NIPES

Advances in Engineering Design Technology

Journal homepage: www.nipesjournals.org.ng

Comparison Validation of Some Compaction Prediction Models

Ethelbert O. Mezie
Nnamdi Azikiwe University, Awka, Nigeria.
Email: eo.mezie@unizik.edu.ng

ARTICLE INFORMATION

Article history:
Received 21 August 2024
Revised 01 September 2024
Accepted 24 September 2024
Available online 28 September 2014

Keywords:
Compaction, index properties, prediction models, fine-grained soils, coarse-grained soils.

https://doi.org/10.5281/zenodo.13852010

ABSTRACT

Compaction is a very important aspect of civil engineering practice but is a very laborious and time-consuming process that equally consumes enormous amounts of natural resources. The effort to reduce the time, labour, and resources wastage associated with traditional compaction processes led to the effort in the development of some models to predict compaction parameters: maximum dry unit weight (MDUW) and optimum moisture content (OMC), however, there has been no unified model developed for this purpose. Each developed model presents a lot of limitations largely bordering on data used for calibrating the models and region of sample collection. In this work, 18 soil models developed for fine-grained and coarse-grained models were comparatively validated using neutral compaction soil data. The study showed the developed models and the most important soil index properties influencing each model for fine-grained and coarse-grained soils. It also showed the biases in each model and the comparative performance of the model against one another while adducing the possible reasons behind that. Recommendations were made on the most important considerations and decisions to be taken in subsequent efforts toward developing more unified models for the prediction of soil compaction parameters of soils.

1. Introduction

Soil compaction is an essential aspect of geotechnical engineering practice that is employed in most civil engineering construction projects to improve important engineering properties of soils. Soil compaction has been used in construction works as far back as the ancient times of earth road and rammed earth construction works for buildings [1] and most recently in the construction of dams, buildings, bridges, roads, embankments, offshore platforms, modern rammed earth construction, etc [2]. Soil compaction aims to improve the shear strength and the bearing capacity of the foundation soil in the field as well as reduce soil compressibility through adequate control of compaction parameters: maximum dry unit weight (MDUW) and optimum moisture content (OMC).

While compaction is a beneficial process that helps to achieve more dense interlocking of soil particles in the field, it can also conduce to loss of particle size and strength due to breakages, hence it is expedient that the required degree of compaction is achieved and not beyond. This means that compaction control in the field is essential. Field compaction control in most cases depends on the outcome of laboratory compaction processes using one of the popular methods such as Proctor, British standard methods, etc. In some laboratory compaction operations, the determination of MDUW and OMC for field compaction control is the main aim, however, compaction operations are equally pre-requisite for the determination of California bearing ratio (CBR) of soils, soil permeability tests and laboratory determination of soil shear strength parameters, hence compaction is a widespread and continuous operation in civil engineering practice.

Laboratory compaction processes are laborious, time-rending, and resources-consuming indispensable geotechnical processes [3], [4], [5], [6], [7]. The time and cost implications of obtaining compaction parameters: MDUW and OMC especially at the early stages of a project such as testing the suitability of a borrow pit [8] led to ongoing efforts to develop model equations that can be used to obtain these parameters from the easily measured index or physical properties of the same soil. This is owing to the observed relationship between MDUW/OMC as functions of soil type, grain-size distribution, index properties, specific gravity, and mineralogical content of the soil [9], [10], [11], [12], [13], [14], [15].

The development of models for prediction of MDUW/OMC has been ongoing for over 4 decades. Pioneers in this field include [9], [10], [16], [17], [18], [19]. Although several models have been developed covering fine-grained soils and coarse-grained soils, there have been limitations in their applications largely bordering on calibration, the number of samples used in the model development, the variability of the samples used, etc.[3], [20]. Evidence from literature shows that there are more models for fine-grained soils when compared to coarse-grained soils.

Further research has continued in this field with researchers employing more sophisticated techniques to effectively capture the theoretical relationship between MDUW/OMC and index properties, however, there is yet no agreement on a unified model or approach which is invariably the ultimate target of these model development efforts. Many techniques that are broadly classified into graphical methods, statistical methods, and soft computing techniques [21] has been applied in the prediction of compaction parameters. Many of these models are specific for either fine-grained or coarse-grained soils while some attempted to develop unified models for both soil groups. These techniques include but are not limited to Simple Linear Regression (SLR) [22]; Multi-linear Regression (MLR) [2], [4], [5], [6], [23], [24], [25], [26], [27]; numerical analysis [28], [29], Artificial Neural Network (ANN) [5], [6], [30], [31], [32], [33], [34], Least Squares – Support Vector Machine (LS-SVM) [30], [31], Multivariate Adaptive Regression Splines (MARS) [30], [31]; Support Vector Regression (SVR) [5], [35], [36], Multi-gene Expression Programming [3], [6], [37]; Gene Expression Programming (GEP) [37]; ANN and Grey Wolf Optimizer (GWO) [38]; Gaussian Process Regression (GPR)[20], [32]; Support Vector Machine (SVM) [32], [33] Least Squares Boost Randon Forest (LSBoostRF)[32] Long Short-term Memory (LSTM) [32]; non-linear regression [8]; Deep Neural Network (DNN) [27]; Group Method of Data Handling (GmDH) type Neural Network (NN) [39]; Genetic Algorithm (GA) and particle swarm optimization (PSO) relevance vector machine [40]; Random Forest (RF)[33], Extreme Gradient Boosting Tree (XGBoost) [33]; AdaBoost[34], and Tree[34]. The techniques outlined above have been applied singly, complementarily, simultaneously, or comparatively to develop prediction equations, however, there is no agreement on a unified prediction model with subsequent authors employing new models often citing limitations in previous models [7], [20].

This study is intended to assess how the different models perform against each other in prediction. The study tends to examine models where two or more dependent variables are used [15]. The plan of study is first, to outline the methods employed in the development of the models, together with the most significant index properties for the developed models based on the author's views. Second, to collect soil experimental data comprising compaction parameters and soil index properties from literature to make comparative validation of prediction models developed for MDUW/OMC for both fine-grained and coarse-grained soils. The essence of using comparative validation is to mimic strategies used by authors to judge the performance of their models. Through the outcome of the comparative validation, comments would be made on strategies to develop more unified compaction prediction models. It is intended that the outcome of the study would provide good reference material for development of subsequent prediction models and practical applications.

1.1 Model Development Tools

Statistical and soft computing techniques have been employed in model development for compaction parameters. Two statistical tools in the literature are simple linear regression (SLR) and multi-linear regression (MLR). The difference between the two lies in the number of independent variables. While SLR takes only one independent variable, MLR takes 2 or more dependent variables. Various authors have shown that using more than one independent variable is more beneficial in developing compaction prediction models [21], [29], [37], [41].
Soft computing techniques are generally built on artificial intelligence or machine learning. Many authors have made an effort to classify soft computing techniques, however, in each of the classifications, variants of fuzzy logic, evolutionary intelligence, swarm intelligence, artificial neural network, chaos theory, and non-linear process [42], [43] are found. Soft computing techniques are employed to seek solutions to complex problems through approximation of functions, partial truth, and uncertainty as well as learning and adaptation to new data [44], [45]. In many compaction prediction model developments, many soft computing techniques have been employed, however, the commonest among them are the ANN, SVR, MARS, SVM, LS-SVM, MEP, and GPR.

1.2 Robustness and Statistical Significance Evaluation Tools

Several statistical tools are used to evaluate the accuracy or significance of model equations developed with statistical or soft computing tools. The commonest among these tools is the coefficient of determination, otherwise known as (R^2). Other statistical accuracy tools employed are the coefficient of correlation (R), the root mean square error (RMSE), the mean absolute error (MAE), the mean squared error (MSE), and the standard error of estimate (SEE).
According to [20], R^2 measures the goodness of fit of a regression model with values between 0 and 1; 1 represents perfect fit while 0 represents no fit. The value of R ranges between 1 and -1 with 1 indicating strong positive correction and -1 indicating strong negative correlation. An R-value of 0 indicates no correlation. RMSE gives the average distance between actual and predicted values while MAE measures the average magnitude of errors with no significant input of directions. The values of RMSE and MAE range from 0 to ∞ with values closer to 0 indicating high accuracy [20]. The SEE performs a similar function by giving an idea of the variance between the output experimental and predicted data [23]. Although many of these techniques have some input on informing the viability of a regression model, [46] highlighted that R^2 is more informative than all because it generates a high positive score only if the model performs excellently in predicting the ground truth elements or positive low score or negative score if the model performs poorly. In lieu of this, the R^2 was used to explain the results obtained from the comparative validation.

2. Prediction Model Equations

This section presents a summary of some model equations found in the literature which can aid reference purposes. In the development of model equations, different researchers use different notations for index properties of soils (independent parameters), however, for the sake of uniformity, Table 1 outlines the dependent and independent parameters used in this study, and their notations. These notations may differ from the original works; however, they have the same meaning and application.

Table 1: Notations for dependent and independent parameters

Properties	Notations
Maximum dry unit weight	MDUW
Optimum moisture content	OMC
Gravel content	C_G
Sand content	C_S
Fines content	C_F
Clay content	C
Silt content	S
Liquid limit	LL
Plastic limit	PL
Shrinkage limit	SL
Plasticity index	PI
Compaction energy	E
Specific gravity	G_s
Soil type index	STI
Coefficient of uniformity	C_u
Coefficient of curvature	Cc
Percentage finer than 50%	D_{50}
Percentage finer than 10%	D_{10}
Final sediment volume	FSV
Swell percent	S_p
Unified Soil Classification System	USCS

Although efforts towards the development of models for the prediction of compaction parameters of soils have been on as early as 1962, the author would limit the models presented here to 2014 which would in effect capture the efforts towards model development in the last decade and to show the recent progress in that regard. In the reviewed literature the method used in the studies, the number of samples used, and the models that were developed if any were highlighted. Letters A – S were used to identify the different authors (18) whose works were reviewed: 2013 (2), 2014 (2), 2015 (2), 2017 (1), 2018 (1), 2019 (2), 2020 (2), 2021 (3), 2023 (3), 2024 (1). The purpose of introducing the letters (A-S) is for reference purposes. The reviews are outlined alphabetically below.

[A.] Dokovic et al. [41] developed models to estimate soil compaction parameters based on Atterberg limits using the MLR method. Seventy-two (72) samples of fine-grained soils that were obtained from the core of earth dams: Rovni (CH – inorganic clays of high plasticity), Selova (CI-CL – inorganic clays of intermediate-low plasticity), Prvonek (CI-CH), and Barje (CL) in Serbia were used in the study. The regression models were developed using the liquid limit (LL), plastic limit (PL), plasticity index (PI), and LL-PL combination and it was shown the LL-PL models were more viable. The model equations obtained are outlined below:

$$MDUW = 2.14 - 0.007LL - 0.005PL \quad [R^2 = 0.73] \quad (1)$$
$$OMC = 4.18 + 0.16LL + 0.323PL \quad [R^2 = 0.73] \quad (2)$$

The models are basically for fine-grained soils.

[B.] Mujtaba et al. [23] employed MLR to model the relationship between MDUW/OMC and index parameters for 110 samples of sandy soils that belong to groups: SM (silty sands), SP-SM (poorly graded – silty sands interface), SP (poorly graded sands), SW-SM (well-graded – silty sands interface), and SW (well-graded sands). The initial independent parameters used were compaction energy (E), coefficient of uniformity (C_u), coefficient of curvature (C_c), percent finer than 50% (D_{50}), percent finer than 10% (D_{10}), gravel content (C_G), and fines content (C_F). Initial stepwise regression analysis showed that the parameters C_u and E have a more significant effect on MDUW/OMC and hence were used to develop the models outlined below:

$$MDUW = 4.49\ LogC_u + 1.51\ LogE + 10.2 \quad [R = 0.9;\ SEE = 0.51] \quad (3)$$
$$OMC = 10^{(1.67 - 0.193\ LogCu - 0.153\ LogE)} \quad [R = 0.84;\ SEE = 0.042] \quad (4)$$

The authors recommended the limitation of the models to medium-fine sands with gravel content not exceeding 5% and a maximum of 45% of non-plastic to low fines.

[C.] Oren [47] developed models to predict compaction parameters from sediment volume tests using the SLR technique. The study was conducted using sediment volumes of 9 soil samples. The developed models are:

$$MDUW = 19.7 - 2.52FSV \quad [R^2 = 0.88] \quad (5)$$
$$OMC = 5.42 + 9.63FSV \quad [R^2 = 0.87] \quad (6)$$

Where:
FSV = final sediment volume

[D.] Xia [48] applied numerical analysis to simulate soil compaction to predict soil compaction parameters using the Cap model. Though no models were created, he argued that finite element numerical analysis is more suited to capture the dynamic nature of the compaction process which analytical models do not capture properly.

[E.] Farooq et al. [49] developed empirical relationships for estimating MDUW/OMC with fine-grained soils collected from the Punjab province of Pakistan. Sixty-eight (68) soil samples were collected, tested and used in developing the models. The USCS classes of the samples are ML (17), CL-ML (10), CL (36), and CH (5). Stepwise regression analyses were used in developing the models outlined below:

$$MDUW = -0.055LL + 0.014PI + 2.21\ logE + 12.84 \quad [R = 0.89,\ SEE = 0.29] \quad (7)$$
$$OMC = 0.133LL + 0.02PI - 5.99\ logE + 28.60 \quad [R = 0.88,\ SEE = 0.86] \quad (8)$$

[F.] Khuntia et al. [30] consolidated the work of [23] through a comparative analysis using ANN, LS-SVM, and MARS. One hundred and ten (110) soil samples comprising largely of sands (fines content < 50%) were used in the study. Model empirical equations were developed based on the MARS system using C_S, C_F, C_u, and E as input or independent parameters. Optimization in MARS follows two unique processes: the forward selection procedure and the backward pruning procedure. The first procedure iterates with a constant term in the model to determine the best pairs of truncated spline functions (BFs) to improve the global model. BFs are usually complex and overfitted models with poor prediction capabilities. In the second procedure, the lack of fit criterion, such as the generalized cross-validation (GCV) criterion, was used to sort and delete BFs with the lowest contribution to generate the best model. BF values are defined in Table 2.

Table 2: Basis function of BFs for MARS model [24]

Basis functions	**MDUW**	**OMC**
BF_1	Max (0, C_u – 5.33)	Max (0, E - 592)
BF_2	Max (0, 5.33 - C_u)	Max (0, C_u – 3.29) x max (0, C_F – 2)
BF_3	Max (0, 2700 - E)	Max (0, C_u – 3.29) x max (0, 2 – C_F)
BF_4	Max (0, C_F - 5)	BF_2 x max (0, C_u – 3.53)
BF_5	Max (0, 5 - C_F)	BF_2 x max (0, 3.53 - C_u)
BF_6	BF_5 x max (0, C_u – 2.93)	Max (0, C_u – 3.29) x max (0, C_u – 3.64)
BF_7	BF_5 x max (0, 2.93 - C_u)	Max (0, C_u – 3.29) x max (0, 3.64 – C_u)
BF_8	BF_1 x max (0, C_F - 38)	Max (0, 3.29 - C_u) x max (0, C_F – 2)
BF_9	BF_5 x max (0, C_S – 94)	Max (0, 3.29 - C_u) x max (0, 2 - C_F)
BF_{10}	BF_5 x max (0, 94 – C_S)	Max (0, 3.29 - C_u) x max (0, C_S – 95)
BF_{11}	BF_9 x max (0, 2.23 - C_u)	Max (0, 3.29 - C_u) x max (0, 95 – C_S)
BF_{12}	BF_4 x max (0, C_u – 4.17)	
BF_{13}	BF_4 x max (0, 4.17 - C_u)	

The following models were developed by employing the BF values:

$$MDUW = 18.2 + 0.438BF_1 - 0.377BF_2 - 0.000473BF_3 + 0.0327BF_4 + 0.19BF_5 - 0.0857BF_6 - 0.766BF_7 + 0.0411BF_8 + 0.0461BF_9 + 0.603BF_{10} + 0.173BF_{11} - 0.00906BF_{12} - 0.0478BF_{13} \quad [R = 0.94] \quad (9)$$

$$OMC = 13.6 - 0.00131BF_1 - 0.0203BF_2 - 0.516BF_3 + 0.00271BF_4 + 21.7BF_5 - 0.0304BF_6 - 41.2BF_7 + 0.743BF_8 - 1.37BF_9 + 0.592BF_{10} - 0.858BF_{11} \quad [R = 0.9] \quad 10)$$

Higher R^2 and R values were reported for this model when compared to the MLR, ANN, and LS-SVM models. Good prediction ranges of ±4 for MDUW and ±13% for OMC were also observed.

[G.] Saikia et al. [2] developed empirical models for predicting MDUW/OMC using MLR. The tests were conducted specifically for fine-grained soils using 40 samples collected from different parts of Assam (India). The developed models are shown below:

$$MDUW = 21.07 - 0.119LL - 0.02PL \quad [R^2 = 0.90] \quad (11)$$

$$OMC = 0.35LL + 0.163PL + 6.26 \quad [R^2 = 0.86] \quad (12)$$

They reported limitations in the use of the models only for clayey soils of Assam (India).

[H.] Omar et al. [5] made a comparative investigation of the three most common methods recently employed in the process of model development: MLR, ANN, and SVR owing to the conflict on the different methods of developing prediction models bordering on their reliability. The study showed that the same strategy is not suitable for both MDUW and OMC. While SVR was adjudged the best for predicting MDUW with R^2 of 0.84, MSE of 0.36, and accuracy of 73%, ANN of one-hidden layer appeared to be more suitable for predicting OMC with R^2 of 0.85, MSE of 0.45, and accuracy of 74.2%. MLR yielded the least significant results for R^2, and MSE, as well as percentage accuracy for both MDUW/OMC.

[I.] Hasnat et al. [35] developed models for the prediction of MDUW/OMC for some 40 soil samples using SVR and was able to show that correlation using both LL and PL in the equation was more significant. It was also shown that LL correlated more strongly with MDUW than OMC which agrees with the work of [2]. The developed models are outlined below:

$$MDUW = 21.07 - 0.119LL - 0.02PL \quad [R^2 = 0.90] \quad (13)$$

$$OMC = 0.34LL + 0.17PL + 6.3 \quad [R^2 = 0.86] \quad (14)$$

[J.] Karimpour-Fard et al. [6] employed 728 soil samples from compaction tests coming from both fine-grained and coarse-grained soils to predict the MDUW/OMC of soils using ANN and MLR methods. The ANN is known to powerfully correlate independent variables with dependent variables but does not produce any equation for rule-of-thumb applications, hence, MLR was also employed to develop empirical relationships. The ratio of fine-grained soils to coarse-grained soils was 3:2. The independent variables used in the study were C_G, C_S, C_F, LL, PL, PI, G_s, E, and STI. STI (soil type index) was a designation of different soil classes by a numeral scale based on the Unified Soil Classification System (USCS); for instance, well-graded gravel or GW as 1 (see Table 3).

Table 3: Soil type index based on USCS

USCS class	**STI**
Well-graded gravel (GW)	1
Poorly-graded gravel (GP)	2
Well-graded sand (SW)	3
Poorly-graded sand (SP)	4
Silty gravel (GM)	5
Silty sand (SM)	6
Clayey gravel (GC)	7
Clayey sand (SC)	8
Silt (ML)	9
Lean clay (CL)	10
Silt of high plasticity (MH)	11
Fat clay (CH)	12
For boundary conditions, average values of STI is used e.g. CL-ML	$\frac{9+10}{2}$ = 9.5

The essence of the STI was to take into account the variability of the soil type used in the study in the model. This development appears to address the concern of [21] on the effect of soil type on the accuracy and efficiency of developed models, however, using simple numeral values for the scales may not represent the most efficient method to do this. From the model development, they observed that C_F and G_s have the most significant effect on MDUW/OMC while C_G has the least effect. They also observed that using a product of two or more individual parameters instead of single parameters in modelling could generate more efficient models. Though the efficiency of the ANN models was slightly higher than that of MLR at 92% against 89% for MDUW, both compared favourably and the developed models below were recommended by the authors to be used for preliminary investigation.

$$\text{MDUW} = 18.2 - 0.3\text{STI} - 0.1\text{PL} + 1.1\ln\text{E} + 0.4\text{LnPI} + 3.3\left[\text{STI} \times \frac{G_s}{LL}\right] - 0.2\left(\frac{CF}{G_s^2}\right) \quad [\text{RMSE} = 3.86] \tag{15}$$

$$\text{OMC} = 797.9 - 829.1\left(1 - \frac{1}{MDUW}\right) + \left(5 \times 10^{-5}\,(\text{CF. PL})\right) \quad [\text{RMSE} = 11.24] \tag{16}$$

[K.] Chindaprasirt et al. [26] developed empirical relationships for predicting MDUW/OMC of fine-grained soils using the plastic limit water content (w_{PL}) due to the linear correlation between OMC and PL that passed through the origin for all the 7 tested samples collected from the Khon Kaen province, located in Northeast of Thailand. Additional 49 fine-grained soil samples collected from different literature were used in developing the models outlined below:

$MDUW = (1.88 – 0.15 \ln E)\ MDUW_{PL}$ [$R^2 = 0.795$] (17)
$OMC = (0.63 + 0.06 \ln E)\ w_{PL}$ [$R^2 = 0.934$] (18)
Where:
$MDUW_{PL} = \frac{G_S \cdot \gamma_W}{1+ G_S \cdot w_{PL}}$
w_{PL} = plastic limit water content

[L.] Wang and Yin [3] employed MEP to develop models for estimating compaction parameters. In the study, they used 226 soil test results obtained from varying literature covering a wide range of soils [lean clay (CL), silty clay (CL-ML), fat clay (CH), elastic silt (MH), silt (ML), clayey sand (SC), poorly graded sand with clay (SP-SC), well-graded sand with clay (SW-SC), silty sand (SM), clayey gravel (GC), poorly graded gravel with clay (GP-GC), well-graded gravel with clay (GW-GC) and silty gravel (GM)]. This data was integrated into a form of genetic algorithm based on the principles of genetic and natural selection to develop models. The independent soil parameters used to build the models were C_G, C_S, C_F, LL, PL, E, OMC, and MDUW. Through the study, they were able to develop the following models:

$MDUW = \left(\frac{2}{B_3}\right)^{B_3} + \frac{2B_3^2 B_4}{C_F} - \frac{C_F}{B_6} + \frac{4C_S B_3}{C_F\,(C_F+2LL-2E)}$ [MAE = 0.05; RMSE = 0.069; $R^2 = 0.872$] (19)
Where:
$B_1 = B_4 + E + \frac{C_S B_3}{C_F}$
$B_2 = \frac{B_4+E}{B_3} + LL – E$
$B_3 = Log\ (PL)$
$B_4 = Log\ (C_F)$
$B_5 = \frac{C_F}{B_3} - \frac{2C_S B_3^2 B_4}{C_F} - \frac{B_1}{B_2} - PL – E – B_4$
$B_6 = B_3 \left(\frac{B_1}{B_2} - B_2 - B_5\right)$

$OMC = \frac{32}{PL} + \frac{A_1+ A_2}{A_3} + \frac{2A_1}{PL+E}$ [MAE = 1.206; RMSE = 1.574; $R^2 = 0.916$] (20)
Where:
$A_1 = \frac{PL^2}{1+ C_S} - \frac{2A_3}{LL} + C_S – C_F + 6$
$A_2 = \frac{PL+E}{64} + C_F + PL + 9$
$A_3 = (5 + C_F)\ (5 + PL)$
Through the monotonicity analysis, the authors showed that the trend corresponds with the monotonicity of the actual database in laboratory measurements. The authors argued that the models have wide application over a range of soils covering both fine-grained and coarse-grained soils that were used in the development of the models. Although the authors did not report limitations with the application of the models, it is unclear the effect having more fine-grained soils in the database would have on the applicability of the models to coarse-grained soils.

[M.] Fondjo et al. [29] employed MLR to develop models for MDUW/OMC from some 15 fine-grained soil samples belonging to classes, CL and CH. Different correlations were done for MDUW/OMC and index properties: C_u, C_c, C_G, G_s, C_S, C_F, LL, PI. It was shown that LL, PI, C_F, G_s are strong predictors of MDUW/OMC achieving $R^2 > 0.9$, with LL, C_F, and G_s being the most significant input parameters based on a P-value of 0.00 to 0.001. It was unclear the basis for using C_S and C_G further in the models instead of PI which was shown to be a strong predictor of MDUW/OMC. The developed models are shown below:

$$MDUW = 12.246 - 0.101\,LL + 0.137\,C_F + 0.136\,C_S + 0.131\,C_G - 0.874\,G_s \quad [R^2 = 0.9714] \quad (21)$$
$$OMC = 14.6 - 0.052\,LL + 0.622\,C_F + 0.274\,C_S + 0.056\,C_G - 13.675\,G_s \quad [R^2 = 0.9919] \quad (22)$$

[N.] Jalal et al. [37] applied two genetic programming-based algorithms: GEP and MEP to predict the compaction parameters of expansive soils. A total of 195 soil samples were employed in building the model and the input parameters were C_F, PL, PI, and G_s for MDUW and the addition of MDUW and swell percent (S_p) to the other parameters for OMC. More than 84% of the datasets have PI > 20% showing the presence of low to very high expansive soils in the existing database. Though the models developed were quite complex, they generally showed better accuracy and performance than models developed with MLR tools with GEP slightly performing better than MEP, however, the authors still reported calibration limitations on the use of the models.

[O.] Nwaiwu and Mezie [4] developed empirical models for the prediction of MDUW/OMC of coarse-grained lateritic soils using MLR. The soils were largely lateritic soils drawn from different parts of Anambra state, Nigeria. All the soils contain significant fines content (19.79 – 42.43%), sand content (57.57 – 80.21%), and insignificant gravel content of less than 1%. The models developed as outlined below were adjudged to be robust but were limited to soils having fines content – sand content ratio of 0.246 – 0.737 and fines content less than 50%. The first model applies to any compactive energy once the fines content – sand content ratio is known while the second model is more transferable from a known compactive energy to an unknown one:

Models 1:

$$MDUW = \left[1.73\left(\frac{C_F}{C_S}\right) + 1.6\right]\log E + 15.83 - 8.58\left(\frac{C_F}{C_S}\right) \quad [RMSE \leq 0.213\ kN/m^3] \quad (23)$$
$$OMC = \left[3.07\left(\frac{C_F}{C_S}\right) - 5.26\right]\log E + 23.59 - 0.39\left(\frac{C_F}{C_S}\right) \quad [RMSE \leq 0.424\%] \quad (24)$$

Models 2:

$$MDUW_{unknown} = MDUW_{known} + \left[1.73\left(\frac{C_F}{C_S}\right) + 1.6\right]\log\left(\frac{E_{unknown}}{E_{known}}\right) [RMSE \leq 0.371\ kN/m^3] \quad (25)$$
$$OMC_{unknown} = OMC_{known} + \left[3.07\left(\frac{C_F}{C_S}\right) - 5.26\right]\log\left(\frac{E_{unknown}}{E_{known}}\right) [RMSE \leq 0.206\%] \quad (26)$$

[P.] Bardhan and Asteris [38] employed hybrid ANN paradigms built with nature-inspired meta-heuristics to model soil compaction parameters. The hybrid model comprises ANN and grey wolf optimizer (GWO). They employed 126 soil samples belonging to both fine-grained (CL, CH, MH, MI, ML) and coarse-grained (GM, GC, SC) soils collected from different parts of India. The ANN-GWO was found to produce the highest level of accuracy with [R^2 = 0.8103; RMSE = 0.0800]; [R^2 = 0.7273; RMSE = 0.0986]; and [R^2 = 0.9491; RMSE = 0.8415], in the training, testing, and experimental validation phase of OMC prediction. In addition, they produced accuracy with [R^2 = 0.8161; RMSE = 0.0805]; [R^2 = 0.7147; RMSE = 0.1017] and [R^2 = 0.9751; RMSE = 0.1977], in the training, testing, and experimental validation phase for MDUW prediction. The primary limitation of the suggested ANN-GWO is the particle position restriction brought about by the predefined search space of the GWO parameters. As there are no thumb rules, it took several runs—another time-consuming task—to find the optimal suitable searching space for the estimation of final outputs. Moreover, only four soil types from a total of 20 datasets were taken into account for experimental validation, and modeling did not take into account the impact of compaction energy or the parental significance of soils. No particular models were developed in the process.

[Q.] Khatti and Grover [32] performed a comparative study of deep learning and standalone models for the prediction of compaction parameters of fine-grained soils. The methods employed include SVM, GPR, LSBoostRF, ANN, and LSTM. The standalone models are the SVM, GPR, and LSBoostRF while the deep learning models are the LSTM and ANN. A total of 243 soil samples were used for the analysis while C_F, C_S, G_s, LL, and PI were used as input parameters. The soils belong to USCS classes: CH, CL, CI, MH, ML, MI, OL, OI, and OH soils. The sensitivity analysis illustrates that fine content (C_F), specific gravity (G_s), and liquid limit (LL) highly influence the prediction of compaction parameters. Among all the employed methods, the LSTM model stands out for achieving the best performance when compared to other models, being potent for computing the desired compaction parameters of the soil, being easy to implement, achieving better accuracy, performance, and less over-fitting ratio, with the small dataset having moderate multi-collinearity. The research was adjudged the first to study the impact of multi-collinearity on predicting compaction parameters of fine-grained soil. Hence, it recommends that LSTM deep learning tools should be employed in subsequent studies to develop compaction prediction models.

[R.] Soltani et al. [50] promulgated reservations over all the developed models irrespective of the tool or technique used in the development of the model because of their limitations to calibration information. They studied how to convert optimum compaction properties of fine-grained soils between rational energy levels (energy conversion (EC) type models) which they presented as a method that would not be influenced by calibration limitations. However, they reported the possibility of overestimation of MDUW/OMC above the zero-air-voids (ZAV) line, especially outside the calibration zone which is impossible in a normal compaction process. Very large soil databases of 242 samples were used for the study. The developed models are as follows:

$$MDUW_R = MDUW_{SP}\left(\frac{E_R}{E_{SP}}\right)^{0.068} \quad [R^2 = 0.970] \tag{27}$$

$$OMC_R = OMC_{SP}\left(\frac{E_R}{E_{SP}}\right)^{-0.178} \quad [R^2 = 0.975] \tag{28}$$

E_R = arbitrary compaction energy
E_{SP} = compaction energy at standard Proctor
To overcome the possibility of obtaining the degree of saturation at optimum compaction exceeding 100%, the authors proposed a new model for the estimation of MDUW which incorporated the specific gravity of the soils, thus:

$$MDUW_R = \frac{G_S\gamma_W}{1+\left[G_S\left(\frac{\gamma_W}{MDD_{SP}}\right)\right]\left(\frac{E_R}{E_{SP}}\right)^{-0.178}} \quad [R^2 = 0.976] \tag{29}$$

This method appears great but may be more suitable when working between different compaction energies and using the same class of soil.

[S.] Polo-Mendoza et al. [27] employed both mathematical models, MLR and computational model, DNN to predict the MDUW/OMC of 90 coarse-grained soil samples. The MLR model achieved R^2 of 0.735 and 0.726 for MDUW and OMC respectively while the DNN achieved R^2 of 0.9804 and 0.9999 for MDUW and OMC showing that computational models are more precise than mathematical models in making predictions. The developed models are as follows:

$$MDUW = 24.1 - 0.7D_{60} - 6.1A + 0.2B + 2.2C + 1.2D \tag{30}$$

$$\begin{bmatrix} MLR: R^2 = 0.7354; MSE = 0.41619; MSLE = 0.00100 \\ DNN: R^2 = 0.9804; MSE = 0.03326; MSLE = 0.00008 \end{bmatrix}$$

$$OMC = 34.3 - 58.2D_{10} + 38.9A - 3.9C + 3.8D \tag{31}$$

$$\begin{bmatrix} MLR: R^2 = 0.7258; MSE = 1.29188; MSLE = 0.00915 \\ DNN: R^2 = 0.9992; MSE = 0.00381; MSLE = 0.00003 \end{bmatrix}$$

Where:
A = D_{10}^2
B = D_{50}*ln (D_{30} + C_u + C_c + A)
C = $\log_{10}(D_{50} + D_{60})$
D = ln(A)

It has been shown in this section that soft computing techniques are more viable based on values of coefficient of determination (R^2) to develop models for the prediction of compaction parameters of soils. It was also shown that Atterberg limits have a stronger correlation with MDUW/OMC for fine-grained soils while gradational parameters have a stronger correlation with MDUW/OMC of coarse-grained soils which agrees with [21]. In most of the developed models, there were still records of calibration limitations bordering on soil classes, location, model development tools, range of particle sizes, etc. Although some authors tried to evade these limitations by combining coarse-grained and fine-grained soils (eg. [3], [38]) or using soil type index (e.g. [6]), there was no scientific basis for their decisions. For reference purposes, it is necessary to outline the most significant parameters influencing model development for each class of soils (section 3.0). To advise on strategies for subsequent model development efforts, it is necessary to carry out a comparative model validation for the developed models (section 4).

3. Most Important Index Parameters for Model Development

Table 4 presents an attempt to distinguish the index properties of the soils used in the developed models based on the most significant and the least significant ones. The classification was based on the report by the authors of the works that were reviewed. Recall that before developing prediction models, it is usually an important step to investigate the soil index properties that correlated most strongly with MDUW/OMC, hence, the data presented in Table 4 can serve for reference purposes and can be investigated further.

Table 4: Classification of important index properties in developed models

Authors	Classification	Model development tool	Dependent variables tested	Most important	Least important
A	Fine-grained	MLR	LL, PL, PI	LL-PL (OMC), LL-PI (MDUW)	PL (MDUW) PI (OMC)
B	Coarse-grained	MLR	E, C_u, C_c, D_{50}, D_{10}, C_G and C_F	C_u, E	No comment
E	Fine-grained	MLR	LL, PI, C_F, G_s, E	LL, PI, E	G_s, C_F
F	Coarse-grained	ANN, LS-SVM, MARS	C_S, C_F, C_u, E	C_S and C_u	E and C_F
G	Fine-grained	MLR	LL, PL	LL	PL
H	Fine-grained	MLR, ANN, SVR	C_S, C_F, C, S, LL, PL, PI, SL, G_s	LL and PL (for MDUW) and LL, PL and C_F (for OMC)	C
I	Fine-grained	SVR	LL, PL, PI	LL, PL	PI
J	Fine-grained	ANN, MLR	C_G, C_S, C_F, LL, PL, PI, G_s, E, STI	C_F, G_s	C_G

L	Fine-grained and coarse-grained	MEP	C_G, C_S, C_F, LL, PL, E,	PL and C_F	LL
M	Fine-grained	Numerical analysis	C_u, C_c, C_G, G_s, C_S, C_F, LL, PI	LL, PI, C_F, G_s	C_u, C_c, C_G, C_S
N	Fine-grained	GEP, MEP	C_F, PL, PI and G_s for MDUW	C_F, PL, PI (GEP) G_s (MEP)	G_s (GEP) C_F, PL, PI (MEP)
	Fine-grained	GEP, MEP	C_F, PL, PI, S_p, MDUW and G_s for OMC	PL, PI, MDUW (GEP) PI, MDD and G_s (MEP)	S_p and G_s(GEP) PL and C_F (MEP)
O	Coarse-grained	MLR	$\frac{C_F}{C_S}$, log E	No comment	No comment
P	Fine-grained	ANN-GWO	C_F, C_S, C_G, LL, PL, G_s	No comment	No comment
Q	Fine-grained	SVM, GPR, LSBoostRF, ANN, LSTM	C_F, C_S, G_s, LL, PI	C_F, G_s, LL	C_S, PI
S	Coarse-grained	MLR, DNN	D_{10}, D_{30}, D_{50}, D_{60}, C_u, C_c	D_{10}, D_{30}, C_u, C_c	D_{50}, D_{60}

The data shown in Table 4 equally shows that the method of developing prediction models, the amount of input data, the input index properties, and the percentage of fines and sands in soils generally have some level of interaction and influence on the outcome of the model development.

4. Comparative Validation of Developed Models

Wang and Yin [3] and [6] opined that using a larger number of databases and employing more sophisticated soft computing techniques tends to improve the accuracy and efficiency of the prediction models. Although this principle aligns with machine learning techniques, there is yet no comparative validation of such for the physical properties of soils. Even though their developed models showed good prospects in prediction, there is no comparative evidence to show that using a smaller number of databases produces less efficient models. For the study by [6], the usual claim that MLR is not efficient enough was partly invalidated by the fact that the difference in accuracy between prediction by ANN and MLR is quite small (≈ 3%).

It appears that incorporating more independent variables improves the validity of developed models, however, it is expedient that the dependent variables (MDUW and OMC) should have a strong correlation with the independent variables This was evident in the model developed by [2] where the use of two dependent variables of LL and PL yielded RMSE of 2.1% and 7% for MDUW and OMC against 7.4 – 7.5% and 17.5 – 28.2% in the predicted versus measured results, respectively, of authors that used only one independent variable in their models. To critically investigate these scientific opinions, it is necessary to conduct comparative validation for fine-grained and coarse-grained soils using the different models available in the literature.

4.1 Fine-grained Soils

Table 5 shows the validation of the various prediction models of fine-grained soils with neutral soil data obtained from *M*. The data in *M* was chosen because it contains all the essential soil properties applicable to the different models used in the validation. The first two columns show the measured values for MDUW and OMC, respectively, while the other columns show the predicted values for the different model equations. Figures 1 and 2 show the predicted versus measured values for MDUW and OMC for each of the models. Figures 3 and 4 show the degree of the departure of prediction outcomes of the models from the measured values. From Figures 1 and 2, high correlation values were obtained for *A*, *E*, *G*, *J*, *M*, and *R*, showing the importance of using only index properties that correlate with the compaction parameters in the development of the models. In the validation of developed models, it is expected that the predicted values should not only have high correlation coefficients but should be as close as possible to the measured values. These were investigated further in Figures 3 and 4 and it was shown that only *A*, *E*, *K*, *M*, and *R* come closest to the measured values. Checking what may have affected these trends; there was no evidence of the effect of the methods used in developing the models such as MLR, ANN, etc. Models *A* and *E* were developed with MLR but showed good prediction outcomes. Interpretation of the comparative validation outcome has shown that while the use of sophisticated tools may be recommended, subsequent models must consider the most important parameters affecting the soil properties as outlined in Table 4. The compaction energy should be employed in the models and the models should be developed from a similar class of soils or boundary class based on the USCS chart irrespective of where the soils exist. It is not necessary to combine both fine-grained and coarse-grained soils unless there is a more effective way to capture the soil type index in the model, probably by involving the soil sensitivity factor.

Table 5: Validation of models of fine-grained soils

Measured MDUW	Measured OMC	A		E		G		J		K		L		M		R	
		MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC
18.5	18.2	16.76	17.76	16.71	19.11	14.89	26.28	22.03	6.49	16.07	18.07	17.97	17.17	18.47	18.29	18.52	18.14
18.21	19.01	16.57	18.37	16.61	19.43	14.62	27.17	21.62	7.20	15.97	18.87	17.78	16.38	18.26	19.46	18.23	18.95
17.99	20.38	16.25	19.39	16.44	19.97	14.15	28.67	21.32	7.748	15.71	20.23	17.46	15.36	17.98	21.02	18.01	20.31
17.58	20.07	15.86	20.77	16.24	20.57	13.61	30.52	20.92	8.50	15.06	22.46	17.09	14.01	17.45	21.06	17.60	20.00
17.16	22.61	15.66	21.42	16.14	20.89	13.32	31.46	20.79	8.75	14.99	23.32	16.97	13.53	17.29	21.77	17.18	22.53
16.95	23	15.37	22.35	15.98	21.38	12.89	32.83	20.56	9.19	14.79	24.59	16.64	12.86	16.99	23.32	16.97	22.92
16.85	24.03	15.57	21.27	16.06	21.32	13.05	32.06	20.88	8.58	15.61	21.58	17.03	14.97	16.95	23.56	16.87	23.95
16.71	24.58	15.36	21.92	15.95	21.67	12.75	33.04	20.74	8.86	15.50	22.41	16.84	14.46	16.77	24.61	16.73	24.49
16.45	26.05	14.98	23.11	15.74	22.32	12.19	34.83	20.42	9.47	15.18	23.91	16.52	13.89	16.35	26.06	16.47	25.96
16.29	26.14	15.26	21.62	15.85	22.22	12.41	33.76	20.83	8.68	16.09	19.84	17.16	16.75	16.38	25.64	16.31	26.05
16.05	26.52	14.86	22.84	15.62	22.94	11.81	35.68	20.58	9.17	15.89	21.22	16.83	16.16	15.82	26.54	16.07	26.43
15.65	27.75	14.48	24.00	15.42	23.62	11.23	37.52	20.36	9.610	15.70	22.55	16.53	15.67	15.48	27.83	15.67	27.66
19.6	17.23	17.19	17.43	17.02	17.73	15.85	23.79	21.87	6.77	15.00	21.34	17.53	13.14	19.56	20.84	19.63	17.17
19.2	18.13	16.88	18.52	16.86	18.1	15.42	25.23	21.64	7.16	14.73	23.09	17.27	12.18	19.18	19.908	19.23	18.07
18.76	18.24	16.43	20.85	16.63	18.80	14.81	27.31	21.31	7.77	14.29	25.61	16.92	11.09	18.63	18.63	18.79	18.18
R^2		**0.963**	**0.851**	**0.980**	**0.952**	**0.986**	**0.945**	**0.820**	**0.808**	**0.168**	**0.174**	**0.476**	**0.127**	**0.992**	**0.922**	**1**	1

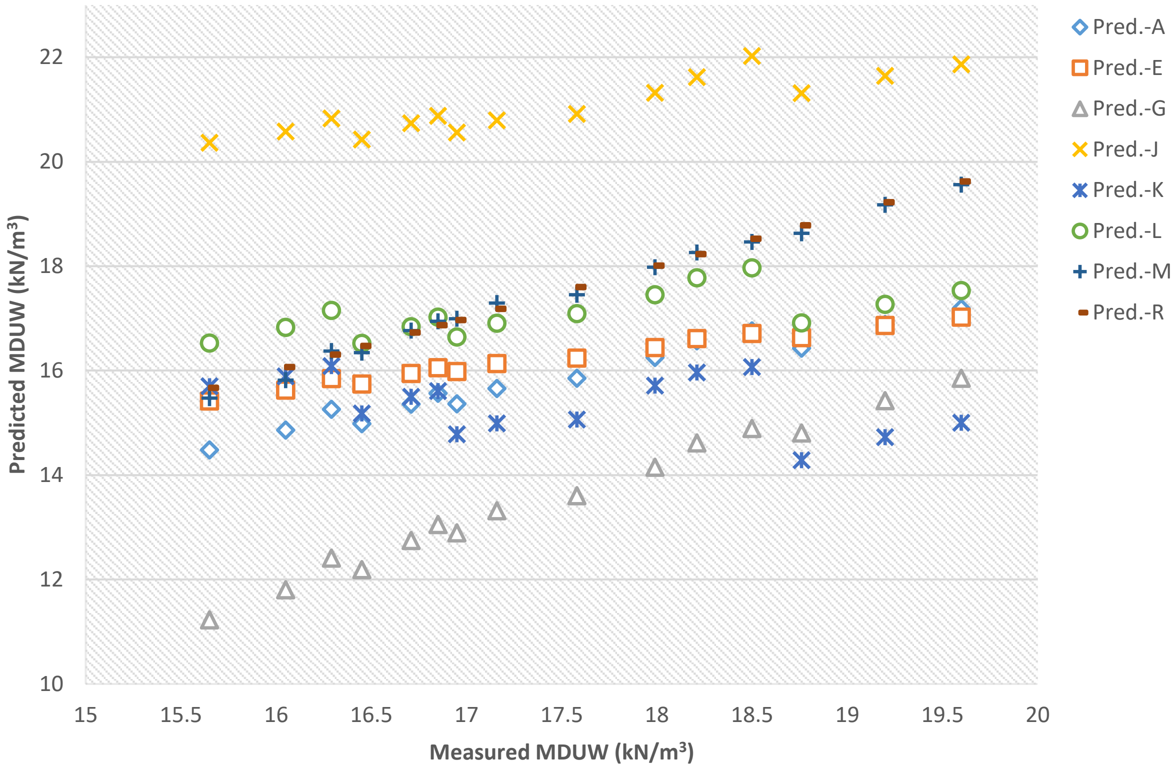

Figure 1: Measured MDUW versus predicted MDUW for all the samples

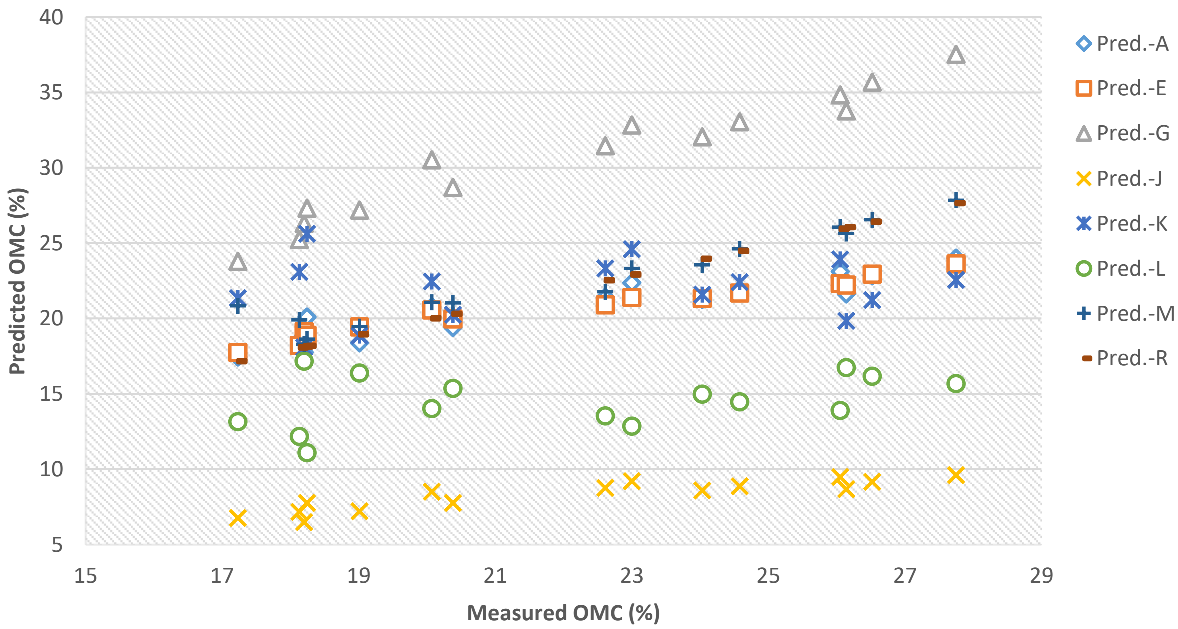

Figure 2: Measured OMC versus Predicted OMC for all samples

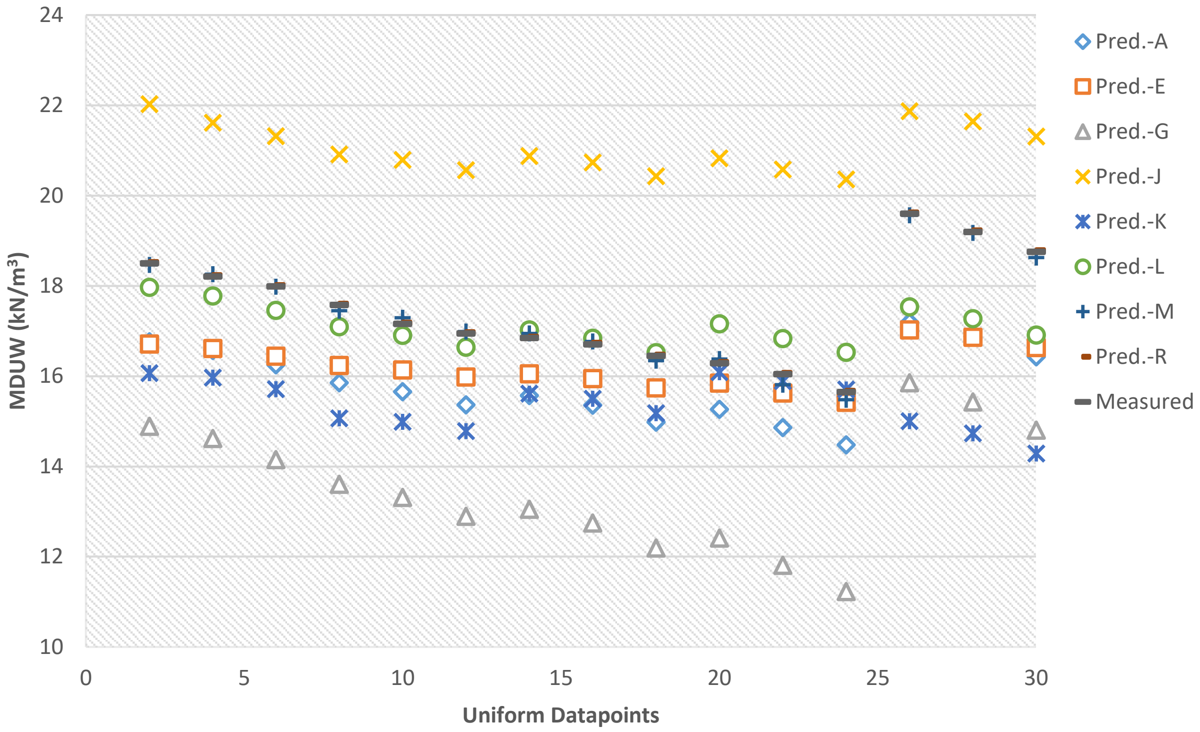

Figure 3: Departure of predicted MDUW from measured values

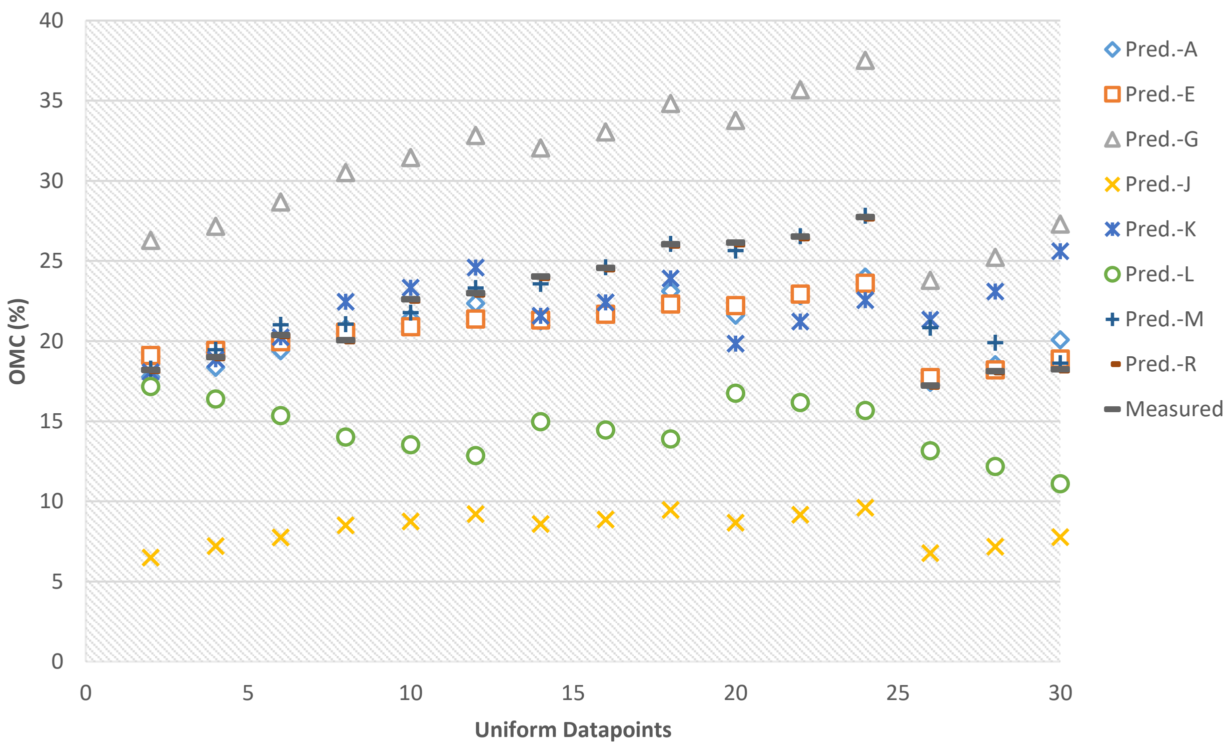

Figure 4: Departure of predicted OMC from measured values

4.2 Coarse-grained soils

Table 6 shows the validation of the various prediction models of coarse-grained soils with neutral soil data obtained from *F*. Similar to 4.1, the data was obtained from sources that contain all the required soil index properties. The first two columns show the measured values for MDUW and OMC, respectively, while the other columns show the predicted values for the different model equations. Figures 1 and 2 show the predicted versus measured values for MDUW and OMC for each of the models. Figures 3 and 4 show the degree of departure of prediction outcomes of the models from the measured values. Predictions *B* and *F* gave favourable outcomes because the data used in validating the models were drawn from their database. The other two models, O and S, didn't perform well, probably because of calibration limitations and the use of diverse soil classes in the prediction models. Similar to fine-grained soils, it is recommended that the most important index properties outlined in Table 4, in addition to E and unified soil class or boundary class should be employed in the development of subsequent models.

Table 6: Validation of models of coarse-grained soils

Measured MDUW	Measured OMC	B		F		O		S	
		MDUW	OMC	MDUW	OMC	MDUW	OMC	MDUW	OMC
17.2	14	16.944	13.632	16.982	13.308	17.955	13.002	14.326	8.864
19.1	12	18.985	11.139	18.818	11.793	17.868	13.189	12.517	3.332
15.4	16.5	15.107	16.350	15.361	13.377	19.444	9.8039	16.299	13.111
15.2	16.5	15.107	16.350	15.458	14.0788	19.489	9.707	16.2996	13.111
17	16	16.708	13.954	16.371	15.627	19.080	10.586	16.285	11.250
17.1	15	17.499	12.904	17.356	13.136	19.098	10.547	15.338	9.255
18.9	11	19.080	11.0344	19.074	11.869	18.261	12.346	14.335	5.859
18.9	11.5	18.946	11.182	18.852	11.795	18.197	12.483	13.781	4.972
19.3	12	19.136	10.973	19.149	11.906	18.305	12.250	13.678	4.597
18.9	12	18.320	11.896	18.732	11.430	16.965	15.129	13.469	6.016
15.4	18	15.151	16.279	14.947	16.351	19.660	9.339	16.532	13.187
15.4	18	14.982	16.555	15.612	15.578	19.577	9.519	16.449	13.341
18.8	12.5	18.472	11.719	18.464	12.682	19.489	9.707	18.149	10.481
15.6	17.5	15.956	15.033	15.874	13.868	19.489	9.707	16.643	12.451
19.5		18.732	11.421	19.574	11.360	16.546	16.032	12.924	4.703
R^2		0.941	0.8915	0.9723	0.8112	0.5853	0.6393	0.4608	0.8451

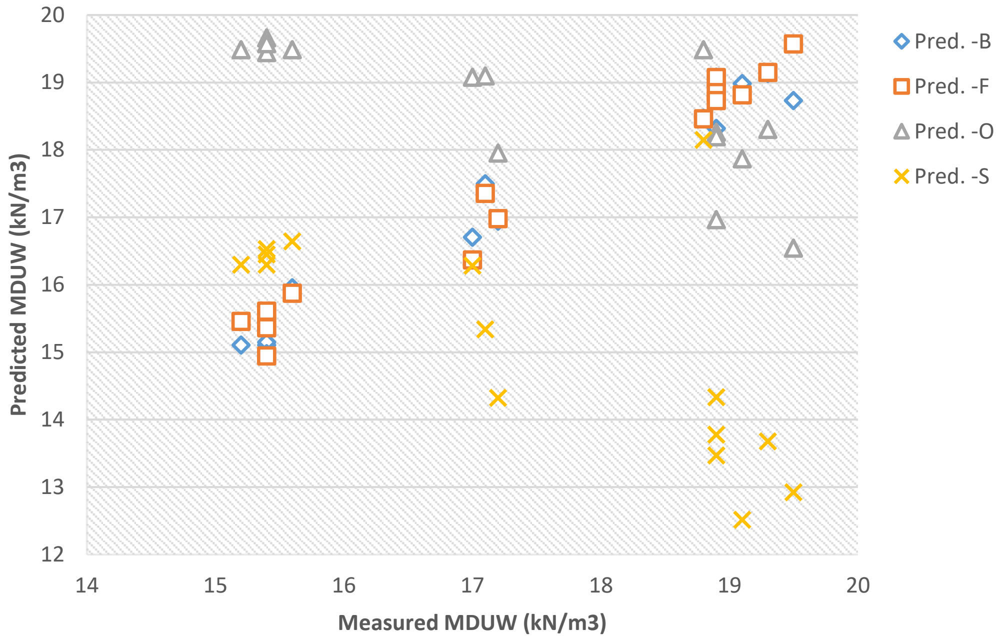

Figure 5: Predicted versus measured for MDUW of coarse-grained soils

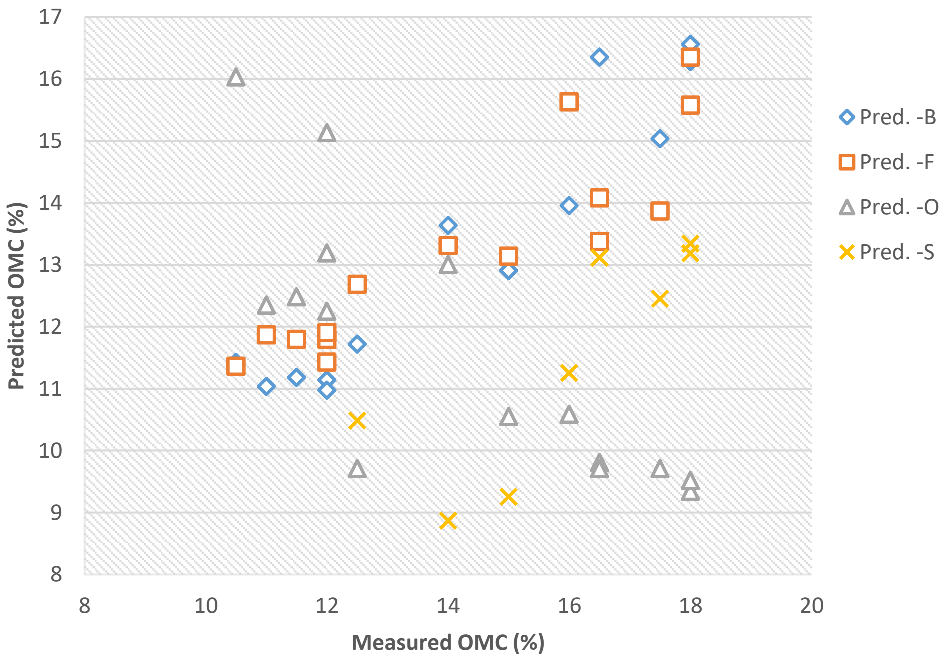

Figure 6: Predicted versus measured for OMC of coarse-grained soils

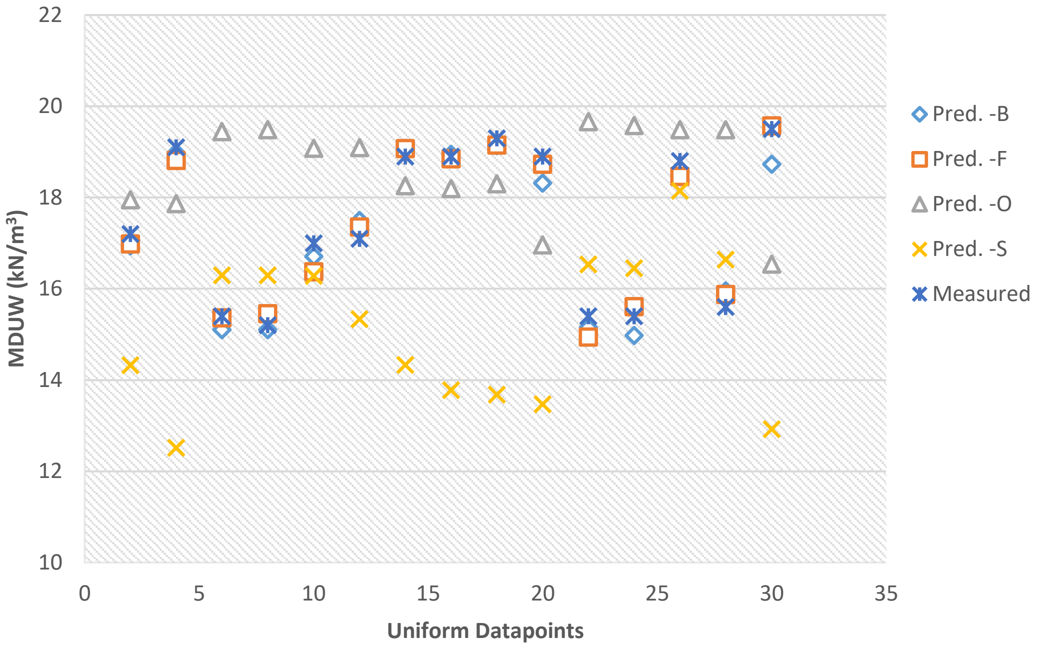

Figure 7: Departure of predicted MDUW from measured values

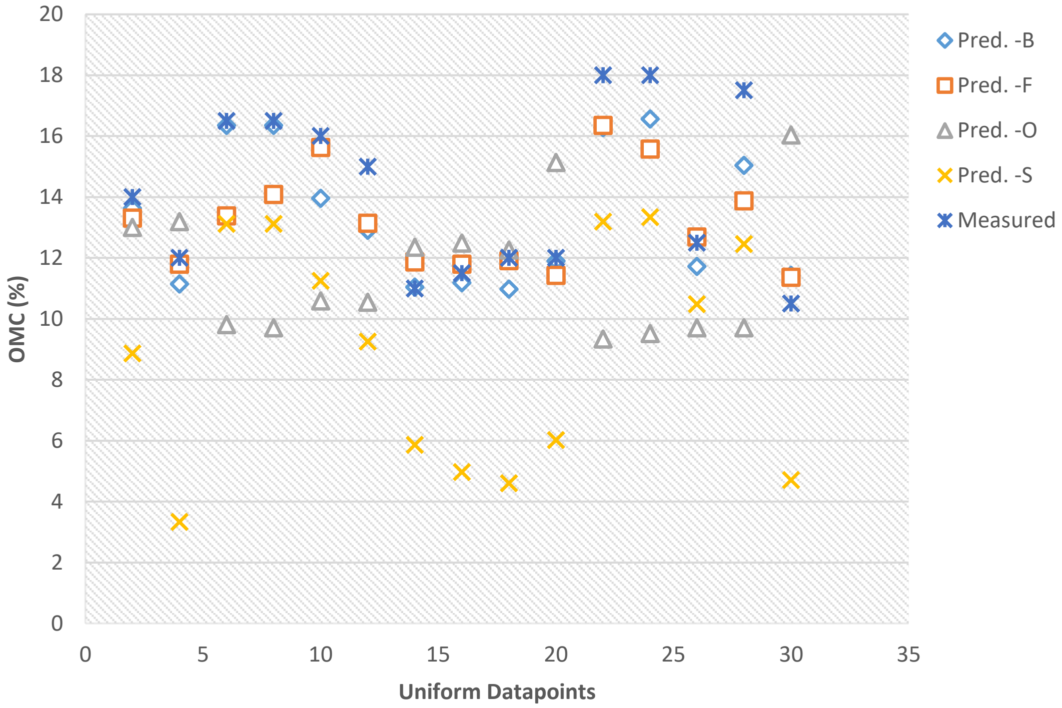

Figure 8: Departure of predicted OMC from measured values

4.3 Discussion

Developed models usually validate soil compaction parameters within the range of the soil parameters used in developing the models. In most of the developed models, after predicting the compaction parameters used in developing the models, the authors often also employ soil parameters that are not used in developing the models and often get good coefficient of determination in comparison to the measured results of the soils used in developing the models. In the instance where the model is applied to an entirely different soil and compared to the prediction outcome from another model, the biases would be more apparent as was shown in preceding sections, hence, gaps still exist on the most important parameters, tools and strategies to be employed in the development of more unified soil prediction models.

In the prediction of fine-grained models, significant coefficient of determination values was obtained for *A, E, K, M,* and *R* even though most of these models were developed with MLR. This may not invalidate the claim that tools such as GEP, ANN, and SVR are better tools, however, it may be necessary to examine what may have contributed to the strength of these models. First, the models identified the soil index properties that correlate well with MDUW and OMC and only used that in the model development; secondly, the models that used particular types and classes of soil seem to have a stronger correlation than those that combined different types and classes of soil. For instance, *J* developed its models with a very large soil database (728) and sophisticated ANN tool but did not achieve much significance probably due to the combination of fine-grained and coarse-grained soil in one model. Even though the soil type index was employed to capture this variability, numeral values do not seem to suffice for the soil type index. In Figures 3 and 4, it can be seen that *J* over-predicted MDUW and under-predicted OMC which seem not acceptable while the models developed around the soil types have better correlation and prediction outcomes closer to the measured values. Collected information has shown that the method of model development seems to not have much significant effect on the performance of the models, however, it seems to have more ability to integrate numerous and varying soil properties into a unified model.

In summary, the work outlined different techniques employed in the development of model prediction equations for soil compaction parameters, some developed model equations and their limitations, the most important index properties for fine-grained and coarse-grained soils, and finally evidence of the limitations of the model equations through comparative validation.

5. Conclusion and Recommendations

In this work, the comparative validation of 18 soil prediction models for MDUW and OMC was studied. This is considered necessary for reference purposes and to investigate how the different models perform comparatively through which to judge the performance of model development tools and strategies.

A brief literature review was done outlining previous efforts towards model developments, key techniques especially soft computing techniques employed in these efforts, model equations that were developed, and the most important/least important index properties. The review also showed that more model equations exist for fine-grained soils when compared to coarse-grained soils. The model equations were further employed in comparative validation.

In all the developed models seen, the USCS class is always present showing that it may have some impact in developing more reliable model equations. Since soil classification is the language of the geotechnical engineer, it is recommended that subsequent models should be developed drawn from a wide range of literature around the world for soils belonging to a particular USCS soil class for instance; model based only on soils belonging to CL soil class or its boundary, CI soil class or its boundary, ML soil class or its boundary, etc for fine-grained soils and then SC or its boundary, SM or its boundary, GM or its boundary for coarse-grained soils. Each soil class should possess similar contents of aggregate percentages, quantities, and similar behaviour irrespective of where available around the world. This can be drawn from a wide range of literature around the world for each class of soil and integrated into models. If the result comes out positive, it will be possible in this way to develop models that have wide application, thus, once the soil class is determined based on index properties tests, the suitable model should be employed to determine the compaction parameters of the soil across any part of the world.

Based on the outcome of this review, the author recommends that further studies should involve soils (as many as possible) belonging to a unified soil class irrespective of where the soil is present around the world in developing models instead of using variable types of soils from one region. This is in agreement with [21] who recommended proper categorization of fine-grained and coarse-grained soils for the improvement of the reliability of the models as well as developing models incorporating more extended ranges of soil index properties. In developing these models, sophisticated tools such as SVR, ANN, or numerical analysis should be employed to examine the correlation of MDUW/OMC and index properties while a minimum of 3 physical properties of soils in addition to the compaction energy should be used for model development. For fine-grained soils, fines content, liquid limit, plastic limit, and compaction energy should be included; for coarse-grained soils, fines content-sand content ratio, sand content, specific gravity, and compaction energy should be employed as these have been shown in most of the reviewed models to have a stronger correlation with MDUW/OMC. In situations where adequate data cannot be obtained for a particular soil class, then the soil type index should be employed for the soil classes available, however, it may be necessary to employ soil sensitivity to capture the soil type index instead of numeral values.
Further research may also be carried out to determine how to integrate the effect of weather conditions, room temperature, sample collection techniques, etc in developing models, thus, authors may need to also state these conditions when reporting their compaction data from different parts of the globe.

Data Availability: The data used in these studies were drawn from reviewed works whose authors have been properly cited.
Funding: This work received no external funding.
Conflict of Interests: There is no conflict of interest associated with the work.

List of Abbreviations (also see Table 1)
CBR - California bearing ratio
SLR - Simple Linear Regression
MLR - Multi-linear Regression
ANN - Artificial Neural Network
LS-SVM - Least Squares – Support Vector Machine
MARS - Multivariate Adaptive Regression Splines
SVR - Support Vector Regression
MEP - Multi-gene Expression Programming
GEP - Gene Expression Programming
GWO - Grey Wolf Optimizer
GPR - Gaussian Process Regression
SVM - Support Vector Machine
LSBoostRF - Least Squares Boost Randon Forest
LSTM - Long Short-term Memory
DNN - Deep Neural Network
GmDH - Group Method of Data Handling
NN - Neural Network
GA - Genetic Algorithm
PSO - Particle Swarm Optimization
RF - Random Forest
XGBoost - Extreme Gradient Boosting Tree

R^2 **-** Coefficient of Determination
R - coefficient of correlation
RMSE - root mean square error
MAE - mean absolute error
MSE - mean squared error
SEE - standard error of estimate
CH – inorganic clays of high plasticity or fat clay
CI – inorganic clays of intermediate plasticity
CL – inorganic clays of low plasticity or lean clay
SM - silty sands
SP - poorly graded sands interface
SW - well-graded sands interface
ML – low plasticity silt
SC – clayey sands
GC – clayey gravel
GM – silty gravel
GP – poorly graded gravel
GC – clayey gravel
MH – high plasticity silt or elastic silt
GCV - generalized cross-validation
w_{PL}– plastic limit water content
S_p – swell percent
E_R - arbitrary compaction energy
E_{SP} - compaction energy at standard Proctor
ZAV – zero air voids line

References

[1] E.O. Mezie; M.N. Ezema & P.T. Mmeka, "Review and Practical Processes on Rammed Earth Wall Construction," *Discovery*, vol. 58, no. 318, pp. 637–653, 2022.

[2] A. Saikia, D. Baruah, K. Das, H. J. Rabha, A. Dutta, and A. Saharia, "Predicting Compaction Characteristics of Fine-Grained Soils in Terms of Atterberg Limits," *Int. J. Geosynth. Gr. Eng.*, vol. 3, no. 2, p. 0, 2017, doi: 10.1007/s40891-017-0096-4.

[3] H. L. Wang and Z. Y. Yin, "High performance prediction of soil compaction parameters using multi expression programming," *Eng. Geol.*, vol. 276, no. May, p. 105758, 2020, doi: 10.1016/j.enggeo.2020.105758.

[4] C. M. O. Nwaiwu and E. O. Mezie, "Prediction of maximum dry unit weight and optimum moisture content for coarse-grained lateritic soils," *Soils and Rocks*, vol. 44, no. 1, pp. 1–10, 2021, doi: 10.28927/sr.2021.054120.

[5] M. Omar, A. Shanableh, O. Mughieda, M. Arab, W. Zeiada, and R. Al-Ruzouq, "Advanced mathematical models and their comparison to predict compaction properties of fine-grained soils from various physical properties," *Soils Found.*, vol. 58, no. 6, pp. 1383–1399, 2018, doi: 10.1016/j.sandf.2018.08.004.

[6] M. Karimpour-Fard, S. L. Machado, A. Falamaki, M. F. Carvalho, and P. Tizpa, "Prediction of Compaction Characteristics of Soils from Index Test's Results," *Iran. J. Sci. Technol. - Trans. Civ. Eng.*, vol. 43, no. s1, pp. 231–248, 2019, doi: 10.1007/s40996-018-0161-9.

[7] T. F. Kurnaz and Y. Kaya, "The performance comparison of the soft computing methods on the prediction of soil compaction parameters," *Arab. J. Geosci.*, vol. 13, no. 4, 2020, doi:

10.1007/s12517-020-5171-9.

[8] A. V. Hohn, R. F. Leme, F. C. da Silva Filho, T. E. Moura, and G. R. A. Llanque, "Empirical models to predict compaction parameters for soils in the state of Ceará, Northeastern Brazil," *Ing. e Investig.*, vol. 42, no. 1, pp. 1–11, 2022, doi: 10.15446/ing.investig.v42n1.86328.

[9] A. Jumikis, "Geology and soils of the Newark (N.J.) Metropolitan Area," *J. Soil Mech. Found. Div., ASCE,* vol. 94, no. 2, 1958.

[10] L.R. Blotz; C.H. Benson & G.P. Boutwell, "Estimating optimum water content and maximum dry unit weight for compacted clays," *J. Geotech. Geoenviron. Eng.*, vol. 124, no. 9, pp. 907–912, 1998.

[11] B. Ramiah; V. Viswanath & H. Krishnamurthy, "Interrealtionship of compaction and index properties," in *In Proceedings of the 2nd South East Asian Conference on Soil Engineering, Rotterdam, the Netherlands*, 1970, pp. 577–587.

[12] G.P. Manikopoulos & C.N. Korfiatis, "Correlation of maximum dry density and grain size," *J. Geotech. Eng. Div.*, vol. 108, no. 9, pp. 1171–1176, 1982.

[13] N.A. Anjita; A.G. Christy & S.V. Krishnankutty, "Prediction of maximum dry density using Genetic Algorithm," *Int. J. Eng. Res. Technol.*, vol. 6, no. 3, pp. 550–552, 2017.

[14] O. Sivrakaya; C. Kayadelen & E. Cecen, "Prediction of the Compaction Parameters for Coarse-grained Soils with Fines Content by MLR and GEP," *Acta Geotech. Slov.*, vol. 2, pp. 29–41, 2013.

[15] H. Hama Ali, "Soft Computing Models to Predict the Compaction Characteristics from Physical Soil Properties," *Eng. Technol. J.*, vol. 41, no. 5, pp. 1–18, 2023, doi: 10.30684/etj.2023.137772.1360.

[16] G. Ring; J. Sallberg & W. Collins, "Correlation of compaction and classification test data," *Highw. Res. Board Bull.*, vol. 325, pp. 55–57, 1962.

[17] A.N. Al-khafaji, "Estimation Soil Compaction Parameters using Atterberg Limits," *Q. J. Eng. Geol.*, vol. 26, pp. 359–368, 1993.

[18] N.S. Pandian; T.S. Nagaraj & M. Manoj, "Re-examination of Compaction Characteristics of Fine Grained Soils," *Geotechnique*, vol. 47, no. 2, pp. 363–366, 1997.

[19] O. G. Ingles & J. B. Metcalf, "Soil Stabilisation," Butterworth pty, Ltd., Australia.

[20] M. H. A. Khan, T. H. Jafri, S. Ud-Din, H. S. Ullah, and M. N. Nawaz, "Prediction of soil compaction parameters through the development and experimental validation of Gaussian process regression models," *Environ. Earth Sci.*, vol. 83, no. 4, pp. 1–20, 2024, doi: 10.1007/s12665-024-11433-4.

[21] G. Verma and B. Kumar, "Prediction of compaction parameters for fine-grained and coarse-grained soils: a review," *Int. J. Geotech. Eng.*, vol. 14, no. 8, pp. 970–977, 2020, doi: 10.1080/19386362.2019.1595301.

[22] A.K. Ören, "Estimating compaction parameters of clayey soils from sediment volume test," *Appl. Clay Sci.*, vol. 101, pp. 68–72, 2014, doi: http://dx.doi.org/10.1016/j.clay.2014.07.019.

[23] H. Mujtaba, K. Farooq, N. Sivakugan, and B. M. Das, "Correlation between gradational parameters and compaction characteristics of sandy soils," *Int. J. Geotech. Eng.*, vol. 7, no. 4, pp. 395–401, 2013, doi: 10.1179/1938636213Z.00000000045.

[24] A. K. Bera and A. Ghosh, "Regression model for prediction of optimum moisture content and maximum dry unit weight of fine grained soil," *Int. J. Geotech. Eng.*, vol. 5, no. 3, pp. 297–305, 2011, doi: 10.3328/IJGE.2011.05.03.297-305.

[25] K. Dokovic; D. Rakic & M. Ljubojev, "Estimation of Soil Compaction Parameters based on the Atterberg Limits," *Min. Metall. Institute, Bor.*, 2013, doi: Doi:10.5937/MMEB1304001D.

[26] P. Chindaprasirt; P. Kampala; A. Arngbunta & S.P. Horpibulsuk, "Prediction of Compaction Parameters of Khon Kaen Loess Soil," in *2nd International Symposium on Construction Innovation Research & PhD Symposium*, Walailak J. Sci & Tech, 2020, pp. 1367–1378.

[27] R. Polo-Mendoza, J. Duque, and D. Mašín, "Prediction of California bearing ratio and modified proctor parameters using deep neural networks and multiple linear regression: A case study of granular soils," *Case Stud. Constr. Mater.*, vol. 20, no. March 2023, 2024, doi: 10.1016/j.cscm.2023.e02800.

[28] K. Xia, "A large deformation finite element model for soil compaction," *Geomech. Geoengin.*, vol. 7, no. 2, pp. 123–137, 2012, doi: 10.1080/17486025.2011.578673.

[29] A.A. Fondjo; E. Theron & R.P. Ray, "Estimation of Optimum Moisture Content and Maximum Dry Unit Weight of Fine-Grained Soils using Numerical Methods," *Walailak J. Sci. Technol.*, vol. 18, no. 16, p. 22792, 2021, doi: https://doi.org/10.48048/wjst.2021.22792.

[30] S. Khuntia, H. Mujtaba, C. Patra, K. Farooq, N. Sivakugan, and B. M. Das, "Prediction of compaction parameters of coarse grained soil using multivariate adaptive regression splines (MARS)," *Int. J. Geotech. Eng.*, vol. 9, no. 1, pp. 79–88, 2015, doi: 10.1179/1939787914Y.0000000061.

[31] Y. Qian, "Maximum dry unit weight and optimum moisture content prediction of lateritic soils using regression analysis," vol. 02, no. 01, pp. 15–26.

[32] J. Khatti and K. S. Grover, "Prediction of compaction parameters for fine-grained soil: Critical comparison of the deep learning and standalone models," *J. Rock Mech. Geotech. Eng.*, vol. 15, no. 11, pp. 3010–3038, 2023, doi: 10.1016/j.jrmge.2022.12.034.

[33] B. Li, Z. You, K. Ni, and Y. Wang, "Prediction of Soil Compaction Parameters Using Machine Learning Models," *Appl. Sci.*, vol. 14, no. 7, 2024, doi: 10.3390/app14072716.

[34] B. T. Pham *et al.*, "A novel approach for classification of soils based on laboratory tests using Adaboost, Tree and ANN modeling," *Transp. Geotech.*, vol. 27, no. July 2020, p. 100508, 2021, doi: 10.1016/j.trgeo.2020.100508.

[35] A. Hasnat, "Prediction of Compaction Parameters of Soil using Support Vector Regression," *Curr. Trends Civ. Struct. Eng.*, vol. 4, no. 1, pp. 1–7, 2019, doi: 10.33552/ctcse.2019.04.000580.

[36] P. Zhu, Y. Zhu, and P. Zhang, "Comparison of SVR models for predicting the compaction properties of lateritic soils as novel hybrid methods," *Eng. Res. Express*, vol. 4, no. 3, 2022, doi: 10.1088/2631-8695/ac87eb.

[37] F.E. Jalal; Y. Xu; M. Iqbal; B. Jamhiri & M.F. Javed, "Predicting the compaction characteristics of expansive soils using two genetic programming-based algorithms," *Transp. Geotech.*, vol. 30, no. 100608, 2021, doi: https://doi.org/10.1016/j.trgeo.2021.100608.

[38] A. Bardhan and P. G. Asteris, "Application of hybrid ANN paradigms built with nature inspired meta-heuristics for modelling soil compaction parameters," *Transp. Geotech.*, vol. 41, no. March, p. 100995, 2023, doi: 10.1016/j.trgeo.2023.100995.

[39] A. Ardakani and A. Kordnaeij, "Soil compaction parameters prediction using GMDH-type neural network and genetic algorithm," *Eur. J. Environ. Civ. Eng.*, vol. 23, no. 4, pp. 449–462, 2019, doi: 10.1080/19648189.2017.1304269.

[40] J. Khatti and K. S. Grover, "Prediction of Compaction Parameters of Soil Using Ga and Pso Optimized Relevance Vector Machine (Rvm)," pp. 2890–2903, 2023, doi: 10.21917/ijsc.2023.0409.

[41] K. Djokovic, D. Rakic, and M. Ljubojev, "Estimation of soil compaction parameters based on the Atterberg limits," *Min. Metall. Eng. Bor*, no. 4, pp. 1–16, 2013, doi: 10.5937/mmeb1304001d.

[42] J. D. N. Dionisio, W. G. Burns, and R. Gilbert, "3D virtual worlds and the metaverse: Current status and future possibilities," *ACM Comput. Surv.*, vol. 45, no. 3, 2013, doi: 10.1145/2480741.2480751.

[43] J. Lin *et al.*, "From ideal to reality: segmentation, annotation, and recommendation, the vital trajectory of intelligent micro learning," *World Wide Web*, vol. 23, no. 3, pp. 1747–1767, 2020, doi: 10.1007/s11280-019-00730-9.

[44] D. Ibrahim, "An Overview of Soft Computing," *Procedia Comput. Sci.*, vol. 102, no. August, pp. 34–38, 2016, doi: 10.1016/j.procs.2016.09.366.

[45] M. Tavana and S. Sorooshian, "A systematic review of the soft computing methods shaping the future of the metaverse," *Appl. Soft Comput.*, vol. 150, no. November 2023, p. 111098, 2024, doi: 10.1016/j.asoc.2023.111098.

[46] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.

[47] A.K. Ören, "Estimating compaction parameters of clayey soils from sediment volume test.," *Appl. Clay Sci.*, vol. 101, pp. 68–72, 2014, doi: http://dx.doi.org/10.1016/j.clay.2014.07.019.

[48] K. Xia, "Numerical prediction of soil compaction in geotechnical engineering," *Comptes Rendus - Mec.*, vol. 342, no. 3, pp. 208–219, 2014, doi: 10.1016/j.crme.2014.01.007.

[49] K. Farooq, U. Khalid, and H. Mujtaba, "Prediction of Compaction Characteristics of Fine-Grained Soils Using Consistency Limits," *Arab. J. Sci. Eng.*, vol. 41, no. 4, pp. 1319–1328, 2016, doi: 10.1007/s13369-015-1918-0.

[50] A. Soltani, M. Azimi, B. C. O'Kelly, and S. Horpibulsuk, "Converting optimum compaction properties of fine-grained soils between rational energy levels," *Transp. Geotech.*, vol. 42, no. July, p. 101096, 2023, doi: 10.1016/j.trgeo.2023.101096.