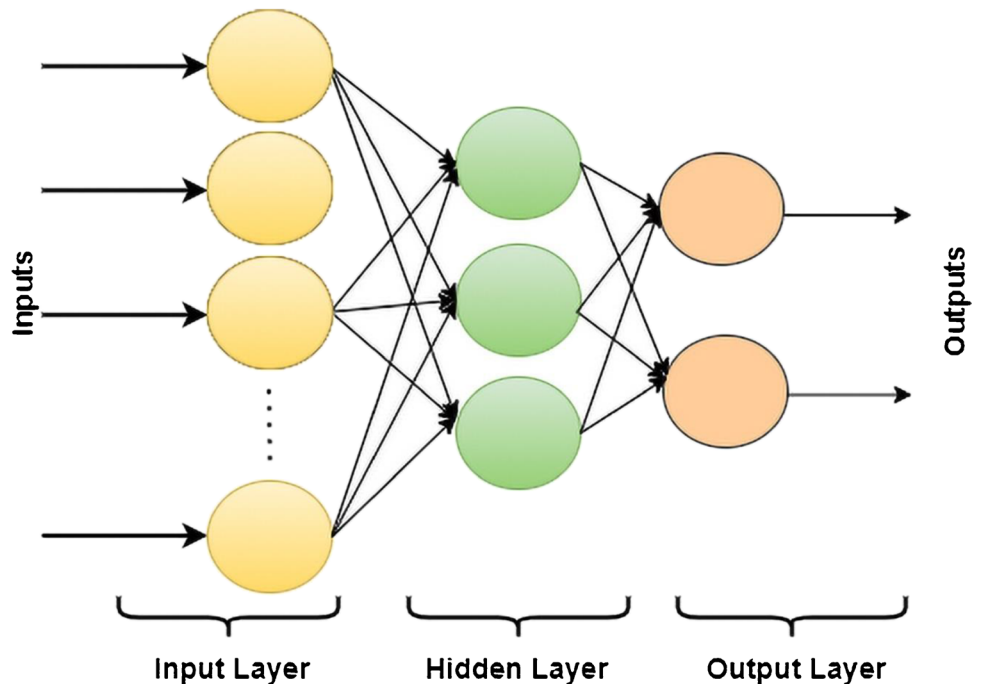
feed-forward network structure was used in our study. Since the network's job is to generate an output for each input, MLP-ANN is based on a supervised learning strategy. There are two phases to MLP-ANN learning. The output of the network is computed first, in the forward calculation step. The weights are determined in the second stage, the backward calculation stage, based on the difference between the estimated output and the output of the networks. Different learning algorithms are used to train the network in the ANN. For the modelling of lateral effective stress, the Levenberg Marquardt (LM) learning algorithm is used, which has a computational speed advantage over the ANN. The detailed information is stated in Moré (1978).

Support vector regression

For classification and regression, support vector machine (SVM) analysis is a popular machine learning method. The statistical learning theory for support vector regression was first introduced by Vapnik (1995). Support vector regression is the use of SVMs in regression (SVR) (Drucker et al. 1997). Since it uses kernel functions, SVR is a nonparametric technique that can balance the trade-off between minimizing empirical error and the complexity of the resulting fitted function. SVR has recently become common in modelling studies, resulting in high performance modelling results.

The SVR algorithm tries to find the best line that separates the two classes. The algorithm allows the line to be drawn to be adjusted in two classes so that it passes the furthest place to its elements as seen in Fig. 4 (Fan et al. 2005; Fan et al. 2006).

Fig. 3 Structure of the multi-layer perceptron



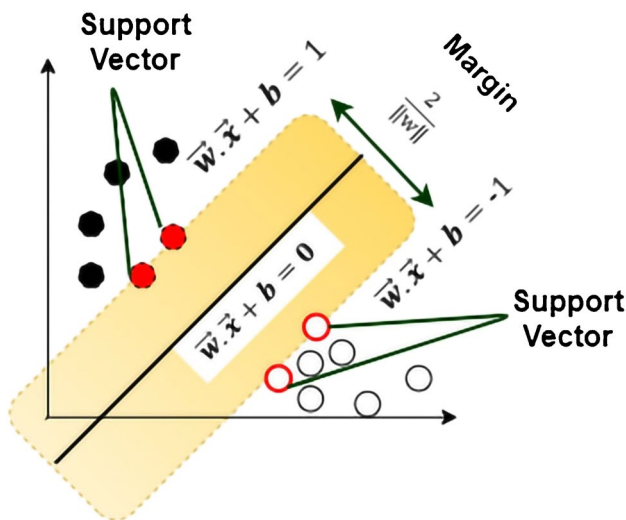


Fig. 4 Representation of support vector

With hyperplanes, the SVR approach attempts to decrease the error rate by maintaining the regression error under a certain threshold value. Assume that the data set satisfies the following criteria.

$$(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), x \in R^D, y \in R$$

$$f(x) = w \cdot x_i + b \tag{7}$$

In this equation, x_i denotes the D-dimensional input vector, y_i denotes the output vectors corresponding to the input vectors and w denotes the normal of the hyperplane, as well as the weight vector, and b denotes the deflection.

The linear relationship between x_i and y_i is assumed in linear support vector regression. The aim is to create a function $f(x)$ that can measure the predicted value y_i at a distance less than or equal to a predetermined value in E (error tolerance) using the actual value x_i , which is each training input data. Errors are ignored in the regression algorithm as long as they are less than E , but any deviation greater than E is not accepted. Equation (8) defines a convex optimization problem.

$$|y_i - (w \cdot x_i + b)| \leq E \tag{8}$$

It's impossible to find a function $f(x)$ that meets this constraint for all data. It is not possible to find a function $f(x)$ that would satisfy such a restriction for all data. For each point, the elasticity variable ξ^+ and ξ^- is used to eliminate this situation. To eliminate this situation, elasticity variable ξ^+ and ξ^- is used for each point ($\xi^+ \geq 0, \xi^- \geq 0$).

$$y_i - (w \cdot x_i + b) \leq E + \xi^+$$

$$(w \cdot x_i + b) - y_i \leq E + \xi^-$$

$$f(x) = C \sum_{i=1}^L (\xi^+ + \xi^-) + \text{minimize} \frac{1}{2} \|w\|^2 \tag{9}$$

where C is a constant value that has a penalty loss effect when an error occurs during training and its value is greater than zero. The following equation is obtained by using the Lagrange multiplier to minimize the error function under constraints.

$$L_p = C \sum_{i=1}^L (\xi^+ + \xi^-) + \frac{1}{2} \|w\|^2 - \sum_{i=1}^L (\mu_i^+ \xi^+ + \mu_i^- \xi^-)$$

$$- \sum_{i=1}^L \alpha_i^+ (E + \xi^+ + y_i - f(x_i)) - \sum_{i=1}^L \alpha_i^- (E + \xi^- - y_i - f(x_i)) \tag{10}$$

In these equation for $\forall_i, \alpha_i^- \geq 0, \alpha_i^+ \geq 0, \mu_i^- \geq 0, \mu_i^+ \geq 0$. The partial derivative of L_p with respect to the variable w, b, ξ^+ and ξ^- is performed to obtain the best solution.

$$\frac{\partial L_p}{\partial w} = 0, \frac{\partial L_p}{\partial b} = 0, \frac{\partial L_p}{\partial \xi^+} = 0, \frac{\partial L_p}{\partial \xi^-} = 0 \tag{11}$$

L_p is maximized with respect to α_i^+ and α_i^- . With respect to Eq. (11), the modelling function is obtained as below equation:

$$f(x) = \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) x_i \cdot x + b$$

$$b = f(x_s) - E - \sum_{m \in S} (\alpha_m^+ - \alpha_m^-) x_m \cdot x_s \tag{12}$$

where S support vectors exist for indices i satisfying the condition $0 \leq \alpha \leq C$ and $\xi^+ = 0$ or $\xi^- = 0$.

Nonlinear regression follows the same steps as linear regression, but with a classifier that cannot be separated linearly.

Data can be moved to the property space or the kernel function can be used to provide a solution. The nonlinear kernel function $K(x_i, x_j) = \varphi(x_i) \varphi(x_j)$ is replaced in Eq. (7) with the dot product $x_i \cdot x_j$ to obtain nonlinear regression. As a result, this is how the modelling function can be written:

$$f(x) = \sum_{i=1}^L (\alpha_i^+ - \alpha_i^-) K(x_i, x) + b \tag{13}$$

Training data is used to build a support vector regression model in the proposed study. Radial basis kernel function is used for the construction of the model. Smola and Schölkopf's sequential minimal optimization (SMO) algorithm was used to optimize SVR parameters during the modelling of the lateral effective stress (Platt 1998).

Random forest algorithm

Random forest (RF) is a tree-based approach that can be used for both regression and classification purposes. Also, it is one of the supervised machine learning methods (Breiman 2001). Leo Breiman developed the RF approach in 2001 at the first time. The main idea of the RF is to build a larger number of decision trees (base learners), and the RF technique is based on a batch-based learning method. Batch classification methods are learning algorithms that generate multiple classifiers instead of a single classifier and then

classify new data based on votes from their predictions. A bootstrap sample of the training data is used to generate each constituent decision tree in the random forest classification and regression process. At each node separation, trees are generated using selected bootstrap samples and m randomly selected estimators during the RF process.

The main stage of RF algorithm is defined as below:

1. The Bootstrap method selects an n -size data set. There are two parts of this data set: training data (inBag) and test data (OOB).
2. The training data set (inBag) is used to produce the largest decision tree (CART), which is not pruned. When dividing each node in this tree, m estimator variables out of a total of p are chosen at random. For branching, the

condition $m < p$ is used. The Gini index determines the value of this variable. This procedure is repeated until no more branches need to be made.

3. A class is allocated to each leaf node. The top of the tree is then the test data set (OOB), and each observation in this data set is allocated to a class.
4. Each stage from the first to the third is repeated N times.
5. Test data that were not used during the creation process are used to test the tree. The classification is performed according to the repetition number of observations.
6. Classification result is obtained with a majority of votes determined over each observation, tree sets.

The flow chart of RF is shown in Fig. 5. Random forest parameters, as shown in the Table 3, are calculated by trial

Fig. 5 The flow chart of random forest

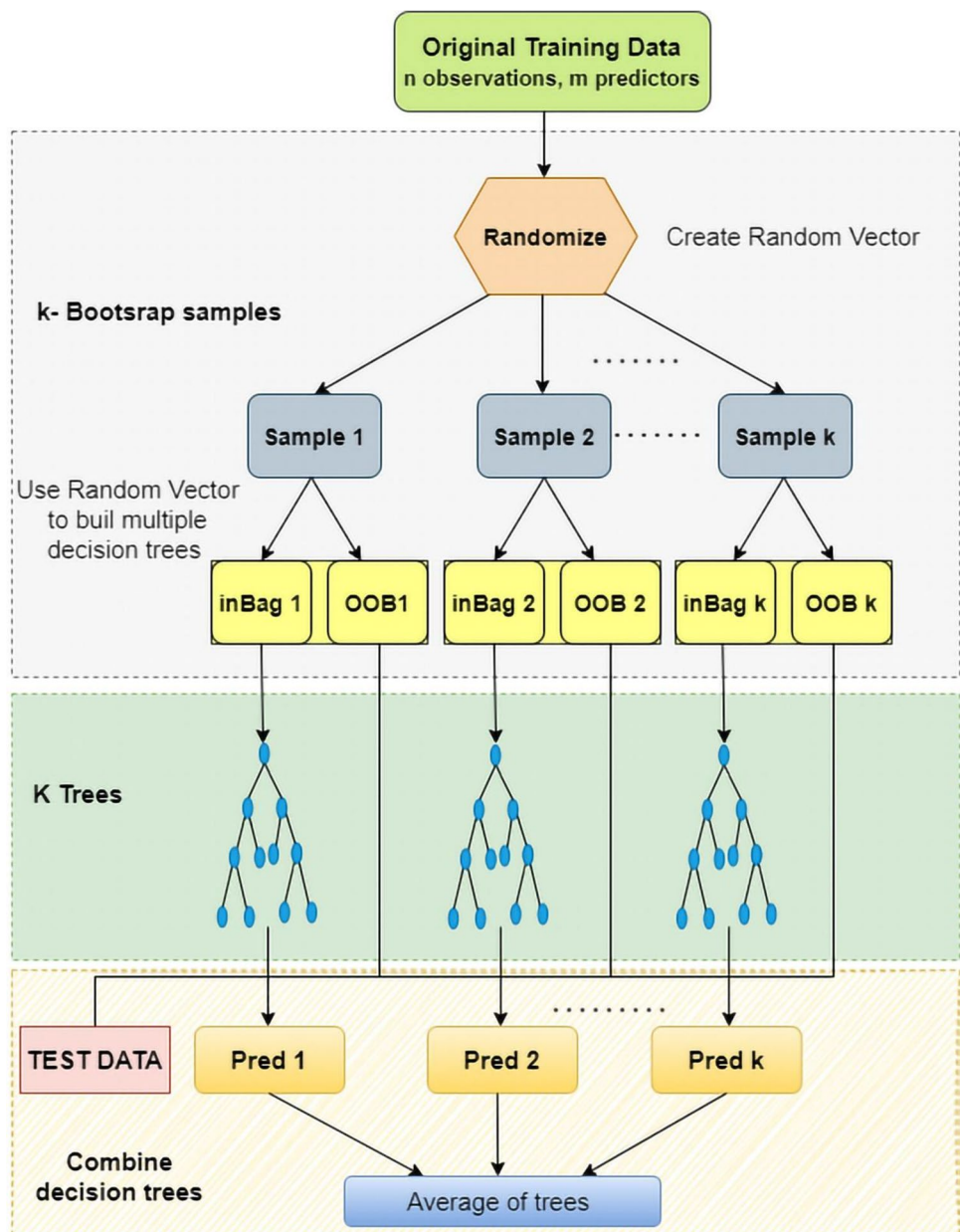


Table 3 Random forest parameters used in the modelling study

Number of iterations	300
Size of each bag, as a percentage of the training set size	80
Seed for random number generator	1
The desired batch size for batch prediction	100

and error during model creation, taking into account calculation time and modelling performance.

Performance evaluation

In this study, the mean absolute error (*MAE*), the mean square error (*MSE*), the correlation coefficient (*R*) and the determination coefficient (*R*²) have been used to show the performance of PSO-ANN, PSO-SVR and PSO-RF models.

The average absolute error measures the variations between observed data and modelled data by the proposed model. The following equation is a summary of *MAE*:

$$MAE = \frac{1}{N} \sum_{i=1}^N |X_{observed,i} - X_{modelled,i}| \quad (14)$$

Mean square error is calculated by squared the average difference across the observed data. The following equation represents the *MSE*:

$$RMSE = \frac{1}{N} \sum_{i=1}^N (X_{observed,i} - X_{modelled,i})^2 \quad (15)$$

The correlation coefficient indicates the degree, direction and significance of the relationship between observed and modelled data. The correlation coefficient, which has a value between [- 1, 1], is represented by the *R*. Below equation is how the *R* value is calculated.

$$R = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{X_{observed,i} - \mu_X}{\sigma_X} \right) \left(\frac{X_{modelled,i} - \mu_{Xe}}{\sigma_{Xe}} \right) \quad (16)$$

In this equation, *X*_{observed,*i*} shows the lateral effective stress data, *μ*_{*X*} is the average and *σ*_{*X*} is the standard deviation of the lateral effective stress data, *X*_{modelled,*i*} is modelled data, and the average of the modelled data is *μ*_{*Xe*} and the standard deviation *σ*_{*Xe*}.

The *R*² coefficient is a commonly used metric for assessing a model’s predictive performance. The range of values for this statistical criterion is - 1 to 1. The findings are outstanding if the *R*² determination coefficient value between the actual and predicted data is one. The following equation is used to measure the *R*² value:

$$R^2 = 1 - \frac{\sum_{i=1}^N [X_{observed,i} - X_{modelled,i}]^2}{\sum_{i=1}^N [X_{observed,i} - \mu_X]^2} \quad (17)$$

K-fold validation for training and testing data

One of the methods for splitting the data set into sections for evaluating and training classification models is k-fold cross validation. If a data set is wanted to be modelled with a simple approach, 75% of the data set is used for training and 25% for testing of the model. However, depending on the distribution of the data, certain deviations (bias) and errors can occur in the training and testing of the model when the data is divided.

Here, k-fold cross validation divides the data into equal parts according to a specified number of k, allowing each part to be used for both training and testing, thus minimizing deviations and errors caused by distribution of the data and dividing of the data. In this study, fivefold cross validation was carried out to obtain training and testing data. It is seen in Fig. 6.



Fig. 6 fivefold cross validation for proposed model

In this study, fivefold cross validation was applied for determination of training and testing data for modelling of lateral effective stress parameter.

Results and discussions

In this study lateral effective stress was modelled using PSO-ANN model and for comparison PSO-SVR and PSO-RF models. The modelling effort was handled in two different ways. Firstly, a data set consisting of six input parameters and one output parameter was used. While the physical properties of the sands such as D_r , D_{10} , D_{60} , quartz mineral percentage, γ and the vertical effective stress values applied in the oedometer tests were selected as input parameters, the lateral effective stress value was chosen as the output parameter. In the second case, since it is not always possible to have information about the quartz mineral percentage, it was excluded from the input parameters. Therefore, the physical properties of the sands such as D_r , D_{10} , D_{60} and γ and the vertical effective stress values applied in the oedometer tests were used as model input parameters, while the lateral effective stress was output parameter. Thus, both the effect of the quartz mineral percentage on the model performance were evaluated. In addition, it was also aimed to develop a model that provides satisfactory predictive performance when there is no information about the quartz mineral percentage.

Firstly, lateral effective stress was calculated using only one feature from two data sets (with quartz mineral percentage and without quartz mineral percentage) based on the parameters defined above. It has been obtained that the vertical effective stress is the most important characteristic selected by the PSO technique in order to estimate the lateral effective stress using the ANN, SVR and RF models in the best performance.

Then, the two, three, four and five most important features were selected from the feature set, which includes quartz mineral percentage, in order to model the lateral effective stress. By the same way, the two, three and four most important features were selected from the data set without quartz mineral percentage using PSO in order to model the lateral effective stress using ANN, SVR and RF.

After the selection of the first most important feature from the feature set containing quartz mineral percentage, the second and the third most important features were determined as relative density and quartz mineral percentage with the proposed optimization-based model. The fourth and fifth most important features were determined as D_{60} and D_{10} parameters, respectively. Using the proposed approach, the second, third and fourth most important features obtained from the data set that do not contain quartz mineral percentage were determined as D_r , D_{60} and D_{10} parameters, respectively, for all models (PSO-ANN, PSO-SVR and PSO-RF).

As seen in Tables 4 and 5, the first and second most important features in both data sets with and without quartz

Table 4 Performance parameters obtained from PSO-ANN, PSO-SVR and PSO-RF models for modelling of lateral effective stress using data set containing quartz mineral percentage

Input number	Features	Models	MSE	MAE	R	R ²	
1	σ'_v	PSO-ANN	0.1453	0.2799	0.9755	0.9517	
		PSO-SVR	0.1695	0.2917	0.9718	0.9444	
		PSO-RF	0.1536	0.2865	0.9725	0.9431	
2	σ'_v D_r	PSO-ANN	0.0132	0.0846	0.9977	0.9953	
		PSO-SVR	0.0253	0.1159	0.9955	0.9911	
		PSO-RF	0.0293	0.1234	0.9950	0.9864	
3	σ'_v D_r Quartz (%)	PSO-ANN	0.0080	0.0658	0.9987	0.9974	
		PSO-SVR	0.0174	0.0985	0.9971	0.9943	
		PSO-RF	0.0199	0.1063	0.9967	0.9927	
4	σ'_v D_r Quartz (%) D_{60} (mm)	PSO-ANN	0.0076	0.0617	0.9987	0.9975	
		PSO-SVR	0.0118	0.0767	0.9981	0.9961	
		PSO-RF	0.0211	0.1087	0.9965	0.9920	
5	σ'_v D_r Quartz (%) D_{60} D_{10}	PSO-ANN	0.0067	0.0585	0.9988	0.9977	
		PSO-SVR	0.0195	0.0920	0.9968	0.9935	
		PSO-RF	0.0183	0.1014	0.9970	0.9933	
6	σ'_v D_r Quartz (%) D_{60} D_{10} γ_s	ANN	0.0078	0.0636	0.9986	0.9972	
		SVR	0.0199	0.0984	0.0969	0.9937	
		RF	0.0185	0.1022	0.9972	0.9938	

Table 5 Performance parameters for modelling of lateral effective stress using data set without quartz mineral percentage

Input number	Features	Models	<i>MSE</i>	<i>MAE</i>	<i>R</i>	<i>R</i> ²	
1	σ'_v	PSO-ANN	0.1453	0.2799	0.9755	0.9517	
		PSO-SVR	0.1695	0.2917	0.9718	0.9444	
		PSO-RF	0.1536	0.2865	0.9725	0.9431	
2	σ'_v D_r	PSO-ANN	0.0132	0.0846	0.9977	0.9953	
		PSO-SVR	0.0253	0.1159	0.9955	0.9911	
		PSO-RF	0.0293	0.1234	0.9950	0.9864	
3	σ'_v D_r D_{60}	PSO-ANN	0.0084	0.0695	0.9986	0.9971	
		PSO-SVR	0.0175	0.0968	0.9969	0.9938	
		PSO-RF	0.0193	0.1045	0.9968	0.9936	
4	σ'_v D_r D_{60} D_{10}	PSO-ANN	0.0076	0.0624	0.9987	0.9973	
		PSO-SVR	0.0120	0.0793	0.9980	0.9960	
		PSO-RF	0.0211	0.1079	0.9965	0.9922	
5	σ'_v D_r D_{60} D_{10} γ_s	ANN	0.0084	0.0673	0.9983	0.9969	
		SVR	0.0128	0.0809	0.9982	0.9963	
		RF	0.0190	0.1033	0.9961	0.9917	

mineral percentage are the same as σ'_v and D_r features. When the third most significant feature is studied, it is discovered that quartz mineral percentage is the third most important feature among the feature set having the quartz mineral percentage, while D_{60} is the third most important feature among the feature set without quartz mineral percentage. The performance parameters are seen in Table 4 for modelling of lateral effective stress using data set containing quartz mineral percentage and in Table 5 for modelling of lateral effective stress using data set without quartz mineral percentage.

In the one-featured lateral effective stress modelling approach, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.1453, 0.2799, 0.9755 and 0.9517 for the PSO-ANN model; 0.1695, 0.2917, 0.9718 and 0.9444 for the PSO-SVM model; and 0.1536, 0.2865, 0.9725 and 0.9431 for the PSO-RF model, respectively.

In the two-featured lateral effective stress modelling approach, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.0132, 0.0846, 0.9977 and 0.9953 for the PSO-ANN model; 0.0253, 0.1159, 0.9955 and 0.9911 for the PSO-SVR model; and 0.0293, 0.1234, 0.9950 and 0.9864 for the PSO-RF model, respectively.

In the three-featured lateral effective stress modelling approach from feature set *having* quartz mineral percentage, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.0080, 0.0658, 0.9987 and 0.9974 for the PSO-ANN model; 0.0174, 0.0985, 0.9971 and 0.9943 for the PSO-SVR model; and 0.0199, 0.1063, 0.9967 and 0.9927 for the PSO-RF model, respectively.

In the three-featured lateral effective stress modelling approach from feature set *without* quartz mineral

percentage, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.0084, 0.0695, 0.9986 and 0.9971 for the PSO-ANN model; 0.0175, 0.0968, 0.9969 and 0.9938 for the PSO-SVR model; and 0.0193, 0.1045, 0.9968 and 0.9936 for the PSO-RF model, respectively.

In the four featured lateral effective stress modelling approach with σ'_v , D_r , quartz mineral percentage and D_{60} features, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.0076, 0.0617, 0.9987 and 0.9975 for the PSO-ANN model; 0.0118, 0.0767, 0.9981 and 0.9961 for the PSO-SVR model; and 0.0211, 0.1087, 0.9965 and 0.9920 for the PSO-RF model, respectively.

In the four featured lateral effective stress modelling approach with σ'_v , D_r , D_{60} and D_{10} features, *MSE*, *MAE*, *R* and *R*² performance parameters were obtained as 0.0076, 0.0624, 0.9987, 0.9973 and 0.9975 for the PSO-ANN model; 0.0120, 0.0793, 0.9980 and 0.9960 for the PSO-SVR model; and 0.0211, 0.1079, 0.9965 and 0.9922 for the PSO-RF model, respectively.

Lateral effective stress estimation performance parameters obtained with ANN, SVR and RF models with five features extracted from PSO algorithm were 0.0067, 0.0585, 0.9988 and 0.9977 for the PSO-ANN model; 0.0195, 0.0920, 0.9968 and 0.9935 for the PSO-SVR model; and 0.0183, 0.1014, 0.9970 and 0.9933 for the PSO-RF model, respectively.

Furthermore, as shown in Tables 4 and 5, when all features were applied to ANN, SVR, and RF models with or without quartz mineral, the 5-featured PSO-ANN model did not increase the predictive performance, regardless of the number of features that was increased.

When these results are examined, it is clearly seen that the performance of the PSO-ANN model is better than the PSO-SVR and PSO-RF models according to *MSE*, *MAE*, *R* and *R²* performance parameters. Figure 7 shows the estimated data for each fold with the PSO-ANN model as an example. As can be seen from Fig. 7, it is obvious that the proposed PSO-ANN model can predict the lateral effective stress parameter with outperform performance.

The order of importance relevance of parameters for estimating the lateral effective stress can clearly be recognized based on the obtained results. For example, an estimation with a coefficient of determination of 0.9517 is obtained with the PSO-ANN model using only σ'_v parameter.

Figure 7 presents the measured lateral effective stresses versus predicted lateral effective stresses by PSO-ANN model with *R²* coefficients for different fold numbers. As

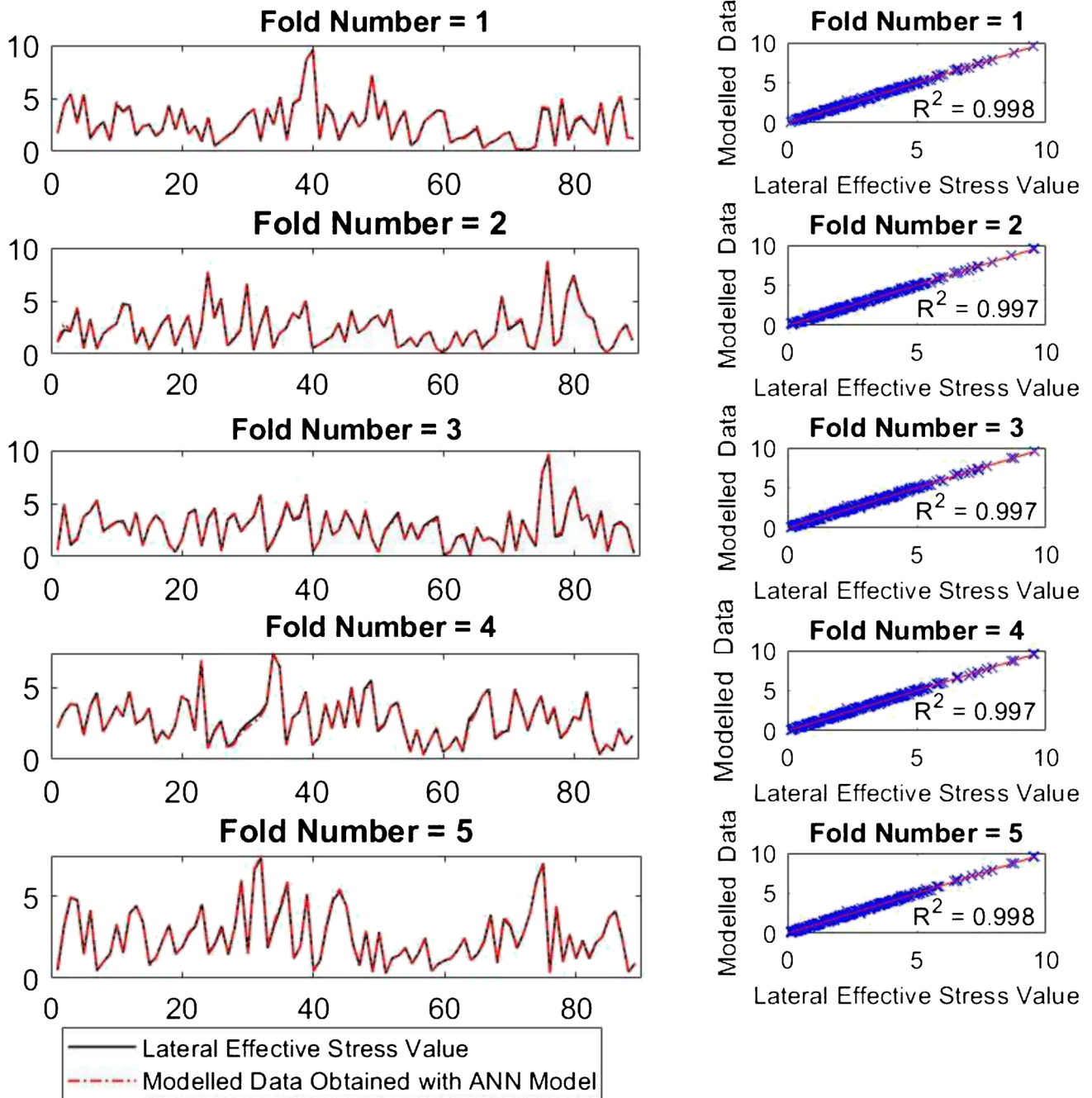


Fig. 7 5 Input PSO-ANN model (*MSE*=0.0067, *MAE*=0.0585, *R*=0.9988, *R²*=0.9977)

seen in Fig. 7, the PSO-ANN model is an effective tool to estimate accurately σ'_h in cohesionless soils.

Taylor diagram investigates the fit of model predictions with measured values and provides the opportunity to make more comparisons between models (Taylor 2001). In this study, Taylor diagram was used to compare PSO-ANN, PSO-SVM and PSO-RF models. Taylor diagram is the graphic which shows the error distributions and model performances with respect to the various performance parameters (Başakın et al. 2021). The Taylor diagram of the PSO-ANN, PSO-SVM and PSO-RF models which includes some of the performance parameters such as correlation coefficient (R), standard deviation (S_d) and root-mean-square deviation ($RMSD$) is illustrated in Fig. 8 as a single chart.

The PSO-ANN model presented in this study was used to estimate the internal friction angles of sands and thus the predictability of the internal friction angle depending on the physical properties of the sand was investigated without experimental studies.

The physical and strength properties of the model sand soils used in the laboratory 1 g model experimental studies in the literature (Quadir 1990; Krabbenhoft et al. 2012; Nasr 2014) have been used with PSO-ANN model developed within this study to predict the lateral effective stress values corresponding to the selected vertical effective stress values. Then, the K_0 coefficient was calculated by the ratio between the lateral effective stress obtained from the PSO-ANN model and vertical effective stress used as input data. Vertical effective stress values vary from 1.0 to 9.0 kg/cm². The values of the internal friction angle corresponding to the calculated K_0 values were obtained using the Jaky (1944) formula by back-calculation. The values of the ϕ angles estimated were compared with the experimental values of the ϕ angles belong to the model sand soils. The experimental ϕ values given in Table 6 have been obtained by performing

triaxial compression tests. The physical and strength properties of sand soils used in model test studies are given in Table 6 with comparative results.

The average absolute difference between the experimental and predicted ϕ values is 2.238°. There is no a distinct relationship between the absolute difference and relative density values.

Triaxial compression tests were carried out on different sand samples under different confinement pressure conditions. The sample sizes used in the experiments were also not the same. However, the experimental results may also include errors during sample preparation and testing.

In any model development process, familiarity with the available data is very important. Generally, different variables comprise different ranges. In the data sets used in the development of the PSO-ANN model, the coefficient of uniformity of the sand soil was between 1.0 and 1.30, while the values of the uniformity coefficient of the sand soils in the experimental studies have been varied between 1.75 and 2.47.

The results indicate that the PSO-ANN model has the ability to predict the internal friction angle indirectly. The ϕ values predicted with this way closely match with the experimental results. It is suggested that the model might serve more generally as a guide to estimate the ϕ values in cohesionless soils. In order to make the prediction more accurate and reliable, some more data would need to be included for different types of sand with various densities and physical properties.

The parameter selection related to the problem has significant effect on the PSO-ANN model performance. It is seen that the quartz mineral percentage has a positive effect on the results obtained considering the MAE and MSE values shown in Tables 4 and 5. However, it is not always possible to know the quartz mineral percentage for all sand samples. For this reason, ϕ angle estimations were made according to the results of the model, which did not include the percentage of quartz mineral, while verifying with the literature data.

As shown in Fig. 9, under normal loading conditions, there is a linear relationship between vertical effective stress and lateral effective stress, and the slope of the line is equal to the K_0 coefficient. Also, the most important parameter controlling the K_0 coefficient under normal loading conditions is the initial void ratio of the sand.

ANNs benefit from their powerful mapping capabilities as well as their naturally parallel and distributed processing features. Due to their flexible nature, ANNs can be considered to be particularly versatile in different classification tasks and having satisfactory modelling performance. However, the more complex the network typology in ANNs, the higher the computation time will be required during the training phases. ANNs are difficult to interpret intuitively

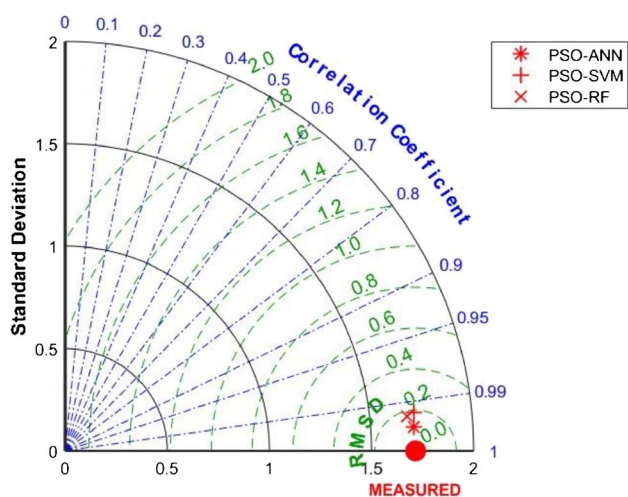


Fig. 8 The Taylor diagram of the PSO-ANN, PSO-SVM and PSO-RF models

Table 6 The physical and strength properties of the model sands with comparative results

References	Sand	D_r	$\phi_{\text{-experimental}}$ (°)	e_{max}	e_{min}	γ_s (gr/cm ³)	c_u	$\phi_{\text{-ANN}}$ (°)	Absolute difference (°)
Mohammed Abdul Quadir (1990)	Jamuna Sand	0.927	38.9	1.150	0.670	2.69	2.47	41.636	2.736
		0.671	37.3					36.734	0.566
		0.581	36.9					35.317	1.583
	Meghna Sand	0.951	39.2	1.100	0.655	2.70	1.75	42.859	3.659
		0.755	38.0					38.697	0.697
		0.566	37.2					35.479	1.727
	Teesta Sand	0.932	39.4	1.040	0.630	2.68	2.18	42.875	3.475
		0.807	38.7					42.742	4.042
		0.644	37.7					40.882	3.182
Krabbenhof et al (2012)	Esbjerg Sand	0.20	31.8	0.733	0.449	2.621	2.32	27.026	4.774
		0.59	38.9					40.271	1.371
		0.84	41.3					40.564	0.736
Nasr (2014)	Commercial Sand	0.35	34.5	0.571	0.399	2.65	1.97	34.953	0.453
		0.78	40.3					37.974	2.326

due to their large number of parameters and complex structure, and parameter tuning requires expert knowledge (Sebastiani 2002).

SVR allows to determine the allowable error is acceptable in the model and will match the data with a suitable line (or hyperplane in higher dimensions). Computing the optimal

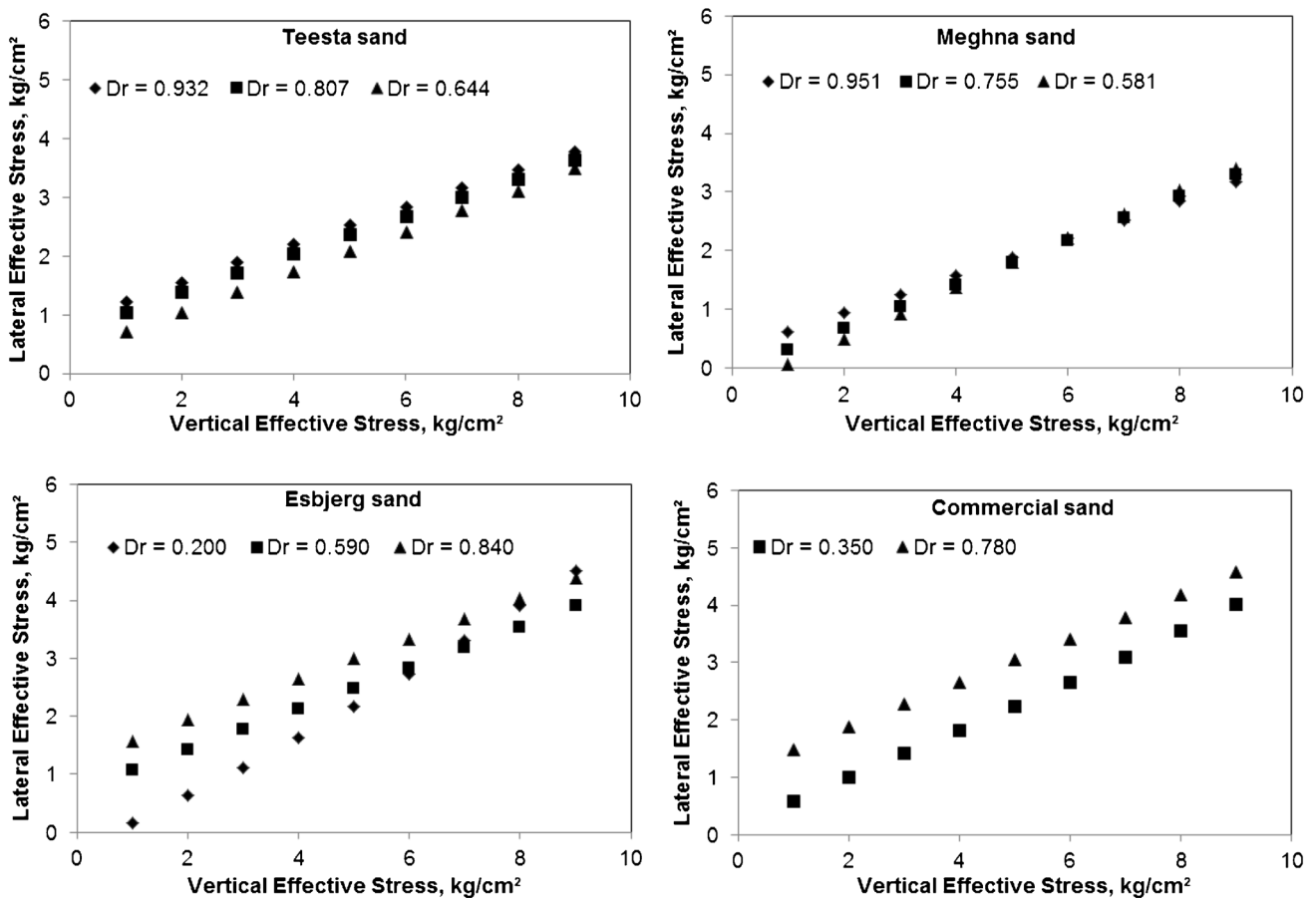


Fig. 9 Vertical effective stresses versus predicted lateral effective stresses by PSO-ANN model

distinctive hyper plane and support vectors's parameters is a convex optimization problem, which can be time-consuming depending on the sample size and number of features (Cortes and Vapnik 1995). SVR has been used to solve a variety of modelling and prediction issues. Their restricted representation, on the other hand, limits their capacity to model nuanced patterns in training data. SVM is also thought to be less prone to overfitting (Joachims 1998).

Because of its hierarchical design, RF is more tolerant to noise and outliers. It can also learn complex relationships between features, perform automated feature selection and model highly nonlinear data. Finally, the RF's training time scales linearly with the ensemble's number of decision trees. The process may be simply paralleled because each tree grows at its own pace (Breiman 2001). This makes RF scalable and computationally efficient, allowing for rapid training of the classifier. It performs satisfactory in terms of classification, contrary not well in terms of regression, since it does not provide exact continuous nature prediction. In addition, small variations in the training set can result in different trees and different predictions for the same validation examples in RF.

In terms of computational costs, the SVR algorithm took significantly longer than the ANN and RF techniques. The computational costs including parameter optimization of the SVR model were more than 28 times than that of the ANN model and more than 120 times than that of the RF model. Furthermore, the ANN algorithm required more time on average than the RF approach. Specifically, the computational costs including parameter optimization of the ANN model were approximately 30 times more than the computational costs of the RF model.

It was observed that when the value and number of used model parameters increased, the computing costs of machine learning models increased as in SVR and ANN models. In addition, longer computing durations do not appear to produce better outcomes. Therefore, the three machine learning models can be rated in terms of general modelling predictive performance as: ANN, followed by SVR, and RF. The performance of the SVR method is similar to that of the RF model but worse than more flexible methods like ANN.

It is suggested that feature selection with PSO and modelling with ANN model would be promising approach in terms of *MSE*, *MAE*, *R* and *R*² and computational efficiency for modelling of lateral earth pressures.

Conclusions

The values of the lateral effective stress σ'_h and the coefficient of lateral earth pressure at rest, K_0 , were investigated using artificial intelligence techniques with the data obtained from oedometer tests on Kilyos, Ayvalık, Yalıköy and Şile

sands. For this purpose, the most important features from the feature set consisting of sand parameters for lateral effective stress estimation were selected using the PSO method and modelled using ANN, SVR and RF models. Based on the investigation the following main conclusions can be drawn.

- Under normal loading conditions, there is a linear relationship between vertical effective stress and lateral effective stress, and the slope of the line is equal to the K_0 coefficient.
- The PSO-ANN model is an effective tool to estimate accurately σ'_h in cohesionless soils.
- The parameter selection related to the problem has significant effect on the PSO-ANN model performance. It is seen that the quartz mineral percentage has a positive effect on the results obtained.
- The results indicate that the PSO-ANN model has the ability to predict the internal friction angle indirectly. The ϕ values predicted with this way closely match with the experimental results. In order to make the prediction more accurate and reliable, more data would need to be included for different types of sand with various densities and physical properties.
- It is clearly seen that the performance of the PSO-ANN model is better than the PSO-SVR and PSO-RF models based on the *MSE*, *MAE*, *R* and *R*² performance parameters.
- It has been obtained that the vertical effective stress is the most important characteristic selected by the PSO technique for predicting the lateral effective stress using the ANN, SVR and RF models in the best performance.

Funding The experimental work for Şile Sand (Özer, 2001) was supported by the Scientific and Technological Research Council of Turkey (TUBITAK grant number 100I025).

Declarations

Conflict of interest The authors declare no competing interests.

References

- Abdelhamid S, Krizek JR (1976) At - rest lateral earth pressure of a consolidating clay. *Journal of The Geotechnical Engineering Division GT7*: 721 - 738.
- Armaghani DJ, Shoib RSNSBR, Faizi K, Rashid ASA (2017) Developing a hybrid PSO-ANN model for estimating the ultimate bearing capacity of rock-socketed piles. *Neural Comput Appl* 28(2):391-405

- Başakın EE, Ekmekcioğlu Ö, Çıtakoğlu H, Özger M (2021) A new insight to the wind speed forecasting: robust multi-stage ensemble soft computing approach based on pre-processing uncertainty assessment. *Neural Computing and Applications* 1–30.
- Braspenning PJ, Thuijisman F, Weijters AJMM (1995) Artificial neural networks: an introduction to ANN theory and practice. Springer Science & Business Media.
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Brooker EW, Ireland HO (1965) Earth pressure at rest related to stress history. *Can Geotech J* 2(1):1–15
- Chu J, Gan CL (2004) Effect of void ratio on K_0 of loose sand. *Geotechnique* 54(4):285–288
- Citakoglu H (2017) Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey. *Theoret Appl Climatol* 130(1):545–556
- Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297
- Drucker H, Burges CJ, Kaufman L, Smola A, Vapnik V (1997) Support vector regression machines. *Adv Neural Inf Process Syst* 9:155–161
- Edil BT, Dhowian WA (1981) At - rest lateral pressure of peat soils. *Journal of The Geotechnical Engineering Division GT2*: 201 – 217.
- Fan RE, Chen PH, Lin CJ (2005) Working set selection using second order information for training support vector machines. *The Journal of Machine Learning Research* 6:1871–1918
- Fan RE, Chen PH, Lin CJ (2006) A study on smo-type decomposition methods for support vector machines. *IEEE Trans Neural Networks* 17:893–908
- Federico A, Elia G (2009) At-rest earth pressure coefficient and Poisson's ratio in normally consolidated soils. In: *Proceedings of the 17th International Conference on Soil Mechanics and Geotechnical Engineering*, M. Hamza et al. (Eds.), pp 7–10. <https://doi.org/10.3233/978-1-60750-031-5-7>.
- Fioravante V, Jamiolkowski M, Lo Presti FCD, Manfredini G, Pedroni S (1998) Assessment of the coefficient of the earth pressure at rest from shear wave velocity measurements. *Geotechnique* 48(5):657–666
- Fukagawa R, Ohta H (1988) Effect of some factors on K_0 value of a sand. *Soils Found* 28(4):93–106
- Ghazvinian H, Mousavi SF, Karami H, Farzin S, Ehteram M, Hossein MS et al (2019) Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction. *PLoS ONE* 14(5):e0217634. <https://doi.org/10.1371/journal.pone.0217634>
- Gronbech GL, Ibsen LB, Nielsen BN (2016) Earth pressure at rest of Sovind Marl – a highly overconsolidated Eocene clay. *Eng Geol* 200:66–74
- Hanna A, Al-Romhein R (2008) At-rest earth pressure of overconsolidated cohesionless soil. *Journal of Geotechnical and Geoenvironmental Engineering* 134(3):408–412
- Hatanaka M, Uchida A (1996) A simple method for the determination of K_0 value in sandy soils. *Soils Found* 36(2):93–99
- Hayashi H, Yamazoe N, Mitachi T, Tanaka H, Nishimoto S (2012) Coefficient of earth pressure at rest for normally and overconsolidated peat ground in Hokkaido area. *Soils Found* 52(2):299–311
- Hayat TM (1992) The coefficient of earth pressure at rest. Dissertation, University of Illinois at Urbana-Champaign.
- Haykin S (1994) *Neural networks: a comprehensive foundation*. Macmillan College Publishing, London
- He Y, Ma WJ, Zhang JP (2016) The parameters selection of PSO algorithm influencing on performance of fault diagnosis. In *MATEC Web of Conferences* 63:02019
- Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. *Neural Netw* 2:359–366
- Hu X, Shi Y, Eberhart R (2004) Recent advances in particle swarm. In *Proceedings of the 2004 Congress on Evolutionary Computation* 1:90–97
- Jaky J (1944) The coefficient of earth pressure at rest. *Journal of the Society of Hungarian Architects and Engineers* 7:355–358
- Joachims T (1998) Text categorization with support vector machines: learning with many relevant features. In: *European Conference on Machine Learning ECML-98* 137–142.
- Kennedy J, Eberhart RC (1995) Particle swarm optimization. In: *IEEE International Conference on Neural Networks*, pp 1942–1948.
- Krabbenhoft S, Clausen J, Damkilde L (2012) The bearing capacity of circular footings in sand: comparison between model tests and numerical simulations based on a nonlinear Mohr failure envelope. *Advances in Civil Engineering*. 1–10. <https://doi.org/10.1155/2012/947276>.
- Krizek JR, Abdelhamid S (1977) Indirect determination of K_0 from multi - stage triaxial compression test. *Geotech Eng* 8:31–52
- Landva AO, Valsangkar AJ, Pelkey SG (2000) Lateral earth pressure at rest and compressibility of municipal solid waste. *Can Geotech J* 37(6):1157–1165
- Lee J, Yun TS, Lee D, Lee J (2013) Assessment of K_0 correlation to strength for granular materials. *Soils Found* 53(4):584–595
- Lee J, Lee D, Park D (2014) Experimental investigation on the coefficient of lateral earth pressure at rest of silty sands: effect of fines. *Geotech Test J* 37(6):1–13
- Levenberg E, Garg N (2014) Estimating the coefficient of at-rest earth pressure in granular pavement layers. *Transportation Geotechnics* 1(1):21–30
- Lirer S, Flora A, Nicotera MV (2011) Some remarks on the coefficient of earth pressure at rest in compacted sandy gravel. *Acta Geotech* 6(1):1–12
- Ly HB, Pham BT (2020) Prediction of shear strength of soil using direct shear test and support vector machine model. *The Open Construction and Building Technology Journal* 14(1):41–50
- Massarch R, Broms BB (1976) Lateral earth pressure at rest in soft clay. *Journal of The Geotechnical Engineering Division GT10*: 1041 - 1047.
- Mayne PW, Kulhawy FH (2003) Discussion on relationship between K_0 and overconsolidation ratio: a theoretical approach. *Geotechnique* 53(4):450–454
- Moré JJ (1978) The Levenberg-Marquardt algorithm: implementation and theory. In: *Numerical Analysis*, pp 105–116.
- Nasr AMA (2014) Experimental and theoretical studies of laterally loaded finned piles in sand. *Can Geotech J* 51:381–393
- Nguyen TA, Ly HB, Jaafari A, Pham TB (2020) Estimation of friction capacity of driven piles in clay using artificial neural network. *Vietnam Journal of Earth Sciences* 42(3):265–275
- Özer TA (2001) Determination of horizontal earth pressure and K_0 coefficient for cohesionless soils by using thin walled oedometer technique and comparison of experimental results with theoretical values. MSc Thesis, Çukurova University, Adana, Turkey (In Turkish)
- Pham BT, Nguyen MD, Ly HB et al. (2019) Development of artificial neural networks for prediction of compression coefficient of soft soil. In *Proceedings of the 5th International Conference on Geotechnics, Civil Engineering Works and Structures*, 54 : 1167–1172. https://doi.org/10.1007/978-981-15-0802-8_187
- Platt J (1998) Sequential minimal optimization: a fast algorithm for training support vector machines, *Advances in Kernel methods, Support Vector Learning*, MIT Press, Boston.
- Puri N, Prasad HD, Jain A (2018) Prediction of geotechnical parameters using machine learning techniques. *Procedia Computer Science* 125:509–517
- Quadir MA (1990) Bearing capacity of strip footing on sand. Dissertation, Bangladesh University.

- Sağlamer A (1973) Kohezyonsuz zeminlerde sükunetteki toprak basıncı katsayısının zemin parametreleri cinsinden ifadesi. Dissertation, Istanbul Technical University.
- Sağlamer A (1975) Soil parameters affecting coefficient of earth pressure at rest of cohesionless soils. Proceedings of the Istanbul Conference on Soil Mechanics and Foundation Engineering 1:9–16
- Sebastiani F (2002) Machine learning in automated text categorization. *ACM Comput Surv* 34(1):1–47
- Sharma LK, Singh R, Umrao RK, Sharma KM, Singh TN (2017) Evaluating the modulus of elasticity of soil using soft computing system. *Engineering with Computers* 33(3):497–507
- Talesnick ML (2012) A different approach and result to the measurement of K_0 of granular soils. *Geotechnique* 62(11):1041–1045
- Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres* 106(D7):7183–7192
- Teerachaiikulpanich N, Okumura S, Matsunaga K, Ohta H (2007) Estimation of coefficient of earth pressure at rest using modified oedometer test. *Soils Found* 47(2):349–360
- Tian Q, Xu Z, Zhou G, Zhao X, Hu K (2009) Coefficients of earth pressure at rest in thick and deep soils. *Min Sci Technol* 19(2):252–255
- Ting CMR, Sills GC, Wijeyesekera DC (1994) Development of K_0 in soft soils. *Geotechnique* 44(1):101–109
- Tong L, Liu L, Cai G, Du G (2013) Assessing the coefficient of the earth pressure at rest from shear wave velocity and electrical resistivity measurements. *Eng Geol* 163:122–131
- Uncuoğlu E, Laman M, Sağlamer A, Kara HB (2008) Prediction of lateral effective stresses in sand using artificial neural network. *Soils Found* 48(2):141–153
- Vapnik VN (1995) *The nature of statistical learning theory*. Springer, New York
- Wang JJ, Yang Y, Bai J, Hao JH, Zhao TL (2018) Coefficient of earth pressure at rest of a saturated artificially mixed soil from oedometer tests. *KSCE J Civ Eng* 22(5):1691–1699
- Wang Y, Akeju OV (2016) Quantifying the cross-correlation between effective cohesion and friction angle of soil from limited site-specific data. *Soils Found* 56(6):1055–1070
- Yun TS, Lee J, Lee J, Choo J (2015) Numerical investigation of the at-rest earth pressure coefficient of granular materials. *Granular Matter* 17(4):413–418
- Zhao X, Zhou G, Tian Q, Kuang L (2010) Coefficient of earth pressure at rest for normal consolidated soils. *Min Sci Technol* 20:406–410

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