AI-Based Prediction of Strength and Tensile Properties of Expansive Soil Stabilized with Recycled Ash and Natural Fibers

Abolfazl Baghbani, Firas Daghistani, Katayoon Kiany and Mohamad Mahdi Shalchiyan

February 19, 2023

AI-Based Prediction of Strength and Tensile Properties of Expansive Soil Stabilized with Recycled Ash and Natural Fibers

Abolfazl Baghbani[1,a *], Firas Daghistani[2,b] , Katayoon Kiany[3,c] and Mohamad Mahdi Shalchiyan[4,d]

[1]School of engineering, Deakin University, Geelong, VIC, Australia

[2]Department of Civil Engineering, La Trobe University, Melbourne, VIC, Australia; Also with Civil Engineering Department, University of Business and Technology, Jeddah, Saudi Arabia

[3]School of Design, University of Melbourne, Melbourne, VIC, Australia

[4]School of Engineering, Guilan University, Rasht, Iran

[a]abaghbani@deakin.edu.au, [b]f.daghistani@latrobe.edu.au, [c]kianyk@student.unimelb.edu.au, [d]mohamad.shalchiyan@gmail.com

* corresponding author

Keywords: Artificial intelligence; Soil stabilizer; Recycled ash, Recycled fiber, PLS, CRRF.

Abstract. This study investigated the uniaxial compressive strength (UCS) and split tensile strength of a mixture of soil and recycled ash and natural fibers using two different methods, partial least squares (PLS) and classification and regression random forest (CRRF). The study analyzed a dataset of 20 sets with five inputs and two outputs, and the importance of the input parameters was evaluated. The performance of the PLS and CRRF models was assessed, and it was found that the CRRF model outperformed the PLS model. The study also revealed the most and least important parameters in predicting the split tensile strength and UCS in both models. The findings of this study have implications for the use of soil and recycled ash mixtures with natural fibers in construction applications.

Introduction

Expansive soils are a type of soil that undergoes significant volume changes due to fluctuations in moisture content [1-3]. These volume changes can cause significant damage to buildings and infrastructure, making the stabilization of expansive soils an essential area of research. In recent years, researchers have been investigating the use of recycled ash and natural fibers as stabilizers for expansive soils to improve their strength and durability [4].

Expansive soils are a type of soil that undergoes significant volume changes due to fluctuations in moisture content. These soils are characterized by high clay content and can cause significant damage to buildings and infrastructure due to swelling and shrinking [1]. The swelling and shrinking of expansive soils are due to the clay particles' ability to absorb and release water, causing significant stresses and strains [5-7].

Recycled ash is a waste material generated from the combustion of coal in thermal power plants [8]. It contains various elements that can be beneficial for soil stabilization, including silica, alumina, and iron oxide. The use of recycled ash in soil stabilization has been studied extensively, and its potential as a soil stabilizer has been recognized [8]. Natural fibers, such as coir, sisal, jute, and kenaf, are obtained from plant sources and have been used in various applications, including soil stabilization [9]. Natural fibers are biodegradable and eco-friendly, making them a sustainable alternative to synthetic materials.

Several recent studies have investigated the use of recycled ash and natural fibers as stabilizers for expansive soil [10-12]. Tiwari et al. [11] investigated the use of recycled ash and natural fibers to control the strength and durability of expansive soil. The soil was stabilized with different percentages of bottom ash and reinforced with coir fibers. The results showed that the use of bottom ash and coir fibers improved the soil's strength and durability, and the approach is a sustainable and economical solution for stabilizing pavement subgrades. Another study by Punthutaecha et al. [10] investigated the use of recycled ash and natural fibers to control the strength and durability of expansive soil. The soil was stabilized with different percentages of bottom ash and reinforced with coir fibers. The study assessed the swelling behavior, mechanical and chemical properties, and durability of the stabilized soil through freeze-thaw cycles. The results showed that the use of bottom ash and coir fibers improved the soil's strength and durability, and the approach is a sustainable and economical solution for stabilizing pavement subgrades.

A variety of factors, including soil density, fly content can affect the strength of soil with recycled ash and natural fibers. By using traditional methods, it is impossible to produce a comprehensive model with sufficient input parameters due to the multiplicity of parameters and their non-linearity. Artificial intelligence is one of the modern solutions to this problem. Using artificial intelligence methods, relationships between parameters can be found with a high degree of accuracy without prior knowledge [13]. In various civil engineering fields, such as slope stability [14-15], road construction and tunnelling [16-17], soil cracking [18-20], artificial intelligence methods have been used successfully. The strength of mixture of soil and recycled ash and natural fibers has not yet been predicted using AI. For the first time, this study examined the uniaxial compressive strength (UCS) and split tensile strength of mixture of soil and recycled ash and natural fibers using a statistical method, i.e., partial least squares (PLS), and an artificial intelligence method, i.e., classification and regression random forest (CRRF). In this study, a 20-sets database with five inputs, including bottom ash, coir fibers (CF) content, electrical conductivity (EC), PH and Calcium content, and two outputs, namely uniaxial compressive strength (UCS) and split tensile strength, were considered. The importance of input parameters has also been evaluated after modelling.

Database Collection and Processing

Experiment and data collection

In order to investigate the effect of recycled ash and natural fibers on soil, a database consisting of 20 tests to determine the soil uniaxial compressive strength (UCS) and split tensile strength of the mixture was collected from the study conducted by Tiwari et al. [11]. Table 1 presents the statistical information from the collected database. The study conducted by Tiwari et al. [11] investigated the effects of combining recycled ash with natural fibers on soil strength and durability. The bottom ash (BA) was used for chemical treatment of the soil and coal fibers (CF) were used as reinforcement against tensile cracks.

Table 1. Statistical information of database

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Split tensile strength (kPa)	20	37.500	180.000	101.450	44.609
UCS (kPa)	20	0	20	140.000	470.000
Bottom Ash (%)	20	0.000	20.000	10.000	7.255
Coir Fibers (CF) (%)	20	0.000	1.000	0.438	0.379
Electrical Conductivity, EC (mS/cm)	20	1.750	5.500	4.100	1.447
PH	20	10.400	13.230	12.186	1.037

Calcium Content (%)	20	0.000	7.500	4.420	2.876

Preparation of the data for AI modelling

A linear normalization of the database was performed before it was used for modelling in AI methods. The linear normalization is represented by Eq. 1. A parameter's maximum value is considered equal to 1, its minimum value is considered equal to 0, and the rest of the values are distributed between 0 and 1. It is important to normalize data in order to improve model accuracy since different parameters have different units.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

where X_{max}, X_{min}, X and X_{norm} are maximum, minimum, actual, and normalized values, respectively.

A database is divided into two parts: training (80%) and testing (20%). A statistical analysis of two databases is presented in Tables 2 and 3. Statistics in two databases are generally similar. AI Model accuracy can be improved as a result of this similarity.

Table 2. Statistical information of training database

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Split tensile strength (kPa)	16	37.500	180.000	97.063	41.751
UCS (kPa)	16	140.000	470.000	265.938	103.073
Bottom Ash (%)	16	0.000	20.000	9.688	7.181
Coir Fibers (CF) (%)	16	0.000	1.000	0.375	0.365
Electrical Conductivity, EC (mS/cm)	16	1.750	5.500	4.047	1.427
PH	16	10.400	13.230	12.162	1.016
Calcum Content (%)	16	0.000	7.500	4.297	2.846

Table 3. Statistical information of testing database

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Split tensile strength (kPa)	4	46.000	175.000	119.000	58.086
UCS (kPa)	4	155.000	430.000	306.250	117.995
Bottom Ash (%)	4	0.000	20.000	11.250	8.539
Coir Fibers (CF) (%)	4	0.250	1.000	0.688	0.375
Electrical Conductivity, EC (mS/cm)	4	1.750	5.500	4.313	1.737
PH	4	10.400	13.230	12.283	1.281
Calcum Content (%)	4	0.000	7.500	4.913	3.388

Data-driven modeling

Partial least squares (PLS)

PLS (partial least squares) is a statistical approach commonly used for multivariate data analysis. This is a type of regression analysis used to model the relationship between independent variables (X) and dependent variables (Y). PLS is particularly useful when there is a large number of predictor variables, which can cause problems with traditional regression methods, such as multicollinearity. It is possible to use PLS for both linear and nonlinear relationships between X and Y. In PLS, latent variables are extracted from the predictor variables (X) and the response variables (Y) that explain the maximum covariance between them. PLS components are linear combinations of the original variables, which are calculated by maximizing the covariance between X and Y. A PLS can be applied to a wide variety of fields, including engineering, chemistry, biology, finance, and marketing. In fields with large amounts of data, this method is used to identify the most significant variables that are driving the relationship between the predictor variables and the response variables.

Classification and regression random forest (CRRF)

The random forest algorithm is a powerful machine learning algorithm that can be used for both classification and regression tasks. Using this method, multiple decision trees are combined to create a more accurate and stable prediction.

In a random forest classification model, multiple decision trees are generated, each trained on a randomly selected subset of training data and a randomly selected subset of predictor variables. To determine the final predicted class, the decision trees are combined by taking a majority vote based on their predictions. The method is effective at reducing overfitting as well as improving the model's accuracy and generalizability.

Additionally, in a random forest regression model, the algorithm creates multiple decision trees, each trained on a randomly selected subset of the training data and a randomly selected subset of the predictor variables. By taking the average of the predictions made by each decision tree, the final predicted value is determined.

- There are several advantages to using random forest models, including:
- In contrast to other machine learning algorithms, they are robust against noise and overfitting.
- Large datasets with a large number of predictor variables can be handled by these programs.
- This allows them to identify important predictor variables that are driving the prediction.

In various fields such as engineering, finance, healthcare, marketing, and natural language processing, random forest models are commonly used. Additionally, they are used in feature engineering, where they can identify the most important variables to include in a model.

Results

Partial least squares (PLS)

After implementing a large number of PLS models and changing various parameters, the best and most optimal PLS model was identified. According to Figs 1 and 2, the actual values are compared to the values predicted by the best PLS model for split tensile strength and UCS, respectively. The PLS model was able to predict both outputs with an almost perfect accuracy based on the obtained results.

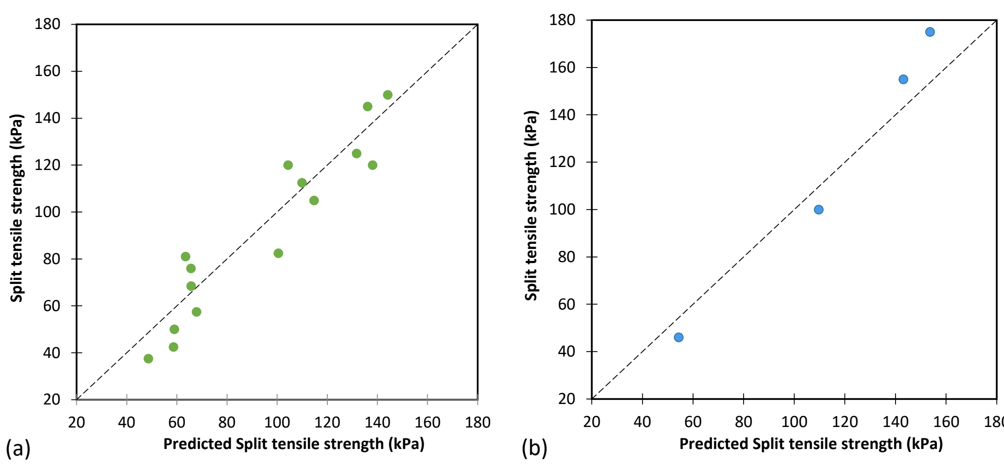

Fig. 1. The results of PLS for predicting split tensile strength

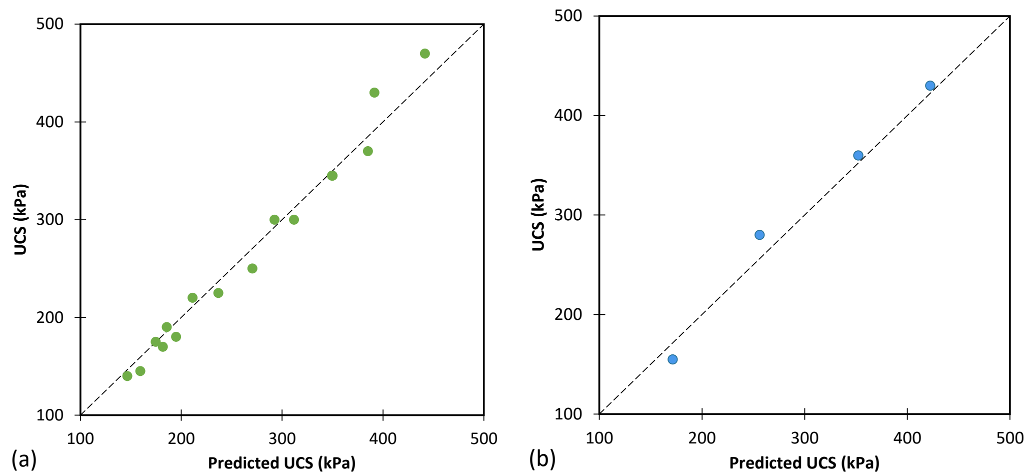

Fig. 2. The results of PLS for predicting UCS

In Table 4, the model's accuracy is demonstrated by the values of performance metrics for predicting split tensile strength and UCS, respectively. The results provide performance metrics of a predictive model on both training and testing data. The PLS model performed well with high accuracy and good generalization to unseen data. The metrics include Mean Absolute Error (MAE) and R^2. According to these metrics, the model performs well both on training and testing databases, with testing data showing slightly worse performance. In general, the PLS model is an effective tool for predicting outputs.

Table 4. The performance of PLS model

Performance metrics	Split tensile strength		UCS	
	Training	Testing	Training	Testing
MAE	10.789	12.808	13.932	12.748
R^2	0.914	0.925	0.977	0.975

Classification and regression random forest (CRRF)

The performance of the CRRF model can be affected by a variety of parameters. The effective parameters of the CRRF model were repeatedly changed to find the best model. Table 5 presents the values of the best CRRF model. Figs 3 and 4 illustrate the actual results of the test compared to the values predicted by the CRRF model. According to the results, the CRRF model is relatively accurate and is capable of predicting both split tensile strength and UCS.

Table 5. The specifications of the best CRRF.

Trees parameters					Forest parameters		
Min. node size	Min. son size	Max depth	Mtry	CP	Sampling	Sample size	Number of trees
2	1	12	2	0.00001	Random with replacement	16	1000

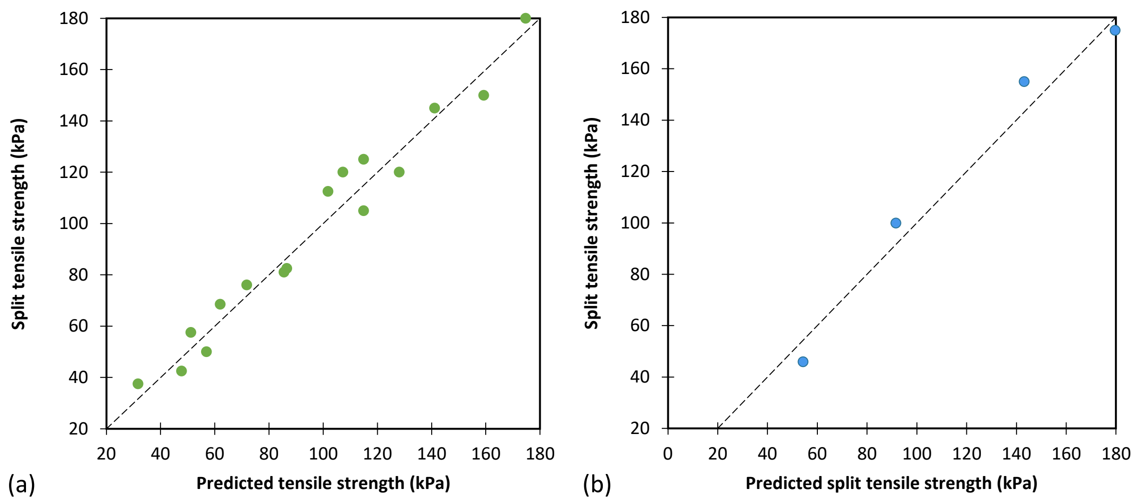

Fig. 3. The results of CRRF for predicting split tensile strength

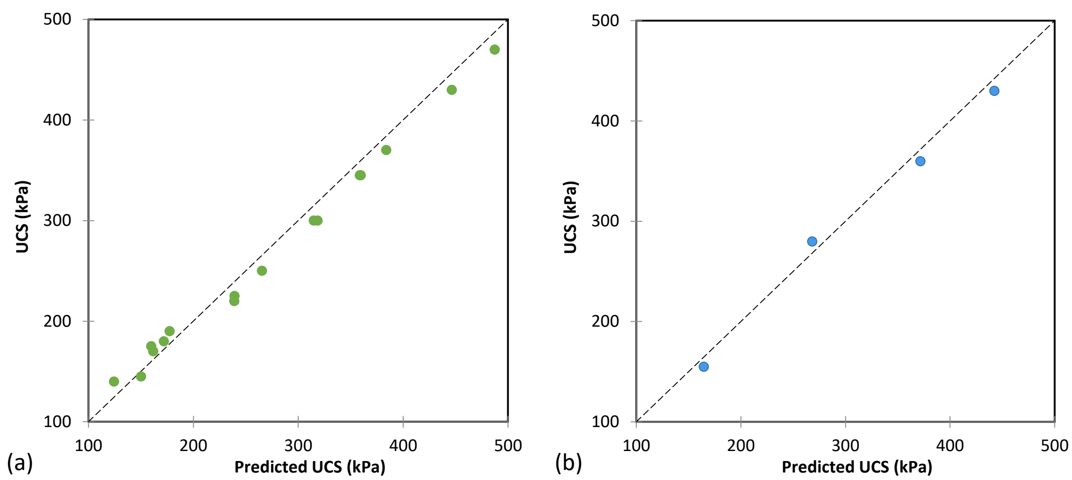

Fig. 4. The results of CRRF for predicting UCS

Table 6 illustrates the accuracy results of the CRRF model based on some parameters. The CRRF model performed well in both training and testing databases with high R^2 values. While the CRRF model slightly overfits on the training dataset, the overall results indicate that the model is accurate and effective in predicting outputs.

Table 6. The specifications of the best CRRF.

Performance metrics	Split tensile strength		UCS	
	Training	Testing	Training	Testing
MAE	11.352	13.880	7.107	9.305
R^2	0.988	0.979	0.965	0.965

The variable importance of input parameters

Modeling artificial intelligence methods requires checking the importance of input parameters. Figs 5 and 6 illustrate the importance of input parameters in determining two outputs in both models. According to Figure 5, based on both models, the most important parameter for predicting split tensile strength was bottom ash content, and the least important was coir fiber content based on the PLS model and EC based on the CRRF model. For both models, for predicting UCS, Calcium content was the most important parameter to predict the UCS and coir fiber was the least important parameter to predict the UCS.

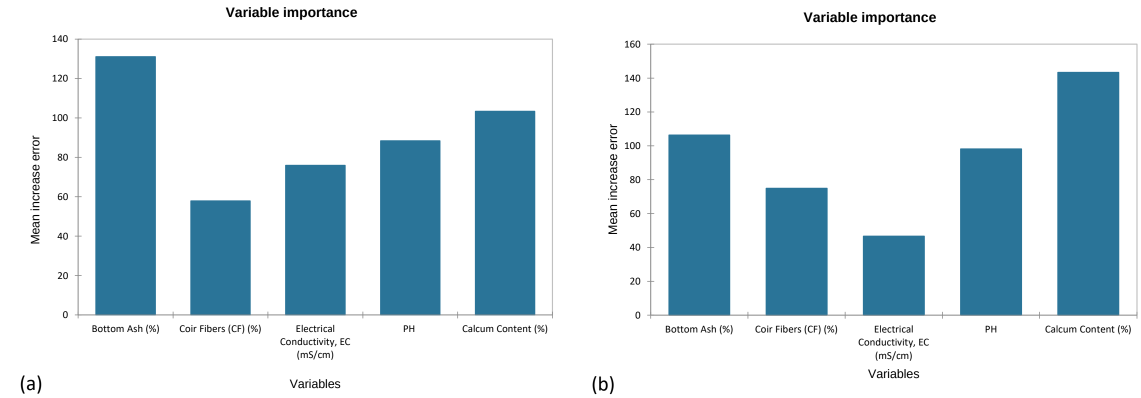

Fig. 5. The importance of parameters to predict split tensile strength based on (a) PLS and (b) CRRF

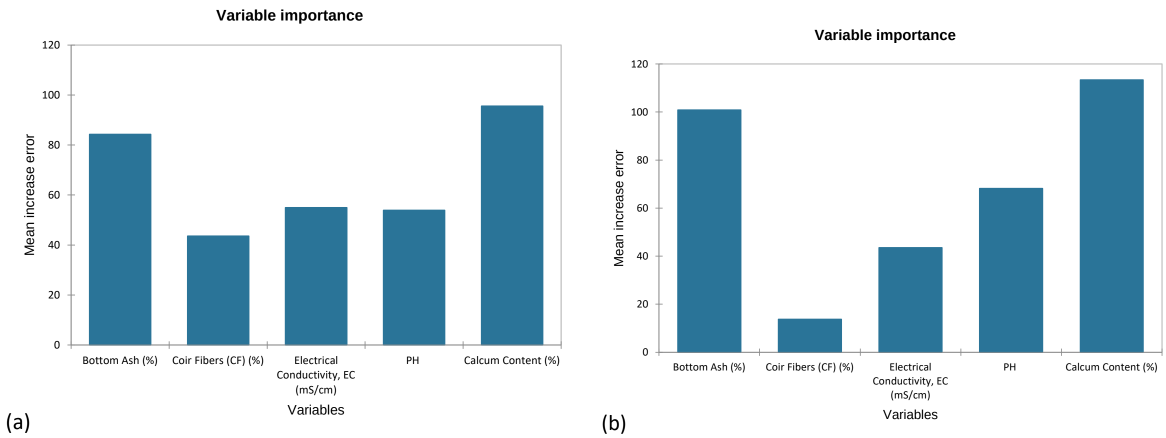

Fig. 6. The importance of parameters to predict UCS based on (a) PLS and (b) CRRF

Conclusion

This study has examined the uniaxial compressive strength (UCS) and split tensile strength of a soil mixture containing recycled ash and natural fibers using two different methods, i.e., partial least squares (PLS) and classification and regression random forest (CRRF). The study has utilized a 20-sets database consisting of five inputs and two outputs, and the importance of input parameters has been evaluated. The results have shown that the best CRRF model has superior performance metrics compared to the PLS model. According to both models, bottom ash content is the most important parameter for predicting split tensile strength, whereas Calcium content is the most significant parameter in predicting UCS. In contrast, coir fiber content is found to be the least important parameter in both models. The results suggest that bottom ash and Calcium content are critical factors that should be considered when designing a soil mixture for achieving higher strength.

References

[1] Puppala AJ, Manosuthikij T, Chittoori BC. Swell and shrinkage characterizations of unsaturated expansive clays from Texas. Engineering Geology. 2013 Sep 17;164:187-94.

[2] Sahebzadeh S, Heidari A, Kamelnia H, Baghbani A. Sustainability features of Iran's vernacular architecture: A comparative study between the architecture of hot–arid and hot–arid–windy regions. Sustainability. 2017 May 4;9(5):749.

[3] Baghbani A, Costa S, O'Kelly BC, Soltani A, Barzegar M. Experimental study on cyclic simple shear behaviour of predominantly dilative silica sand. International Journal of Geotechnical Engineering. 2022 Oct 23:1-5.

[4] Puppala AJ, Congress SS, Banerjee A. Research advancements in expansive soil characterization, stabilization and geoinfrastructure monitoring. InFrontiers in geotechnical engineering 2019 Feb 12 (pp. 15-29). Singapore: Springer Singapore.

[5] Estabragh AR, Pereshkafti MR, Parsaei B, Javadi AA. Stabilised expansive soil behaviour during wetting and drying. International Journal of Pavement Engineering. 2013 Apr 1;14(4):418-27.

[6] Baghbani A, Choudhury T, Samui P, Costa S. Prediction of secant shear modulus and damping ratio for an extremely dilative silica sand based on machine learning techniques. Soil Dynamics and Earthquake Engineering. 2023 Feb 1;165:107708.

[7] Nguyen MD, Baghbani A, Alnedawi A, Ullah S, Kafle B, Thomas M, Moon EM, Milne NA. Experimental Study on the Suitability of Aluminium-Based Water Treatment Sludge as a Next Generation Sustainable Soil Replacement for Road Construction. Available at SSRN 4331275.

[8] Asokan P, Saxena M, Asolekar SR. Coal combustion residues—environmental implications and recycling potentials. Resources, Conservation and recycling. 2005 Feb 1;43(3):239-62.

[9] Leao, A.L., Rowell, R. and Tavares, N., 1998. Applications of natural fibers in automotive industry in Brazil—thermoforming process. Science and Technology of Polymers and Advanced Materials: Emerging Technologies and Business Opportunities, pp.755-761.

[10] Punthutaecha K, Puppala AJ, Vanapalli SK, Inyang H. Volume change behaviors of expansive soils stabilized with recycled ashes and fibers. Journal of materials in Civil Engineering. 2006 Apr;18(2):295-306.

[11] Tiwari N, Satyam N, Puppala AJ. Strength and durability assessment of expansive soil stabilized with recycled ash and natural fibers. Transportation Geotechnics. 2021 Jul 1;29:100556.

[12] Puppala A, Hoyos L, Viyanant C, Musenda C. Fiber and fly ash stabilization methods to treat soft expansive soils. InSoft ground technology 2001 (pp. 136-145).

[13] Baghbani A, Choudhury T, Costa S, Reiner J. Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. Earth-Science Reviews. 2022 May 1;228:103991.

[14] Bardhan, A. and Samui, P., 2022. Application of artificial intelligence techniques in slope stability analysis: a short review and future prospects. International Journal of Geotechnical Earthquake Engineering (IJGEE), 13(1), pp.1-22.

[15] Baghbani, A., Daghistani, F., Naga, H.A. and Costa, S., 2022. Development of a Support Vector Machine (SVM) and a Classification and Regression Tree (CART) to Predict the Shear Strength of Sand Rubber Mixtures. In Proceedings of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR), Newcastle, Australia.

[16] Lin SS, Shen SL, Zhang N, Zhou A. Modelling the performance of EPB shield tunnelling using machine and deep learning algorithms. Geoscience Frontiers. 2021 Sep 1;12(5):101177.

[17] Baghbani A, Costa S, Choundhury T, Faradonbeh RS. Prediction of Parallel Desiccation Cracks of Clays Using a Classification and Regression Tree (CART) Technique. InProceedings of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR), Newcastle, Australia 2022.

[18] Xu JJ, Zhang H, Tang CS, Cheng Q, Tian BG, Liu B, Shi B. Automatic soil crack recognition under uneven illumination condition with the application of artificial intelligence. Engineering Geology. 2022 Jan 1;296:106495.

[19] Baghbani A, Costa S, Choundhury T, Faradonbeh RS. Prediction of Parallel Desiccation Cracks of Clays Using a Classification and Regression Tree (CART) Technique. InProceedings of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR), Newcastle, Australia 2022.

[20] Baghbani A, Costa S, Lu Y, Soltani A, Abuel-Naga H and Samui P. Effects of Particle Shape on Shear Modulus of Sand Using Dynamic Simple Shear Testing. Arabian journal of geosciences. 2023.