Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly

doi: 10.3850/978-981-18-5182-7_08-004-cd

ISGSR 2022

Development of a Support Vector Machine (SVM) and a Classification and Regression Tree (CART) to Predict the Shear Strength of Sand-Rubber Mixtures

Abolfazl Baghbani[1], Firas Daghistani[2], Hossam Abuel Naga[3] and Susanga Costa[4]

[1]School of Engineering, Deakin University, 3086 VIC, Australia.
E-mail: abaghbani@deakin.edu.au
[2]School of Engineering and Mathematical Sciences, La Trobe University, 3086 VIC, Australia
E-mail: f.daghistani@latrobe.edu.au
[3]School of Engineering and Mathematical Sciences, La Trobe University, 3086 VIC, Australia.
E-mail: h.aboel-naga@latrobe.edu.au
[4]School of Engineering, Deakin University, 3086 VIC, Australia.
E-mail: susanga.costa@deakin.edu.au

Abstract: Recycling waste materials such as waste tires in geotechnical projects can greatly contribute to environmental issues. An important feature of sand-rubber mixtures is their shear strength, which depends on many factors such as the size distribution of sand and rubber, density of mixtures etc. Due to the multiplicity of these effective factors, this paper evaluated the performance of two Artificial intelligence (AI) methods, namely a support vector machine (SVM) and a classification and regression tree (CART) algorithm, to predict the shear strength of sand-rubber mixtures. For this purpose, a database with 101 datasets including nine inputs and one output, i.e., the ratio of shear strength to normal stress, was used. The inputs parameters included dry density, mean particle size (D_{50}), coefficient of curvature (C_c) and uniformity coefficient (C_u) of sand, normal stress, rubber percentage, and D_{50}, C_c and C_u of rubber. The results of the best SVM and CART models were also compared with the result of multiple linear regression (MLR) method. The results show that R^2 for the test database was 0.90, 0.90 and 0.55 for the CART, SVM and MLR models, respectively. In addition, the MAE of CART, SVM and MLR methods were 0.013, 0.013 and 0.041, respectively. Therefore, according to the results, both AI methods have a great performance to predict the shear strength of sand-rubber mixtures.

Keywords: Sand-rubber mixtures; Shear strength; Support vector machine; Classification and regression trees.

1 Introduction

Every year, a large volume of scrap tires ends up in landfills (Sahebzadeh et al., 2017, Al-Fakih et al. 2020). In addition to adverse effects of waste tires on the environment, they can increase the fire risks and damage the health of creatures. Therefore, recycling tire waste is one of the important topics in geotechnical engineering research. Scrap tires have engineering properties such as low specific gravity, high durability and flexibility (Al-Fakih et al. 2020) that can be useful in geotechnical engineering applications such as lightweight fillers (Kong et al. 2018), highway construction (Siddique and Naik 2004, Baghbani et al., 2022a), soil reinforcement (Zhang et al. 2020) and soil retaining walls (Reddy and Krishna 2015). Studies (Ahmed 1993; Zornberg et al. 2004; Yoon et al. 2008) have shown that there are several factors affecting the shear strength of gravel- or sand-rubber mixture. For example, in one of the first studies on sand-tire chip mixtures, Ahmed (1993) showed that the percentage of rubber chips, sample compaction method and confinement pressure in the triaxial test were the effective factors on the shear behavior of sand-rubber mixtures (Ahmed 1993). In another study, Zornberg et al. (2004) used triaxial experiments on sand with rectangular tire chips and showed that the best shear performance was related to a mixture of sand-rubber with 35% by weight of tire chips (Zornberg et al. 2004). Yoon et al. (2008) showed that by increasing the percentage of granulated rubber with diameter less than 20 mm, the California bearing ratio (CBR) decreased (Yoon et al. 2008). Studies (Ghazavi 2004; Anvari et al. 2017; Rouhanifar et al. 2021) also showed that with increasing the percentage of rubber in the rubber-sand mixture, the shear strength of the mixture increased.

Due to the complexity of effective factors on the shear strength of sand-rubber mixture, including the size of sands and rubber, the percentage of rubber, normal stress, dry density, etc., there is still no model to predict maximum shear strength of sand-rubber mixtures based on all effective parameters. One of the methods that has been well used in various fields of geotechnical engineering and has shown acceptable results, is the artificial intelligence (AI) methods (Baghbani et al. 2022b). However, there is no published study on the application of AI methods to predict maximum shear strength of sand-rubber mixtures. In this paper, for the first time, using a database (including 101 datasets) from direct shear tests, two AI methods including classification and regression tree (CART) Algorithm and support vector machine (SVM) were modelled to evaluate the effects of nine effective parameters on the shear strength of sand-rubber mixtures parameter.

2 Simulation Model

2.1 Experimental Condition and Parameter Selection

The database was collected from two publications (Anvari et al. 2017; Rouhanifar et al. 2021). In these two papers, as shown in Figure 1, different rubbers and sands with different distributions were considered. For sample preparation, mixtures with different weight percentages of rubber were prepared. Then, using wet tamping method (Ladd 1974), the samples were compacted in three layers to get the desired density. Finally, the direct shear tests were conducted on samples with dimensions of 10*10 cm and height of 3.5 cm with different normal stresses. The displacement rate of the direct shear test was 0.5 mm/min.

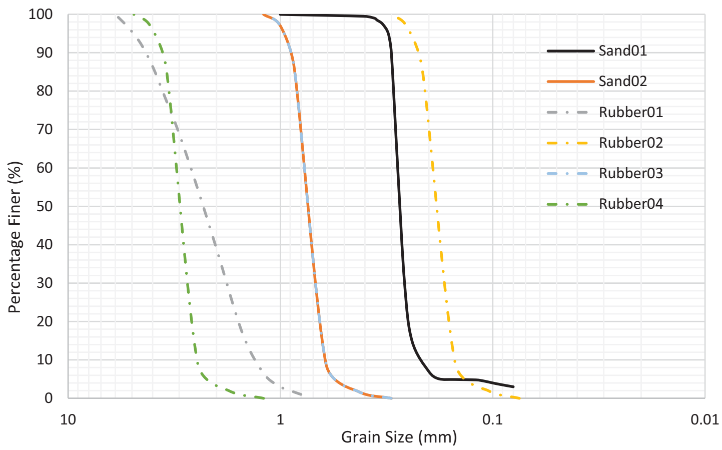

Figure 1. The grain size distribution curves for used sands and rubbers.

2.2 Database Collection and processing

To use different methods of AI, a database with a large number of datasets is needed. In total, the obtained database from direct shear experiments had nine inputs including relative median diameter (D_{50}), coefficient of uniformity (C_u), coefficient of curvature (C_c) of both sand and rubber, the percentage of rubber, normal stress, mixture dry density, and one output, i.e., the ratio of maximum shear strength to normal stress in direct shear tests. A summary of the database is shown in Figure 2.

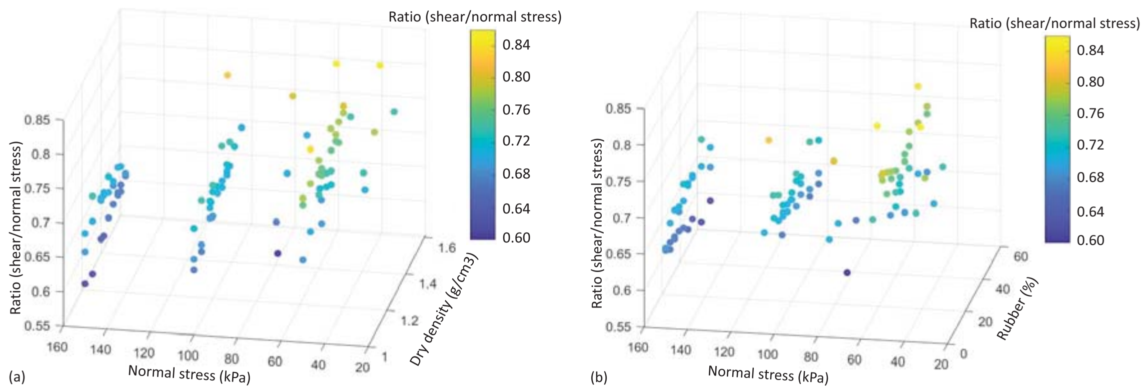

Figure 2. The distribution of collected database.

To increase the accuracy of modelling, the collected database was normalized by Equation 1. Normalization caused that in AI methods, all the parameters involved in the model were treated equally, and the error due to the accuracy of different parameters was reduced.

$$x_{normal} = (x - x_{min}) / (x_{max} - x_{min}) \tag{1}$$

where x_{normal}, x, x_{min} and x_{max} are the normalized, actual, minimum, and maximum parameter value, respectively. In AI models, the database was divided randomly into two databases: training (including 80% of the total database) and testing (including 20% of the total database) database. Tables 1 and 2 show the statistical information of these two databases. As can be seen, the important statistical information, including minimum, maximum, mean, standard standards, of both databases were almost same.

Table 1. Statistical information of training database.

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Ratio (Shear/Normal)	81	0.595	0.833	0.701	0.046
D50-Sand	81	0.270	0.730	0.650	0.175
Cu-sand	81	1.260	1.450	1.293	0.072
Cc-sand	81	0.950	1.080	0.972	0.049
D50-Rubber	81	0.190	2.920	1.640	1.213
Cu-Rubber	81	1.260	1.800	1.353	0.205
Cc-Rubber	81	0.710	0.950	0.909	0.091
Rubber (%)	81	0.000	50.000	20.062	15.380
Dry density	81	1.000	1.610	1.270	0.154
Normal stress (kPa)	81	34.500	150.000	96.457	41.028

Table 2. Statistical information of testing database.

Variable	Observations	Minimum	Maximum	Mean	Std. deviation
Ratio (Shear/Normal)	20	0.644	0.835	0.714	0.050
D50-Sand	20	0.270	0.730	0.592	0.216
Cu-sand	20	1.260	1.450	1.317	0.089
Cc-sand	20	0.950	1.080	0.989	0.061
D50-Rubber	20	0.190	2.920	1.242	1.145
Cu-Rubber	20	1.260	1.800	1.422	0.254
Cc-Rubber	20	0.710	0.950	0.878	0.113
Rubber (%)	20	0.000	50.000	19.500	14.409
Dry density	20	1.000	1.460	1.288	0.127
Normal stress (kPa)	20	34.500	150.000	81.350	38.055

2.3 Multiple linear regression (MLR)

Multiple linear regression (MLR) is a statistical method for predicting an output variable based on several independent input variables. This method is a developed method from linear regression (OLS) method that has only one input variable and one output variable (see Figure 3).

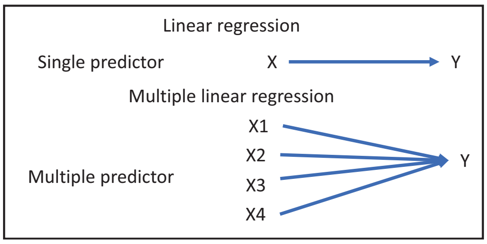

Figure 3. Multiple linear regression and linear regression.

In MLR, the relationship between input parameters and output parameter is considered linearly. Also in this method, the data is used in a normalized way.

$$y = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n + \varepsilon \quad (2)$$

Where y is the predicted value, β_0 is y-intercept when all other parameters are equal to 0, X_1 and X_n are first and last independent variables, β_1, β_n are regression coefficient of first and last independent variables, and ε is model error.

In MLR, to achieve the best line, regression coefficients are selected so that the model has the least error. In this study, before modelling AI methods, MLR as a one of the simplest regression methods, was used to check the accuracy of the simple linear regression.

To determine the accuracy of the proposed models, the coefficient of determination (R^2), and the mean absolute error (MAE) between the predicted and measured values have been determined. Equations 2 and 3 show the definition of the coefficient of determination (R^2), and the mean of the absolute error (MAE).

$$R^2 = \frac{\sum_N (X_m)^2 - \sum_N (X_m - X_p)^2}{\sum_N (X_m)^2} \quad (3)$$

$$MAE = \frac{\sum_N \left| (X_m - X_p) \right|}{N} \quad (4)$$

Where X_m, X_p, $\overline{X_m}$, $\overline{X_p}$ are actual values, predicted values, the average of actual values and the average of predicted values, respectively and N is the number of datasets. The best model is a model that has the coefficient of determination (R^2) of 1 and mean of the absolute error (MAE) equal 0.

2.4 Classification and Regression Tree Algorithm (CART)

Classification and Regression Tree Algorithm (CART) is one of the well-known methods of artificial intelligence (AI) which was first introduced by Breiman (1996). CART has the ability to find the relationship between input and output variables without any presuppositions and for various purposes, such as prediction (i.e., regression tree) and classification (i.e., classification tree). Classification trees are generally used for continuous variables in order to find the group of the target variable that is most likely to fall. Regression trees are also used to predict continuous variables. This method is used to construct a decision tree based on input variables using classification and regression. The result of CART is a tree-like structure with different nodes and branches. One of the advantages of the CART method compared to other AI methods is that it is a white box method, and its final tree structure can be used easily to predict output parameters based on the input variables. This is while most AI methods, such as artificial neural network (ANN), are known as a black box. Figure 4a shows the typical tree structure of a CART model, which is comprised of nodes (i.e., root node, internal node, and leaf node), rules, and branches. In the CART algorithm, each node is divided into two sub-nodes with left and right branches.

2.6 Support vector machine (SVM)

The support vector machine (SVM) method is one of the AI methods used for regression and classification. In this method, which was introduced by Boser et al. (1992), a hyperplane is used to separate the data input nodes using mathematical equations. Figure 4b shows a typical diagram of the SVM method. The performance of this method is determined by the location of the hyperplane. A hyperplane works best when it achieves the largest positive vectors and separates the most data nodes.

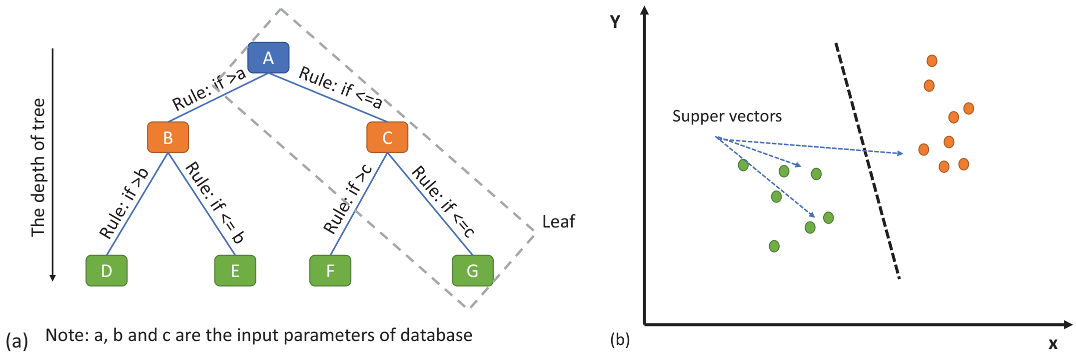

Figure 4. The typical structure of a (a) CART and (b) SVM.

3 Results and discussion

3.1 Multiple linear regression

The MLR results using MATLAB software for the 101 direct shear tests are shown in Figure 5. The results show that for the training database, the R^2 and MAE are equal to 0.481 and 0.034, respectively, and for the testing database, the R^2 and MAE are 0.553 and 0.041, respectively. These values show that simple regression methods, such as MLR, cannot achieve high accuracy in predicting the shear strength ratio (shear/normal stress) and other methods like AI methods should be examined.

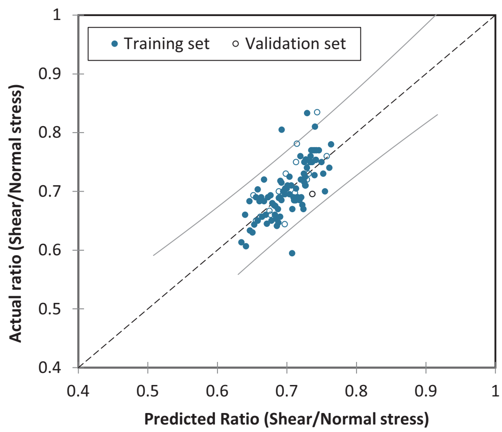

Figure 5. The results of multiple linear regression (MLR) model.

3.2 CART Model

Figure 6 shows the best tree for predicting the shear strength ratio (shear/normal stress). This tree is very easy to use and according to the rule of each leaf, model prediction for the shear strength ratio (shear/normal stress) can be achieved. For example, in Figure 6, in node 8:

- Left branch: If the dry density is less than or equal to 1.44 g/m3,
- Right branch: If the dry density is greater than 1.44 g/m3.

As Figure 6 shows, the predicted shear strength ratio (shear/normal stress) in nodes 16 and 17 are:

- Left branch (node 16): 0.76
- Right branch (node 17): 0.83

As a result, with increasing dry density of soil, the predicted shear strength ratio (shear/normal stress) is greater.

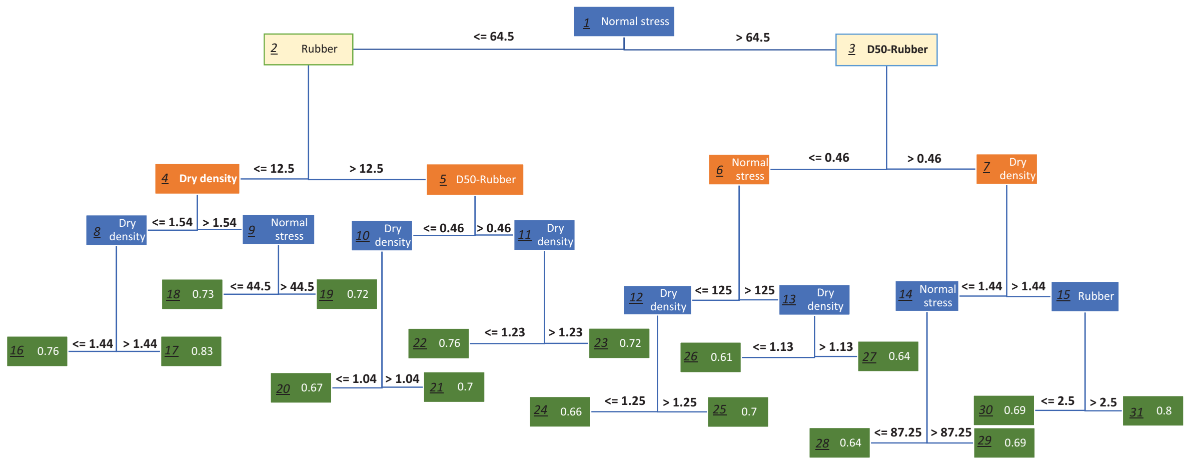

Figure 6. Tree structure of the best CART model.

Figure 7 represents the results for the best CART. The results show that the R^2 and MAE of the best CART for the training database are 0.817 and 0.015, respectively. Also, for the testing database, the best CART has the R^2 and MAE of 0.897 and 0.013, respectively.

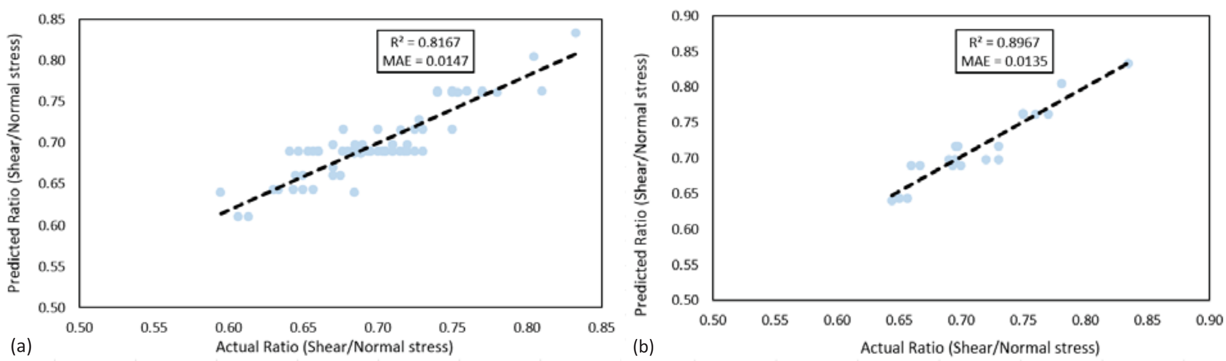

Figure 7. The result of best CART model for (a) training database, (b) testing database.

3.3 SVM model

After examining different SVM models, the best model was obtained to predict shear strength ratio (shear/normal stress). The results in Figure 8 show that the R^2 and MAE of the model were 0.942 and 0.008, respectively for training database, and 0.895 and 0.013 respectively, for testing database. This accuracy indicates the proper performance of the SVM method in predicting the shear strength ratio (shear/normal stress).

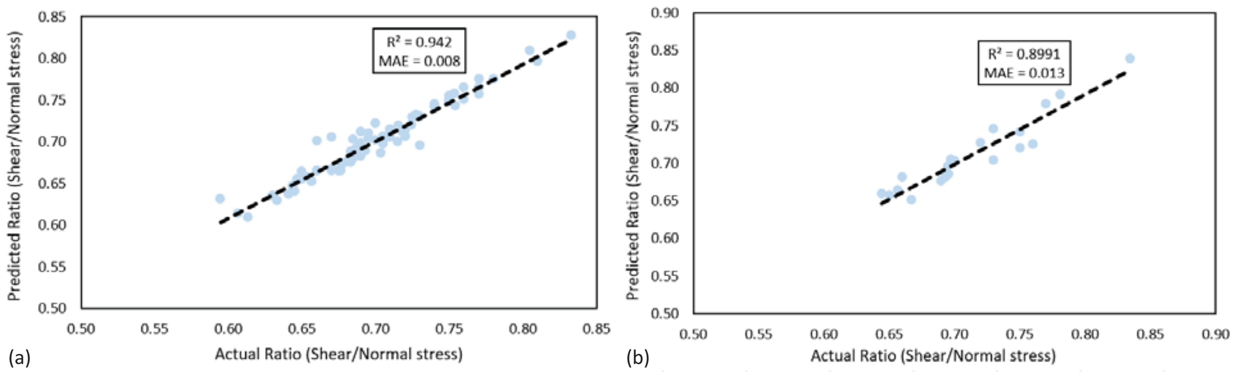

Figure 8. SVM results for (a) training and (b) testing databases.

4 Conclusions

In this paper, for the first time, two methods of artificial intelligence (AI), namely CART and SVM, were used to predict the shear strength ratio (shear/normal stress) of sand-rubber mixtures. For this purpose, a database of 101 sets from direct shear tests were collected. During modelling, 80% of the collected database was randomly allocated for model training and 20% for model testing. The database consisted of nine input parameters, including dry density, mean particle size (D_{50}), coefficient of curvature (C_c) and uniformity coefficient (C_u) of sand and rubber, normal stress, rubber percentage, and one output parameter, i.e., the shear strength ratio (shear/normal stress). The results showed that for the testing database, the R^2 of the best CART, SVM and MLR were equal to 0.90, 0.90 and 0.55, respectively. Also, the MAE of the CART, SVM and MLR were equal to 0.013, 0.013 and 0.041. These results show the good performance of AI methods in predicting the shear strength ratio (shear/normal stress) of sand-rubber mixtures compared to simple regression methods such as MLR.

References

Ahmed I. (1993). Laboratory study of properties of rubber-soils. *Final report. Purdue University, Indiana Department of Transportation.*

Al-Fakih, A., Mohammed, B.S. and Liew, M.S. (2020). Tires rubber as a useable material in civil engineering applications: an overview. *Int. J. Adv. Res. Eng. Technol*, 11(11), 315-325.

Anvari, S.M., Shooshpasha, I. and Kutanaei, S.S. (2017). Effect of granulated rubber on shear strength of fine-grained sand. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(5), 936-944.

Baghbani, A., Baghbani H., Shalchiyan MM. and Kiany K. (2022a). Utilizing artificial intelligence and finite element method to simulate the effects of new tunnels on existing tunnel deformation. *Journal of computational and Cognitive Engineering 2022*, vol 1(3), 152-162

Baghbani, A., Choudhury, T., Costa, S. and Reiner, J. (2022b). Application of artificial intelligence in geotechnical engineering: A state-of-the-art review. *Earth-Science Reviews*, 103991.

Boser, B.E., Guyon, I.M. and Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers. *In Proceedings of the fifth annual workshop on Computational learning theory*, 144-152.

Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123-140.

Ghazavi, M. (2004). Shear strength characteristics of sand-mixed with granular rubber. *Geotechnical & Geological Engineering*, 22(3), 401-416.

Kong, D., Deng, M., Liu, Y. and Wang, X. (2018). The environmental study on consolidated undrained triaxial compression tests on lightweight soil mixed with rubber chips of scrap tires. *Ekoloji*, 27(106), 1503-1510.

Ladd, R.S. (1974). Specimen preparation and liquefaction of sands. *Journal of the Geotechnical Engineering Division*, 100(10), 1180-1184.

Reddy, S.B. and Krishna, A.M. (2015). Recycled tire chips mixed with sand as lightweight backfill material in retaining wall applications: an experimental investigation. *International Journal of Geosynthetics and Ground Engineering*, 1(4), 1-11.

Rouhanifar, S., Afrazi, M., Fakhimi, A. and Yazdani, M. (2021). Strength and deformation behaviour of sand-rubber mixture. *International Journal of Geotechnical Engineering*, 15(9), 1078-1092.

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

Siddique, R. and Naik, T.R. (2004). Properties of concrete containing scrap-tire rubber–an overview. *Waste management*, 24(6), 563-569.

Yoon Y.W., Heo, S.B., Kim, K.S. (2008). Geotechnical performance waste tire for soil reinforcement from chamber tests. *Geotextiles and Geomembranes*, 26(1), 100.

Zhang, H., Yuan, X., Liu, Y., Wu, J., Song, X. and He, F. (2020). Experimental study on the pullout behavior of scrap tire strips and their application as soil reinforcement. *Construction and Building Materials*, 254, 119288.

Zornberg, J.G., Cabral, A.R., Viratjandr, C. (2004). Behaviour of tire shred-sand mixtures. *Canadian Geotechnical Journal*, 41(2), 227.