Proceedings of the XVII ECSMGE-2019
Geotechnical Engineering foundation of the future
ISBN 978-9935-9436-1-3

doi: 10.32075/17ECSMGE-2019-0422

XVII European Conference
ECSMGE 2019
Reykjavik
on Soil Mechanics
and Geotechnical Engineering

Comparison of multiple regression, and artificial neural networks in estimating compaction characteristics

Comparaison de la régression multiple et des réseaux de neurones artificiels dans l'estimation des caractéristiques de compactage

C. Atalar
Near East University, Nicosia, North Cyprus, Mersin 10, Turkey

E. Adu-Parkoh

WBHO GH LTD, Accra, Ghana

B. M. Das
California State University Sacramento, Sacramento, USA

A. Özyankı

Near East University, Nicosia, North Cyprus, Mersin 10 Turkey

ABSTRACT: One of the most important techniques that can be used for soil improvement is the compaction technique. Laboratory experiments of compaction properties; moisture dry density and optimal water contents are laborious and time-consuming processes. In order to determine these properties, predictive models can be developed by using Atterberg limits and grain-size distribution. Statistical techniques such as multiple linear regression analysis and artificial neural network can be used as predictive models. In this study, standard proctor test, Atterberg limits, and grain-size distribution techniques were used on 77 lateritic soil samples from Ghana. Prediction models were constructed using the liquid limit, the plasticity index, and the percentage of fine grain content, to estimate the moisture dry density, and optimal water contents. Generated models were tested and compared with multiple regression analysis and artificial neural networks. As a result, the predicted values of regression analysis showed between 65%-%90, and artificial neural networks models showed between 80%-%90 compatibility. These results show that artificial neural networks gave better results than multiple linear regression analysis to predict the compaction properties.

RÉSUMÉ: L'une des techniques les plus importantes utilisés pour améliorer la qualité des sols est le compactage. Les expériences de laboratoire sur les propriétés de compactage telles que la densité sèche, d'humidité et les teneurs optimales en eau sont des processus laborieux et fastidieux. Afin de déterminer ces propriétés, à l'aide des Limites d'Atterberg et des distributions granulométriques, des modèles prédictifs peuvent être développés. Des techniques statistiques telles que : l'analyse par régression linéaire multiple et par réseau neuronal artificiels peuvent être utilisées comme modèles prédictifs. Dans cette étude, le test standard Proctor, les limites d'Atterberg et les techniques de distribution granulométrique ont été utilisés sur 77 échantillons de sol latéritique provenant du Ghana. Les modèles de prévision ont été construits à l'aide de la limite de liquidité, de l'indice de plasticité et du pourcentage de teneur en grains fins, afin d'estimer la densité sèche et la teneur optimale en eau. Les modèles générés ont été testés et comparés avec l'analyse de régression multiple et les réseaux de neurones artificiels. En conséquence, les valeurs prédites de l'analyse de régression montraient une compatibilité entre 65% et 90% et les modèles de réseaux de neurones artificiels montraient une compatibilité entre 80% et 90%. Ces résultats montrent que afin de prédire les propriétés de compactage des sols, les réseaux de neurones artificiels procurent des résultats meilleurs que ceux obtenus par analyse de régression linéaire multiple..

Keywords: Standard proctor test; multiple regression analysis; artificial neural network; Lateritic soils; Ghana

1 INTRODUCTION

P The mechanical processing in which we increase the soil density by reducing the volume of air in the soil is called soil compaction (Holtz et al., 2011). This process causes variations in particle distribution, pore size, and soil durability. By increasing compressibility, and soil porosity, increasing soil strength, and stiffness is the main purpose of compression (Rollings and Rollings, 1996). Some of the most important parts of the compaction process are soil class and grain size as it causes a decrease in the pore space in the soil, and increases the density of the soil. The density of the clay and silty soils is lower than the coarse grained soils because of high pore space. In order to explain compaction characteristics theory, we can use compaction curve which obtained from laboratory test or field compaction (Hausmann, 1990).

Many researchers have tried to estimate compaction parameters from various soil parameters. One of the early researches, Joslin (1958) defined the curves known as the Ohio compaction curves with the curves obtained from his study. Torrey (1970) initiated a new discussion in his research by saying that there was a mathematical relationship between Atterberg limits and compaction parameters. The compaction characteristics of lateritic soils were estimated using empirical equations (Adu-Parkoh et al., 2016).

In the studies of civil engineering and geotechnical engineering, Artificial Neural Network (ANN) has been widely used since early 1990 (Rafiq et al., 2001; Basma and Kallas, 2004). In the previous studies, it is observed that ANN is frequently used in estimating the compaction, uplift of pile foundations, axial, and lateral load capacities (Goh, 1996; Hanna et al., 2004), drilled pole (Shahin and Jaksa, 2009), foundation settlements (Sivakugan et al., 1998), and anchors embedment (Rahman et al., 2001; Shahin and Jaksa, 2006).

The ANN method is applied to other applications in earth sciences; retaining walls (Ghaleini et al., 2018), dams (Stojanovic et al., 2016), earthquake (Dindar et al., 2017), geographical information systems (Aslantaş and Kurban, 2007), mining (Afram et al., 2017), geoenvironmental engineering (Shang et al., 2004), petroleum engineering (Kulga et al., 2018), and rock mechanics (Kanungo et al., 2014).

2 STUDY AREA

Ghana, is a country with an area of about 240,000 km^2 on West Africa's Gulf of Guinea. The geology of Ghana is primarily very ancient crystalline basement rock, volcanic belts and sedimentary basins, affected by periods of igneous activity and two major orogeny mountain building events. Ghana is underlain partly by what is known as the Basement complex. It comprises a wide variety of Precambrian igneous and metamorphic rocks which covers about 54% of the country's area; mainly the southern, and western parts of the country (**Error! Reference source not found.**). The primary components are gneiss, phyllites, schists, migmatites, granite-gneiss, and quartzites. The rest of the country is underlain by Paleozoic consolidated sedimentary rocks referred to as the

Voltaian Formation consisting mainly of sandstones, shale, mudstone, sandy, pebbly beds, and limestones (Dapaah-Siakwan and Gyau-Boakye, 2000).

Laterites or residual soils occurs in tropical and sub-tropical countries of different parts of Africa. These soils are used in the construction of roads, earth dams, etc. Their colour is red, their mineralogical composition is different. They are formed at temperature above of 25°C and at an annual rainfall about 1200mm. The soil samples were obtained from depths of about 300 mm to 2 metres during the construction of Tailings Storage Facility, TSF dam for a gold mine in Tarkwa, Ghana (Adu-Parkoh, 2016). Figure 2 shows the site plan of the Tailings Storage Facility, TSF dam.

3 ANALYSIS METHODS

In this study, multiple linear regression (MLR), and ANN methods were applied separately to estimate moisture dry density (MDD) and optimal water contents (OMC) values. The two methods were compared with each other in estimating ground compaction properties. The coefficient of determination (R^2) is the parameter used in our comparison of these two methods (Equation 1).

$$R^2 = 1 - \frac{SSE}{SST} \tag{1}$$

Where; SSE= $\sum(y_i - \bar{y})^2$ is the residual sum of square, and SST is the total sum of squares, $\sum y_i^2$.

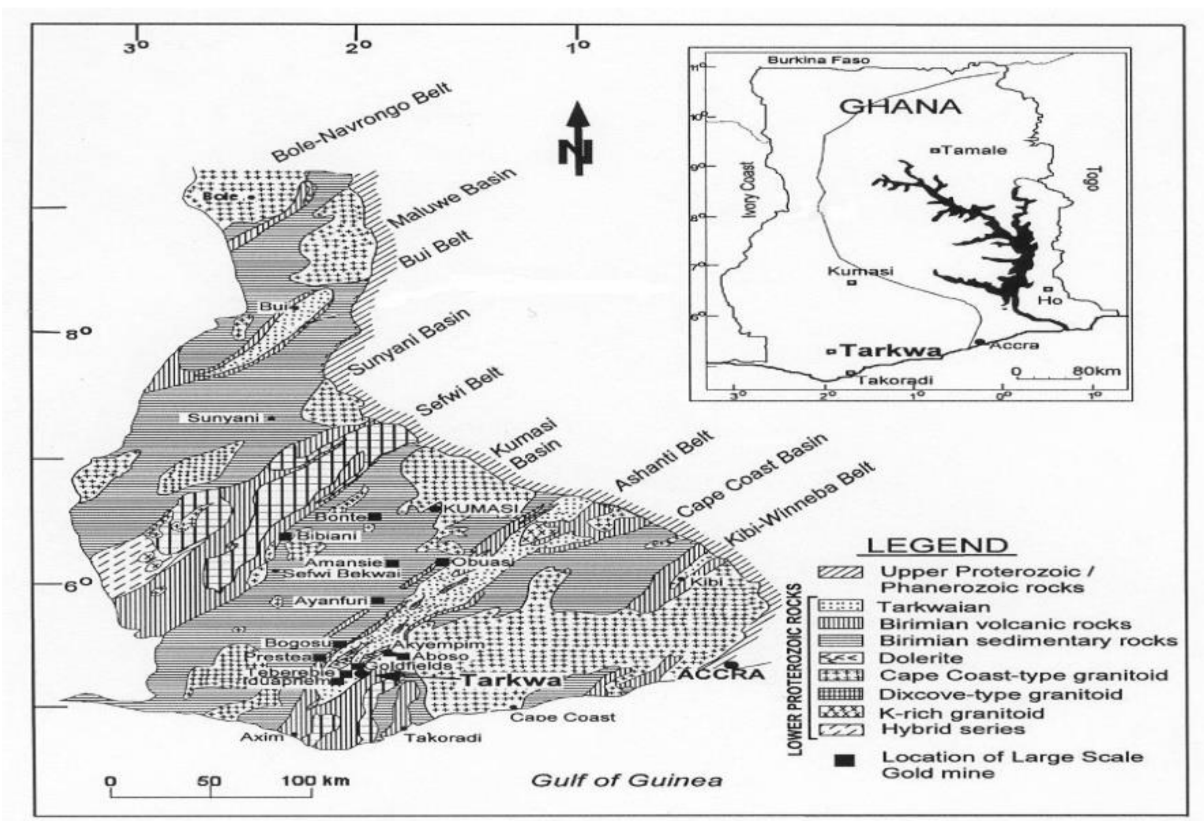

Figure 1 Simplified geological map of southwest Ghana (modified from Kuma, 2004)

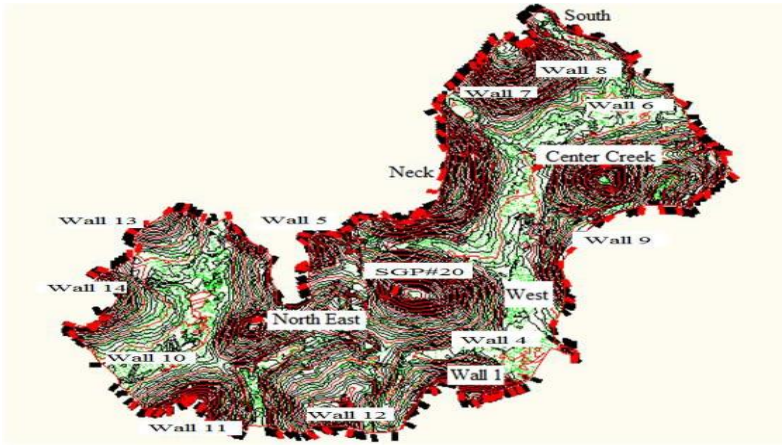

Figure 2 Site layout of the TSF dam, Tarkwa (Adu-Parkoh, 2016)

3.1 Multiple linear regression

In this study, we tried to obtain mathematical definitions for dry unit weight and optimum water content values by using multiple linear regression analysis method which is made by using variables plastic limit (I_P), plasticity index (w_L), and fine grain content (FC) values of 77 samples. Of these data, 64 were used in regression and 13 were used in test the regression formulation. The range of our data is shown in Table 1.

As a result of multiple linear regression analysis, two different mathematical correlations were obtained for MDD (Equation 2) and OMC (Equation 3) values.

$$MDD=19.02-0.017FC-0.077w_L-0.215I_p \quad (2)$$
R^2=0.743

$$OMC=19.30+0.039\ FC +0.114\ w_L -0.399\ I_p \quad (3)$$
R^2=0.649

The comparison between predicted and measured MDD are shown in Figure 3, and the comparison between predicted and measured OMC are shown in Figure 4.

Table 1 Statistical description of data

	FC %	w_L %	I_p %	MDD kN/m3	OMC %
N	77	77	77	77	77
Range	56.3	31.8	32.3	7.9	13.6
Min.	8.1	19.6	.7	16.3	8.8
Max.	64.4	51.4	32.9	24.3	22.4
Mean	37.7	38.7	20.6	19.9	14.9
Std. Dev.	14.85	7.54	9.37	1.62	3.37
Var.	220.5	56.8	87.8	2.6	11.3
Skew.	.081	-.33	-.93	-.37	.63
Kurt.	-1.00	-.66	-.40	-.13	-.59

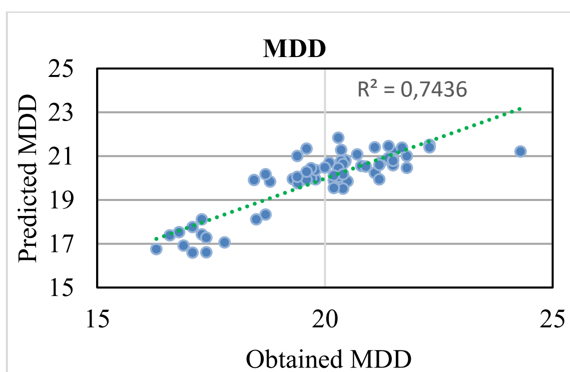

Figure 3 Comparison between Predicted, and obtained MDD

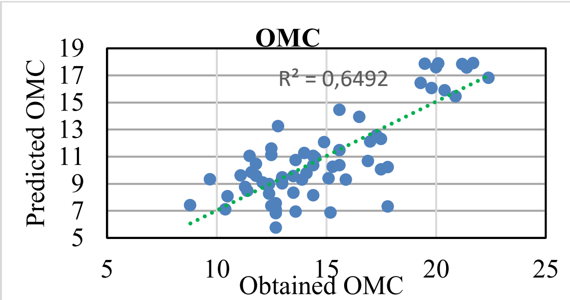

Figure 4 Comparison between Predicted, and measured OMC

The mathematical model obtained as a result of multiple linear regression analysis was tested on 13 data. According to the results of this test, the comparison between real data, and calculated data is given in Figures 5, and Figure 6.

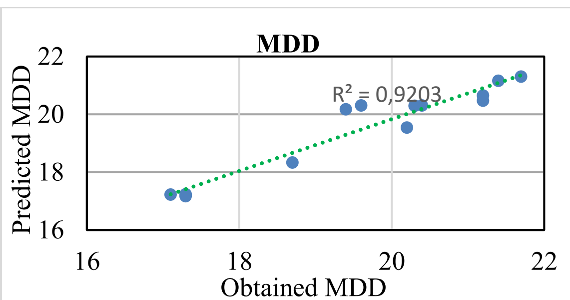

Figure 5 Testing of equation 1 with measured, and predicted MDD values

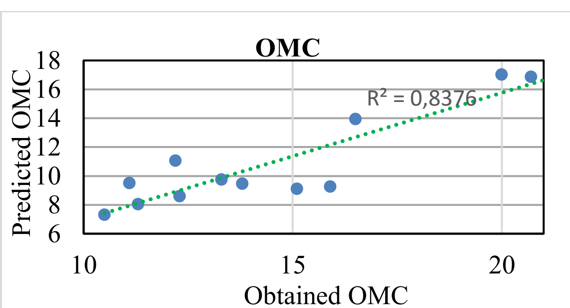

Figure 6 Testing of equation 1 with measured, and predicted MDD values

3.2 Artificial neural networks

ANN are computer-based modeling, and statistical techniques that mimic the human brain's thinking, and acting characteristics. This system consists of the input layer, hidden layers, the output layer, weights (w), and bias (b) as shown in Figure 4.4. Input values for the ANN model are given in the system, and multiplied by the corresponding weights. After that, entries with weight sums from all input sources are added to the hidden layers. After data is generated, the hidden layer transfer function is activated, and it is calculated as the input layer. This process maintains until the output layer is obtained.

A total of 77 data is devoted to trained, and simulate an ANN model. Of these data, 64 were used in training, and 13 were used in simulation. I_P, w_L, and FC data were used as input parameters for comparison of ANN, and MLR analysis. The MDD, and OMC were used as output parameters separately. While creating an ANN model, normalization process (4) was applied to values in order to group the data in a certain order, and a certain range (0-1). Another advantage of this process is to increase processing time, and reliability.

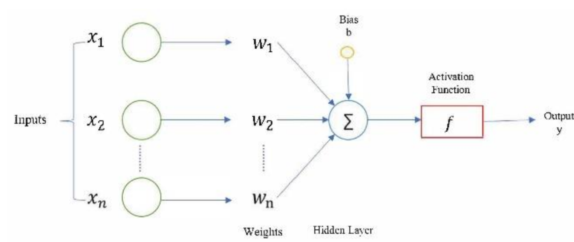

Figure 7 The generalized ANN model

A total of 77 data is devoted to trained, and simulate an ANN model. Of these data, 64 were used in training, and 13 were used in simulation. I_P, w_L, and FC data were used as input parameters for comparison of ANN, and MLR analysis. The MDD, and OMC were used as output parameters separately. While creating an ANN model, normalization process (Equation 4)

was applied to values in order to group the data in a certain order, and a certain range (0-1). Another advantage of this process is to increase processing time, and reliability.

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4}$$

The ANN regression analysis results are shown in Figure 8, and Figure 9.

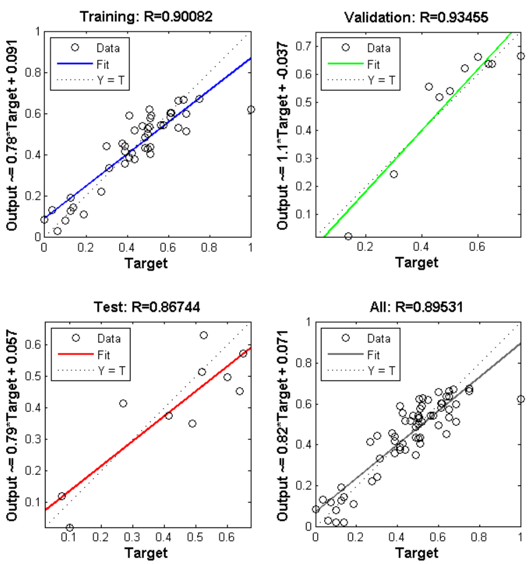

Figure 8 ANN regression analysis result for MDD values.

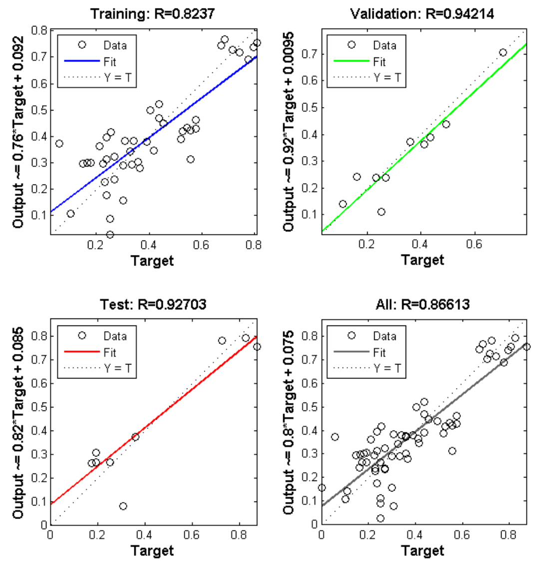

Figure 9 ANN regression analysis result for OMC values.

After checking the predicted values, and real data relationship we calculated an R^2 value (Figure 10, and Figure 11).

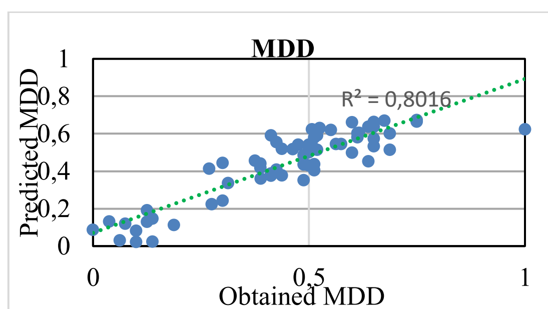

Figure 10 Comparison between predicted, and measured MDD using ANN.

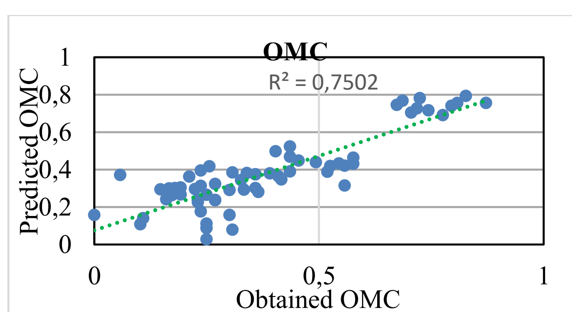

Figure 11 Comparison between predicted, and measured OMC using ANN

The model obtained as a result of ANN was tested on 13 data. According to the results of this test, the comparison between real data, and calculated data is given in Figures 12, and Figure 13.

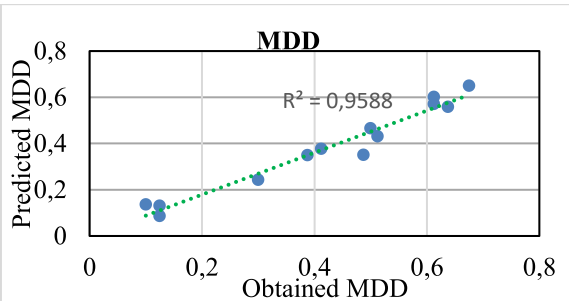

Figure 12 Testing data comparison between predicted, and measured MDD using ANN

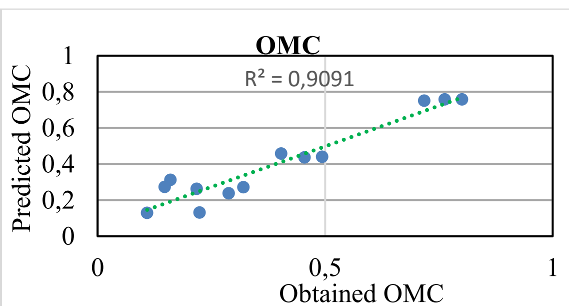

Figure 13 Testing data comparison between predicted, and measured OMC using ANN

4 RESULTS

R^2 values of these relations were calculated as 0.743 for MDD, and 0.649 for OMC. These values indicate that the estimation process can be performed with high accuracy. The regression analysis results were testing by some data for validate equations. For this purpose, R^2 values were calculated with 13 real data, and calculated data. The results are 0.920 for MDD values, and 0.834 for OMC values.

As a result of the training of the model, R^2 for MDD was calculated as 0.802 in training, 0.959 in simulation, and R^2 for OMC was calculated as 0.750 in testing, and 0.909 in simulation.

The comparison with MLR, and ANN results are given in Table 2.

Table 2 Comparison with MLR, and ANN results

	R^2			
	Training		**Testing**	
	MDD	**OMC**	**MDD**	**OMC**
MLR	0.74	0.65	0.92	0.84
ANN	0.80	0.75	0.96	0.91

5 CONCLUSIONS

In this study, multiple linear regression analysis and artificial neural network methods were used for estimation moisture dry density and optimal water content data. According to the results of the study, both methods had high R^2 values. Besides, it was found that artificial neural network method gives better results in estimation of both moisture dry density and optimal water content values.

6 REFERENCES

Adu-Parkoh, E. 2016. Prediction of compaction characteristics of lateritic soils in Ghana, Master Thesis, Near East University, North Cyprus.

Adu-Parkoh, E., Das, B.M. & Atalar, C. 2016. Estimation of compaction characteristics of lateritic soils, *ZM2016: Proceedings, 4th ICNDSMGE Conference* (Eds: Atalar, C., Çinicioğlu, F., Das, B.M., Sağlamer, A. & Toğrol, E.), 719-728. Near East University Press, Lefkoşa, North Cyprus.

Afram, A., Janabi-Sharifi, F., Fung, A. S. & Raahemifar, K. 2017. Artificial neural network (ANN) based model predictive control (MPC), and optimization of HVAC systems: A state of the art review, and case study of a residential HVAC system.

Energy, and Buildings, *141*, 96–113.

Aslantaş, V. & Kurban, R. 2007. Artificial neural networks, and applications in geographical information systems. In *TMMOB Harita ve Kadastro Mühendisleri Odası Ulusal Coğrafi Bilgi Sistemleri Kongresi*. Trabzon. Retrieved from https://web.itu.edu.tr/tahsin/cbs2007/bildiri/P_45.pdf

Basma, A. A. & Kallas, N. 2004. Modeling soil collapse by artificial neural networks. *Geotechnical & Geological Engineering*, *22*(3),427–438.

Dapaah-Siakwan, S. & Gyau-Boakye, P. 2000. Hydrogeologic framework, and borehole yields in Ghana. *Hydrogeology Journal*, *8*(4), 405–416.

Dindar, H., Dimililer, K., Özdağ, Ö. C., Atalar, C., Akgün, M. & Özyankı, A. 2017. Vulnerability Index Assessment Using Neural Networks (Viann): A Case Study Of Nicosia, Cyprus. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, *4,* 189-195.

Ghaleini, E. N., Koopialipoor, M., Momenzadeh, M., Sarafraz, M. E., Mohamad, E. T. & Gordan, B. 2018. A combination of artificial bee colony, and neural network for approximating the safety factor of retaining walls. *Engineering with Computers*, 1–12.

Goh, A. T. C. 1996. Pile driving records reanalyzed using neural networks. *Journal of Geotechnical Engineering*, *122*(6), 492–495.

Hanna, A. M., Morcous, G. & Helmy, M. 2004. Efficiency of pile groups installed in cohesionless soil using artificial neural networks. *Canadian Geotechnical Journal*, *41*(6), 1241–1249.

Hausmann, M. R. 1990. *Engineering Principles of Ground Modification*. New York: McGraw-Hill.

Holtz, R. D., Kovacs, W. D. & Sheahan, T. C. 2011. *An introduction to geotechnical engineering* (2nd ed.). Upper Saddle River (NJ): Pearson.

Joslin, J. G. 1958. Ohio's typical moisture-density curves. *Symposium on Application of Soil Testing in Highway Design, and Construction*, *239*, 111–118.

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*(3), 439–456.

Kulga, B., Artun, E. & Ertekin, T. 2018. Characterization of tight-gas sand reservoirs from horizontal-well performance data using an inverse neural network. *Journal of Natural Gas Science, and Engineering*, *59* (May), 35–46. https://doi.org/10.1016/j.jngse.2018.08.017

Kuma, J.S. 2004. Is groundwater in the Tarkwa gold mining district of Ghana potable. Environmental Geology, 45, 391-400

Rafiq, M. Y., Bugmann, G. & Easterbrook, D. J. 2001. Neural network design for engineering applications. *Computers, and Structures*, *79*(17), 1541–1552.

Rahman, M. S., Wang, J., Deng, W. & Carter, J. P. 2001. A neural network model for the uplift capacity of suction caissons. *Computers, and Geotechnics*, *28*(4), 269–287.

Rollings, M. P. & Rollings, R. S. 1996. *Geotechnical materials in construction*. New York: McGraw-Hill.

Shahin, M. A. & Jaksa, M. B. 2006. Pullout capacity of small ground anchors by direct cone penetration test methods, and neural networks. *Canadian Geotechnical Journal*, *43*(6), 626–637.

Shahin, M. A. & Jaksa, M. B. 2009. Intelligent computing for predicting axial capacity of drilled shafts. In *Contemporary Topics in In Situ Testing, Analysis, and Reliability of Foundations*, 26–33.

Shang, J. Q., Ding, W., Rowe, R. K. & Josic, L. 2004. Detecting heavy metal contamination in soil using complex permittivity, and artificial neural networks. *Canadian Geotechnical Journal*, *41*(6), 1054–1067.

Sivakugan, N., Eckersley, J. & Li, H. 1998. Settlement predictions using neural networks. *Australian Civil Engineering Transactions*, *40*, 49.

Stojanovic, B., Milivojevic, M., Milivojevic, N. & Antonijevic, D. 2016. A self-tuning system for dam behavior modeling based on evolving artificial neural networks. *Advances in Engineering Software*, *97*, 85–95.

Torrey, V. H. 1970. *Analysis of field compaction data. Report 2: Littleville Dam, Westfield River, Massachusetts*. Army Engineer Waterways Experiment Station Vicksburg Miss.