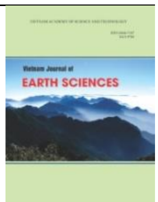
Vietnam Journal of Earth Sciences, 43(2), 189-198, https://doi.org/10.15625/2615-9783/15926

Vietnam Academy of Science and Technology

Vietnam Journal of Earth Sciences

http://www.vjs.ac.vn/index.php/jse



Estimation of shear strength parameters of soil using Optimized Inference Intelligence System

Binh Thai Pham[1*], Mahdis Amiri[2], Manh Duc Nguyen[3], Trinh Quoc Ngo[1], Kien Trung Nguyen[1], Hieu Trung Tran[1], Hoang Vu[1], Bui Thi Quynh Anh[1], Hiep Van Le[1], Indra Prakash[4]

[1]*University of Transport Technology, Hanoi, Vietnam*

[2]*Department of Watershed & Arid Zone Management, Gorgan University of Agricultural Sciences & Natural Resources, Gorgan 4918943464, Iran*

[3]*University of Transport and Communications, Hanoi, Vietnam*

[4]*DDG (R) Geological Survey of India, Gandhinagar 382010, India*

Received 8 December 2020; Received in revised form 2 February 2021; Accepted 10 March 2021

ABSTRACT

In recent years, machine learning techniques have been developed and used to build intelligent information systems for solving problems in various fields. In this study, we have used Optimized Inference Intelligence System namely ANFIS-PSO which is a combination of Adaptive Neural-Fuzzy Inference System (ANFIS) and Particle Swarm Optimization (PSO) for the estimation of shear strength parameters of the soils (Cohesion “C” and angle of internal friction “φ”). These parameters are required for designing the foundation of civil engineering structures. Normally, shear parameters of soil are determined either in the field or in the laboratory which require time, expertise and equipments. Therefore, in this study, we have applied a hybrid model ANFIS-PSO for quick and cost-effective estimation of shear parameters of soil based on the other six physical parameters namely clay content, natural water content, specific gravity, void ratio, liquid limit and plastic limit. In the model study, we have used data of 1252 soft soil samples collected from the different highway project sites of Vietnam. The data was randomly divided into 70:30 ratios for the model training and testing, respectively. Standard statistical measures: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Correlation Coefficient (R) were used for the performance evaluation of the model. Results of the model study indicated that performance of the ANFIS-PSO model is very good in predicting shear parameters of the soil: cohesion (RMSE = 0.075, MAE = 0.041, and R = 0.831) and angle of internal friction (RMSE = 0.08, MAE = 0.058, and R = 0.952).

Keywords: Adaptive Neural-Fuzzy inference system; particle swarm optimization; shear strength; soft soil; Vietnam.

1. Introduction

In general, shear strength is the capability of the soil to sustain shear stress (Das, 2021). One of the important applications of soil shear strength in geotechnical engineering is for designing and construction of civil engineering structures to withstand static and dynamic loads (Tan et al., 2019). Therefore,

*Corresponding author, Email: binhpt@utt.edu.vn

determination of the shear parameters (Cohesion "C" and angle of internal friction "φ") is required for proper designing of the civil engineering structures. Measurement of soil shear strength is usually done in two ways, direct and indirect methods. In the first group, there are a number of tools that can measure directly soil shear strength, such as shear blades, conical penetration tool, torsional shear boxes, straight shear boxes and Zhang's system, which are performed in the laboratory (Das, 2021). The procedure of checking the shear strength of the soil of an area directly is very time consuming and costly. On the other hand, it has low accuracy due to human errors and device errors. It should be noted that the applicability of most of these methods is very difficult in large areas in addition to being time consuming. For these reasons, several studies have been carried out to establish relationships of soil shear parameters with other physical properties such as plastic index, liquid limit, moisture content, amount of clay (Das, 2021). Nowadays, Arificial Intelligence (AI) and Machine Learning (ML) methods are being widely used in many scientific fields of engineering, including geotechnics (Foong, Moayedi, Lyu, 2020; Samui et al., 2019). Many researchers have used AI or ML as an advanced tool for data analysis to build models for predicting soil shear strength (Bui, Hoang, Nhu, 2019; Ly, Pham, 2020; Pham, Hoang, Nguyen, Bui, 2018). This is because models based on ML performed excellently in nonlinear modeling. We can also select a large number of independent variables for processing, which predict soil shear strength (Nhu et al., 2020; Pham et al., 2020). On the other hand, AI/ML models are also very flexible which can predict the results according to the input data (Wei Chen et al., 2017; Zhou et al., 2020). In the soil mechanics field, Artificial Neuron Network (ANN) method has been used by many researchers for the estimation of shear parameters based on the other physical parameters of soil (Sharma et al., 2017). Support Vector Regression (SVR) and ANN have also been used and compared to estimate shear parameters (Kuo et al., 2009). In addition, other ML algorithms such as Classification and Regression Tree (CART) analysis, a generalized linear (GL) model, Chi-squared Automatic Interaction Detection (CHAID), were used to identify factors affecting shear strength (Kanungo et al., 2016). Random Forest (RF) algorithm was also used for predicting shear strength (Breiman, 2001; Weiting Chen et al., 2014). Other algorithms used in solving geotechnical engineering problems and shear strength are: Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Ant Colony Optimization (ACO), Firefly Algorithm (FA), and Artificial Bee Colony (ABC) (Armaghani et al., 2015; Kalatehjari et al., 2014; Salehin, 2017).

In general, development and applications of AI/ML methods are continuous process. Therefore, in this study, we have applied a hybrid model ANFIS-PSO which is a combination of Adaptive Neural-Fuzzy Inference System (ANFIS) and Particle Swarm Optimization (PSO) in estimating shear parameters based on the other six physical parameters of the soils namely clay content, natural water content, specific gravity, void ratio, liquid limit and plastic limit. The data soil parameters were collected from various highway projects of Vietnam. The main difference of this study compared with the published works is that it is the first time ANFIS was combined with PSO for the prediction of shear strength parameters of soil. Weka software was used for the model development and data analysis.

2. Materials and methods

2.1. Data used

In this study, a total of 1252 soft soil samples data was collected from various highway projects of Vietnam namely Riviera

Point complex project (700 samples), Da Nang - Quang Ngai expressway project (145 samples), Ha Noi - Hai Phong national highway project (251 samples), and Hai Phong - Ninh Binh costal highway project (154 samples). Out of these data, two dependent variables "C" and "φ" determined from the direct shear tests were used as "outputs", and six independent variables: clay content, natural water content, specific gravity, void ratio, liquid limit, plastic limit determined at laboratory were used as input variables in the model study for the estimation of shear parameters as per several published works (Ly & Pham, 2020; Nguyen et al., 2021). Table 1 shows the summarized values of soil parameters used in this study.

Normalization or scaling of the data was done to minimize information clutter and error in the model study. As a part of normalization process the values of numeric columns in the soil dataset were changed to a common scale, without distorting differences in the ranges of values that is between 0 and 1. The normalization of the data in columns was performed by the following equation:

$$X^{scaled} = x^{raw} - \beta / \alpha - \beta \tag{1}$$

where α and β are the most (maximum) and low (minimum) values of the parameter x. Splitting of the soil data was randomly done in 70:30 ratios for training (70% data) and testing (30% data) for the model study. The ratio of 70:30 of splitting the data was selected based on the experience of researchers in similar studies (Nguyen et al., 2021).

Table 1. Minimum, maximum, average and Standard Deviation (StD) values of soil parameters determined in the laboratory

Parameters	Abbreviation	Unit	Minimum	Maximum	Average	StD
Natural water content	W	%	15.9	163.3	41.346	25.126
Void Ratio	e	-	0.462	4.313	1.168	0.65
Specific Gravity	Gs	g/cm^3	2.53	2.75	2.674	0.032
Liquid limit	LL	%	18	156.87	43.762	21.293
Plastic Limit	PL	%	9.82	65.82	24.593	9.933
Clay content	-	%	0.2	82.4	23.053	15.335
Internal friction angle	φ	radian	0	0.55	0.23	0.138
Cohesion	c	kPa	0.04	59.6	7.899	8.273

2.2. Methods used

2.2.1. ANFIS

The combination of fuzzy inference systems based on logical rules and the method of Artificial Neural Networks (ANNs) that have the ability to extract knowledge from numerical information, leads to the presentation of an Adaptive Neuro-Fuzzy Inference System (ANFIS) (Nwobi-Okoye et al., 2019). ANFIS uses neural network algorithms and fuzzy logic to design a nonlinear mapping between the input and output space (Jaypuria et al., 2019). As a powerful tool, this system has the ability to predict results using existing numerical data. In the algorithm, the first layer is the input nodes. In this layer, the degree of membership of the input nodes (the extent to which each input belongs) to different fuzzy intervals is determined by the user using the membership function. Modeling operations are performed in the second to fourth layer. By multiplying the input values of each node by each other, the weight of each rule in the second layer is obtained. In the third layer, the relative weights of the rules are calculated. In the fourth layer, each node has a node function and is connected to all inputs and a node in the third layer. The last layer is the network output, which aims to summarize all the output of the rules (Zhang et al., 2021).

2.2.2. PSO

Particle Swarm Optimization (PSO) algorithm is an evolutionary computational method of collective intelligence. The algorithm is inspired by the social behaviors of a group of birds and a group of fish in finding food (Eberhart & Kennedy, 1995). The basis of this algorithm is to repeat the search in the problem space by a random population, in each iteration, the objective function is evaluated and then the best position of each particle and the best position of all particles are determined as the best local position and the best general position, respectively (Wang et al., 2020). In fact, particle motion in this algorithm depends on two factors: individual motion and collective motion, and the combination of these two motions leads to the creation of an efficient model to find the best target point in optimization problems. As mentioned, the particle swarm algorithm is affected by both cognitive and social component (Guo et al., 2020). Each particle, with two vectors of velocity and position, represents an answer in the next D space of the problem. In this regard, according to the two parameters, the best condition met by the particle *pbest* and the best state encountered in all particles *gbest* is determined by the motion of each particle in the search space (Cockshott & Hartman, 2001):

In this study, PSO was used to optimize the bias and weights of ANFIS to create the hybrid model namely ANFIS-PSO for prediction of soil shear parameters. Hyper-parameters used in the model include: the number of cluster (10), inertia Weight (0.4), the number of population (30), the number of iterations (500).

2.2.3. Validation indicators

In the present study, in order to evaluate performance of the ANFIS-PSO algorithm in predicting shear parameters, statistical measures: Root Mean Square Error (RMSE), Mean Absolue Error (MAE) and Correlation Coefficient (R) were used. R is the correlation coefficient between two independent and dependent variables, assuming that all independent variables affect the dependent variable (Qasim et al., 2020). RMSE is the most important statistical quantity for evaluating models, which is sensitive to outlier data and indicates non-systematic errors (mistakes) (Li & Heap, 2014; Panem et al., 2020). The closer this quantity is to zero, the lower the error of the model used. MAE should ideally be zero, that, positive and negative values indicating overestimation and underestimation, respectively, of the actual value. This parameter represents the accuracy of the method and the average amount of error (Li & Heap, 2014; Panem et al., 2020). In general, if all the predicted values are equal to the measured values, then the R index is equal to 1 and the RMSE and MAE indices are equal to zero (Amaro et al., 2021). All three evaluation indicators (RMSE, MAE and R) are calculated according to the following equations:

$$RMSE= \sqrt{\sum(z^*(X_i)- z(X_i)^2/n} \quad (2)$$

$$MAE= 1/n\sum|z^*(X_i)- z(X_i)| \quad (3)$$

$$R= \sum(P_i- \bar{P})(Y_i- \bar{Y})/\sqrt{\sum(P_i- \bar{P})}\sqrt{\sum(Y_i- \bar{Y})} \quad (4)$$

In these relations (RMSE and MAE), n is the number of samples, $z^*(X_i)$ is the estimated value, and $z(X_i)$ is the value measured at the known point. In the R equation, Y_i and $\bar{Y}$ are the measure and average amounts of soil parameters shear strength, P_i and $\bar{P}$ are efficiency amounts from the model respectively.

3. Results and discussion

3.1. Prediction of internal friction angle of soil using ANFIS-PSO model

In this section, ANFIS-PSO method was used to predict value of "φ" (Output/ Target) based on six parameters (clay content, natural water content, specific gravity, void ratio, liquid limit and plastic limit) as inputs in the model. Cost function analysis of the model using statistical measures for 500 iteration of model is presented in Fig. 1. Results indicate that RMSE (0.08) and MAE (0.06) values are near zero and R value (0.95) near 1. The closer to zero values of the RMSE and MAE and closer the value of R to 1 indicate higher prediction accuracy of the model. The correlation analysis results of the model for training and testing data are 0.957 and 0.952, respectively, which indicated a high correlation between actual and predicted values (Fig. 2). We have also done error analysis of ANFIS-PSO model using training dataset (Fig. 3) showing the values of RMSE and MAE: 0.07 and 0.052, respectively, whereas for the testing dataset, these values 0.08 and 0.058, respectively. On the other hand, the Mean Error (ME) are (-) 0.0005 and (-) 0.0085 for training and testing datasets, respectively. Also, the value of error StD are, 0.086 and (-) 0.0085 for training and testing datasets, respectively. These errors for the model are very low and thus performance of the model is excellent.

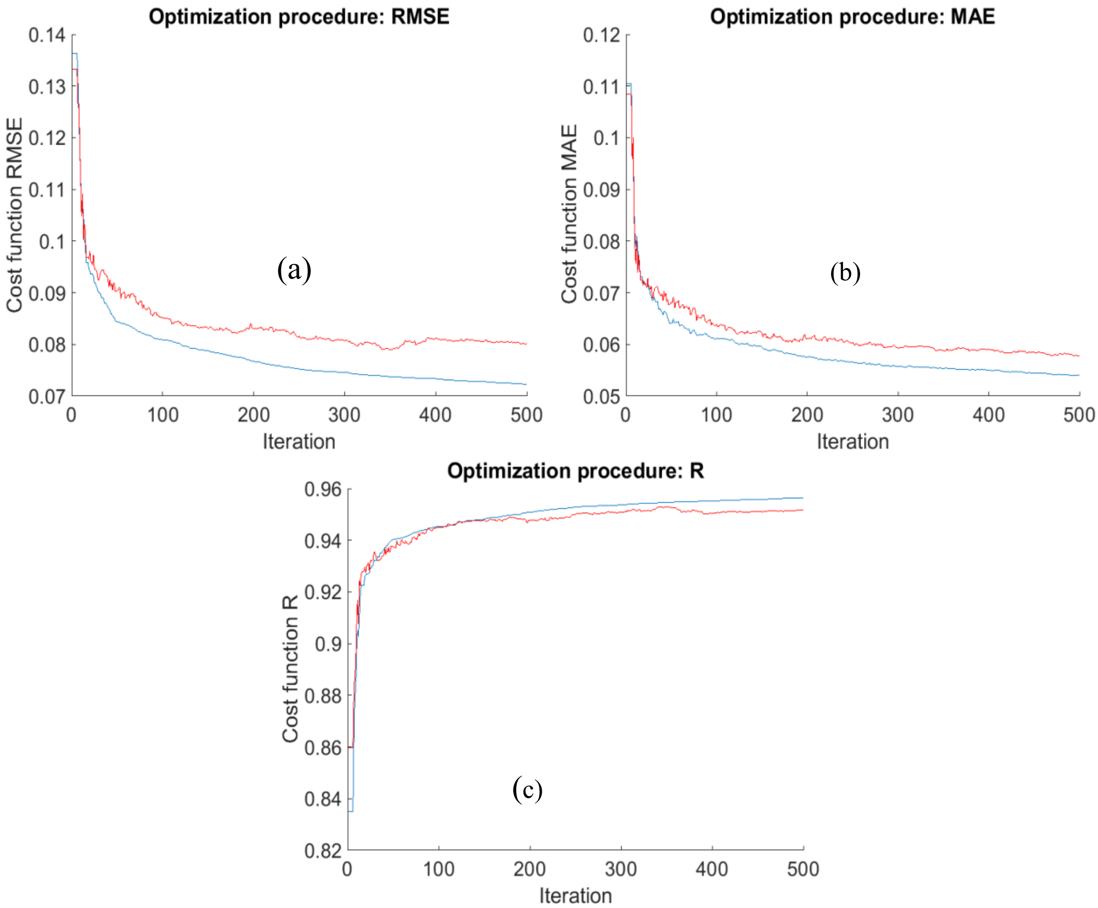


Figure 1. Cost function analysis of ANFIS-PSO for prediction of internal friction angle of soil using (a) RMSE, (b) MAE, (c) R

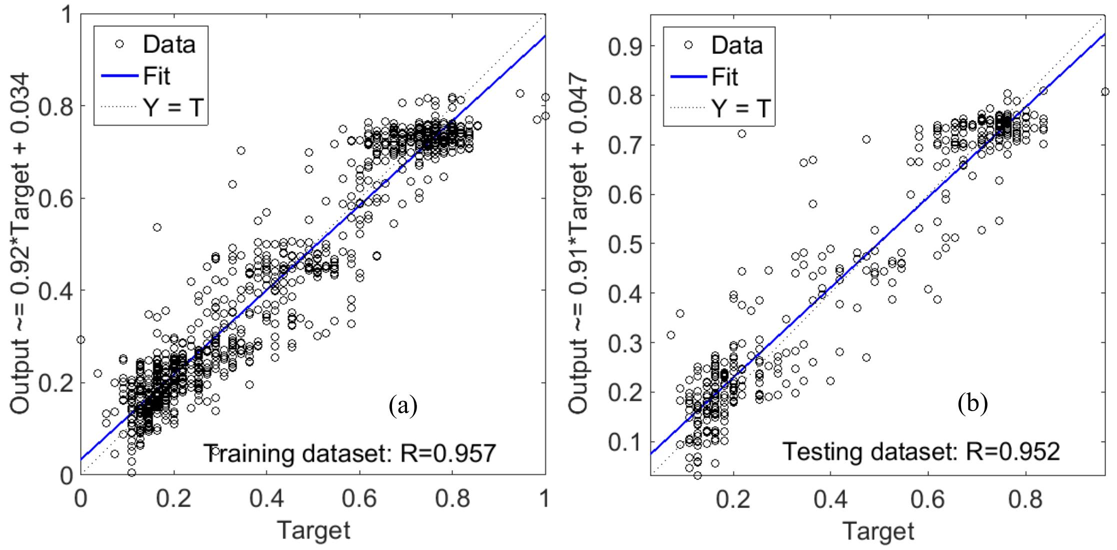


Figure 2. Correlation analysis of actual and predicted outputs using ANFIS-PSO for prediction of internal friction angle of soil: (a) training dataset and (b) testing dataset

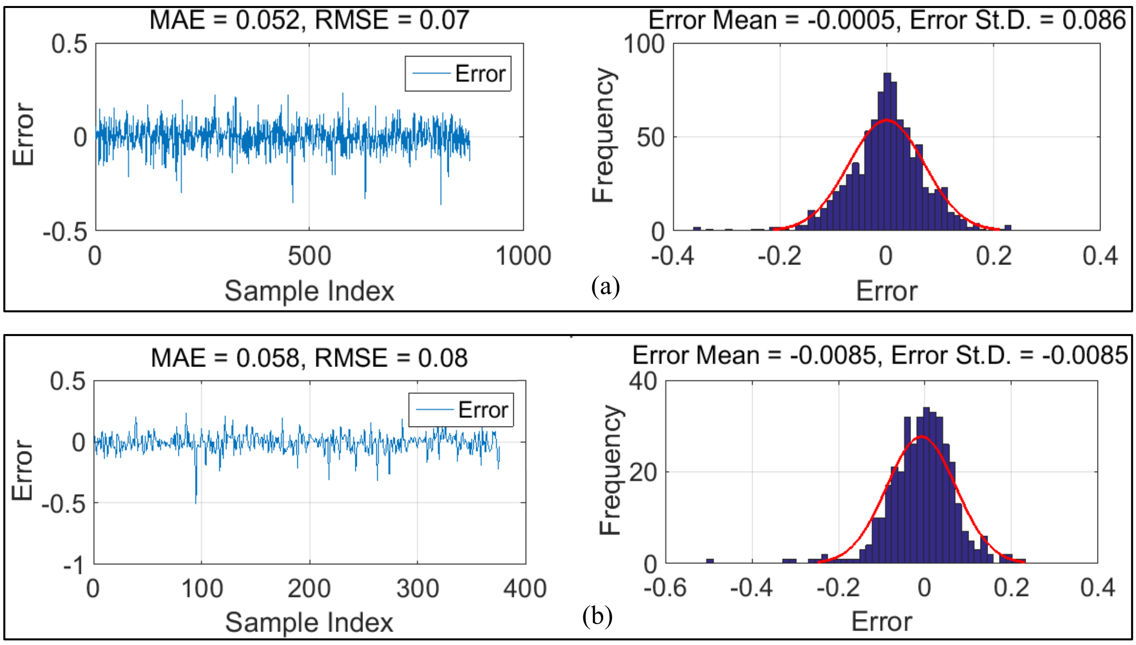


Figure 3. Error analysis of ANFIS-PSO for prediction of internal friction angle of soil: (a) training dataset and (b) testing dataset

3.2. Prediction of cohesion of soil using ANFIS-PSO model

In the model study, clay content, natural water content, specific gravity, void ratio, liquid limit and plastic limit were used as input parameters and ‘C’ as target or output parameter. Cost function analysis for 500 iteration of the model using statistical measures is presented in Fig. 4. Results indicated that RMSE (0.078) and MAE

(0.043) values are near zero and R value (0.88) near 1. Thus, accuracy of model in predicting the "C" value is excellent. The correlation analysis results of the model for training and testing data are 0.871 and 0.831, respectively, which indicate very good correlation between actual and predicted values (Fig. 5). Error analysis (Fig. 6) shows the Mean Error (ME) value is 0.0031 and (-) 0.0016 for training and testing datasets, respectively and error StD is 0.075 and (-) 0.0016 for training and testing datasets, respectively. These errors for the model are very low and thus performance of the studied model is very good to excellent.

In general, performance of the ANFIS-PSO model is good and excellent for the prediction of shear strength parameters of the soil. It is reasonable as the advantage of using ANFIS-PSO model is that it cannot fall into the optimal local trap by using PSO algorithm and increase the accuracy and global search capability for ANFIS training (Noushabadi et al., 2020). In addition, the ANFIS has both advantages of the fuzzy principle (smoothness property) and the neural networks training structure (adaptability property), which can enhance its predictive capability in engineering applications (Walia et al., 2015). The results of this study are comparable with other published works, which stated that ANFIS and ANFIS-PSO are great tools for prediction problems (Besalatpour et al., 2012; Ghanizadeh & Tavana Amlashi, 2018).

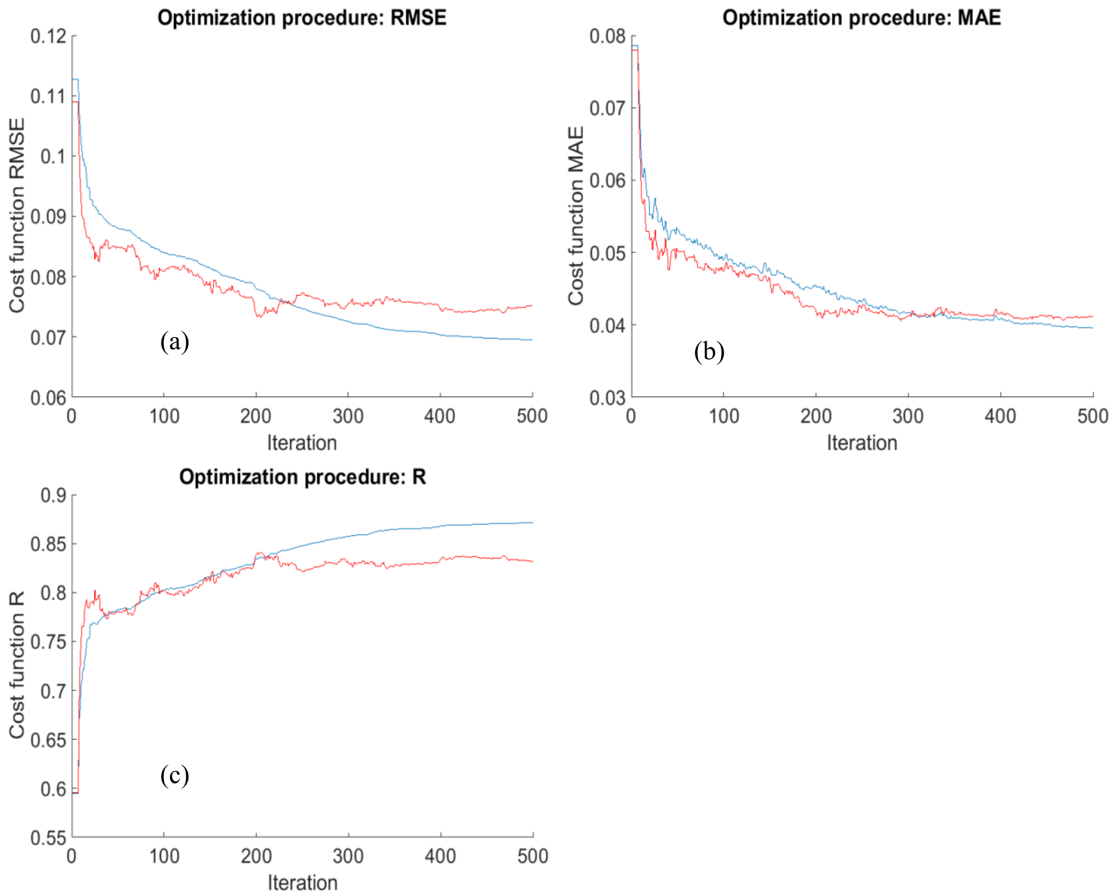


Figure 4. Cost function analysis of ANFIS-PSO for prediction of cohesion of soil using (a) RMSE, (b) MAE, and (c) R

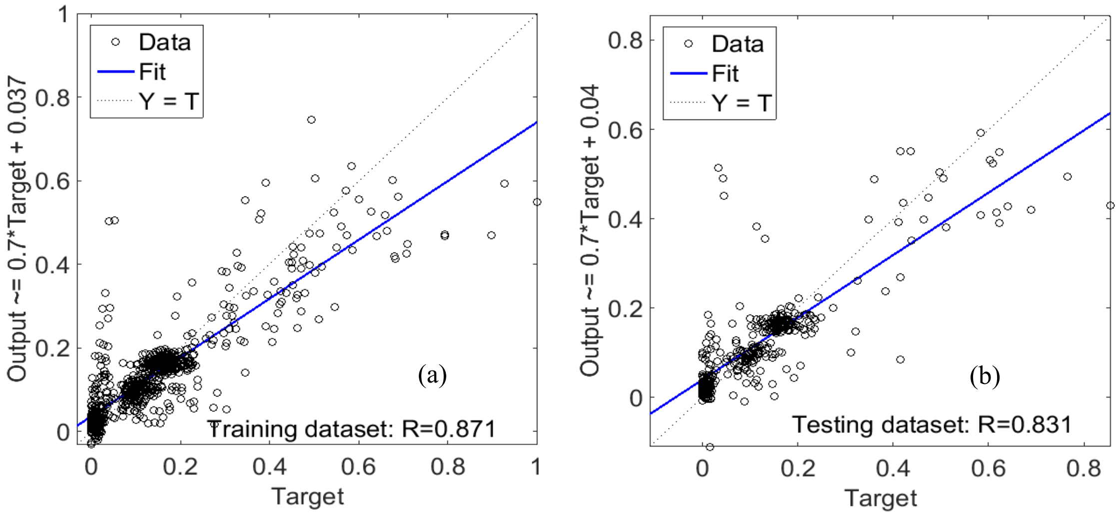


Figure 5. Correlation analysis of actual and predicted outputs using ANFIS-PSO for prediction of cohesion of soil: (a) training dataset and (b) testing dataset

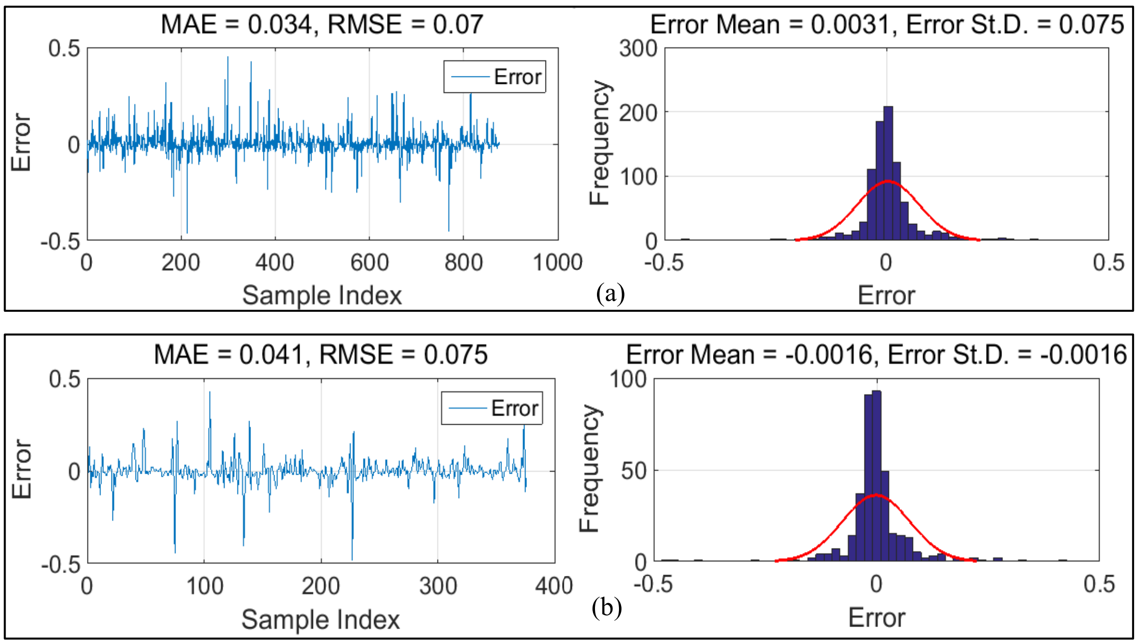


Figure 6. Error analysis of ANFIS-PSO for prediction of cohesion of soil: (a) training dataset and (b) testing dataset

4. Conclusions

In this study we have applied a hybrid model ANFIS-PSO for the estimation of shear parameters ("C" and "φ") based on the six physical parameters: clay content, natural water content, specific gravity, void ratio, liquid limit and plastic limit which can be determined in the laboratory relatively easily and with less cost. Model results indicated that performance of the ANFIS-PSO algorithm in predicting shear parameters of soil is very good to excellent. Therefore, this hybrid model can be used for the accurate estimation of "C" and "φ" values of soft soil for designing and safe construction of the civil engineering structures without directly measuring these parameters. The limitation of this study is that the predictive capability of the ANFIS-PSO model was validated on selected types of soil. It would be better to evaluate performance of this model on various types of soil with different combination of input variables. In addition, the model development is continuous process and thus there is always scope in the improvement of accuracy in determining different parameters based on new algorithms. In the future, we will compare the results of ANFIS-PSO model with other models developed and applied for the estimation of soil strength parameters to improve performance, if required, by considering more soil parameters in the modeling.

Acknowledgements

This research is funded by Ministry of Transport, project title "Building Big Data and development of machine learning models integrated with optimization techniques for prediction of soil shear strength parameters for construction of transportation projects" under grant number DT 203029.

References

Amaro J., Rosado D.J.M., Mendiburu A.Z., dos Santos L.R., de Carvalho Jr J.A., 2021. Modeling of syngas composition obtained from fixed bed gasifiers using Kuhn-Tucker multipliers. Fuel, 287, 119068.

Armaghani D.J., Tonnizam Mohamad E., Momeni E., Monjezi M., Sundaram Narayanasamy M., 2015. Prediction of the strength and elasticity modulus of granite through an expert artificial neural network. Arabian Journal of Geosciences, 9, 48.

Besalatpour A., Hajabbasi M., Ayoubi S., Afyuni M., Jalalian A., Schulin R., 2012. Soil shear strength prediction using intelligent systems: artificial neural networks and an adaptive neuro-fuzzy inference system. Soil science and plant nutrition, 58, 149-160.

Breiman L., 2001. Random Forests. Machine Learning, 45, 5-32.

Bui D.T., Hoang N.-D., Nhu V.-H., 2019. A swarm intelligence-based machine learning approach for predicting soil shear strength for road construction: a case study at Trung Luong National Expressway Project (Vietnam). Engineering with Computers, 35, 955-965.

Chen W., Panahi M., Pourghasemi H.R., 2017. Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. CATENA, 157, 310-324.

Chen W., Wang Y., Cao G., Chen G., Gu Q., 2014. A random forest model based classification scheme for neonatal amplitude-integrated EEG. Biomed Eng Online, 13 Suppl 2, S4-S4.

Cockshott A.R., Hartman B.E., 2001. Improving the fermentation medium for Echinocandin B production part II: Particle swarm optimization. Process Biochemistry, 36, 661-669.

Das B.M., 2021. Principles of geotechnical engineering. Cengage learning.

Eberhart R., Kennedy J., 1995. A new optimizer using particle swarm theory, MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science. Publishing, 39-43.

Foong L.K., Moayedi H., Lyu Z., 2020. Computational modification of neural systems using a novel stochastic search scheme, namely evaporation rate-based water cycle algorithm: an application in geotechnical issues. Engineering with Computers.

Ghanizadeh A.R., Tavana Amlashi A., 2018. Prediction of fine-grained soils resilient modulus using hybrid ANN-PSO, SVM-PSO and ANFIS-PSO methods. Quarterly Journal of Transportation Engineering, 9, 159-181.

Guo X., Ji M., Zhao Z., Wen D., Zhang W., 2020. Global path planning and multi-objective path control for unmanned surface vehicle based on modified particle swarm optimization (PSO) algorithm. Ocean Engineering, 216, 107693.

Jaypuria S., Ranjan Mahapatra T., Jaypuria O., 2019. Metaheuristic Tuned ANFIS Model for Input-Output Modeling of Friction Stir Welding. Materials Today: Proceedings, 18, 3922-3930.

Kalatehjari R., Ali N., Kholghifard M., Hajihassani M., 2014. The effects of method of generating circular slip surfaces on determining the critical slip surface by particle swarm optimization. Arabian Journal of Geosciences, 7, 1529-1539.

Kanungo D.P., Sharma S., Pain A., 2014. Artificial Neural Network (ANN) and Regression Tree (CART) applications for the indirect estimation of unsaturated soil shear strength parameters. Frontiers of Earth Science, 8, 439-456.

Kiran S., Lal B., Tripathy S., 2016. Shear strength prediction of soil based on probabilistic neural network. Indian Journal of Science and Technology, 9.

Kuo Y.L., Jaksa M.B., Lyamin A.V., Kaggwa W.S., 2009. ANN-based model for predicting the bearing capacity of strip footing on multi-layered cohesive soil. Computers and Geotechnics, 36, 503-516.

Li J., Heap A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, 173-189.

Ly H.-B., Pham B.T., 2020. Prediction of shear strength of soil using direct shear test and support vector machine model. The Open Construction and Building Technology Journal, 14.

Nguyen Q.H., Ly H.-B., Ho L.S., Al-Ansari N., Le H.V., Tran V.Q., Prakash I., Pham B.T., 2021. Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. Mathematical Problems in Engineering.

Nhu V.-H., Hoang N.-D., Duong V.-B., Vu H.-D., Bui D.T., 2020. A hybrid computational intelligence approach for predicting soil shear strength for urban housing construction: a case study at Vinhomes Imperia project, Hai Phong city (Vietnam). Engineering with Computers, 36, 603-616.

Noushabadi A.S., Dashti A., Raji M., Zarei A., Mohammadi A.H., 2020. Estimation of cetane numbers of biodiesel and diesel oils using regression and PSO-ANFIS models. Renewable Energy, 158, 465-473.

Nwobi-Okoye C.C., Ochieze B.Q., Okiy S., 2019. Multi-objective optimization and modeling of age hardening process using ANN, ANFIS and genetic algorithm: Results from aluminum alloy A356/cow horn particulate composite. Journal of Materials Research and Technology, 8, 3054-3075.

Panem C., Gad V.R., Gad R.S., 2020. Sensor's data transmission with BPSK using LDPC (Min-Sum) error corrections over MIMO channel: Analysis over RMSE and BER. Materials Today: Proceedings, 27, 571-575.

Pham B.T., Hoang T.-A., Nguyen D.-M., Bui D.T., 2018. Prediction of shear strength of soft soil using machine learning methods. Catena, 166, 181-191.

Pham B.T., Nguyen-Thoi T., Ly H.-B., Nguyen M.D., Al-Ansari N., Tran V.-Q., Le T.-T., 2020. Extreme learning machine based prediction of soil shear strength: a sensitivity analysis using Monte Carlo simulations and feature backward elimination. Sustainability, 12, 2339.

Qasim M., Park S., Kim J.-O., 2020. Empirical Reynolds number model for drag coefficient determination of ballasted flocs. Journal of Water Process Engineering, 101803.

Salehin S., 2017. Investigation into engineering parameters of marls from Seydoon dam in Iran. Journal of Rock Mechanics and Geotechnical Engineering, 9, 912-923.

Samui P., Hoang N.-D., Nhu V.-H., Nguyen M.-L., Ngo P.T.T., Bui D.T., 2019. A New Approach of Hybrid Bee Colony Optimized Neural Computing to Estimate the Soil Compression Coefficient for a Housing Construction Project. Applied Sciences, 9, 4912.

Sharma L.K., Singh R., Umrao R.K., Sharma K.M., Singh T.N., 2017. Evaluating the modulus of elasticity of soil using soft computing system. Eng. with Comput., 33, 497-507.

Tan H., Chen F., Chen J., Gao Y., 2019. Direct shear tests of shear strength of soils reinforced by geomats and plant roots. Geotextiles and Geomembranes, 47, 780-791.

Walia N., Singh H., Sharma A., 2015. ANFIS: Adaptive neuro-fuzzy inference system-a survey. International Journal of Computer Applications, 123.

Wang X., Ting D.S.K., Henshaw P., 2020. Mutation particle swarm optimization (M-PSO) of a thermoelectric generator in a multi-variable space. Energy Conversion and Management, 224, 113387.

Zhang Z., Peng B., Luo C.-H., Tai C.-C., 2021. ANFIS-GA system for three-dimensional pulse image of normal and string-like pulse in Chinese medicine using an improved contour analysis method. European Journal of Integrative Medicine, 101301.

Zhou W., Dong H., Liang Y., 2020. The deterministic dendritic cell algorithm with Haskell in earthquake magnitude prediction. Earth Science Informatics, 13, 447-457.