Machine learning techniques versus classical statistics in strength predictions of eco-friendly masonry units

Conference Paper · November 2021

DOI: 10.1109/ICECCO53203.2021.9663760

2 authors:

Assia Aboubakar Mahamat
Nile University of Nigeria

Moussa Mahamat Boukar
Nile University of Nigeria

Machine learning techniques versus classical statistics in strength predictions of eco-friendly masonry units

Assia Aboubakar Mahamat
Department of Civil Engineering
Nile University of Nigeria
Abuja, Federal Capital Territory, Nigeria

Moussa Mahamat Boukar
Faculty of Natural Applied Science
Nile University of Nigeria
Abuja, Federal Capital Territory, Nigeria

Abstract**—** Earth-based materials demonstrated promising characteristics in the development of eco-friendly, low cost and sustainable construction materials. However, their unconventional utilization in construction makes the assessment of their properties very difficult and inaccurate because they are assessed based on conventional materials procedures. Hence, the properties of earth-based materials are not well understood. The assessment of earth-based materials properties for sustainable construction is time-consuming, expensive, and inaccurate. To obtain more accurate properties, an artificial neural network and statistical linear regression analysis were used to predict the compressive strength of alkali-activated soil. Statistical linear regression analysis was carried out to compare the efficiency of the machine learning technique with the classical statistics model. Parameters such as Si/Al, activator level, curing temperature, water absorption, and weight were used as input parameters to predict the target variable. The coefficient of determination was used to examine the performance of the models. The results depict that artificial neural network outperformed statistical linear regression analysis with R^2 =0.74, RMSE=0.119 and R^2 =0.48, RMSE=0.466 respectively. This indicates that statistical linear regression analysis is inefficient for prediction of the strength in alkali activated soils.

Keywords**—** *machine learning, artificial neural network, statistical linear regression, eco-friendly masonry bricks, compressive strength.*

I. INTRODUCTION

Artificial intelligence (AI) techniques are like human brains and capable of solving very complex variables. Machine learning (ML) is a subset of AI and it is used to anticipate and evaluate various properties in different fields like engineering, hydrology and building materials [1][2]. However, the predictive model is derived by learning algorithms that search for hidden structure and patterns in the training data set [3]. Nowadays, artificial neural network (ANN) is being significantly employed to predict the compressive strength of concrete. The ANN is a mimic of the natural neural system, using computer software and electronic components which makes it more useful than the classical statistics methods where it's difficult to obtain efficient prediction when the correlation between input and output is non-linear. ANN models were used by Chithra to predict the compressive strength of high-performance concrete (HPC) containing nano silica and copper slag as partial cement and fine aggregate replacement respectively [2]. It was proved that the ANN-based strength prediction model could be successfully used to predict the strength of concrete for various mix ingredients [4] scaling from conventional materials [5] to new materials [6]. Moreover, statistical linear regression (SLR) is a prediction method that operates based on the dataset. The value of a variable (output) can be estimated based on another given value (input). This can be accomplished by estimating the output from a least-squares curve that fits the dataset. The use of a linear regression curve of input over output, is aimed to replacing the vertical deviations in the definition of the least-squares curve. SLR is a basic and commonly used type of predictive statistical analysis used to explain the relationship between one dependent variable and one or more independent variables. It is a useful method for solving binary classification problem [1]. Namyong have used the regression equations for prediction of in-situ concrete compressive strength and for this purpose they have used the information of mixture proportions of ready-mixed concrete and test results of compressive strength from construction sites [7]. On the other hand, Eco-friendly construction materials refer to construction materials that are "green" (manufactured or naturally occurring) and not environmentally harmful throughout their life cycle. In addition to their "green" manufacturing process, natural resources protection plays a significant role in the new requirements of the construction industry [5]. Earth-based materials fulfill the "green" requirements of the construction industry to replace conventional materials [8]. In the last decades researchers have explored the use of various earth-based materials produced through environmental-friendly processes [9][10] to overcome the challenges facing by the construction field in terms of CO_2 emissions. However, the procedure of selecting appropriate component and investigate their properties for buildings construction is time consuming and not cost-effective. Specially the development of new materials for construction is complex as the properties and state-of-art have not been analyzed previously. Subsequently, testing the performance of these new materials is not properly carried out because they are inexistent in the regulations and standards. Based on published literature, the prediction of independent variables based on dependent variables was successful with the application ANN and SLR techniques [1]. Moreover, in the application of ML in construction field, most researchers use secondary data collected from published literatures. However, the correlation between the various constituent during the prediction, the training and testing data were prepared from experimental primary datasets carried out during this study. This study compares the efficiency of ANN and SLR methods in the prediction of eco-friendly materials' properties in construction's application

II. MATERIALS AND METHODS

A. *Materials*

The vegetal soil was collected from the excavated soil in a construction field in N'djamena, Chad. The particle sizes of the soil were determined via mechanical sieving and hydrometer sedimentation in accordance with BS 1377:2 [11]. It showed that 56% of particles finer than 75µm. In accordance with BS1377:2, Atterberg limits of the soil were performed as seen in TABLE *1*.

TABLE 1. SOIL'S CHARACTERISTICS

Physical Properties	Particle size distribution 56% smaller than 75 µm Moisture content 2.45% Specific gravity 2.61 Dry density 0.458 g/cm3 Optimum moisture content 14% Liquid limit 34.17% Plastic limit 21.08% Plasticity index 13.09%
Chemical Composition (%)	$\mathbf{SiO_2}$ (20.64), $\mathbf{Al_2O_3}$ (8.51), $\mathbf{Fe_3O_4}$ (8.53), $\mathbf{K_2O}$ (2.65)

Potassium alum or potash was obtained from local market in N'djamena, Chad. It is very common because of its utilization for water filtration by the local population. Therefore available, cheap, environmental-friendly compared to the synthetic potash and has satisfactory performances. Its processing is like the described process elsewhere [9]. Microstructural characterizations were performed on the raw materials and mix design.

B. *Methods*

In the bricks production sequence, the mix designs were prepared with 1, 3 and 5wt% of the activator and desired quantity of water was determined. The precursor and activator were mixed dry in a laboratory mixer for 5 min before addition of water at room temperature (27ºC) and mixed for additional 5 min before pouring the paste into molds of dimensions 160 x 40 x 40mm in accordance with ASTM C 1161[12]. The bricks were cured at room temperature for 90 days prior to mechanical testing. The mixture was compressed with a Zwick press to optimize the bricks strength and cured at room temperature for 90 days prior to mechanical testing. The results of the mechanical testing were used as datasets. Microstructural analysis (Scanning electron microscopy-Energy dispersive spectroscopy (SEM-EDS)) was performed on the sample after mechanical failure. That analysis gives information about the chemical component and morphology of the specimens to detect the bond between the particles at the different activation's level. A Carl Zeiss Model EVO LS10 instrumented with an EDX system that can detect elements between sodium (Na) and uranium (U) with high resolution was used to carry out that analysis.

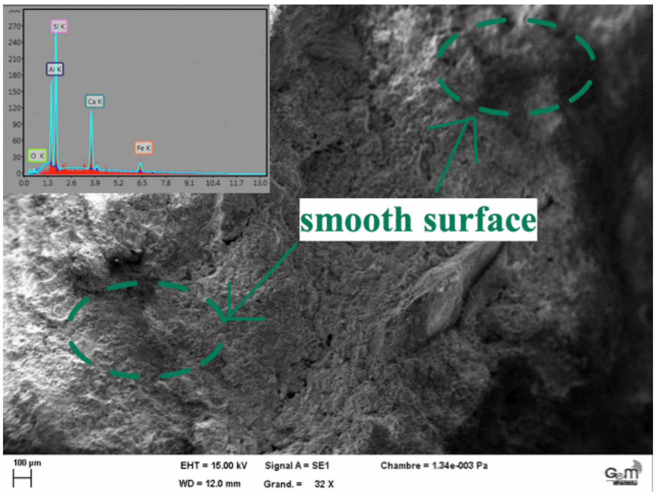
ANN was used because of its capability to learn non-linear models and its capability to learn models in real-time. Therefore, it can learn a non-linear function approximator for either classification or regression. In this study, the validation methods used are the coefficient of determination (R^2) and root mean square (RMSE). SLR analysis was used mainly to establish the comparison between the ANN and classical statistics analysis methods. The process of the analysis with the SLR consisted of evaluating the regression statistics before running the analysis of variance (ANOVA) one way to obtain the intercepts and standard deviations. The predicted/actual values and residuals were used for plotting and fitting the SLR curve. The experiment was implemented in Python 2.7.12 and Minitab for the ANN and SLR respectively. The hardware configuration used for the implementation environment was Intel Core (TM) i5-4790 CPU, 3.60 GHz and 4GB RAM. TABLE *2* contain information's about the datasets.

TABLE 2. DATASETS USED THE TRAINING OF THE PREDICTION'S MODELS

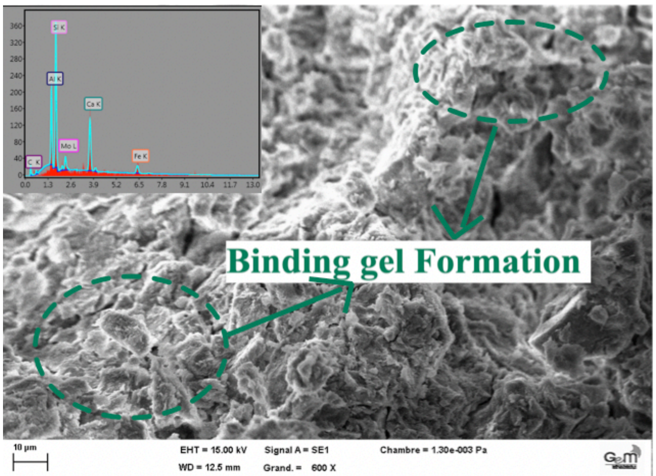
Inputs variables					Output
Al/Si	Potash (%)	Curing Temperature (ºC)	Water absorption (%)	Weight (Kg)	Strength (Mpa)
1.85	0.03	27	10.68	0.191	0
1.91	0.03	27	13.36	0.189	0.1076
1.43	0.03	27	0.74	1.86	0
1.57	0.03	27	5.09	1.87	1.5796
1.72	0.03	60	17.73	0.185	0.5264
1.83	0.03	60	11.91	0.183	2.516
1.85	0.03	27	10.86	0.191	0
1.91	0.03	27	13.33	0.189	0.6084
1.43	0.03	27	0.71	1.86	0
1.57	0.03	27	4.88	1.87	2.7144
1.72	0.03	60	17.07	0.185	0.246
1.83	0.03	60	11.09	0.183	2.252
1.85	0.03	27	10.52	0.191	1.4156
1.91	0.03	27	13.28	0.189	1.42
1.43	0.03	27	0.79	1.86	0.4488
1.57	0.03	27	5.03	1.87	0.7076
1.72	0.03	60	18.01	0.185	0.588
1.83	0.03	60	12.58	0.183	0.1944
1.85	0.03	27	10.48	0.191	0.05184
1.91	0.03	27	13.55	0.189	2.3132
1.43	0.03	27	0.82	1.86	0.696
1.57	0.03	27	5.99	1.87	0.6688
1.72	0.03	60	18.25	0.185	0.48
1.83	0.03	60	12.05	0.183	0.0876
1.31	0.05	60	1.49	0.189	3.215
1.35	0.05	27	2.53	1.86	3.431
1.99	0.05	27	2.45	1.87	0.78
1.62	0.05	27	0.04	0.185	1.512
1.31	0.05	60	1.51	0.183	4.628
1.99	0.05	27	1.89	0.189	0.612
1.62	0.05	27	0.13	1.86	1.416
2.39	0.01	60	10.2	1.87	1.98
3.19	0.01	27	1.58	0.185	0.668
2.29	0.01	27	0.56	1.87	0.844
2.39	0.01	60	9.98	0.185	2.147
3.19	0.01	27	1.74	0.183	0.58
2.29	0.01	27	0.57	0.191	0.839

III. RESULTS *AND DISCUSSIONS*

The Scanning electron micrographs of the raw materials and mix designs give the information's regarding the chemical composition of the specimens. Chemical composition imparts significantly to the strength of construction materials. In alkali activated materials the chemical composition governs the formation of the gel responsible of particles' bonding (Fig. *1*) Duxson has shown the importance of the Si/Al in strength determination of alkaline activated materials [13] and the results obtained from the analysis of Si/Al in this study align with the previous finding.

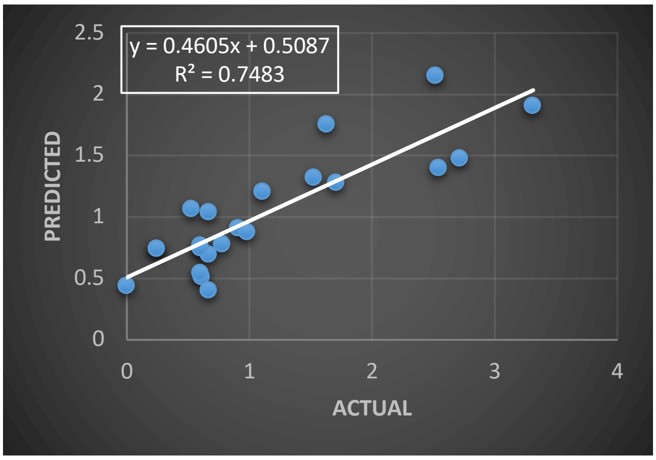


a)

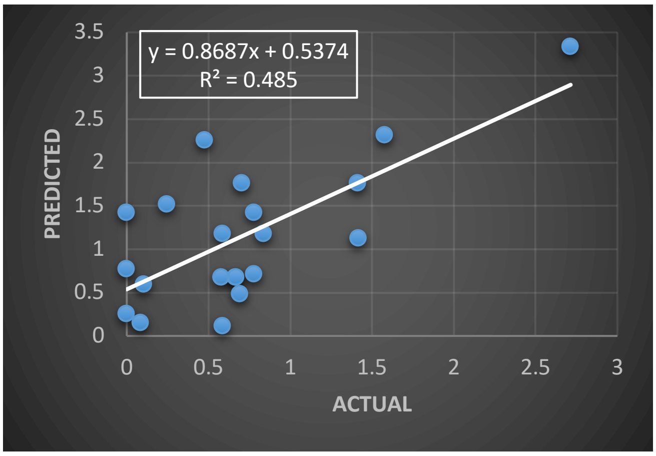


b)

Fig. 1. SEM-EDX micrographs of specimens at activation level of a) 1% and b) 5%



a)



b)

Fig. 2. Prediction vs actual graphs of a) ANN and b) SLR

Examination of the results obtained from the prediction models, shows that the ANN demonstrated higher performance in terms or R^2 (*Fig. 2*) and very low RMSE (0.119). The advantages of ANNs are the unrestricted number of inputs and outputs [14]. Previous works have also shown the efficient performance of ANN in predicting the strength of non-conventional materials [4].

The SLR has the lower performance in terms of R^2 and high RMSE (0.466) because the coefficients estimate for SLR rely on the interdependence of the input variables and output [2]. When features are correlated and the columns of the design matrix X have an approximate linear dependence, the design matrix becomes close to singular and as a result, the least-squares estimate becomes highly sensitive to random errors in the observed output, producing a large variance which could result in a low coefficient of determination (R^2) [1].

IV. CONCLUSION

In this study alkali activation technique and mechanical compaction were used as Eco-friendly process to optimize efficiently the properties of earth-based construction materials for potential application in masonry. The comparative study of the ML technique and SLR was effective, and the following conclusions can be drawn:

1. A comparison between ANN and LR methods can be used to select effectual the most performant prediction method. It depicts that ANN outperformed SLR in predicting the strength of the alkali activated soil.

2. The results obtained from ANOVA, in terms of prediction versus experimental values showed a scarce plot proving the inefficiency of using ANOVA to predict this relationship. Additionally, the RMSE is used to evaluate the error that emerged during the training, testing and validation. The ANOVA results showed higher RMSE compared to ANN.
3. SLR primary limitation is its inability to effectively predict nonlinear relationship between the components. Therefore, its modeling performance is reportedly poor proving the nonlinearity of the inputs and target correlation. Indeed, alkali activated materials are novel materials, the connection between their properties is not adequately simulated.

4. ANN displayed very insignificant difference between predicted and actual value. And the coefficient of determination (R^2) used to evaluate the feasibility of the ANN methods was close to 1, demonstrating a good fitting of the model. The values obtained for the coefficient of determination (R^2) was of 0.74 and 0.48 for ANN and SLR respectively. While the RMSE results were 0.119 and 0.466 for ANN and SLR respectively demonstrating the error level for the models.
5. Results have indicated that ANN models are practical and highly efficient for predicting the strength of alkali activated soil. Therefore, by adopting ANN models, there is no need to go through time-consuming and costly laboratory tests to obtain properties of alkali activated msterials.

Acknowledgment

The authors would like to acknowledge the Multifunctional Materials Research Group at the African University of Science and Technology, Abuja-Nigeria, the staff of the civil engineering laboratory of the Nile University of Nigeria, Abuja-Nigeria, and the staff of the Institut de Recherche en Génie Civil et Mécanique (GeM) équipe de recherche Interactions Eau-Géomatériaux (IEG), Saint Nazaire-France for their care and useful suggestions.

References

[1] A. A. Mahamat et al., "Machine learning approaches for prediction of the compressive strength of alkali activated termite mound soil," Applied Sciences (Switzerland), vol. 11, no. 11, 2021, doi: 10.3390/app11114754.
[2] S. Chithra, S. R. R. S. Kumar, K. Chinnaraju, and F. Alfin Ashmita, "A comparative study on the compressive strength prediction models for High Performance Concrete containing nano silica and copper slag using regression analysis and Artificial Neural Networks," Construction and Building Materials, vol. 114, pp. 528–535, 2016, doi: 10.1016/j.conbuildmat.2016.03.214.
[3] J. M. Marangu, "Prediction of Compressive Strength of Calcined Clay Based Cement Mortars Using Support Vector Machine and Artificial Neural Network Techniques," Journal of Sustainable Construction Materials and Technologies, vol. 5, no. 1, pp. 392–398, Apr. 2020, doi: 10.29187/jscmt.2020.43.
[4] A. M. Al-Swaidani and W. T. Khwies, "Applicability of Artificial Neural Networks to Predict Mechanical and Permeability Properties of Volcanic Scoria-Based Concrete," Advances in Civil Engineering, vol. 2018, no. i, 2018, doi: 10.1155/2018/5207962.
[5] H. Naderpour, A. H. Rafiean, and P. Fakharian, "Compressive strength prediction of environmentally friendly concrete using artificial neural networks," Journal of Building Engineering, vol. 16, no. January, pp. 213–219, 2018, doi: 10.1016/j.jobe.2018.01.007.
[6] I. I. Obianyo et al., "Multivariate regression models for predicting the compressive strength of bone ash stabilized lateritic soil for sustainable building," Construction and Building Materials, vol. 263, Dec. 2020, doi: 10.1016/j.conbuildmat.2020.120677.
[7] J. Namyong, Y. Sangchun, and C. Hongbum, "9 Journal of Asian Architecture and Building Engineering/ Prediction of Compressive Strength of In-Situ Concrete Based on Mixture Proportions," 2004.
[8] A. A. Mahamat, N. Linda Bih, O. Ayeni, P. Azikiwe Onwualu, H. Savastano, and W. Oluwole Soboyejo, "Development of Sustainable and Eco-Friendly Materials from Termite Hill Soil Stabilized with Cement for Low-Cost Housing in Chad," Buildings, vol. 11, no. 3, p. 86, Feb. 2021, doi: 10.3390/buildings11030086.
[9] A. A. Mahamat et al., "Alkali activation of compacted termite mound soil for eco-friendly construction materials," Heliyon, vol. 7, no. 3, Mar. 2021, doi: 10.1016/j.heliyon.2021.e06597.
[10] B. Ngayakamo, A. M. Aboubakar, C. G. Komadja, A. Bello, and A. P. Onwualu, "Eco-friendly use of eggshell powder as a bio-filler and flux material to enhance technological properties of fired clay bricks," Metallurgical and Materials Engineering, vol. 27, no. 3, pp. 371–383, Jun. 2021, doi: 10.30544/628.
[11] BS 1377, "idoc.pub_bs-1377-part-2-methods-of-test-for-soils-for-civil-en80cpdf," British Standard Institutions, 2020.
[12] "ASTM C1161-02, Standard Test Method for Flexural Strength of Advanced Ceramics at Ambient Temperature," ASTM International, West Conshohocken, PA, 2002, Accessed: Jul. 08, 2021. [Online]. Available: www.astm.org Highlight
[13] P. Duxson, S. W. Mallicoat, G. C. Lukey, W. M. Kriven, and J. S. J. van Deventer, "The effect of alkali and Si/Al ratio on the development of mechanical properties of metakaolin-based geopolymers," Colloids and Surfaces A: Physicochemical and Engineering Aspects, vol. 292, no. 1, pp. 8–20, Jan. 2007, doi: 10.1016/j.colsurfa.2006.05.044.
[14] J.-S. Chou, C.-K. Chiu, M. Farfoura, and I. Al-Taharwa, "Optimizing the Prediction Accuracy of Concrete Compressive Strength Based on a Comparison of Data-Mining Techniques," Journal of Computing in Civil Engineering, vol. 25, no. 3, pp. 242–253, May 2011, doi: 10.1061/(asce)cp.1943-5487.0000088.