Rocznik Ochrona Środowiska

Volume 26 Year 2024 ISSN 2720-7501 pp. 424-431

https://doi.org/10.54740/ros.2024.041 open access

Received: July 2024 Accepted: September 2024 Published: September 2024

Machine Learning and Linear Regression Approach to Model Unconfined Compressive Strength of Ceramic Waste Modified Soil as Subgrade Pavement Material

Meshel Qablan Alkahtani

Department of Civil Engineering, King Khalid University, Abha, Saudi Arabia

corresponding author's e-mail: mgalqhtani@kku.edu.sa

Abstract: An effective application of artificial intelligence involves artificial neural networks. Artificial neural networks and linear regression models were developed to simulate the effects of using discarded ceramic waste as a subgrade for pavement. The ceramic waste was used at 2.5%, 5%, 7.5%, 10%, 12.5%, and 15%. A sample with 0% ceramic waste was tested to serve as a reference sample. The dataset was produced from laboratory experimentation findings used to train, test, and evaluate the model. A training set, a target set, and a prediction set were created from the dataset. The artificial neural network MSE was 0.42-1.40, while the linear regression model range was 1.74 to 3.63 for ceramic modified samples. The R^2 range for the ANN model was 0.85-0.92, and the linear regression model exhibited a range of 0.71-0.78. The ANN model was more accurate than the linear regression model. Future studies are required to compare different machine-learning approaches for predicting soil mechanical properties.

Keywords: ceramic waste, artificial neural network, unconfined compression strength, California bearing ratio

1. Introduction

With the advent of artificial intelligence and machine learning, pavement construction can be completed more quickly and precisely (Nouhi & Pour 2021). Monitoring road performance parameters is required to give users the necessary safety levels and driving comfort (Sollazzo et al. 2017). The performance of the subgrade material is equally as important as the structural performance of the pavement (Gong et al. 2018). It is the last part of any pavement design and is in charge of distributing traffic loads to the ground and preventing surface failure. Due to their affordability and environmental friendliness, waste materials are increasingly being considered for use as subgrade materials (Kumar et al. 2022).

Ceramic waste is generated from its production unit and turns into waste after its use, such as building construction tiles, cookware, pottery, etc. This has prompted the use of ceramic waste as a possible building material. Deboucha et al. (2020) have evaluated the effects of cement, marble dust and ceramic waste in the pavement sub-base layer. Liu et al. (2024) examined the unconfined compressive strength of rice straw fibre-reinforced soil and modelled it using a finite element approach. Chen et al. (2024) have employed decision tree, linear regression, and machine learning models to predict the unconfined compressive strength of geopolymer stabilised soil. Taseer et al. (2024) predicted the unconfined compressive strength of soil contaminated with heavy metal using gene-expressing programming, extreme machine learning machines, random forest and multiple linear regression. Zhao et al. (2024) modelled the unconfined compressive strength of low-strength waste soil using optimised support vector machine models.

From the published literature, it can be inferred that unconfined compressive strength modelling has recently gained attention for various soil types. However, modelling the performance of subgrade material modified by ceramic waste has yet to gain researchers' attention. Hence, this study was carried out to address the existing research gap. The objectives of this study are: 1. To assess the accuracy of artificial neural networks to predict unconfined compressive strength, and 2. To compare the performance of the ANN model and linear regression model for their performance.

2. Data and Method Used

2.1. Laboratory Data

Laboratory testing for subgrade material was carried out to obtain actual data. The subgrade material was developed with zero ceramic waste as reference data. The ceramic waste specimens were prepared with varying percentage of 2.5% for each sample type. The varying percentage of ceramic waste was 2.5%, 5%, 7.5%, 10%, 12.5% and 15%. The range has been selected based on previously published data from which it was observed that beyond 15%, there was no improvement in the properties of subgrade material (Cabalar et al. 2016, Deboucha et al. 2020, Oluwaseun 2018).

2.2. Artificial Neural Network model

ANN application in pavement research is wide and has successfully predicted various aspects and properties. Traffic speed deflections (Mabrouk et al. 2021), pavement acoustic longevity (Cao et al. 2020), viscoelastic behaviour (Sadat Hosseini et al. 2021), optimisation of asphalt mixes (Sebaaly et al. 2018) and thermal and mechanical properties of pavement material comprising of demolition waste (Ghorbani et al. 2021) are among some of the studies employing ANN for modelling.

Fig. 1 presents this study's artificial neural network model for modelling pavement subgrade material properties modified using ceramic waste tiles. This study employed a multi-feed-forward neural network, which is also among the most employed in previous studies (Sollazzo et al. 2017). This study also employed the same as the parameters evaluated in this study are affected by various other soil properties. The typical neural network model comprises three layers: the input layer, hidden layer, and out or target layer, which are interconnected by layers of neurons. The function of neurons is to process input data based on a d-defined function and produce output results through a defined network topology. Specific weight (w_i) is defined for each connection based on which the input data is altered (reduced or amplified). The relationship between input (a_i), output (b_i) and single neuron is presented in the equation below:

$$f(X) = \frac{1}{1+x^{-1}} \quad (1)$$

where X is the sum of $w_i\ a_i$ of weighted a_i input resulted from previous neurons.

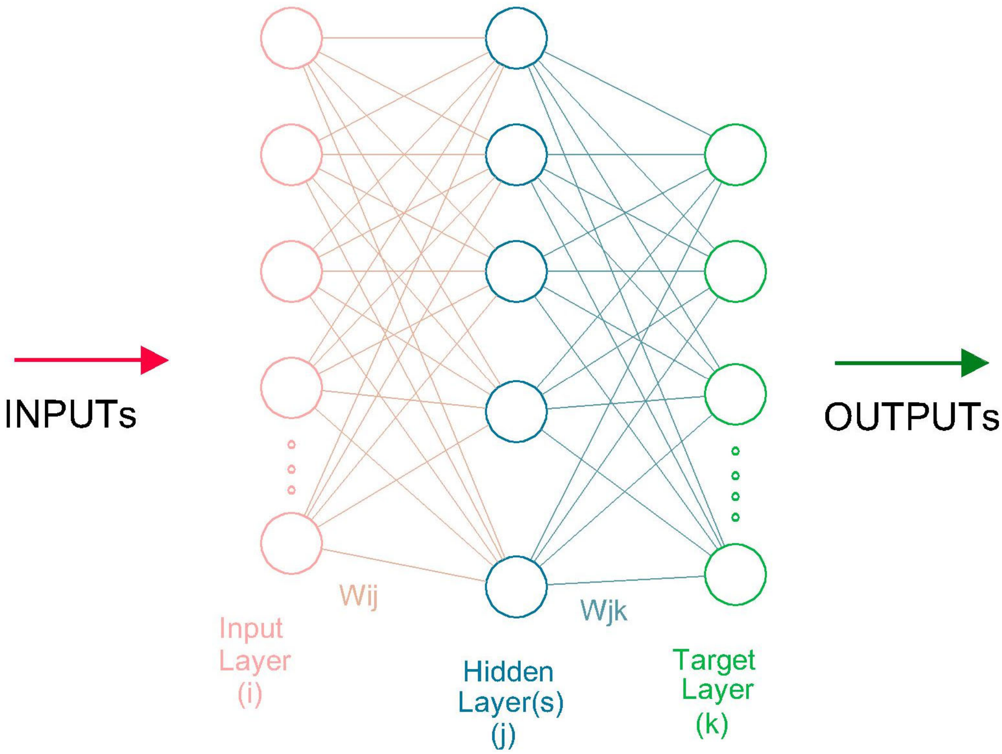

Fig. 1. ANN typical model consisting of three-layer prediction

2.3. Linear regression model

The linear prediction model performs well when the relationship between input and output variables is linear (Barua & Zou 2021). The linear regression model has been actively used in modelling pavement materials performance with additives (Sadat Hosseini et al. 2021). The linear regression model is among the most used tools for statistical analysis (Sadat Hosseini et al. 2021). The conventional linear regression model concerning identical and independent observations (a_i, b_i) can be presented in the form of Equation 2 as follows:

$$fb_i = a^x{}_1\alpha + €_I \quad (2)$$

where, m x 1 vector represents the unknown α, while independent and identically distributed value from a is presented by $€_I$. α is generally estimated using equation 3 based on the ordinary least squares approach.

$$\sum_{i=1}^{n}(b_i - a_i^x\ \alpha)^2 \quad (3)$$

2.4. Accuracy of Model

Mean square error (MSE) and coefficient of determination (R^2) were employed to determine the model's accuracy. Pérez-Acebo et al. (2021) have modelled IRI of semi-rigid pavement with single carriageway roads and used R^2 as a model accuracy measurement tool. Setyawan et al. (2015) employed R^2 to evaluate the performance of the regression model and predict the remaining life service of asphalt pavement based on the pavement condition index. The coefficient of determination is the proportion of variation between input and output variable which can be predicted from the independent variable. Typical R^2 values range in between 0 to 1. Values of R^2 closer to 0 infer that model is not capable of predicting the target values from the given input values. While values closer to 1 indicate high precision prediction can be obtained with minimum error. R^2 values equal to one infers that the model prediction performance is with zero error that means 100% accurate. On the contrary if the values are in negative it indicates worse fit model which exhibits the tendency of data itself to predict better than deploying the model for prediction. The R^2 values were obtained base on the equation as follows:

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(a_i - a^\wedge{}_i)^2}{\frac{1}{n}\sum_{i=1}^{n}(a_i - a\prime_i)^2} \tag{4}$$

Where, a_i is the measured value, a'_i is an average value and $a^\wedge{}_i$ is the predicted value. Gong et al. (2018) used both measures (R^2 & MSE) to evaluate the accuracy of the random forest regression model in predicting asphalt pavement IRI. Sollazzo et al. (2017) employed MSE to evaluate the accuracy of the ANN model to correlate structural and roughness performance of asphalt pavement. Xiao et al. (2020) also used MSE to evaluate the accuracy of multilayer perceptron for predicting surface roughness. Mean square error values are totally in contrast concerning R^2 values inference. The lower the value, the better the model performance, and the higher the value, the model performance has no accuracy. Mean square error, when based on N number of data, it can be estimated using equation 5:

$$\text{Mean Square Error (MSE)} = \frac{1}{N}\sum_{i=n}^{i}(a_i - p_i)^2 \tag{5}$$

Where, p_i is the predicted values obtained from the model and t_i is the expected output.

3. Results and Discussion

ANN model identifies the optimum model based on input data. The optimised model represents the best interrelationship of input and output data. Overfitting is one of the vital issues faced during the optimisation of the model, as the optimised model should be more generalised. Hence, the dataset is divided into two sets, one on which the neural network is trained and another on which the trained algorithm is tested. Several studies suggest the division between 70% and 30% for training and testing data, respectively. However, the range of testing data has also been given as between 15-30% (Ghorbani et al. 2020). The training data provides the base for the neural model to establish the behaviour of variables in the dataset. The test data set evaluates the model's accuracy to predict actual values based on unseen datasets. In this study, 85% of the dataset was identified as training data and 15% was taken as testing data. The model accuracy was measured in terms of R^2 and MSE. The testing data model performance revealed that the ANN model was performing better than the LR approach. The R^2 value was observed to be 0.91 for the ANN model, and the MSE value was 0.47. For the LR model, the R^2 value was observed to be 0.72, and the MSE value was 4.09. The prediction results in testing data for the ANN and LR models are presented in Figure 2. On the right side of the figure, ANN models prediction versus actual is presented starting from top to bottom for 0%, 2.5%, 5%, 7.5%, 10%, 12.5% and 15%, respectively. Similarly, from top to bottom, LR models are presented on the left side for respective ceramic waste percentages. The R^2 value for the 0% ceramic waste sample for the ANN prediction model was 0.9, while for the LR model, it was 0.71. For 2.5%, 5%, 7.5%, 10%, 12.5% and 15% ceramic waste samples, the R^2 value was between 0.85-0.92 for the ANN model and the LR model, the R^2 range was observed to be 0.71-0.78. The highest accuracy was observed for soil samples with 7.5% ceramic waste. This can be attributed to minimum variation and reduced error in the unconfined compressive strength and predicted value, respectively. Wakjira et al. (2024) reported an R^2 value of 0.90-0.99 while using a hybrid machine-learning approach for predicting ultra-high-performance concrete. Afolagboye et al. (2023) observed an R^2 value of 0.90-0.95 for predicting the unconfined compressive strength of rocks using a machine-learning approach. Ahmadi Sheshde & Cheshomi (2015) found an R^2 value of more than 0.9 for predicting unconfined compressive rocks using a modified point load force approach. Ahenkorah et al. (2023) analysed enzyme-induced carbonate precipitation and microbial-induced carbonate precipitation ground improvement technique and employed evolutionary polynomial regression to model the unconfined compressive strength of sand specimens with an MSE value of 0.075. Table 1 presents the comparison of model accuracy in comparison to this study.

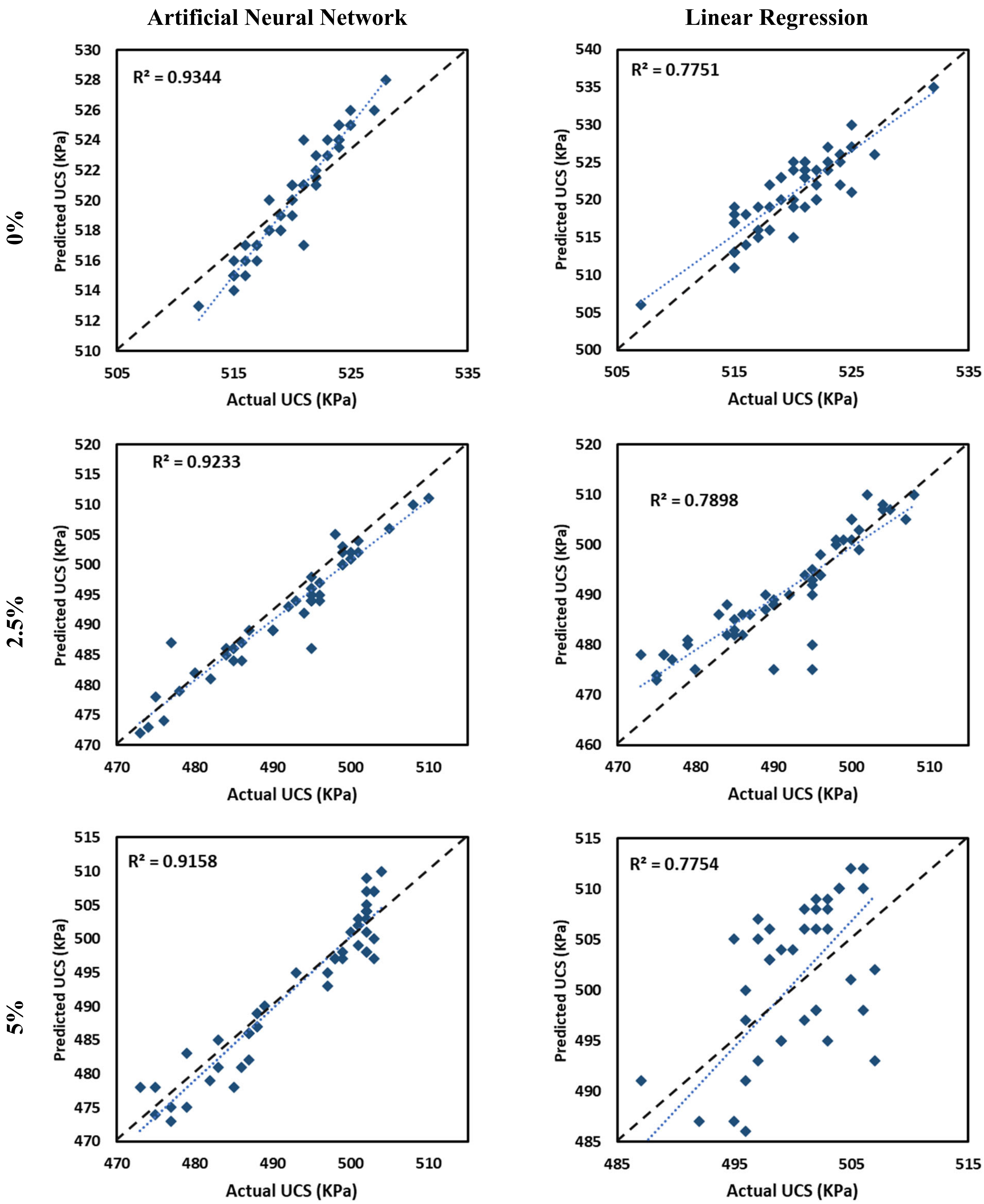

Fig. 2a. Actual vs predicted unconfined compressive strength of soil sample tested at 0%, 2.5%, and 5% for ANN (left side from top to bottom) and LR (right side from top to bottom), respectively

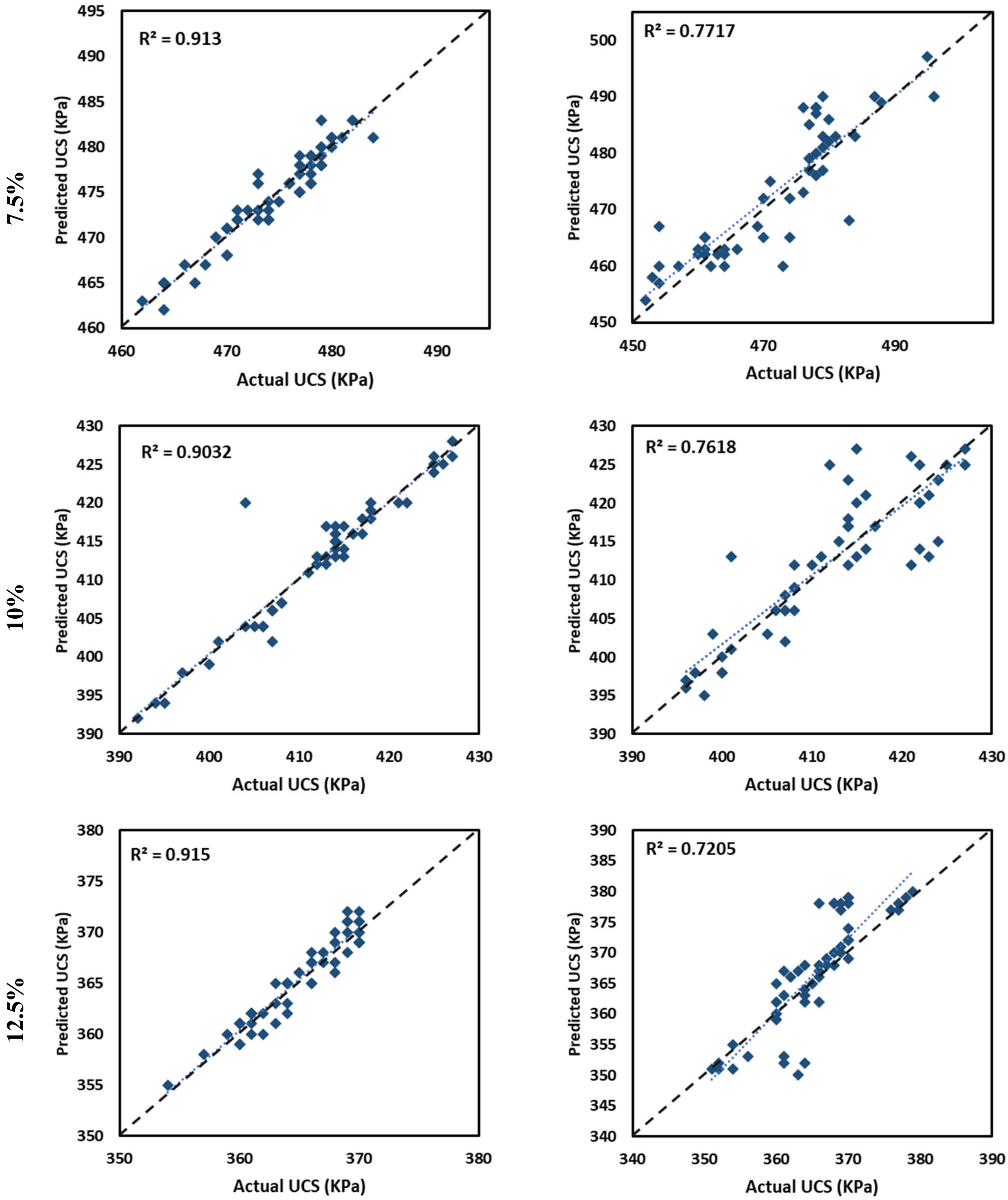

Fig. 2b. Actual vs predicted unconfined compressive strength of soil sample tested at 7.5%, 10%, and 12.5% for ANN (left side from top to bottom) and LR (right side from top to bottom), respectively

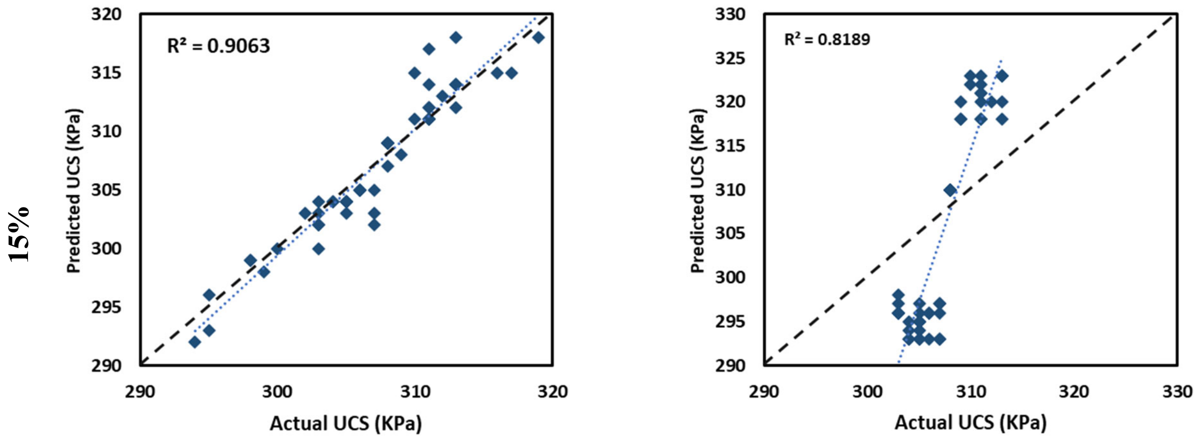

Fig. 2c. Actual vs predicted unconfined compressive strength of soil sample tested at 15% for ANN (left side from top to bottom) and LR (right side from top to bottom) respectively

Table 1. Studies predicting unconfined compressive strength of soil using different prediction approach and their accuracy in comparison with this study's results

Material	Soil Property	Model/Accuracy	Reference	This study result
Geopolymer stabilised soil	Unconfined compressive strength	deep learning method /R^2 = 0.9966	(Yao et al. 2024)	R^2 = 0.91, RMSE = 0.437, MAE = 0.78
Low strength waste soil		SVR model/ R^2 = 0.909, MSE = 0.011, RMSE = 0.105, MAE = 0.085 and MAPE = 15.502%	(Zhao et al. 2024)	
One-part geopolymer stabilised soil		PSO-XGBOOST & PSO-ET/ R^2 = 9964 & R^2 = 0.9928	(Chen et al. 2024)	
Heavy metal contaminated soil		Extreme learning machine/ R^2 = 0.964	(Taseer et al. 2024)	
Microbial/enzyme induced carbonate precipitation treated sand		Genetic Algorithm/ MSE = 0.075, RMSE = 0.273	(Ahenkorah et al. 2023)	
Fly ash treated alkali soil		GEP-tree based AI approach/ R>0.8, MAE = 24.19 MPa, RMSE = 33.15 MPa	(Ashfaq et al. 2022)	

4. Conclusion

This study investigated the potential modelling of the unconfined compressive strength of soil modified with ceramic waste to be used as sub-grade material for pavement. The ceramic waste was added at 0%, 2.5%, 5%, 7.5%, 10%, 12.5% and 15%. The artificial neural network and linear regression approach were used to model the unconfined compressive strength of specimens. The results of the developed model indicated that the artificial neural network was more accurate than the linear regression model. The model performance was validated based on mean square error and coefficient of determination. The artificial neural network MSE was 0.42-1.4, while the linear regression model range was 1.74 to 3.63 for ceramic-modified samples. The R^2 range for the ANN model was 0.85-0.92, and the linear regression model exhibited a range of 0.71-0.78. The ANN model was more accurate than the linear regression model. Future studies are required to compare different machine-learning approaches for predicting soil mechanical properties. This study was limited to the unconfined compressive strength of concrete. To provide a comprehensive modelling approach, more studies are required to model UCS along with shear strength, California bearing ratio, and soaked and wet condition variations.

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/253/45.

Reference

Abed, M.A., Taki, Z.N.M., Abed, A.H. (2021). Artificial neural network modeling of the modified hot mix asphalt stiffness using Bending Beam Rheometer. *Materials Today: Proceedings*, *42*, 2584-2589. https://doi.org/10.1016/j.matpr.2020.12.583

Afolagboye, L.O., Ajayi, D.E., Afolabi, I.O. (2023). Machine learning models for predicting unconfined compressive strength: A case study for Precambrian basement complex rocks from Ado-Ekiti, Southwestern Nigeria. *Scientific African*, *20*. https://doi.org/10.1016/j.sciaf.2023.e01715

Ahenkorah, I., Rahman, M.M., Karim, M.R., Beecham, S. (2023). Unconfined compressive strength of MICP and EICP treated sands subjected to cycles of wetting-drying, freezing-thawing and elevated temperature: Experimental and EPR modelling. *Journal of Rock Mechanics and Geotechnical Engineering*, *15*(5), 1226-1247. https://doi.org/10.1016/j.jrmge.2022.08.007

Ahmadi Sheshde, E., Cheshomi, A. (2015). New method for estimating unconfined compressive strength (UCS) using small rock samples. *Journal of Petroleum Science and Engineering*, *133*, 367-375. https://doi.org/10.1016/j.petrol.2015.06.022

Aju, D.E., Onyelowe, K.C., Alaneme, G.U. (2021). Constrained vertex optimisation and simulation of the unconfined compressive strength of geotextile reinforced soil for flexible pavement foundation construction. *Cleaner Engineering and Technology*, *5*, 100287. https://doi.org/10.1016/j.clet.2021.100287

Ashfaq, M., Iqbal, M., Khan, M.A., Jalal, F.E., Alzara, M., Hamad, M., Yosri, A.M. (2022). GEP tree-based computational AI approach to evaluate unconfined compression strength characteristics of Fly ash treated alkali contaminated soils. *Case Studies in Construction Materials*, *17*(September), e01446. https://doi.org/10.1016/j.cscm.2022.e01446

Barua, L., Zou, B. (2021). Planning maintenance and rehabilitation activities for airport pavements: A combined supervised machine learning and reinforcement learning approach. *International Journal of Transportation Science and Technology*, *11*(2), 423-435. https://doi.org/10.1016/j.ijtst.2021.05.006

Cabalar, A.F., Hassan, D.I., Abdulnafaa, M.D. (2017). Use of waste ceramic tiles for road pavement subgrade. *Road Materials and Pavement Design*, *18*(4), 882-896. https://doi.org/10.1080/14680629.2016.1194884

Cao, R., Leng, Z., Hsu, S.C., Hung, W.T. (2020). Modelling of the pavement acoustic longevity in Hong Kong through machine learning techniques. *Transportation Research Part D: Transport and Environment*, *83*(May), 102366. https://doi.org/10.1016/j.trd.2020.102366

Deboucha, S., Mamoune, A., Mohammed, S., Universitaire, C., Temouchent, A., Sail, Y. (2020). *Effects of Ceramic Waste, Marble Dust, and Cement in Pavement Sub-base Layer*. https://doi.org/10.1007/s10706-020-01211-x

Deboucha, S., Ziani, H., Amriou, A. (2023). Effect of Ceramic Waste and Lime in Road Sub-Base Layer. *Advances in Transdisciplinary Engineering*, *43*, 879-884. https://doi.org/10.3233/ATDE230810

Ghafari, S., Ehsani, M., Moghadas Nejad, F. (2022). Prediction of low-temperature fracture resistance curves of unmodified and crumb rubber modified hot mix asphalt mixtures using a machine learning approach. *Construction and Building Materials*, *314*(PB), 125332. https://doi.org/10.1016/j.conbuildmat.2021.125332

Ghorbani, B., Arulrajah, A., Narsilio, G., Horpibulsuk, S., & Bo, M. W. (2020). Development of genetic-based models for predicting the resilient modulus of cohesive pavement subgrade soils. *Soils and Foundations*, *60*(2), 398-412. https://doi.org/10.1016/j.sandf.2020.02.010

Ghorbani, B., Arulrajah, A., Narsilio, G., Horpibulsuk, S., Win Bo, M. (2021). Thermal and mechanical properties of demolition wastes in geothermal pavements by experimental and machine learning techniques. *Construction and Building Materials*, *280*, 122499. https://doi.org/10.1016/j.conbuildmat.2021.122499

Gong, H., Sun, Y., Shu, X., Huang, B. (2018). Use of random forests regression for predicting IRI of asphalt pavements. *Construction and Building Materials*, *189*, 890-897. https://doi.org/10.1016/j.conbuildmat.2018.09.017

Huang, J., Shiva Kumar, G., Ren, J., Zhang, J., Sun, Y. (2021). Accurately predicting dynamic modulus of asphalt mixtures in low-temperature regions using hybrid artificial intelligence model. *Construction and Building Materials*, *297*, 123655. https://doi.org/10.1016/j.conbuildmat.2021.123655

Issa, A., Samaneh, H., Ghanim, M. (2021). Predicting pavement condition index using artificial neural networks approach. *Ain Shams Engineering Journal*, *13*(1), 101490. https://doi.org/10.1016/j.asej.2021.04.033

Kumar, C.V., Swaminathen, A.N., Vittalaiah, A., Kumar, C.R. (2022). Materials Today : Proceedings Pavement Subgrade Stabilized with Waste coal Ash and Geosynthetics : An Experimental Study and Multiple Regression Modelling. *Materials Today: Proceedings*, *52*(3), 1543-1550. https://doi.org/10.1016/j.matpr.2021.11.233

Li, Y., Chen, J., Dan, H., Wang, H. (2022). Probability prediction of pavement surface low temperature in winter based on bayesian structural time series and neural network. *Cold Regions Science and Technology*, *194*(August 2021), 103434. https://doi.org/10.1016/j.coldregions.2021.103434

Mabrouk, G.M., Elbagalati, O.S., Dessouky, S., Fuentes, L., Walubita, L.F. (2021). Using ANN modeling for pavement layer moduli backcalculation as a function of traffic speed deflections. *Construction and Building Materials*, *315*, 125736. https://doi.org/10.1016/j.conbuildmat.2021.125736

Mohd, M.H., Koting, S., Hashim, H., Mo, K.H., Aziz, S.A. (2024). Utilising tile waste as an additive to enhance lime-based subgrade stabilisation. *Case Studies in Construction Materials*, *20*, e03342. https://doi.org/10.1016/j.cscm.2024.e03342

Nouhi, S., Pour, M. (2021). Prediction of surface roughness of various machining processes by a hybrid algorithm including time series analysis, wavelet transform and multi view embedding. *Measurement: Journal of the International Measurement Confederation*, *184*, 109904. https://doi.org/10.1016/j.measurement.2021.109904

Oluwaseun, P. (2018). *Mechanical and microstructural characterisation of ceramic-laterized concrete composite*. *May*.

Pérez-Acebo, H., Gonzalo-Orden, H., Findley, D.J., Rojí, E. (2021). Modeling the international roughness index performance on semi-rigid pavements in single carriageway roads. *Construction and Building Materials*, *272*. https://doi.org/10.1016/j.conbuildmat.2020.121665

Piryonesi, S.M., El-Diraby, T. (2021). Climate change impact on infrastructure: A machine learning solution for predicting pavement condition index. *Construction and Building Materials*, *306*(November 2020), 124905. https://doi.org/10.1016/j.conbuildmat.2021.124905

Saber, S. A., Iravanian, A. (2022). Using Waste Ceramic Dust in Stabilization of Clay Soils. *International Journal of Sustainable Construction Engineering and Technology*, *13*(1), 68-79. https://doi.org/10.30880/ijscet.2022.13.01.007

Sadat Hosseini, A., Hajikarimi, P., Gandomi, M., Moghadas Nejad, F., Gandomi, A.H. (2021). Optimised machine learning approaches for the prediction of viscoelastic behavior of modified asphalt binders. *Construction and Building Materials*, *299*(July), 124264. https://doi.org/10.1016/j.conbuildmat.2021.124264

Sebaaly, H., Varma, S., Maina, J.W. (2018). Optimising asphalt mix design process using artificial neural network and genetic algorithm. *Construction and Building Materials*, *168*, 660-670. https://doi.org/10.1016/j.conbuildmat.2018.02.118

Setyawan, A., Nainggolan, J., Budiarto, A. (2015). Predicting the remaining service life of road using pavement condition index. *Procedia Engineering*, *125*, 417-423. https://doi.org/10.1016/j.proeng.2015.11.108

Sivabalaselvamani, D., Kulanthaivel, P., Yogapriya, J., Dhanoa, I.S. (2022). Study on engineering strength properties of ceramic waste powder stabilised expansive soil using machine learning algorithms. *Journal of Ceramic Processing Research*, *23*(6), 902-911. https://doi.org/10.36410/jcpr.2022.23.6.902

Sollazzo, G., Fwa, T.F., Bosurgi, G. (2017). An ANN model to correlate roughness and structural performance in asphalt pavements. *Construction and Building Materials*, *134*, 684-693. https://doi.org/10.1016/j.conbuildmat.2016.12.186

Wakjira, T.G., Abushanab, A., Alam, M.S. (2024). Hybrid machine learning model and predictive equations for compressive stress-strain constitutive modelling of confined ultra-high-performance concrete (UHPC) with normal-strength steel and high-strength steel spirals. *Engineering Structures*, *304*(October 2023), 117633. https://doi.org/10.1016/j.engstruct.2024.117633

Xiao, G., Xing, J., Zhang, Y. (2020). Surface roughness prediction model of GH4169 superalloy abrasive belt grinding based on multilayer perceptron(MLP). *Procedia Manufacturing*, *54*(Det 2020), 269-273. https://doi.org/10.1016/j.promfg.2021.07.042