Abstract: This study aims to develop an intelligent modeling approach for accurately predicting compaction characteristics of cohesive soils across compaction energy (CE) levels. A comprehensive database of 1001 observations falling within the theoretical bounds was created through experimental investigation encompassing sieve analysis, hydrometer analysis, liquid limit (w_L), plastic limit (w_P), specific gravity (G_s), and compaction tests on natural soil samples and literature review, encompassing diverse cohesive soils, CE levels, and compaction characteristics. Multiple machine learning techniques, including Extreme Gradient Boosting (XGBoost), Random Forest (RF), Gene Expression Programming (GEP), Multi Expression Programming (MEP), Artificial Neural Networks (ANN), and Multiple Linear Regression (MLR), were applied to develop predictive models. XGBoost demonstrated superior performance in predicting maximum dry density (γ_{dmax}) and optimum moisture content (w_{opt}) as evaluated by statistical indicators and external validation and compared with existing models in the literature. The models effectively captured the influence of key parameters, highlighting the primary role of CE and w_L, the secondary role of plastic limit (w_P), the tertiary role of plasticity index (I_P) and fines activity (A_F), and the quaternary role of soil gradation in predicting and influencing the compaction characteristics of cohesive soils. This approach enables accurate global modeling of cohesive soil compaction across varying CE levels, providing a valuable tool for geotechnical engineers and researchers to determine compaction characteristics for a known CE level using basic soil properties used for soil classification.

Keywords: Machine learning; XGBoost; Compaction; Cohesive soils; Big geotechnical data

1. Introduction

Compaction control is a crucial factor that governs the engineering properties of fine-grained soils also referred to as cohesive soils, across various earthwork projects, including the construction of foundation structures, roads, compacted clay liners, earthen dams, nuclear repositories, and embankments, among others (Wang and Yin, 2020; Zhu et al., 2024). To attain desired engineering characteristics through optimal densification in a project, the soil density and moisture content are benchmarked against the maximum dry density (γ_{dmax}) and optimum water content (w_{opt}) obtained through a compaction test (Horpibulsuk et al., 2009). The field compaction parameters are benchmarked against the parameters obtained at specified CE levels linked with the standard and modified proctor tests to obtain the degree of compaction (D_c) (Ran et al., 2024). Thus, compaction parameters are crucial in designing and commencing the compaction process in earthwork, while simultaneously establishing a standard for quality control and assurance (QA/QC) during field operations (J. Li et al., 2024; Ma et al., 2024; Tarantino and De Col, 2008). However, the testing methodologies required to obtain these parameters are characterized by labor-intensive and tedious procedures, necessitating a substantial quantity of representative material for the examination, factors linked to escalating project time and cost (Teramoto et al., 2024; Wang et al., 2020). Furthermore, this challenge is exacerbated by the adoption of test protocols standardized across various conceptual compaction energy (CE) levels, i.e., standard or modified, which may or may not align with actual field conditions (Alzubaidi et al., 2024; Jia et al., 2024). In practical applications, compaction energy (CE) levels can vary significantly, either adhering to established conceptual levels or being set at project-specific levels. This variability underscores the need for efficient methods to determine compaction characteristics across a range of CE levels. Recent research by Miller and Vahedifard (2024) and Spagnoli and Shimobe (2020) has highlighted this necessity, emphasizing the importance of adaptable compaction assessment techniques. Furthermore, there is a growing demand for methods to accurately determine the compaction state at arbitrary CE levels in the field. This is crucial for assessing the D_c achieved under various field conditions, which often differ from laboratory settings. In large-scale projects, to mitigate errors, field values of compaction characteristics are also increasingly being evaluated using samples retrieved from different field locations at arbitrary CE levels during the compaction process (Blanco, 2015). This underscores a need for an expedited method for the prediction of compaction parameters at arbitrary CE levels to reduce the testing requirements for QA/QC (Deng et al., 2024; Miller and Vahedifard, 2024).

For quick estimation of compaction characteristics, several researchers have established methods to determine γ_{dmax} and w_{opt} of cohesive soils using Atterberg's limits, i.e., liquid limit (w_L), plastic limit (w_P), and plasticity index (I_P) (Al-Khafaji, 1993; Farooq et al., 2016; Ito and Komine, 2008; Khalid and Rehman, 2018; Sridharan and Nagaraj, 2005; Tilahun et al., 2024; Wang and Yin, 2020). Most of the past correlations were established for a singular conceptual CE designated by the ASTM, i.e., standard (592.3 kN-m/m^3) or modified (2693.3 kN-m/m^3) or either unspecified CE; thereby, their application for an arbitrary CE or even to the conceptual CE designated by other standards, e.g., BS which involve a slight variation in CE than ASTM, is restricted (Blotz et al., 1998; Gurtug and Sridharan, 2004; Prakash et al., 2024). Attempts have been made to incorporate CE alongside other input parameters to predict the compaction characteristics of cohesive soils. Nevertheless, some models, e.g., Farooq et al. (2016) have treated CE as a constant rather than an independent input parameter, thereby compromising their predictive capability across various CE levels. Meanwhile, certain models developed for multiple CE levels not only highlight limitations in the range of CE levels in their dataset but also yield a dataset featuring CE application across inconsistent soil samples, thereby undermining the actual sensitivity of CE as a parameter within the models (Gurtug and Sridharan, 2004; Wang and Yin, 2020). Addressing this issue necessitates a dataset characterized by not only a broad spectrum of CE levels but also the systematic application of multiple CE levels to various homogeneous soil samples to broaden the compaction modeling horizon vis-a-vis CE (Najdi, 2023). Additionally, previous models have often utilized a limited set of input parameters such as w_L and w_P for predicting compaction parameters, potentially neglecting output variability imparted by other influencing factors (Ali et al., 2024; Spagnoli and Shimobe, 2020). Moreover, in different models, disparities arise in the assigned relative sensitivity of the input parameters, highlighting the necessity for more expansive incorporation of diverse parameters within the modeling framework to enhance generalizability. Nonetheless, for the sake of convenience in prediction, these input parameters should ideally be derived from simplified procedures. Furthermore, different existing models are typically applicable only within a limited range of cohesive soil plasticity, owing to their development based on a restricted dataset (Associates et al., 2016; Wang et al., 2023). Consequently, they lack the requisite generality to encompass the full spectrum of cohesive soil plasticity from low to high ranges (Agus et al., 2010; Ito and Komine, 2008). Therefore, it is imperative to devise global models that are founded upon expansive datasets incorporating a comprehensive array of plasticity indices, CE levels, grain size distribution parameters, and pertinent compaction attributes. By explicitly accounting for CE, models can bridge the gap between controlled laboratory conditions and variable field applications, improving their practical utility and adaptability to different geotechnical scenarios (Dialmy et al., 2024). This approach enables more precise predictions of γ_{dmax} and w_{opt} across a spectrum of energy levels, contributing to more efficient compaction processes and enhanced quality control in field applications (Miller and Vahedifard, 2024). Meanwhile, conventional regression-based modeling strategies used in existing models often yield diminished predictive efficacy, especially on a large dataset owing to their incapacity to effectively account for the antagonistic interactions or complex interplay among multiple parameters, underscoring the need for a more robust modeling framework.

Machine learning (ML) has emerged as a potent tool for predictive modeling, particularly in handling large datasets and input parameters (Rehman et al., 2024; Zhang et al., 2021). Various approaches have been devised, including black box, grey box, and white box techniques, which have evolved alongside advancements in neural networks, evolutionary, and decision tree-based ML algorithms (Khatti and Grover, 2023; ur Rehman et al., 2022). These ML-based methodologies have found extensive application in the modeling of geotechnical or geological databases. However, the inherent high variability within these datasets owing to antagonist response and complex interplay between different geological properties to define the response of geomaterials often leads to disparate predictive performances when employing specific ML techniques (Shi and Wang, 2021). To establish dependable models for geotechnical or geological databases, it is imperative to employ multiple techniques and robust modeling procedures to identify the most effective predictive models (Karpatne et al., 2019). Furthermore, when utilizing ML algorithms, careful consideration must be given to ensuring that the models adhere to the defined theoretical geotechnical framework. This is particularly crucial when modeling compaction characteristics, as not only high predictive accuracy is indispensable as these parameters serve as guidelines for QA/QC, but it also demands accurate prediction of parameter combinations, such as γ_{dmax} and w_{opt}, within the theoretical boundary of 100% degree of saturation ($S_{\gamma dmax}$), aspects often neglected in past models (Spagnoli and Shimobe, 2020; Tarantino and De Col, 2008; Tatsuoka and Gomes Correia, 2018). These aspects demand a systematic and extensive database, a robust modeling framework, and critical evaluation of developed models not only based on conventional key performance indicators (KPIs) but also against typical and customized external validation indices (EVIs).

Considering the preceding discourse, the current study establishes an extensive database for modeling the compaction characteristics of soils through a customized testing regimen and comprehensive literature review. The compilation of this database involves subjecting identical soil samples to multiple CEs, ensuring adherence to specified geotechnical theoretical bases to facilitate effective modeling frameworks. Further, by employing a suite of modeling techniques, including two variants of decision tree-based ML technique (i.e., Extreme Gradient Boosting (XGBoost) and Random Forest (RF)) encompassing enhanced black box technique, two variants of evolutionary ML technique (i.e., Genetic Expression Programming (GEP) and Multi Expression Programming (MEP)) encompassing gray box technique, along with two conventional methods such as Artificial Neural Networks (ANN) featuring standard black box ML technique and Multi-Linear Regression (MLR) featuring white box technique, a series of models were developed. These models underwent rigorous statistical assessment utilizing established KPIs to identify the most robust ones, subsequently validated against independent datasets using both typical and customized EVIs, and benchmarked against the existing models. Subsequently, the most optimal model was proposed and scrutinized for its internal modeling mechanisms, offering insights of interest to researchers and practitioners in the field. By providing a modeling strategy applicable to any CE level, including specified levels like standard and modified proctor, the current study presents a more versatile and potentially global solution for predicting compaction characteristics in place of doing laborious testing, ultimately benefiting the engineering community in optimizing soil compaction practices (Najdi, 2023). The predicted γ_{dmax} at standard or modified Proctor effort can serve as the reference for calculating the D_c for typical earthworks while the estimated w_{opt} for the anticipated field compaction effort can guide moisture conditioning of fill materials (Miller and Vahedifard, 2024). By predicting local γ_{dmax} values across a site, spatial variations in achievable density can be assessed, allowing for more targeted compaction efforts. Furthermore, estimated compaction parameters can inform initial earthwork planning and specifications potentially streamlining the project timeline and reducing preliminary costs.

2. Database Descriptions and Analysis

The establishment of the database was based on a comprehensive testing regimen and a systematic review of pertinent literature. The criteria delineating the inclusion of observations are derived from the specific objectives delineated within the ambit of this study. Each observation admitted into the modeling database necessitates the application of at least two distinct CE levels to homogenous soil specimens. Furthermore, to ensure the theoretical integrity of the modeling data, a criterion is enforced whereby the combination of the γ_{dmax} and w_{opt} is selected such that the $S_{\gamma dmax}$ does not surpass the theoretical threshold of 100%. Moreover, to maintain technical coherence within the modeling dataset, cohesive soils are operationally defined in accordance with ASTM D2487 specifications, which mandate the fines content (F_{200}) equal to or exceeding 50%. The calculation of CE levels adheres to prescribed standards, with a concerted effort expended towards encompassing a broad spectrum of such levels through an exhaustive data search. Additionally, the database is constructed primarily utilizing natural soil samples, the majority of which exhibit w_L and I_P below 100%, although a subset exceeding this threshold is also incorporated to enrich the modeling horizon. It is imperative to underscore that, for the validation dataset sourced from literature, a judicious selection process was employed, wherein datasets adhering to both the aforementioned limits and those deviating from them were considered, thereby facilitating an assessment of the models' applicability across a wider array of soil compositions and characteristics.

2.1. Test results data

To construct the database, a comprehensive testing program was developed, involving the collection of 156 soil samples from diverse locations within the Indus Plain, obtained from depths ranging between 1-2 meters. This selection aimed to encompass a broad spectrum of fine-grained soils for analysis. The characterization of these soil samples involved grain size distribution analysis conducted through both sieve analysis (ASTM D422) and hydrometer analysis (ASTM D7928), alongside Atterberg's limit tests adhering to ASTM D4318-95A. The outcomes of the grain-size distribution analysis revealed that all examined soil samples were classified as cohesive soils having fines (F_{200}) in a range of 50-100%, with grains such as gravel, sand, silt, and clay distributed within defined ranges: 0% to 10%, 0% to 50%, 42% to 90%, and 10% to 58%, respectively. Atterberg's limit testing illustrated w_L and w_P spanning from 15% to 78% and 9% to 26.6%, respectively, and I_P ranging from 1.63% to 60% (Fig. 1). These parameters collectively facilitated the categorization of soils into distinct subcategories, delineated as ML, CL-ML, CL, and CH according to ASTM D2487 (Fig. 1). Notably, the ML subset exhibited w_L values ranging from 15% to 29% and I_P values from 0% to 5%, whereas the CL-ML subset demonstrated w_L values between 19% and 28%, with I_P values ranging from 4% to 7%. The CL subset presented w_L and I_P ranges of 30% to 48% and 7.7% to 26.5%, respectively, while the CH subset displayed ranges of 52% to 78% for w_L and 27.5% to 60% for I_P. Furthermore, the clay activity (A_c), defined as the ratio of I_P to clay fraction, exhibited a range of 0.16 to 1.3 across all samples. This parameter has been utilized in previous models as an input; however, its determination necessitates the hydrometer test, a labor-intensive and time-consuming procedure coupled with Atterberg's limit test. Consequently, in the interest of modeling simplicity and circumventing reliance on the hydrometer test, the current study introduces an alternative parameter, namely fines activity (A_F), defined as the ratio of I_P to the fraction passing through the F_{200}. This parameter can be swiftly determined via sieve analysis and Atterberg's limit test and was observed to range from 0.017 to 0.625 for the tested samples. Thus, the testing dataset illustrates the collection of a diverse array of cohesive soil samples for this investigation.

Subsequently, these soil samples underwent soil compaction tests using standard CE of 593 kN-m/m^3 as per ASTM D-698 and modified CE of 2700 kN-m/m^3 as per ASTM D-1557. The relationship between the γ_d and w was plotted as a compaction curve, from which the γ_{dmax} and w_{opt} were determined as shown in Figure 2. Standard compaction tests, yielded γ_{dmax} and w_{opt} within a range of 13.5 kN/m^3 to 19.5 kN/m^3 and 10% to 25%, respectively. For CL, ML, CL-ML, and CH soils, γ_{dmax} values ranged from 14.3 kN/m^3 to 18.8 kN/m^3, 15.9 kN/m^3 to 19.3 kN/m^3, 16.3 kN/m^3 to 19.5 kN/m^3, and 13.5 kN/m^3 to 16.7 kN/m^3, respectively. Correspondingly, w_{opt} values for CL soils ranged from 11% to 21.8%, while for ML, CL-ML, and CH soils, w_{opt} ranged from 10.5% to 19.9%, 10% to 18.9%, and 14.5% to 25%, respectively (Fig. 2(a)). Furthermore, modified compaction tests, as illustrated in Figure 2b, yielded compaction parameters, including γ_{dmax} and w_{opt}, which fell within the ranges of 16 kN/m^3 to 20.4 kN/m^3 and 8.4% to 16.6%, respectively. These parameters were further delineated for each soil subgroup, with γ_{dmax} values ranging from 16.6 kN/m^3 to 20.4 kN/m^3 for CL, 17.8 kN/m^3 to 20 kN/m^3 for ML, 18 kN/m^3 to 20.4 kN/m^3 for CL-ML, and 13.5 kN/m^3 to 16.7 kN/m^3 for CH. Corresponding w_{opt} values ranged from 9% to 16.6% for CL, 8.4% to 13.5% for ML, 8.4% to 13% for CL-ML, and 12.8% to 16% for CH. Thus, a total of 312 sets of observations were yielded for the database based on this experimental program. Additionally, specific gravity (G_s) tests were conducted following ASTM D-854, revealing values ranging from 2.7 to 2.83. The G_s values were subsequently utilized to ascertain the zero air void line and $S_{\gamma dmax}$ as follows:

$$S_{\gamma dmax} = \frac{\gamma_{dmax}}{G_s \cdot \gamma_w - \gamma_{dmax}} \times (w_{opt} \cdot G_s) \quad (1)$$

where γ_w is the unit weight of water. The results indicated that all combinations of γ_{dmax} and w_{opt} exhibited $S_{\gamma dmax}$ lesser than the theoretical range of 100% saturation and nearly over 80%. It is important to note that compaction tests, such as the standard and modified Proctor tests, generally target an optimal saturation range of close to 80%–90% for cohesive soils, with around 80% often being the minimum for desirable engineering properties (e.g., low permeability) (Miller and Vahedifard, 2024; USDA, 2000). However, the actual saturation range mainly depends on the soil type and CE level.

2.2. Literature-based data

The research methodology encompassed a meticulous review of existing literature encompassing over 500 publications to select the most suitable one for data extraction. This process commenced by pinpointing relevant scholarly works using tailored keywords corresponding to the study's aims, such as maximum dry density, optimum moisture content, compaction energy, standard and modified tests as well as variations like reduced modified/standard, plasticity, and grain size distribution, etc. Databases like Web of Science, Scopus, and Google Scholar were extensively searched to locate pertinent published literature. Initially, abstracts of the selected studies were screened to determine their relevance, followed by a comprehensive evaluation vis-a-vis the current study's objectives through in-depth reading. Additionally, references cited within the searched studies were scrutinized to ensure any overlooked but relevant literature was included. Finally, as mentioned earlier, the extracted data underwent thorough scrutiny to identify and remove any outliers that exceeded the theoretical conditions and other predefined thresholds established for the current study. As a result, 681 data observations were finalized to be included in the modeling dataset mainly from a diverse range of literature from 1959-2020 (i.e., (Agus et al., 2010; Aldaood et al., 2015; Associates et al., 2016; Bello, 2013; Benson and Trast, 1995; Blotz et al., 1998; Daniel and Trautwein, 1994; Farooq et al., 2023, 2016; Gurtug and Sridharan, 2004; Horpibulsuk et al., 2009, 2008; Hussain, 2017; Ito and Komine, 2008; McRae, 1959; Mehmood et al., 2011; Miller et al., 2002; Osinubi and Nwaiwu, 2005; Osinubi et al., 2006; Osinubi and Bello, 2011; Perez N et al., 2013; Prasanna et al., 2020; Sabat, 2015; SAPEI et al., 1996; Sengupta et al., 2017; Sridharan and Gurtug, 2004; Taha and Kabir, 2005; Tinjum et al., 1997; Tripathy et al., 2005; Woon-Hyung Kim and Daniel, 1992; Yilmaz et al., 2016; Yogeshraj Urs and Prasanna, 2023;

Yusoff et al., 2017)). Atterberg's limit testing illustrated w_L and w_P ranging from 15.3% to 608% and 6% to 48.3%, and I_P spanning from 1.6% to 570.1% (Fig. 1). The soil classification was ranged in CL, ML, CH, MH and CL-ML in these samples as per ASTM D2487. Meanwhile, Sand and F_{200} were observed to be in the range of 0% to 50% and 50% to 100%, respectively, with A_F spanned from 0.001 to 11.4. Meanwhile, vast CE levels were procured in the dataset from literature ranging from 202 kN-m/m^3 to 10832.1 kN-m/m^3. Meanwhile, the γ_{dmax} and w_{opt} were in the range of 10.9 % to 22.5 % and 5.2% to 45%, respectively, with no datapoint exceeding theoretical $S_{\gamma dmax}$ over 100%. The description of the modeling dataset by amalgamating the testing dataset and the aforementioned literature dataset is presented in Table 1 and Figure 3. Moreover, within the context of literature, 139 datasets were obtained that met all the specified selection criteria for this study, with the exception of having undergone at least two distinct CE levels on identical samples and some observations having F_{200} less than 50%. These data points were excluded from the modeling database used for both the model training and testing processes. However, they were reserved for the external validation of the models due to their ability to provide a diverse validation horizon. The compilation of this dataset primarily originated from studies spanning the years 1993 to 2022 (e.g., (Akcanca and Aytekin, 2012; Al-Hussaini, 2017; Al-Khafaji, 1993; Burton et al., 2015; Clariá and Rinaldi, 2007; Daniel and Wu, 1993; Delage et al., 2011; Duong et al., 2014; Fox et al., 2014; Günaydın, 2009; Inci et al., 2003; Khalid et al., 2019; Kiliç et al., 2016; Lim and Miller, 2004; Othman and Benson, 2011; Rehman et al., 2017; Sawangsuriya et al., 2009; Shafiee, 2008; Shelley and Daniel, 1993; Taïbi et al., 2008; Vassallo et al., 2007; Vega and McCartney, 2015; Wang et al., 2017; Yang et al., 2022; Zhang et al., 2015). A detailed statistical description of the validation database sourced from the literature is presented in Table 1.

2.3. Data and input analyses

The database comprising 1001 observations was established through the amalgamation of testing data and literature-based data for model testing and training, encompassing a wide range of soil classification and compaction-related parameters (Fig. 3). The selection of input parameters was conducted based on theoretical relevance, ease of testing and statistical significance. Initially, four sets of input parameters were identified in the database for the input analysis based on theoretical relevance including Atterberg's limits (w_L, w_P, and I_P), grain size (Sand fraction (Sand), F_{200}, A_c, Silt, Clay, and A_F), physical property (G_s) and compaction (CE).

The w_L can be obtained through standard tests outlined in ASTM D4318-95A, specifically through Method A, i.e., Casagrande apparatus or Method B, i.e., fall cone apparatus. Similarly, the w_P can be determined through plastic limit test following the ASTM D4318-95A, also allowing for the calculation of the I_P as the difference between w_L and w_P. The Sand and F_{200} can be assessed through sieve analysis testing in accordance with ASTM D6913. Additionally, the A_F is calculated by dividing the I_P by F_{200}. For CE, the standard laboratory formula is given by $\mathrm{CE} = \frac{\text{Number of blows} \times \text{Number of layers} \times \text{Weight of hamer} \times \text{Hieght of drop}}{\text{Volume of mold}}$, as specified in ASTM D697 and ASTM D1557 and for field a different formula can be employed $C\mathrm{E} = \frac{\text{Drawbar pull} \times \text{ Number of passes} \times \text{Number of lifts for 0.31 m depth}}{\text{Drawbar width}}$ as per Johnson and Sallberg (1960). Thus, the selection of these parameters limits the testing requirements for obtaining input parameters to only sieve analysis and Atterberg limit tests, since CE does not require additional testing and can be calculated using formulas based on the specific compaction task. Further, these parameters, being essential to soil classification, offer engineers an initial, interpretable insight into soil properties, which is critical for envisaging the engineering behavior of soils. Parameters such as Silt, Clay, and A_c which require hydrometer analysis and G_s, were not further considered, as their incorporation would necessitate testing in excess of the basic requirements of USCS and AASHTO soil classification systems for model input despite their potential relevance to compaction characteristics. However, considering the theoretical relevance of G_s for soil compaction, it was indirectly used to determine $S_{\gamma dmax}$, which is included to evaluate the theoretical validity of the compaction parameters and various EVIs.

The statistical examination of the dataset utilized for model training and testing has revealed considerable variability across all pertinent input parameters, namely w_L, w_P, I_P, Sand, F_{200}, A_F, and CE, as delineated in Table 1 and Figure 4. Kernel Density Estimation (KDE) plots illustrate the probability density (P_d) distribution, portraying a consistent trend of P_d elevation from low to high plasticity ranges, which are representative of prevalent natural cohesive soils worldwide. Furthermore, a congruent trend in P_d is observed across parameters w_L, A_F, and I_P. Sand and F_{200} parameters also demonstrate notable variability, with heightened P_d evident at both lower and higher value ranges for sand, and predominantly at higher values for F_{200}. CE exhibits considerable variance, characterized by two instances of elevated P_d levels. Conversely, output parameters, γ_{dmax} and w_{opt}, demonstrate extensive variance within defined theoretical bounds, with heightened P_d occurring at central values for both. Subsequent analysis employing excess kurtosis indicates that all input and output parameters exhibit Leptokurtic distributions, except for sand and F_{200}, which manifest slightly Platykurtic distributions (Table 1). Moreover, skewness analysis reveals right-skewed distributions for all parameters, except for F_{200} and γ_{dmax}, which exhibit slight left-skewedness. This analysis underscores a high degree of symmetry in the distribution patterns of input and output parameters, albeit with minor exceptions.

Furthermore, a meticulous correlation analysis was conducted, involving the generation of both correlational and Pearson correlation coefficient (PC) matrices to scrutinize the interrelationships among the input and output variables, as illustrated in Figure 5. This comprehensive examination unveiled that no individual input parameter exhibits a significantly elevated or negligible correlation i.e., PC<0.1, as quantified by the PC, with the output parameter. This observation suggests that no single parameter singularly possesses the capacity to yield a substantial correlation with the output parameter, indicating the presence of a multifaceted relationship. Moreover, the assessment of correlations among the input variables revealed moderate to low PC values, indicative of a diminished likelihood of multicollinearity among the input parameters. Notably, γ_{dmax} and w_{opt} demonstrated a significantly high PC and a linear relationship, underscoring a discernible synchronization between these parameters, thereby reflecting the coherence and reliance of the dataset. Based on these insights, it was ascertained that all input parameters designated for inclusion exhibited reasonable statistical robustness. Additionally, in the context of ML-based modeling, the analysis of input parameters necessitates consideration vis-a-vis the number of data points, as a larger dataset provides a broader modeling horizon, while a greater number of statistically suitable input parameters contributes to enhanced predictive accuracy. However, a balanced ratio of at least 5 data points per input parameter is deemed desirable (ur Rehman et al., 2022); for the present dataset, this ratio stands at 143, exceeding the prescribed threshold value.

3. Modeling and Evaluation Methods

Figure 6 presents the overall methodology of the current study. Drawing upon input analysis that accounts for both physical and statistical significance, the input parameters employed comprise sieve analysis, Atterberg's limits, and compaction-based parameters. The output parameters can be expounded with input parameters as follows.

$$\gamma_{dmax} = f(w_L, w_P, I_P, \text{Sand}, F_{200}, A_F, \text{CE}) \quad (2)$$

$$w_{opt} = f(w_L, w_P, I_P, \text{Sand}, F_{200}, A_F, \text{CE}) \quad (3)$$

Furthermore, in order to construct machine learning (ML) based models, the database underwent partitioning into training and testing sets at a ratio of 70% to 30%. Additionally, an independent dataset constituting approximately 14% of the total data was reserved for external validation, distinct from the modeling process. A comprehensive array of ML methodologies was employed, encompassing three variants of black-box techniques, including the enhanced decision tree-based modeling approaches including Boosting Programming (i.e., XGBoost) and RF, alongside a conventional feed-forward ANN. Grey-box techniques were represented by two variants: Genetic Programming (i.e., GEP) and Evolutionary Programming (i.e., MEP). In addition, multiple linear regression (MLR) which is a white-box technique, was included as a baseline for comparison with other techniques.

Figures 7-10 present the algorithm architecture of different ML algorithms designed for this study. Extreme Gradient Boosting (XGBoost) is an advanced decision tree-based machine learning algorithm known for its efficiency and effectiveness in solving supervised learning problems to develop predictive models (Fig. 7). By utilizing a gradient-boosting framework, XGBoost sequentially combines weak learners, typically decision trees, to form a strong ensemble model, making it a superior variant compared to other decision tree-based techniques (Chen and Seo, 2023). It incorporates regularization methods like Lasso (L1) and Ridge (L2), which help prevent overfitting and enhance generalization by applying penalties to the coefficients' magnitudes. Additionally, XGBoost is optimized for high performance, offering parallel and distributed computing capabilities that enable fast training on large datasets and help capture complex data patterns. It also provides customizable parameters and feature importance scores, making it a flexible and interpretable choice for various domains, especially geotechnical databases. The XGboost algorithm developed for the current study can be mathematically expounded as follows.

$$O_i = \sum_{i=1}^{n} L(y_{\mathrm{i}}, \hat{y}_{\mathrm{i}}) + \sum_{k=1}^{K} \Omega(f_{\mathrm{k}}) \quad (4)$$

$$\Omega(f_{\mathrm{k}}) = \gamma T + L_1 + L_2 \quad (5)$$

where O_i is the objective, $\Omega(f_k)$ is the regularization function, $L(y_i, \hat{y}_i)$ is the loss function, with n representing the number of training samples, K is the number of trees and γT defines the thickness of the decision tree. Random Forest (RF) is another powerful ensemble learning algorithm that constructs multiple decision trees during training, each trained on a random subset of the data and features using a bagging technique (Fig. 8). Similar to XGBoost, RF offers insights into feature importance and robustness against overfitting through the aggregation of predictions. However, unlike XGBoost, which uses gradient boosting and regularization, RF relies on the diversity of decision trees and bagging to achieve predictive power. Gene Expression Programming (GEP) is a gray-box machine learning method inspired by genetic evolution, utilizing regression and neural techniques to solve complex problems. It iteratively adjusts parameters to optimize performance and employs expression trees to represent nonlinear entities (Fig. 9). GEP has shown promising results in handling complex geotechnical data, making it a suitable choice for the current modeling study. Multi-Expression Programming (MEP), another evolutionary modeling approach, uses fixed-length binary strings to encode multiple computer programs within a single chromosome. MEP applies a fitness evaluation process to generate the most optimal solution through recombination and mutation of chromosomes (Fig. 10). Artificial Neural Networks (ANN) are black-box models where data flows through interconnected nodes across layers, learning complex patterns from input data via activation functions applied to the weighted sum of inputs. A feed-forward ANN model is utilized in this study to test various black-box methods on the geotechnical dataset. Meanwhile, Multiple Linear Regression (MLR) is a widely used statistical method that models the relationship between a dependent variable and multiple independent variables; it estimates the coefficients that best fit the data.

This selection facilitated a comprehensive exploration of ML modeling paradigms on the established geotechnical database for compaction modeling. Multiple models were iteratively developed through algorithmic refinement and parameter optimization across the spectrum of employed techniques. The models exhibiting superior statistical performance were identified through a meticulous evaluation process. The most refined models derived from each ML technique further underwent meticulous examination and comparison, predicated upon a variety of absolute and relative statistical KPIs. The ML-based model's performance underwent evaluation through distinct categories of key performance indices (KPIs). This includes error indices such as root mean squared error (RMSE), relative root mean square error value (RRMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Additionally, correlation and effeciency indices, comprising Nash-Sutcliffe efficiency coefficient (NSE), Kling-Gupta efficiency coefficient (KGE), correlation coefficient (R^2), residual sum of square (RSS), norm of residual (NR) and adjusted correlation coefficient (adj R^2), and PC were analyzed as per ur Rehman et al., (2022). These indices have been chosen to encompass relative and absolute statistical perspectives, facilitating a comprehensive model performance analysis. Additionally, variance analysis (ANOVA) was performed on the developed models to assess their statistical integrity including F-value and prob>F.

Subsequently, the performance of these optimized models, as determined by the designated criteria, was compared with that of existing models in the literature, utilizing the modeling dataset. Additionally, both the optimized and the literature-based models underwent further validation procedures utilizing an independent validation dataset, and their performance was assessed using standard and customized EVIs. To further validate the credibility of the proposed model, an additional dataset comprising 139 observations sourced from existing literature was utilized. This dataset served to assess the predictive efficacy of the proposed model, comparing its performance with that of established models documented in prior literature some of which even used this validation dataset in their modeling process. Various standard EVIs, including the coefficient of determination (R), the correlation coefficient (R_m), and the slopes of regression lines (k and k'), along with their corresponding indices (R_0 and R_0'), were methodically assessed in this comparative analysis as per Rehman et al., (2024). Additionally, customized EVIs were developed to assess the compaction characteristics, including Normalized Absolute Percentage Error (NAPE) and Outlier $S_{\gamma dmax}$ ($S_{\gamma dmax(ol)}$). NAPE is defined mathematically as follows:

$$\text{NAPE} = \left| \frac{E_i - M_i}{((E_i + M_i)/2)} \right| \times 100 \qquad (6)$$

NAPE was further customized for the w_{opt} and γ_{dmax} in accordance with AS 1289.5.1.1 and AS 1289.5.2.1 standards, which stipulate that NAPE should not exceed 20% and 4% for the former and latter, respectively, for more than 5% of prediction samples in a validation dataset, with certain exceptions in high plastic clays (Spagnoli and Shimobe, 2020). Furthermore, $S_{\gamma dmax(ol)}$ is defined when Eq. 1 for any model prediction for the combination of w_{opt} and γ_{dmax} yields a value exceeding 100% or less than 0 (Tarantino and De Col, 2008). This ensures the theoretical validity of the developed model by verifying its adherence to the theoretical bounds of the soil saturation state. $S_{\gamma dmax(ol)}$ and the outlier NAPE ($NAPE_{ol}$) for w_{opt} and γ_{dmax} can be defined as follows:

$$\text{NAPE}_{ol} = \frac{k_{NAPE(ol)}}{n} \times 100 \qquad (7)$$

$$\text{S}_{\gamma\text{dmax(ol)}} = \frac{k_{S\gamma dmax(ol)}}{n} \times 100 \qquad (8)$$

where $k_{(ol)}$ denotes the count of outliers for each index, taking into account the pertinent theoretical bounds.

Finally, the most sophisticated model was proposed, marking the culmination of a rigorous process in selecting the optimal ML-based models for this study. Furthermore, the modeling mechanism of the proposed model was subjected to thorough analysis through Shapley additive explanations (SHAP) analysis on the modeling dataset and parametric and sensitivity evaluations on the assumed ranged dataset.

4. Results and Discussion

4.1. Model Development

4.1.1. XGBoost model

An extensive XGBoost modeling approach was utilized to construct a robust predictive model for predicting γ_{dmax} and w_{opt} parameters. Model training procedures were conducted to attain optimal settings that yield high predictive accuracy while mitigating model complexity. By meticulously fine-tuning the model parameters and iteratively refining the algorithmic settings, a level of predictive accuracy was achieved that closely mirrors experimental values. This optimization process was achieved by establishing specific thresholds for model complexity parameters and varying algorithmic parameters within these bounds. Optimal settings were determined wherein the number of trees was limited to 100, each with a depth of 6, and a learning rate of 0.30 was employed. Additionally, the fraction of features and regularization constant were constrained to 1.00. Through iterative refinement, the best-performing model was identified based on these algorithmic constraints. The optimal model for γ_{dmax} showed a high level of accuracy in the prediction aligning with experimental observations (Fig. 11a). Both training and testing data points and trend lines exhibited minimal deviation from the ideal 45-degree line remaining within the ±5% reference error band with R^2 of 0.99 and 0.99, respectively (Table 2), indicative of strong model performance. Similarly, the predictive accuracy of w_{opt} was demonstrated through XGBoost modeling, with the predicted values closely matching experimental observations mainly remaining within the ±5% reference error band of the 45-degree line with R^2 of 0.98 and 0.98 for testing and training databases, respectively (Fig. 12a). Notably, trend lines for both predicted and experimental data points remained consistent with the 45-degree line, further affirming the model's reliability. The close alignment between the trend lines of testing and training datasets for both models and the 45-degree line suggests that the model generalizes well to unseen data and model is interpretable and aligns well with domain knowledge. This also indicates that the model has learned the underlying patterns in the data without overfitting the training set. These findings and statistical KPIs as manifested in Table 2 underscore the high predictive accuracy and interpretability afforded by the developed XGBoost models in predicting γ_{dmax} and w_{opt} parameters. XGBoost model is highly applicable when prediction accuracy of compaction characteristicscs is critical, and the data shows complex, non-linear relationships.

4.1.2. RF model

An RF modeling approach was undertaken to develop models for γ_{dmax} and w_{opt}, involving meticulous fine-tuning of model parameters and iterative refinement of algorithmic settings. Through this process, optimal models were obtained based on their predictive performance and complexity considerations. Notably, the optimal model for γ_{dmax} exhibits minimal deviations between predicted and experimental values, as evidenced by their proximity to the 45-degree line within ±5% reference error band across both testing and training datasets with R^2 of 0.93 and 0.83, respectively. Additionally, while slight deviations are observed at extreme values, the trend lines for both datasets largely align with the 45-degree line (Fig. 11b). Similarly, the optimal RF model for w_{opt} demonstrates a comparable performance, with minimal deviation between predicted and experimental values across testing and training datasets yielding R^2 of 0.95 and 0.92, respectively, and trend lines consistent with the 45-degree line (Fig. 12b). These analyses and statistical KPIs of the developed models in Table 2 highlight a good performance of the optimal RF models for γ_{dmax} and w_{opt}. The RF model's performance also suggests a good balance between accuracy and generalizability for the prediction of γ_{dmax} and w_{opt}. Its ensemble nature allows it to capture non-linear relationships while maintaining robustness against overfitting. This makes the RF model particularly suitable for scenarios where both predictive power and model generalizability are desired.

4.1.3. GEP model

A wide-ranging GEP modeling methodology was employed to optimize genetic operators, constants, and general parameters. This optimization procedure was facilitated through a stepwise GEP modeling strategy, iteratively transitioning from simpler to more complex algorithmic parameters to enhance predictive performance. The attainment of the optimal model configuration necessitated meticulous adjustment of mutation, inversion, genetic recombination, uniform recombination, and constant fine-tuning rates, resulting in respective values of 0.00138, 0.00546, 0.000277, 0.00755, and 0.00206. Upon thorough evaluation, the optimal GEP model for predicting γ_{dmax} exhibited a moderate degree of dispersion between predicted and experimental values across both testing and training datasets centered around the 45-degree line with R^2 of 0.77 and 0.71, respectively. It is noteworthy that the trend lines for both datasets displayed deviations, characterized by underestimation at higher values and overestimation at lower values (Fig. 11c). A similar trend was observed in the estimation of w_{opt} with R^2 of 0.78 and 0.71 for training and testing datasets, respectively (Fig. 12c). These observations and statistical KPIs (Table 2) suggest that the GEP models predict γ_{dmax} and w_{opt}, reasonably aligning with experimental observations while manifesting a moderate degree of discrepancies. The performance of the GEP model indicates its potential for capturing the underlying patterns in the data. However, the observed deviations in γ_{dmax} and w_{opt} predictions suggest that the model may be struggling to fully represent the complexity of the relationships between variables in the current dataset. This model could be particularly useful in situations where a balance between interpretability and predictive power is required, and insights into the data structure are needed, especially when a more flexible model is desired, albeit with some limitations in accuracy.

4.1.4. MEP model

MEP modeling was employed to achieve optimal models for γ_{dmax} and w_{opt} by varying the algorithmic parameters and applying multiple evolutionary iterations. The configuration of the MEP algorithm involved the establishment of two subpopulations, each composed of 1000 iterations. It implemented a crossover probability of 0.9 and incorporated a tournament size of 2. Based on the statistical KPIs, the best-performing MEP models for γ_{dmax} and w_{opt} were finalized. The optimal MEP model predictions for γ_{dmax} showed a considerable level of scatteredness in comparison to experimental values for both testing and training datasets around the 45-degree line, with R^2 values of 0.65 and 0.61, respectively (Fig. 11d). The trend lines of both testing and training datasets underscore considerable underestimation at higher values and overestimation at lower values. A similar trend was observed for the w_{opt} prediction by the optimal MEP model, with slightly improved R^2 values of 0.71 and 0.66, respectively, for the training and testing datasets (Fig. 12d). These analyses and statistical KPIs of both optimal MEP models for γ_{dmax} and w_{opt} (Table 2) demonstrate considerably low prediction accuracy, generalizability, and interpretability. However, this model offers simplicity, ease of interpretation, and an evolutionary prediction stream, its applicability may be limited to scenarios where slightly less accurate estimates of γ_{dmax} and w_{opt} are sufficient.

4.1.5. ANN model

The feedforward ANN model was constructed with a variable number of neurons ranging from 10 to 100 per hidden layer. The selection of this range aimed to strike a balance between model complexity and efficacy, as excessively high neuron counts tend to yield intricate and burdensome models. Optimal performance was attained when employing 100 neurons, coupled with an α-value of 10^{-4}. The ANN model developed for predicting γ_{dmax} exhibited a discernible discrepancy between predicted and experimental values, as evidenced by respective R^2 values of 0.70 for both the training and testing datasets (Fig. 11e). Notably, trend line analysis revealed a tendency towards overestimation at lower values and underestimation at higher values. A similar pattern was observed in the prediction of w_{opt} by the optimal ANN model, yielding R^2 values of 0.71 and 0.66 for the training and testing datasets, respectively (Fig. 12e). These analyses, along with statistical KPIs presented in Table 2, collectively suggest that the feedforward ANN models exhibit relatively low accuracy, generalizability, and interpretability, particularly concerning predictions of γ_{dmax} and w_{opt}. The model's performance and its tendency to overestimate at lower values and underestimate at higher values suggest that it may be particularly useful in mid-range predictions of γ_{dmax} and w_{opt}. In practical applications, this ANN model could be employed in preliminary soil compaction assessments or as part of a larger ensemble of models to provide a comprehensive view of soil behavior.

4.1.6. MLR model

MLR was also utilized in this study, as it stands as the most prevalent technique in the existing literature for developing predictive models concerning γ_{dmax} and w_{opt}. However, the MLR approach yielded models displaying substantial deviations in the prediction of γ_{dmax}, as indicated by relatively low R^2 values of 0.55 and 0.54 for the training and testing datasets, respectively (Fig. 11f). Notably, both the training and testing dataset trend lines exhibited significant overestimation and underestimation in comparison to the 45-degree line at low and high values. A parallel pattern emerged in the prediction of w_{opt} using MLR, albeit with marginally enhanced R^2 values of 0.63 and 0.71 for the training and testing datasets, respectively (Fig. 12f). These analyses, in conjunction with statistical KPIs (Table 2), collectively underscore the limited accuracy and generalizability exhibited by MLR models in predicting γ_{dmax} and w_{opt}. This suggests that while MLR is simple and interpretable, its applicability in the current context is limited. The model's performance could potentially be improved through feature engineering but more complex ML models are better suited for this particular problem.

4.2. Performance analysis

4.2.1. Performance analysis of developed models

The performance of the best models from all the applied ML techniques was evaluated to identify the most suitable models for further analysis, based on the statistical KPIs presented in Figure 13 and Table 2. Regarding correlation and efficiency indices (i.e., R^2, PC, adj R^2, KGE, and NSE), the model performance showed a significant improvement in the order of MLR, MEP, ANN, GEP, RF, and XGBoost for both γ_{dmax} and w_{opt}. All models exhibited better performance on the training dataset compared to the testing dataset for both categories of models. XGBoost demonstrated the smallest reduction in performance from the training to the testing dataset for both γ_{dmax} and w_{opt}, followed by ANN, MLR, RF, GEP, and MEP for γ_{dmax} (Fig. 13a), and RF, GEP, MEP, and ANN for w_{opt} (Fig. 13b). These correlation and efficiency indices clearly indicate that the XGBoost model outperforms other models (Table 2), as all these indices are close to unity, demonstrating high prediction accuracy (Figure 13). The RF model also performed excellently in terms of correlation and efficiency indices. However, all other models showed values below 0.9 for these indices. These results highlight that black-box decision-tree-based ML techniques effectively approximate predicted values to experimental values in the current geotechnical database (Shi and Wang, 2021).

The performance analysis was further conducted using multiple error indices (i.e., MAE, MAPE, MSE, RMSE, and NRMSE), covering both absolute and relative statistical perspectives. The models' performance was consistent with the results from correlation and efficiency indices for both testing and training datasets for γ_{dmax} and w_{opt} (Table 2 and Fig. 13). The models' performance improved in terms of error indices in the order of MLR, MEP, ANN, GEP, RF, and XGBoost for both γ_{dmax} and w_{opt}. The XGBoost model outperformed all the models, yielding error indices close to zero (Table 2 and Figure 13). The RF model also showed relatively low error indices compared to other models except XGBoost. Higher error indices were observed for GEP, MEP, ANN, and MLR. Furthermore, XGBoost exhibited exceptionally low NR and RSS, with RF showing the second-best performance, while other models showed high values indicating larger residuals in predictions. These analyses emphasize the extremely low prediction error by XGBoost and the reasonable performance of RF on the current database. Nevertheless, all models demonstrated satisfactory performance, as evidenced by high F-values and low Prob>F in the ANOVA, with XGBoost showing the highest F-value followed by RF. Based on the comprehensive performance across all statistical KPIs, XGBoost exhibited exceptionally high prediction accuracy outperforming all the models. At the same time, the RF model also showed reasonable performance compared to other models except XGBoost for both γ_{dmax} and w_{opt}. Thus, XGBoost and RF models are shortlisted for further evaluation.

4.2.2. Performance analysis of shortlisted and existing models

The performance of existing models in the extant literature (e.g., (Al-Khafaji, 1993; Blotz et al., 1998; Farooq et al., 2016; Günaydın, 2009; Gurtug and Sridharan, 2004; Ito and Komine, 2008; Sridharan and Nagaraj, 2005; Wang and Yin, 2020)) for predicting γ_{dmax} and w_{opt}, either with CE as an input parameter or without specifying bounds for any particular CE level, was evaluated using the current training and testing datasets used for the model development to compare with the proposed ML models, i.e., XGBoost and RF, developed in this study. All existing models exhibited substantial deviations between predicted and experimental γ_{dmax} and w_{opt} values from the 45-degree line, with these deviations significantly exceeding the reference ±5% error band, indicating inferior performance compared to XGBoost and RF (Figs. 12a, b; 13 a, b; 14a; 15a). This finding suggests that the applicability of existing models is limited to certain thresholds but to a larger spectrum of cohesive soils and CE levels. Furthermore, performance comparisons based on correlation indices (R^2 and PC) and error indices (MAE, RMSE) revealed suboptimal performance of all the existing models and excellent performance of the proposed models for predicting γ_{dmax} and w_{opt}, when combining the training and testing datasets used previously (Figs. 14b, c; 15b, c). XGBoost outperformed all other models based on this big data showing a wider application range. The Taylor diagram analysis further confirms that XGBoost predicts γ_{dmax} and w_{opt} with the least deviation and the highest correlation strength, followed by RF (Fig. 16). In contrast, all existing models demonstrated a larger deviation from experimental values in their predictions, with a standard deviation exceeding 2 (Fig. 16). This indicates a low fidelity of these models in predicting accurate results. The superior performance of XGBoost and RF in comparison to existing models underscores their robustness and reliability in capturing the underlying patterns in this big dataset, as evidenced by their proximity to the experimental values on the Taylor diagram. It is noteworthy that while all MLR-based models demonstrated a high degree of deviation, the MEP-based model by Wang and Yin (2020) outperformed the other existing models but still exhibited inferior performance to the MEP model developed in the current study demonstrating the effectiveness of training the model on a big database (Figs. 14 and 15 and Table 2). The superior performance of XGBoost and RF compared to existing models can be attributed to their training on a larger dataset encompassing a broad spectrum of soil plasticity, classification, and CE levels (Chen and Seo, 2023). It is also important to note that the part of the database used in this performance analysis also serves as a parent database, necessitating further verification of these models as presented in the following section.

4.3. External validation

The shortlisted models, namely XGBoost and RF, along with existing models from the current literature, were validated using an independent dataset of 139 observations obtained from previous studies. This dataset, which spans a wide range of statistical input and output parameters (Table 1), was not used in the model testing and training phases, thus termed as an independent dataset. It was observed that XGBoost predicted the γ_{dmax} and w_{opt} values within a

±5% reference error band, whereas all other models exhibited greater deviations (Fig. 17). Taylor diagram analysis further confirmed these findings, indicating that XGBoost predicted γ_{dmax} and w_{opt} values closest to the experimental results, followed by the RF and the MEP model by Wang and Yin (2020) (Fig. 18). Most existing models showed significant deviations in γ_{dmax} and w_{opt}, highlighting their lower fidelity in prediction accuracy. The external validation was extended to assess the models' ability to yield a combination of γ_{dmax} and w_{opt} that breaches the threshold $S_{\gamma dmax}$ value, i.e., higher than 100% or equal to or lower than 0%. It was observed that all models tended to produce some $S_{\gamma dmax}$ values outside theoretical bounds at certain instances, except for XGBoost, which consistently predicted γ_{dmax} and w_{opt} values close to the experimental data without breaching theoretical limits (Fig. 19). This indicates that other models are prone to unacceptable results, while XGBoost effectively captures minimal prediction error and remains within the desired bounds of 70% < $S_{\gamma dmax}$ < 100% as of experimental values (Fig. 19). This performance is mainly attributed to the screened database used for model development, which excluded observations yielding unacceptable $S_{\gamma dmax}$ combined with the sophisticated training ability of XGBoost demonstrated on the current database to maintain minimal error indices and close proximity with experimental values.

Further external validation was conducted by assessing the performance of the models on the independent dataset using standard Evaluation Indices (EVIs) (i.e., R, R_m, k, k', R_0, and R_0') and customized EVIs ($NAPE_{ol}$ and $S_{\gamma dmax(ol)}$) for γ_{dmax} and w_{opt}, as presented in Tables 3 and 4, respectively. The results indicate that XGBoost maintains all the EVIs within desirable ranges set to demonstrate high prediction accuracies, whereas the other models struggle to do so, failing in some or all EVIs. This demonstrates the superior prediction accuracy of XGBoost and its robust capability to perform global predictions of cohesive soil compaction within acceptable limits. In contrast, the other models exhibit certain limitations, failing to consistently achieve desirable EVI metrics. These findings highlight the efficacy of XGBoost in modeling complex soil compaction parameters, reinforcing its potential for broader applications in geotechnical engineering. The comprehensive evaluation underscores the model's ability to generalize well across diverse datasets, thus providing reliable predictions. This superior performance can be attributed to the advanced algorithmic structure of XGBoost, which effectively captures nonlinear relationships and interactions within the data coupled with training on a big and systematic dataset, resulting in higher fidelity and precision (Chen and Seo, 2023; Shi and Wang, 2021). Consequently, XGBoost stands out as a reliable tool for predicting γ_{dmax} and w_{opt}, offering significant improvements over traditional and other ML models in this domain.

4.4. Model explanation and modeling mechanism analysis

4.4.1. Proposed model explanation

The XGBoost model is a black box, making it difficult to interpret it in terms of mathematical formulations for predictions. To improve interpretability, SHAP analysis was used, based on modeling data used in this study. SHAP assigns rank values to input features for individual predictions, offering insight into how input features, i.e., Atterberg's limit, gradation parameters and compaction effort affect predictions of γ_{dmax} and w_{opt}. (X. Li et al., 2024). The SHAP summary plots for the prediction of the XGBoost model proposed for γ_{dmax} and w_{opt} in this study are presented in Figures 20 (a) and (b), respectively. In SHAP summary plots, the horizontal axis represents SHAP values, which indicate the extent to which each input feature pushes the prediction higher (positive SHAP values) or lower (negative SHAP values). Each row corresponds to an input feature, and each dot represents a SHAP value for a particular observation. The color gradient from blue (low feature values) to red (high feature values) shows how the magnitude of each input feature's value affects the model's output.

The CE feature has a particularly strong impact on γ_{dmax} and w_{opt}. Low values significantly reduce the prediction of γ_{dmax} and increase the prediction of w_{opt}, as indicated by a cluster of blue dots with negative and positive SHAP values, respectively. Conversely, high values of CE push the prediction upward and downward for γ_{dmax} and w_{opt}, respectively, indicating that this feature is a key driver for increasing γ_{dmax} and decreasing w_{opt}. w_L also shows a substantial influence on both γ_{dmax} and w_{opt}. High values of w_L increase the prediction of γ_{dmax} and low values decrease it; meanwhile, a reverse response was observed for w_{opt}. This feature has a relatively wide distribution for both γ_{dmax} and w_{opt}, indicating interactions with other features or nonlinear effects. w_P displays a balanced spread around zero for both γ_{dmax} and w_{opt}, with high values associated with a slight increase in γ_{dmax} and low values linked to a decrease in the prediction of γ_{dmax}; meanwhile, a reverse trend was observed for w_{opt}. This distribution pattern suggests that w_P has a moderate but consistent effect on both γ_{dmax} and w_{opt} since clustering is closer to zero SHAP value as compared CE and w_L. Moreover, CE has a consistent impact on both γ_{dmax} and w_{opt}. In contrast, w_L and w_P have slightly varying degrees of impact on γ_{dmax} and w_{opt}, with a stronger influence on w_{opt} than on γ_{dmax}. I_P has a also mixed influence on compaction characteristics, with high values leaning toward increasing γ_{dmax}, while low values slightly decrease; a reverse trend was observed for w_{opt}. However, the distribution pattern of I_P shows an interaction with other features and more clustering close to zero showing less impact than CE, w_L and w_P. A_F, on the other hand, shows a less pronounced impact for both γ_{dmax} and w_{opt}; the values cluster around zero, meaning it slightly affects the predictions. The Sand and F_{200} features also show a los impact on γ_{dmax} and w_{opt}, as most SHAP values are concentrated near zero, indicating these features have little effect on the models' output in comparison to CE, w_L, w_P, and I_P. Moreover, the occasional and frequent occurrence of red and blue dots across negative and positive SHAP values for input features suggests that the impact of each feature is influenced by its interactions with other features for different instances in predicting γ_{dmax} and w_{opt}, thereby helping to bridge gaps in prediction and enhance prediction accuracy.

Overall, w_L and CE stand out as the most influential features playing a primary role in predicting γ_{dmax} and w_{opt}. w_P has a secondary role, with moderate effects that mirror w_L's trend in predicting compaction parameters. I_P and A_F, have a tertiary and F_{200}, and Sand exhibits a quaternary role. This shows that the proposed XGBoost aligns with geotechnical principles, as increased CE allows particles to rearrange efficiently at lower moisture levels and decrease the void spaces significantly resulting in high γ_{dmax}. w_L shows a strong positive correlation with γ_{dmax} and w_{opt}, reflecting that soils with higher w_L require more water for optimal compaction due to their greater water-holding capacity and undergo less densification while w_P and I_P also mirror this impact in a comparatively less pronounced manner. The small impact of F_{200} and Sand suggests that grain size distribution shows low influence over compaction characteristics compared to Atterberg's limit and CE for cohesive soils but their impact somewhat influences the impact of other peoperties.

4.4.2. Parametric analysis

A parametric analysis was conducted by maintaining all parameters at their minimum values and varying individual input parameters to observe the impact on the model's performance, as illustrated in Figure 21. This analysis revealed that in the XGBoost model, the w_{opt} increases while the γ_{dmax} decreases with increases in w_L, w_P, and I_P, with w_L exerting the most significant influence among them. Specifically, the impact of w_L is pronounced at approximately 100%, after which its influence becomes negligible. The variables Sand and F_{200} demonstrated inverse effects relative to each other and slightly impacted compaction parameters. This effect was significant only when Sand was below 50% and F_{200} exceeded 50%, highlighting that the model accounts for the impact within the fine-grained soil boundary defined by ASTM D2487, given the data limitations for cohesive soils. Furthermore, the CE increased γ_{dmax} and decreased w_{opt}, consistent with basic soil mechanics principles. However, this impact was significant only for CE values below 5000 kN-m/m³, after which it became less significant. The variable A_F exhibited a slight impact on γ_{dmax} and w_{opt}, primarily at lower values. The XGBoost model consistently performed without anomalies, whereas the RF model exhibited some irregularities. For instance, with CE values, the RF model showed a decrease in w_{opt} up to 2500 kN-m/m³, followed by an increase up to 5000 kN-m/m³. Similar irregularities were observed for F_{200} and Sand in the RF model. These irregularities can be regarded as the reason for the inferior performance of RF than XGBoost.

Additionally, the model's efficacy was evaluated by examining the combination of predicted γ_{dmax} and w_{opt} to assess its impact on $S_{\gamma dmax}$, by varying G_s within the theoretical bounds of 2.65 and 2.9 for cohesive soils (Fig. 22). The XGBoost model consistently yielded theoretically valid results, did not produce a combination that resulted in $S_{\gamma dmax}$ exceeding 100% for any instance. In contrast, the RF model frequently produced $S_{\gamma dmax}$ values surpassing 100%. This demonstrates the robust underlying modeling mechanism of the XGBoost model, enabling it to accurately predict γ_{dmax} and w_{opt} while adhering to theoretical constraints.

4.4.3. Sensitivity analysis

The sensitivity analysis using the assumed database used in parametric analysis revealed that in the XGBoost model, all input parameters possess a certain S_i, validating the input analysis conducted for this study (Figure 23). Notably, w_L and CE emerged as the most sensitive parameters for γdmax and w_{opt}, similar to the SHAP analysis. Consequently, the XGBoost model is suitable for application across a range of cohesive soil classifications and varying CE levels, accommodating a wide range of variations in these two most significant input features. The S_i for the XGBoost model varied in the order of w_L, CE, w_P, I_P, A_F, Sand and F_{200} for γ_{dmax}, and CE, w_L, w_P, I_P, A_F, F_{200}, and Sand for w_{opt} (Fig. 22). This indicates a robust modeling process employed by the XGBoost algorithm, which mathematically bounds different parameters with distinct schemes for γ_{dmax} and w_{opt}, thereby accurately predicting these values.

5. Practical Application

The proposed method in this study for predicting γ_{dmax} and w_{opt} at arbitrary CE levels has significant implications for field applications in geotechnical engineering and construction projects. In large-scale earthwork operations, where soil properties can vary considerably across the site, this method offers a more efficient alternative to traditional laboratory and field testing for each location. By utilizing readily available soil parameters such as w_L and w_P, field engineers can quickly estimate location-specific compaction characteristics, potentially reducing the time and cost associated with extensive laboratory testing. This approach is particularly valuable in projects where discrepancies between laboratory-derived and field-observed compaction parameters are common. Furthermore, the method allows for a more nuanced approach to compaction control by enabling the redefinition of the compaction degree based on site-specific predicted γ_{dmax} values that account for variable field CE levels. This redefinition could lead to a more accurate assessment of compaction adequacy across heterogeneous sites, potentially optimizing compaction efforts and improving quality control processes. Moreover, the ability to estimate compaction parameters at arbitrary CE levels also opens up possibilities for adaptive compaction strategies, where equipment settings and pass numbers can be adjusted in real time based on local soil conditions and predicted optimal compaction parameters (Miller and Vahedifard, 2024). A particularly significant application of this method is in achieving efficient compaction by adjusting the water content of fill material to match the w_{opt} at the field CE level. The proposed method allows for the prediction of w_{opt} corresponding to the estimated field CE, enabling more precise moisture content adjustments. This capability can lead to optimized compaction efforts and improved soil performance. However, it's important to note that to fully leverage this advantage, it is necessary to develop reliable means of evaluating the field CE level, albeit different methods are available in literature as proposed by Johnson and Sallberg, (1960). Additionally, this method could facilitate the development of more sophisticated specifications for earthwork projects, allowing for variable compaction requirements that are tailored to specific areas within a site based on their unique soil properties and the estimated field CE level. Such an approach could result in more efficient use of resources, improved overall compaction quality, and potentially reduced construction time and costs. However, implementation of this method in field practice would require careful calibration and validation against traditional methods which can be done in future research. Also, the proposed XGBoost model exposes the complex feedback loops that govern soil compaction. For instance, the model shows that increasing CE boosts γ_{dmax} by facilitating particle rearrangement and reducing void spaces at lower moisture levels. Meanwhile, the parametric study also uncovers a threshold effect, where CE becomes less effective beyond a certain point. Similarly, while soils with higher w_L require more moisture to achieve optimal compaction, their increased water-holding capacity paradoxically limits their ability to densify effectively. The model trained on a global dataset also provides insights into the categorization of soil properties to influence the compaction characteristics of the cohesive soil useful for controlling compaction in fields based on soil properties.

6. Conclusions,

The current study provides a comprehensive approach for the global modeling of compaction parameters at arbitrary CE levels using an extensive database established through testing and literature surveys. The following are the main findings of this study.

- Different ML techniques showed inconsistent performance with the current database. The optimal XGBoost model had the lowest prediction deviation for γ_{dmax} and w_{opt}, followed by the RF model. In contrast, GEP, MEP, ANN, and MLR models had higher deviations. Statistical KPIs confirmed these results, reflecting the complexity of the geotechnical data and the varying effectiveness of different modeling techniques. This suggests that no single ML method universally performs best for geotechnical databases, and tailored performance analysis may be needed for different datasets.

- The XGBoost model outperformed existing models and RF in predicting γ_{dmax} and w_{opt}, both using the established database and an independent dataset, meeting all evaluation criteria. Its superior performance highlights the effectiveness of training on a comprehensive database and XGBoost's robust framework, resulting in minimal prediction error within theoretical bounds. Predicted γ_{dmax} and w_{opt} values for $S_{\gamma dmax}$ remained consistently within 100% across different CE levels, while other models exceeded this threshold, indicating the need for careful consideration when using them.
- By combining SHAP analysis with real modeling data, the XGBoost model reveals how CE, soil plasticity and gradation influence γ_{dmax} and w_{opt}. The analysis emphasizes the primary roles of CE and w_L, with CE increasing γ_{dmax} and decreasing w_{opt}, and w_L having the opposite effect. w_P play a secondary role in prediction, with moderate impacts that align with the trends of w_L. I_P and A_F play a tertiary role, and F_{200} and Sand show a small impact, confirming their quaternary role in prediction. Meanwhile, all these parameters combined result in high prediction accuracy, helping to fill the gaps in the predictions. The model shows insignificant irregularities from these trends during parametric analysis across a wide range of input feature values. Sensitivity analysis further confirms the model's adaptability across various soil plasticity ranges and CE levels. The current modeling approach allows engineers to focus on the most influential factors and thresholds, effectively optimizing compaction for cohesive soils.

The proposed method for predicting γ_{dmax} and w_{opt} at varying CE levels offers a promising solution to enhance field compaction practices, especially in large-scale projects, by providing a more efficient alternative to traditional, labor-intensive measurement techniques. Predicting compaction parameters at arbitrary CE levels is particularly useful in the field, as it allows for more flexible and accurate adjustments to compaction efforts based on site-specific conditions. This approach, based on readily available soil properties, could improve compaction quality, optimize resource usage, and offer a more accurate assessment of compaction in heterogeneous field conditions. Meanwhile, future work should focus on developing reliable methods for estimating field CE, validating this modeling approach in the field, and expanding it to different soil types, i.e., granular soils.

Acknowledgment

NUST Risalpur is acknowledged for technical support.

Competing interests

The authors declare that they have no competing interests

Availability of code

The codes of the proposed models are available at https://github.com/ZiaurRehman16/Soil-Compaction-Model.git.

References

Agus, S.S., Schanz, T., Fredlund, D.G., 2010. Measurements of suction versus water content for bentonite-sand mixtures. Can. Geotech. J. 47, 583–594. https://doi.org/10.1139/T09-120/ASSET/IMAGES/LARGE/T09-120F9.JPEG

Akcanca, F., Aytekin, M., 2012. Effect of wetting-drying cycles on swelling behavior of lime stabilized sand-bentonite mixtures. Environ. Earth Sci. 66, 67–74. https://doi.org/10.1007/S12665-011-1207-5/FIGURES/7

Al-Hussaini, O., 2017. Volume change behaviour of some geomaterials under combined influence of freeze-thaw and wet-dry cycles:An experimental investigation.

Al-Khafaji, A.N., 1993. Estimation of soil compaction parameters by means of Atterberg limits. Q. J. Eng. Geol. 26, 359–368. https://doi.org/10.1144/GSL.QJEGH.1993.026.004.10

Aldaood, A., Bouasker, M., Al-Mukhtar, M., 2015. Soil–Water Characteristic Curve of Gypseous Soil. Geotech. Geol. Eng. 33, 123–135. https://doi.org/10.1007/S10706-014-9829-5/TABLES/3

Ali, H.F.H., Omer, B., Mohammed, A.S., Faraj, R.H., 2024. Predicting the maximum dry density and optimum moisture content from soil index properties using efficient soft computing techniques. Neural Comput. Appl. 36, 11339–11369. https://doi.org/10.1007/S00521-024-09734-7/TABLES/6

Alzubaidi, L., Chlaib, H.K., Fadhel, M.A., Chen, Y., Bai, J., Albahri, A.S., Gu, Y., 2024. Reliable deep learning framework for the ground penetrating radar data to locate the horizontal variation in levee soil compaction. Eng. Appl. Artif. Intell. 129, 107627. https://doi.org/10.1016/J.ENGAPPAI.2023.107627

Associates, G., Angel Reyes Rodríguez, M., García Cruz, H., Pérez Luis, R., Científico Natalia Pérez García, A., 2016. Compaction and mechanical properties of soils compacted in the gyratory compactor. Rev. Infraestruct. Vial 18, 20–29.

Bello, A.A., 2013. Hydraulic conductivity of three compacted reddish brown tropical soils. KSCE J. Civ. Eng. 17, 939–948. https://doi.org/10.1007/S12205-013-0155-X/METRICS

Benson, C.H., Trast, J.M., 1995. Hydraulic conductivity of thirteen compacted clays. Clays Clay Miner. 43, 669–681. https://doi.org/10.1346/CCMN.1995.0430603/METRICS

Blanco, L.M.P., 2015. Compaction quality control on site of earthworks - A comparative study.

Blotz, L.R., Benson, C.H., Boutwell, G.P., 1998. Estimating Optimum Water Content and Maximum Dry Unit Weight for Compacted Clays. J. Geotech. Geoenvironmental Eng. 124, 907–912. https://doi.org/10.1061/(ASCE)1090-0241(1998)124:9(907)

Burton, G.J., Sheng, D., Campbell, C., 2015. Bimodal pore size distribution of a high-plasticity compacted clay. https://doi.org/10.1680/geolett.14.00003 4, 88–93. https://doi.org/10.1680/GEOLETT.14.00003

Chen, C., Seo, H., 2023. Prediction of rock mass class ahead of TBM excavation face by ML and DL algorithms with Bayesian TPE optimization and SHAP feature analysis. Acta Geotech. 18, 3825–3848. https://doi.org/10.1007/S11440-022-01779-Z/TABLES/12

Clariá, J.J., Rinaldi, V.A., 2007. Shear Wave Velocity of a Compacted Clayey Silt. Geotech. Test. J. 30, 399–408. https://doi.org/10.1520/GTJ100655

Daniel, D.E. (David E., Trautwein, S.J., 1994. Hydraulic conductivity and waste contaminant transport in soil 611.

Daniel, D.E., Wu, Y.K., 1993. Compacted Clay Liners and Covers for Arid Sites. J. Geotech. Eng. 119, 223–237. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:2(223)

Delage, P., Audiguier, M., Cui, Y.J., Howat, M.D., 2011. Microstructure of a compacted silt. https://doi.org/10.1139/t96-030 33, 150–158. https://doi.org/10.1139/T96-030

Deng, N., An, Z., Hao, J., Zhang, Q., 2024. Few shot intelligent assessment method for compaction quality of earthwork considering uncertainty. Expert Syst. Appl. 256, 125000. https://doi.org/10.1016/J.ESWA.2024.125000

Dialmy, A., Rguig, M., Meliani, M., 2024. Quantification and Optimization of Compaction Energy Used in Earth Construction: Case of Static and Dynamic Compaction. Int. J. Eng. Res. Africa 68, 67–84. https://doi.org/10.4028/P-U3RDPK

Duong, T.V., Cui, Y.J., Tang, A.M., Dupla, J.C., Canou, J., Calon, N., Robinet, A., 2014. Investigating the mud pumping and interlayer creation phenomena in railway sub-structure. Eng. Geol. 171, 45–58. https://doi.org/10.1016/J.ENGGEO.2013.12.016

Farooq, K., Khalid, U., Mujtaba, H., 2016. Prediction of Compaction Characteristics of Fine-Grained Soils Using Consistency Limits. Arab. J. Sci. Eng. 41, 1319–1328. https://doi.org/10.1007/S13369-015-1918-0/METRICS

Farooq, K., ur Rehman, Z., Shahzadi, M., Mujtaba, H., Khalid, U., 2023. Optimization of Sand-Bentonite Mixture for the Stable Engineered Barriers using Desirability Optimization Methodology: A Macro-Micro-Evaluation. KSCE J. Civ. Eng. 27, 40–52. https://doi.org/10.1007/S12205-022-2088-8/METRICS

Fox, P.J., Thielmann, S.S., Stern, A.N., Athanassopoulos, C., 2014. Interface Shear Damage to a HDPE Geomembrane. I: Gravelly Compacted Clay Liner. J. Geotech. Geoenvironmental Eng. 140, 04014039. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001132/ASSET/437587EF-977F-48D7-BB28-D33AB914EE25/ASSETS/IMAGES/LARGE/FIGURE19.JPG

Günaydın, O., 2009. Estimation of soil compaction parameters by using statistical analyses and artificial neural networks. Environ. Geol. 57, 203–215. https://doi.org/10.1007/S00254-008-1300-6/FIGURES/12

Gurtug, Y., Sridharan, A., 2004. Compaction Behaviour and Prediction of its Characteristics of Fine Grained Soils with Particular Reference to Compaction Energy. Soils Found. 44, 27–36. https://doi.org/10.3208/SANDF.44.5_27

Horpibulsuk, S., Katkan, W., Apichatvullop, A., 2008. AN APPROACH FOR ASSESSMENT OF COMPACTION CURVES OF FINE GRAINED SOILS AT VARIOUS ENERGIES USING A ONE POINT TEST. SOILS Found. 48, 115–125. https://doi.org/10.3208/SANDF.48.115

Horpibulsuk, S., Katkan, W., Naramitkornburee, A., 2009. Modified Ohio's Curves: A Rapid Estimation of Compaction Curves for Coarse- and Fine-Grained Soils. Geotech. Test. J. 32, 64–75. https://doi.org/10.1520/GTJ101659

Hussain, S., 2017. Effect of Compaction Energy on Engineering Properties of Expansive Soil. Civ. Eng. J. 3, 610–616. https://doi.org/10.28991/CEJ-030988

Inci, G., Yesiller, N., Kagawa, T., 2003. Experimental Investigation of Dynamic Response of Compacted Clayey Soils. Geotech. Test. J. 26, 125–141. https://doi.org/10.1520/GTJ11328J

Ito, H., Komine, H., 2008. Dynamic compaction properties of bentonite-based materials. Eng. Geol. 98, 133–143. https://doi.org/10.1016/J.ENGGEO.2008.01.005

Jia, F., Zhang, J., Shen, J., Wu, L., Ma, S., 2024. Compaction quality assessment of road subgrades using explainable deep graph learning framework. Comput. Geotech. 176, 106795. https://doi.org/10.1016/J.COMPGEO.2024.106795

Johnson, A.W., Sallberg, J.R., 1960. Factors that influence field compaction of soils: compaction characteristics of field equipment. Highw. Res. Board Bull.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V., 2019. Machine Learning for the Geosciences: Challenges and Opportunities. IEEE Trans. Knowl. Data Eng. 31, 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006

Khalid, U., Rehman, Z. ur, 2018. Evaluation of compaction parameters of fine-grained soils using standard and modified efforts. Int. J. Geo-Engineering 9, 1–17. https://doi.org/10.1186/S40703-018-0083-1/FIGURES/13

Khalid, U., Rehman, Z. ur, Liao, C., Farooq, K., Mujtaba, H., 2019. Compressibility of Compacted Clays Mixed with a Wide Range of Bentonite for Engineered Barriers. Arab. J. Sci. Eng. 44, 5027–5042. https://doi.org/10.1007/S13369-018-03693-7/METRICS

Khatti, J., Grover, K.S., 2023. A Scientometrics Review of Soil Properties Prediction Using Soft Computing Approaches. Arch. Comput. Methods Eng. 2023 313 31, 1519–1553. https://doi.org/10.1007/S11831-023-10024-Z

Kiliç, R., Küçükali, Ö., Ulamiş, K., 2016. Stabilization of high plasticity clay with lime and gypsum (Ankara, Turkey). Bull. Eng. Geol. Environ. 75, 735–744. https://doi.org/10.1007/S10064-015-0757-2/TABLES/5

Li, J., Wang, X., Li, Jiale, Zhang, J., Ma, G., 2024. A generative adversarial learning strategy for spatial inspection of compaction quality. Adv. Eng. Informatics 62, 102791. https://doi.org/10.1016/J.AEI.2024.102791

Li, X., Liu, W., Xu, C., Liu, N., Feng, S., Zhang, X., Li, Y., Hao, J., 2024. Explainable machine learning-based prediction model for dynamic resilient modulus of subgrade soils. Transp. Geotech. 49, 101415. https://doi.org/10.1016/J.TRGEO.2024.101415

Lim, Y.Y., Miller, G.A., 2004. Wetting-Induced Compression of Compacted Oklahoma Soils. J. Geotech. Geoenvironmental Eng. 130, 1014–1023. https://doi.org/10.1061/(ASCE)1090-0241(2004)130:10(1014)

Ma, Y., Lu, W., Luan, Y. cheng, Ma, T., Wang, C. le, 2024. Research on global optimization mechanism of intelligent compaction parameters of soil subgrade based on difference method. J. Build. Eng. 82, 108381. https://doi.org/10.1016/J.JOBE.2023.108381

McRae, J.L., 1959. Index of Compaction Characteristics. ASTM Spec. Tech. Publ. STP 239, 119–127. https://doi.org/10.1520/STP39323S

Mehmood, E., Ilyas, M., Farooq, K., 2011. Effect of Initial Placement Conditions on Swelling Characteristics of Expansive Soils 2397–2403. https://doi.org/10.1061/41165(397)245

Miller, C.J., Asce, M., Yesiller, ; N, Asce, A.M., Yaldo, ; K, Merayyan, S., 2002. Impact of Soil Type and Compaction Conditions on Soil Water Characteristic. J. Geotech. Geoenvironmental Eng. 128, 733–742. https://doi.org/10.1061/(ASCE)1090-0241(2002)128:9(733)

Miller, K.C., Vahedifard, F., 2024. A New Saturation-Based Framework for Compaction Quality Control. Int. J. Geomech. 24. https://doi.org/10.1061/IJGNAI.GMENG-9084

Najdi, A., 2023. Experimental and numerical analysis of the unsaturated soil shrinkage and swelling behaviour under different compaction conditions. TDX (Tesis Dr. en Xarxa). https://doi.org/10.5821/DISSERTATION-2117-402165

Osinubi, K. ‘J., Nwaiwu, C.M., 2005. Hydraulic Conductivity of Compacted Lateritic Soil. J. Geotech. Geoenvironmental Eng. 131, 1034–1041. https://doi.org/10.1061/(ASCE)1090-0241(2005)131:8(1034)/ASSET/2C2C5F7D-E054-4882-BAC0-208D0EE0F059/ASSETS/IMAGES/8.JPG

Osinubi, K.J., Asce, M., Nwaiwu, C.M.O., Asce, A.M., 2006. Compaction Delay Effects on Properties of Lime-Treated Soil. J. Mater. Civ. Eng. 18, 250–258.

https://doi.org/10.1061/(ASCE)0899-1561(2006)18:2(250)

Osinubi, K.J., Bello, A.A., 2011. Soil-Water Characteristics Curves for Reddish Brown Tropical Soil. Electron. J. Geotech. Eng. 16, 1–25.

Othman, M.A., Benson, C.H., 2011. Effect of freeze–thaw on the hydraulic conductivity and morphology of compacted clay. https://doi.org/10.1139/t93-020 30, 236–246. https://doi.org/10.1139/T93-020

Perez N, Garnica P, Mendoza I, Reyes MA, 2013. Behavior of fine-grained soils compacted with high shear stresses Comportement des sols fins compactés avec des niveaux de cisaillement élevés, in: Proceedings of the 18 Th International Conference on Soil Mechanics and Geotechnical Engineering, Paris. pp. 407–410.

Prakash, K., Sridharan, A., Jayachandra, G. V., 2024. Compaction Characteristics of Soils: A User Friendly Approach. Int. J. Geosynth. Gr. Eng. 10, 1–7. https://doi.org/10.1007/S40891-024-00535-5/TABLES/5

Prasanna, H.S., Basavaraju, Chaitra, A.R., 2020. Characterization of Compacted Fine-Grained Soils. Lect. Notes Civ. Eng. 85, 1–19. https://doi.org/10.1007/978-981-15-6086-6_1

Ran, Y., Bai, W., Kong, L., Fan, H., Yang, X., Li, X., 2024. Prediction of soil degree of compaction based on machine learning: a case study of two fine-grained soils. Eng. Comput. (Swansea, Wales) 41, 46–67. https://doi.org/10.1108/EC-06-2023-0304/FULL/PDF

Rehman, Z. ur, Aziz, Z., Khalid, U., Ijaz, N., Rehman, S. ur, Ijaz, Z., 2024. Artificial intelligence-driven enhanced CBR modeling of sandy soils considering broad grain size variability. J. Rock Mech. Geotech. Eng. https://doi.org/10.1016/J.JRMGE.2024.05.048

Rehman, Z.U., Khalid, U., Farooq, K., Mujtaba, H., 2017. Prediction of CBR Value from Index Properties of different Soils. Tech. J. Univ. Eng. Technol. Taxila 22.

Sabat, A.K., 2015. Effect of Compaction Energy on Engineering Properties of Fly Ash – Granite Dust Stabilized Expansive Soil. Artic. Int. J. Eng. Technol.

SAPEI, A., DHALHAR, M.A., NAKANO, M., 1996. The Effects of Soil Compaction on the Physical Properties of Indonesian Latosol. J. Irrig. Eng. Rural Plan. 1996, 108–116. https://doi.org/10.11408/JIERP1982.1996.108

Sawangsuriya, A., Edil, T.B., Bosscher, P.J., 2009. Modulus-Suction-Moisture Relationship for Compacted Soils in Postcompaction State. J. Geotech. Geoenvironmental Eng. 135, 1390–1403. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000108

Sengupta, A., Mukherjee, S., Ghosh, A., 2017. Improvement of Bearing Ratio of Clayey Subgrade Using Compacted Flyash Layer. Geotech. Geol. Eng. 35, 1885–1894. https://doi.org/10.1007/S10706-017-0179-Y/FIGURES/8

Shafiee, A., 2008. Permeability of compacted granule-clay mixtures. Eng. Geol. 97, 199–208. https://doi.org/10.1016/J.ENGGEO.2008.01.002

Shelley, T.L., Daniel, D.E., 1993. Effect of gravel on hydraulic conductivity of compacted soil liners. J. Geotech. Eng. 119, 54–68. https://doi.org/10.1061/(ASCE)0733-9410(1993)119:1(54)

Shi, C., Wang, Y., 2021. Development of Subsurface Geological Cross-Section from Limited Site-Specific Boreholes and Prior Geological Knowledge Using Iterative Convolution XGBoost. J. Geotech. Geoenvironmental Eng. 147, 04021082. https://doi.org/10.1061/(ASCE)GT.1943-5606.0002583/ASSET/7DF3B390-888B-4434-

B61B-3AF0B421CDE4/ASSETS/IMAGES/LARGE/FIGURE19.JPG

Spagnoli, G., Shimobe, S., 2020. An overview on the compaction characteristics of soils by laboratory tests. Eng. Geol. 278, 105830. https://doi.org/10.1016/J.ENGGEO.2020.105830

Sridharan, A., Gurtug, Y., 2004. Swelling behaviour of compacted fine-grained soils. Eng. Geol. 72, 9–18. https://doi.org/10.1016/S0013-7952(03)00161-3

Sridharan, A., Nagaraj, H.B., 2005. Plastic limit and compaction characteristics of fine-grained soils. Gr. Improv. 9, 17–22. https://doi.org/10.1680/GRIM.9.1.17.58544

Taha, M.R., Kabir, M.H., 2005. Tropical residual soil as compacted soil liners. Environ. Geol. 47, 375–381. https://doi.org/10.1007/S00254-004-1160-7/TABLES/4

Taïbi, S., Fleureau, J.M., Hadiwardoyo, S., Kheirbek-Saoud, S., 2008. Small and large strain behaviour of an unsaturated compacted silt. Eur. J. Environ. Civ. Eng. 12, 203–228. https://doi.org/10.1080/19648189.2008.9693010

Tarantino, A., De Col, E., 2008. Compaction behaviour of clay. Geotechnique 58, 199–213. https://doi.org/10.1680/GEOT.2008.58.3.199/ASSET/IMAGES/SMALL/GEOT58-199-F25.GIF

Tatsuoka, F., Gomes Correia, A., 2018. Importance of controlling the degree of saturation in soil compaction linked to soil structure design. Transp. Geotech. 17, 3–23. https://doi.org/10.1016/J.TRGEO.2018.06.004

Teramoto, S., Ito, S., Kobayashi, T., 2024. Deep learning-based soil compaction monitoring: A proof-of-concept study. J. Terramechanics 111, 65–72. https://doi.org/10.1016/J.JTERRA.2023.10.001

Tilahun, Y., Qinghua, X., Ashongo, A.A., Han, X., 2024. Determination of Compaction Parameters of Cement-Lime Soils: Boosting-Based Ensemble Models. Int. J. Numer. Anal. Methods Geomech. https://doi.org/10.1002/NAG.3846

Tinjum, J.M., Benson, C.H., Blotz, L.R., Members, /, 1997. Soil-Water Characteristic Curves for Compacted Clays. J. Geotech. Geoenvironmental Eng. 123, 1060–1069. https://doi.org/10.1061/(ASCE)1090-0241(1997)123:11(1060)

Tripathy, S., Leong, E.C., Rahardjo, H., 2005. Suction of compacted residual soils. Unsaturated Soils Exp. Stud. 111–122. https://doi.org/10.1007/3-540-26736-0_9

ur Rehman, Z., Khalid, U., Ijaz, N., Mujtaba, H., Haider, A., Farooq, K., Ijaz, Z., 2022. Machine learning-based intelligent modeling of hydraulic conductivity of sandy soils considering a wide range of grain sizes. Eng. Geol. 311, 106899.

USDA, 2000. Basic Principles of Compaction for Minimal Permeability. 210-VI WA State ENG Tech Note No. WA-4

Vassallo, R., Mancuso, C., Vinale, F., 2007. Effects of net stress and suction history on the small strain stiffness of a compacted clayey silt. Can. Geotech. J. 44, 447–462. https://doi.org/10.1139/T06-129

Vega, A., McCartney, J.S., 2015. Cyclic heating effects on thermal volume change of silt. Environ. Geotech. 2, 257–268. https://doi.org/10.1680/ENVGEO.13.00022

Wang, H.L., Cui, Y.J., Lamas-Lopez, F., Dupla, J.C., Canou, J., Calon, N., Saussine, G., Aimedieu, P., Chen, R.P., 2017. Effects of inclusion contents on resilient modulus and damping ratio of unsaturated track-bed materials. https://doi.org/10.1139/cgj-2016-0673 54, 1672–1681. https://doi.org/10.1139/CGJ-2016-0673

Wang, H.L., Yin, Z.Y., 2020. High performance prediction of soil compaction parameters using multi expression programming. Eng. Geol. 276, 105758. https://doi.org/10.1016/J.ENGGEO.2020.105758

Wang, L., Li, X., Tsang, D.C.W., Jin, F., Hou, D., 2020. Green remediation of Cd and Hg contaminated soil using humic acid modified montmorillonite: Immobilization performance under accelerated ageing conditions. J. Hazard. Mater. 387, 122005. https://doi.org/10.1016/J.JHAZMAT.2019.122005

Wang, X., Dong, X., Li, J., Zhang, Z., Zhang, J., Ma, G., 2023. Developing an advanced ANN-based approach to estimate compaction characteristics of highway subgrade. Adv. Eng. Informatics 56, 102023. https://doi.org/10.1016/J.AEI.2023.102023

Woon-Hyung Kim, B., Daniel, D.E., 1992. Effects of Freezing on Hydraulic Conductivity of Compacted Clay. J. Geotech. Eng. 118, 1083–1097. https://doi.org/10.1061/(ASCE)0733-9410(1992)118:7(1083)

Yang, L., Ren, Q., Zheng, K., Jiao, Z., Ruan, X., Wang, Y., 2022. Migration of heavy metals in the soil-grape system and potential health risk assessment. Sci. Total Environ. 806, 150646. https://doi.org/10.1016/J.SCITOTENV.2021.150646

Yilmaz, Y., Kheirjouy, A.B., Akgungor, A.P., 2016. Investigation of the Effect of Different Saturation Methods on the Undrained Shear Strength of a Clayey Soil Compacted with Standard and Modified Proctor Energies. Period. Polytech. Civ. Eng. 60, 323–329. https://doi.org/10.3311/PPCI.8891

Yogeshraj Urs, C., Prasanna, H.S., 2023. Parametric Study on Compaction Characteristics of Clay Sand Mixtures. Lect. Notes Civ. Eng. 296, 141–152. https://doi.org/10.1007/978-981-19-6513-5_12

Yusoff, S.A.N.M., Bakar, I., Wijeyesekera, D.C., Zainorabidin, A., Azmi, M., Ramli, H., 2017. The effects of different compaction energy on geotechnical properties of kaolin and laterite. AIP Conf. Proc. 1875, 30009. https://doi.org/10.1063/1.4998380/755978

Zhang, P., Yin, Z.Y., Jin, Y.F., 2021. State-of-the-Art Review of Machine Learning Applications in Constitutive Modeling of Soils. Arch. Comput. Methods Eng. 2021 285 28, 3661–3686. https://doi.org/10.1007/S11831-020-09524-Z

Zhang, X., Mavroulidou, M., Gunn, M.J., 2015. Mechanical properties and behaviour of a partially saturated lime-treated, high plasticity clay. Eng. Geol. 193, 320–336. https://doi.org/10.1016/J.ENGGEO.2015.05.007

Zhu, Y., Zhang, W., Chen, F., Ma, T., Ma, Y., Fang, Z., 2024. A state-of-the-art review of intelligent compaction measurement values (ICMVs) for subgrade and pavement: Advances and challenges. Constr. Build. Mater. 438, 137174. https://doi.org/10.1016/J.CONBUILDMAT.2024.137174

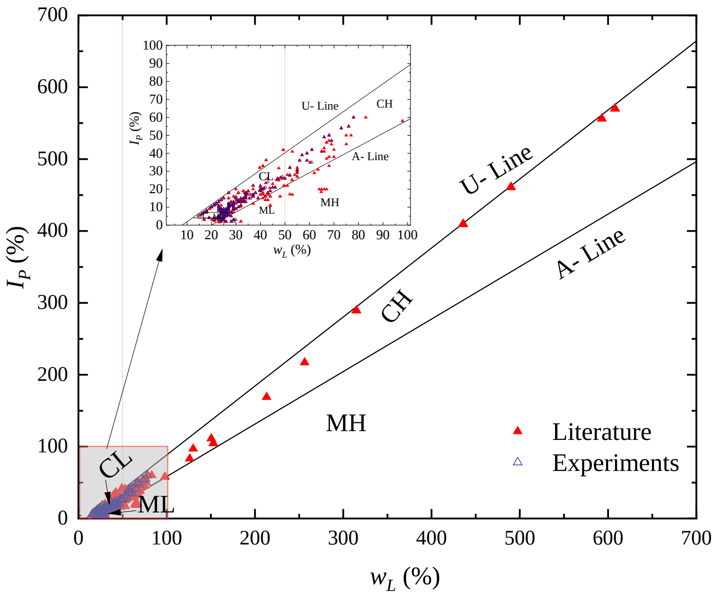


Fig. 1: Soil classification data orientation on A-line plasticity chart

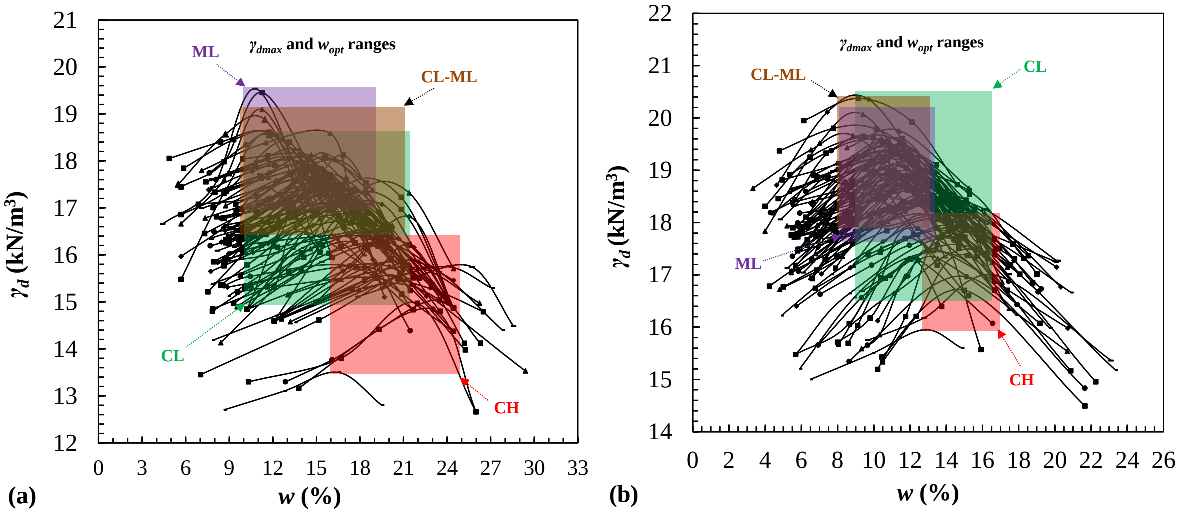
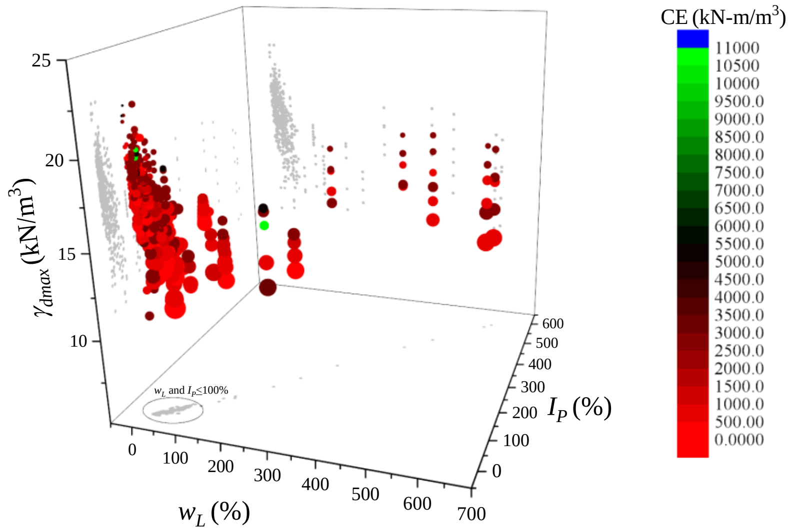


Fig. 2: Relationship between γ_d and w of tested samples (a) standard compaction tests; (b) modified compaction tests



Symbol size proportional to w_{opt} (%): scaling factor 0.5

Fig. 3: Modeling data orientation

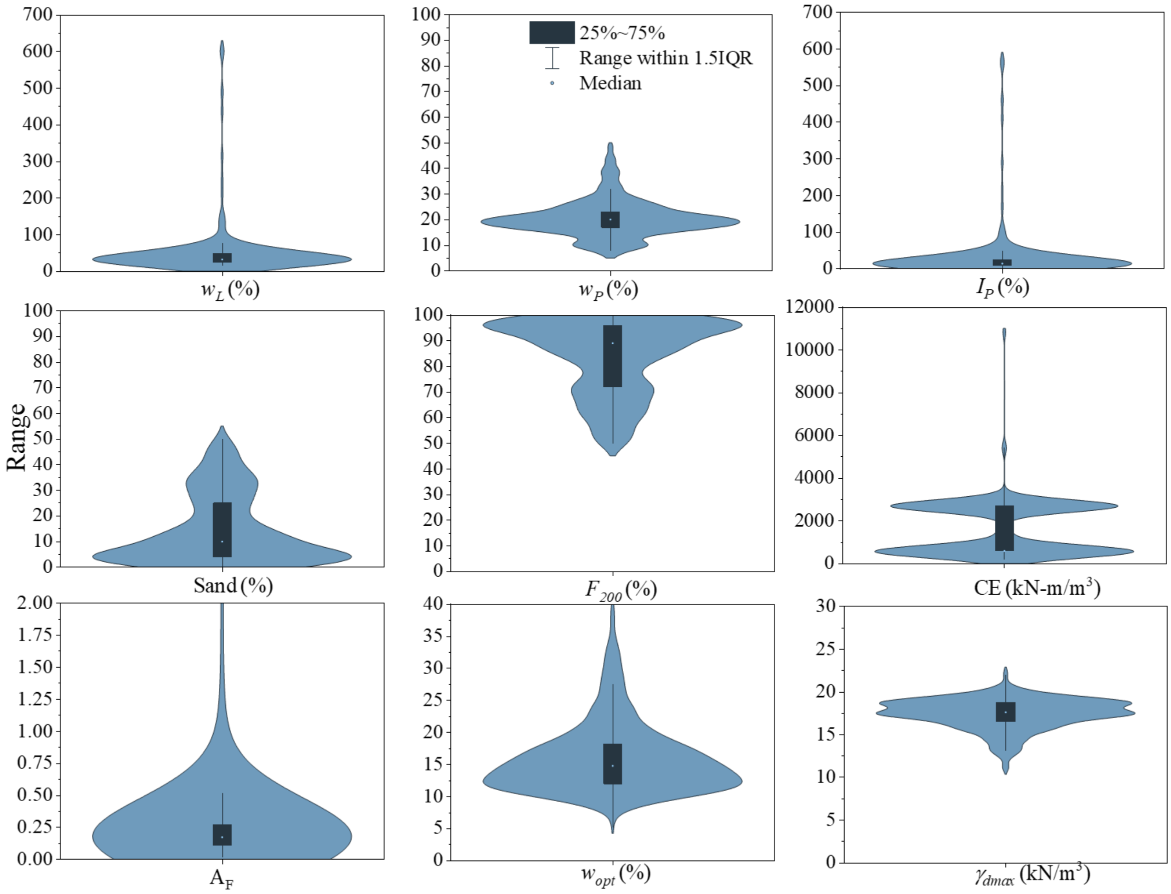


Fig. 4: Violin plots of the established database for modeling

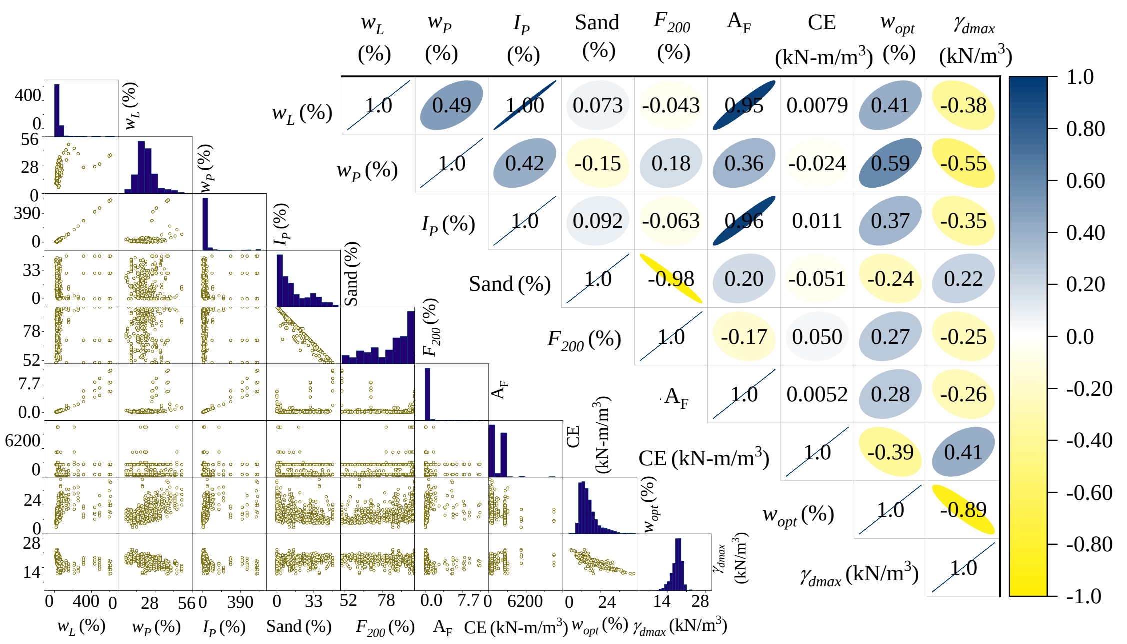


Fig. 5: Correlational and Pearsons correlation coefficient matrices of input and output parameters

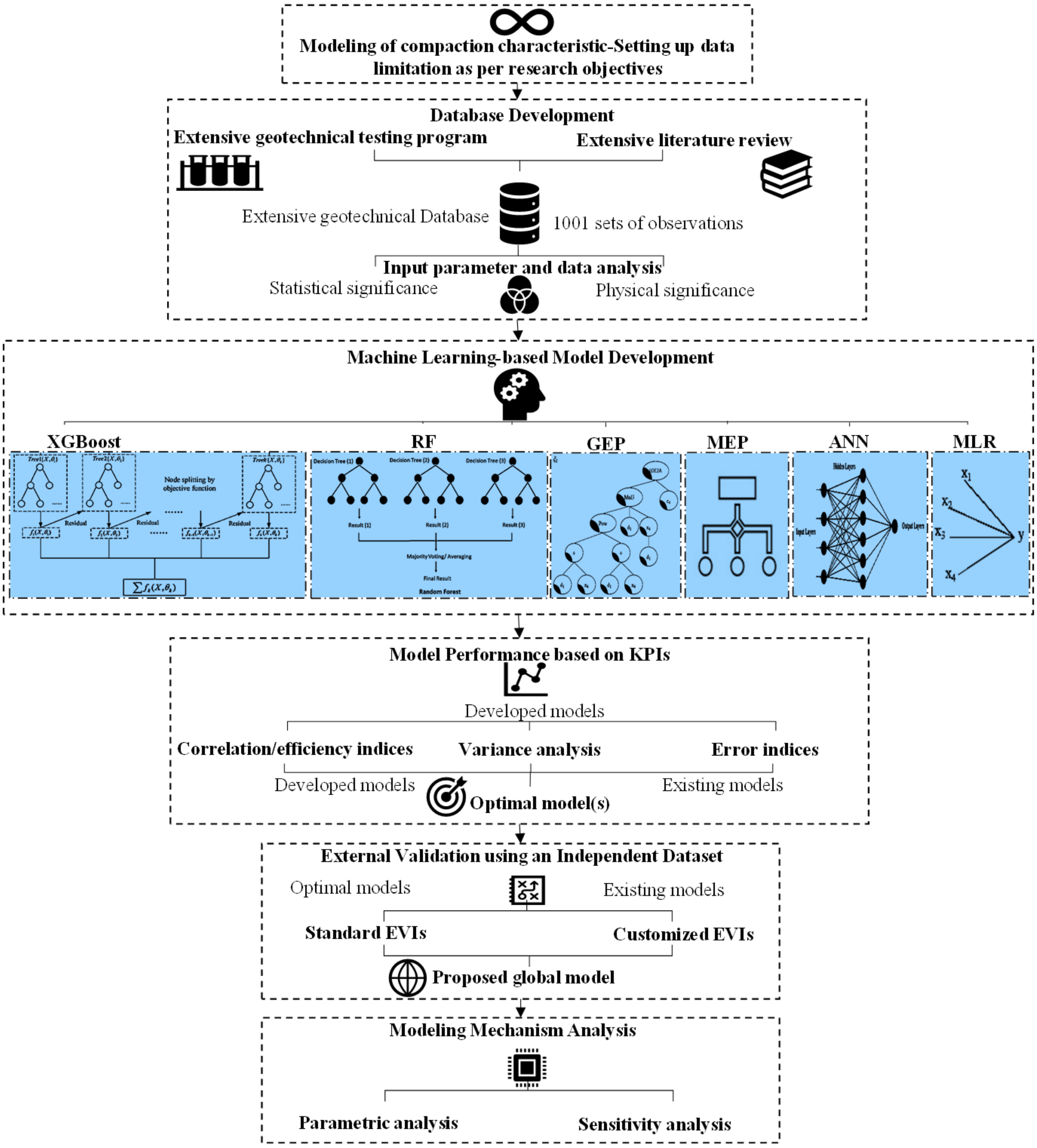


Fig. 6: Methodology of the current study

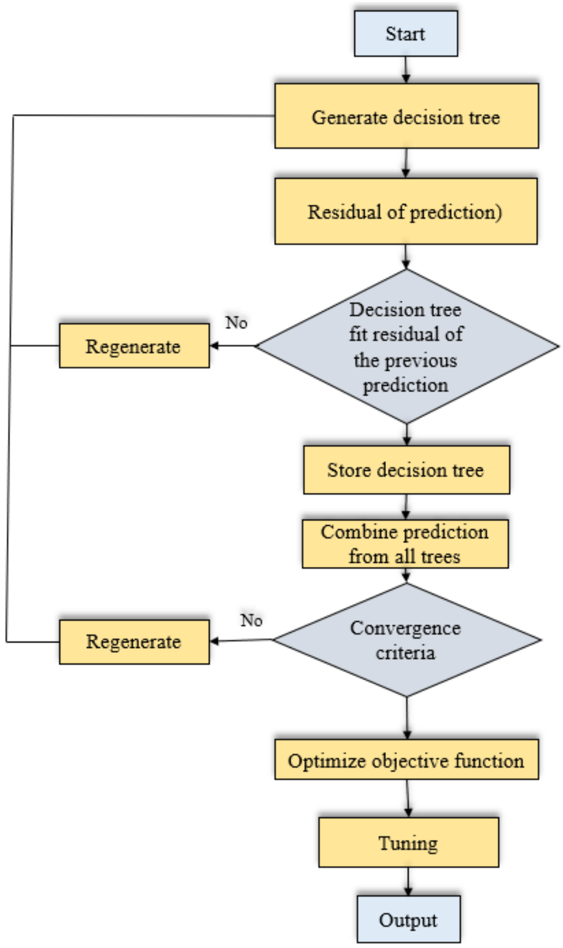


Fig. 7: XGBoost algorithm architecture

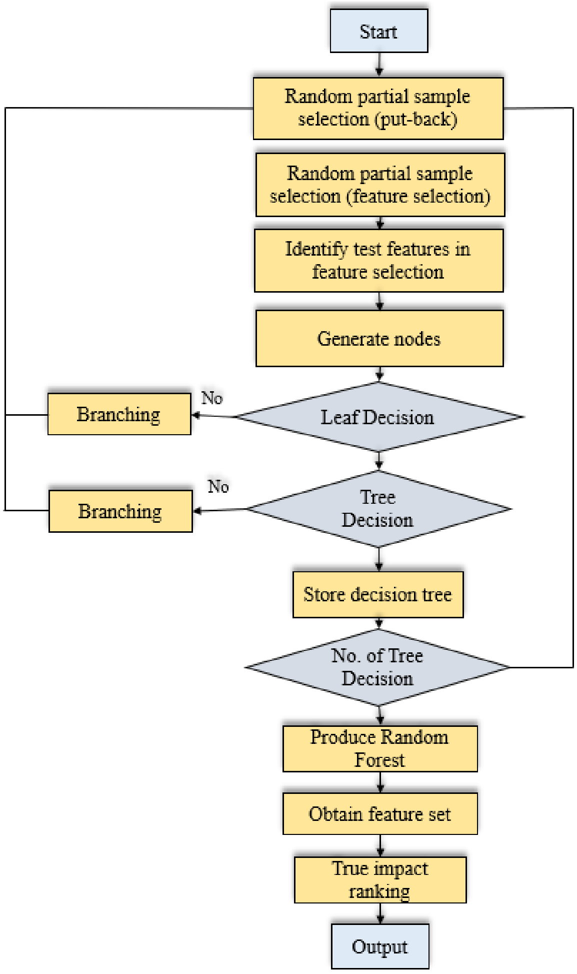


Fig. 8: RF algorithm architecture

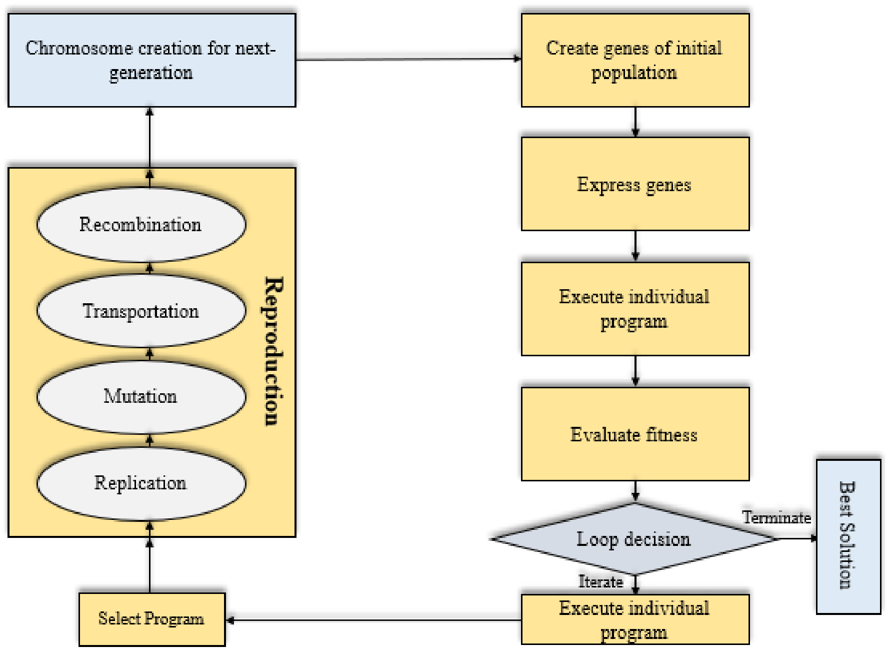


Fig. 9: GEP algorithm architecture

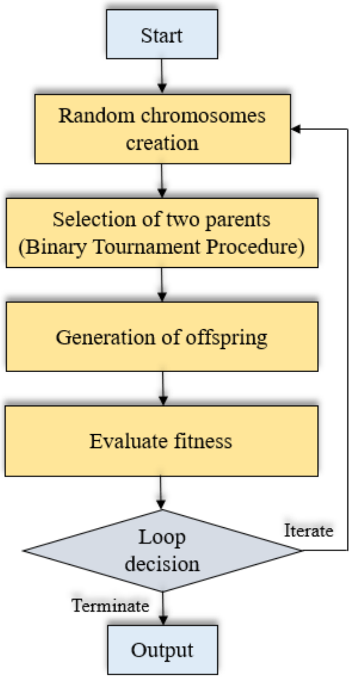


Fig. 10: MEP algorithm architecture

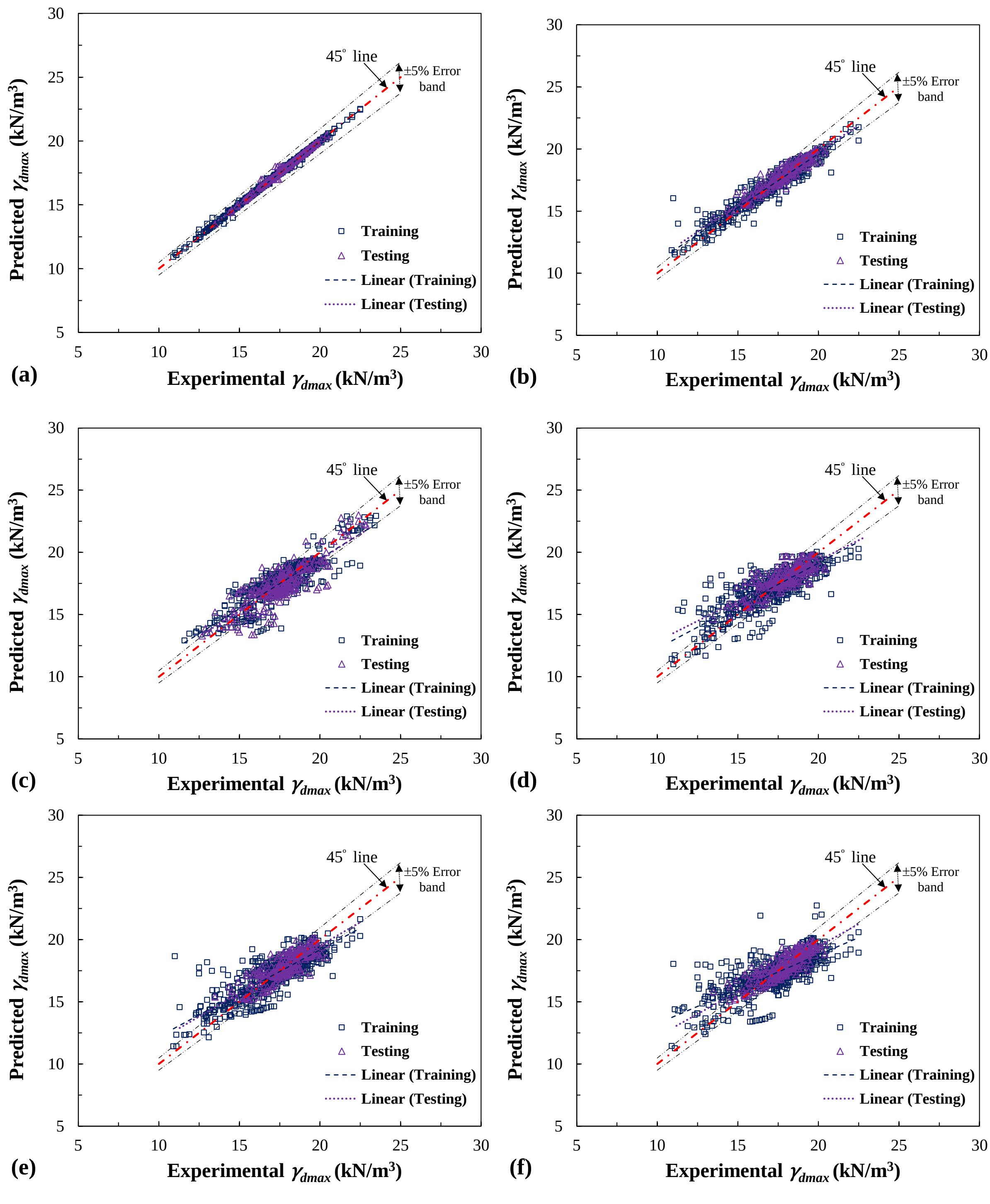


Fig. 11: γ_{dmax} model training and testing (a) XGBoost; (b) RF; (c) GEP; (d) MEP; (e) ANN; (f) MLR

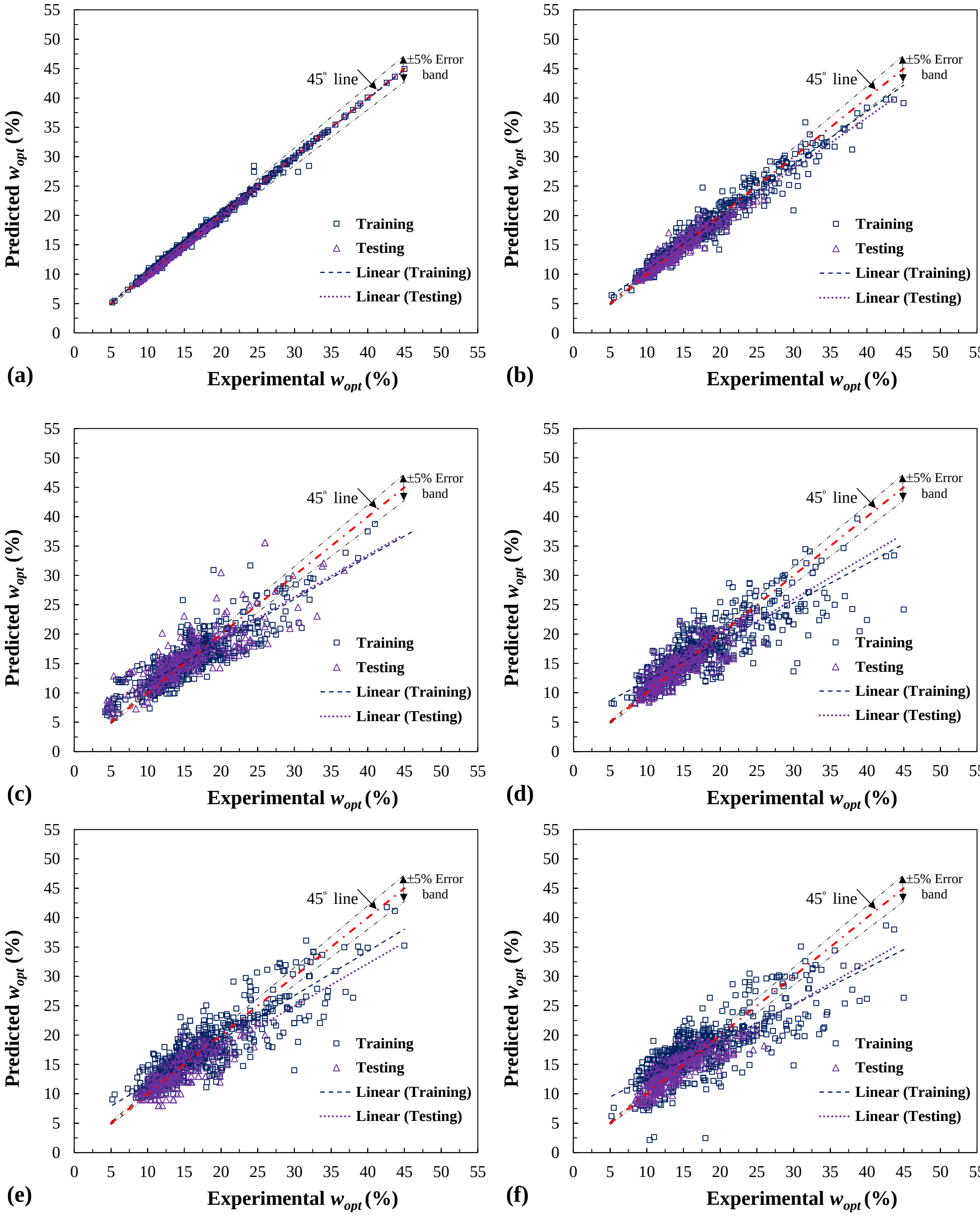


Fig. 12: w_{opt} model training and testing (a) XGBoost; (b) RF; (c) GEP; (d) MEP; (e) ANN; (f) MLR

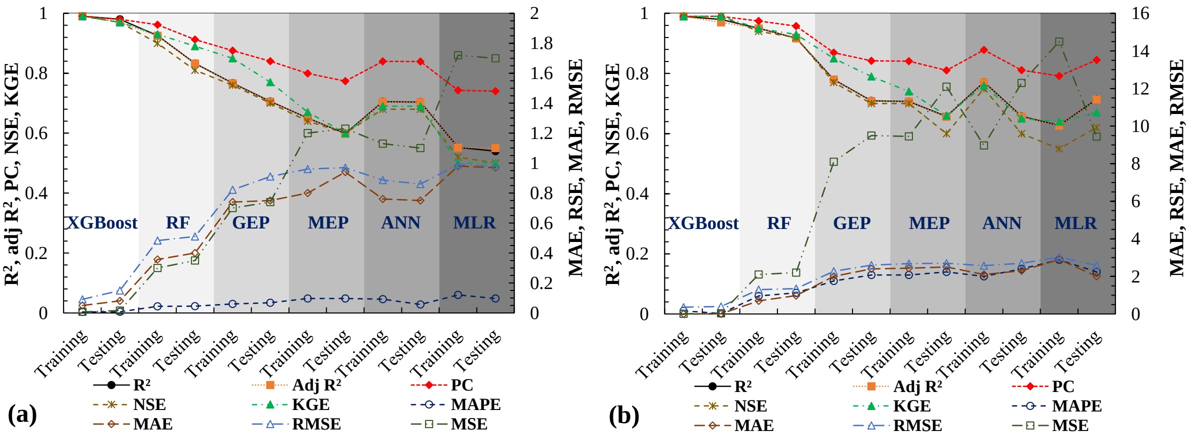


Fig. 13: Performance comparison of developed models (a) for γ_{dmax}; (b) for w_{opt}

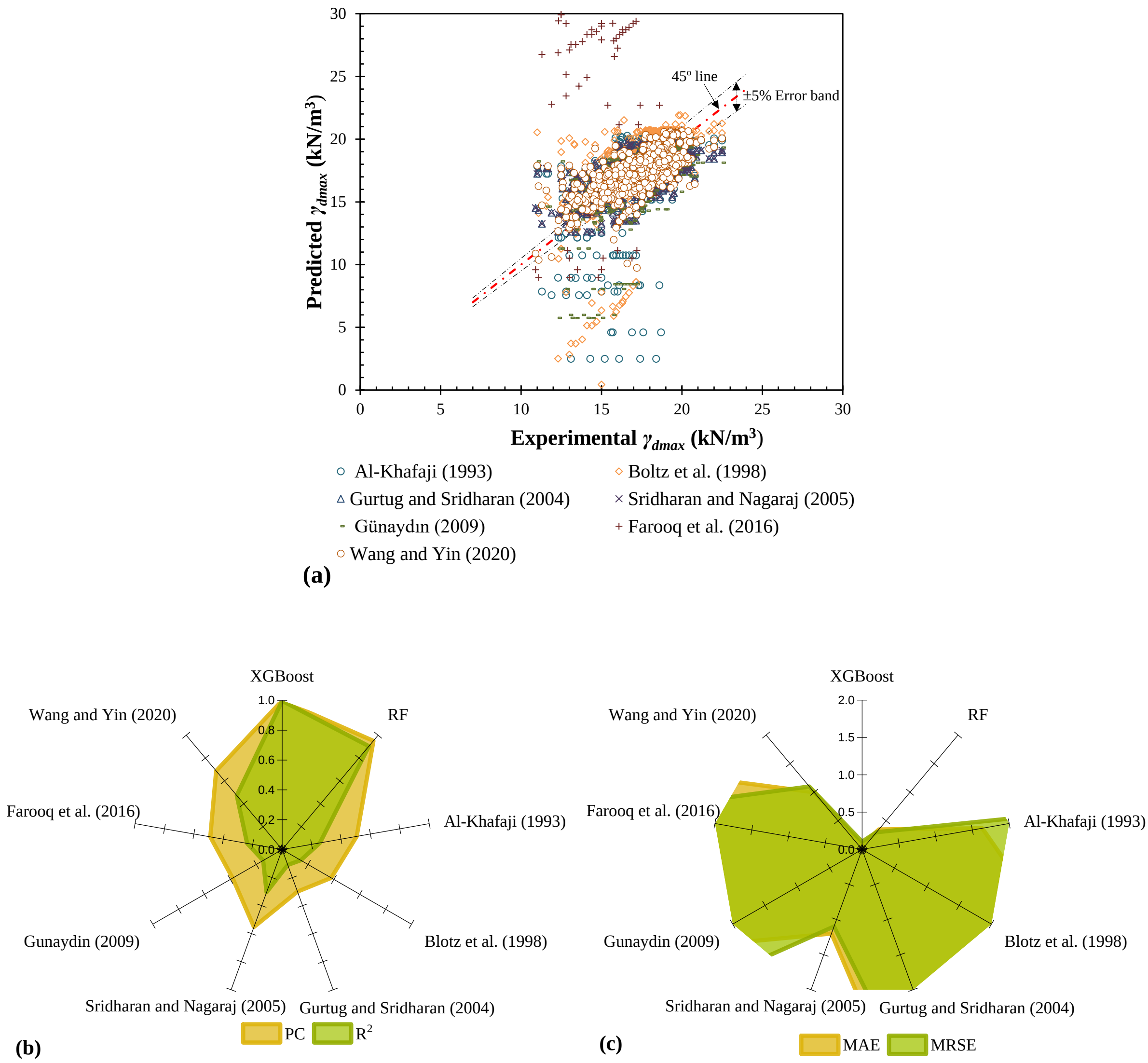


Fig. 14: Performance comparison of shortlisted models with existing models for γ_{dmax} (a) scatter plot; (b) correlation indices; (c) error indices

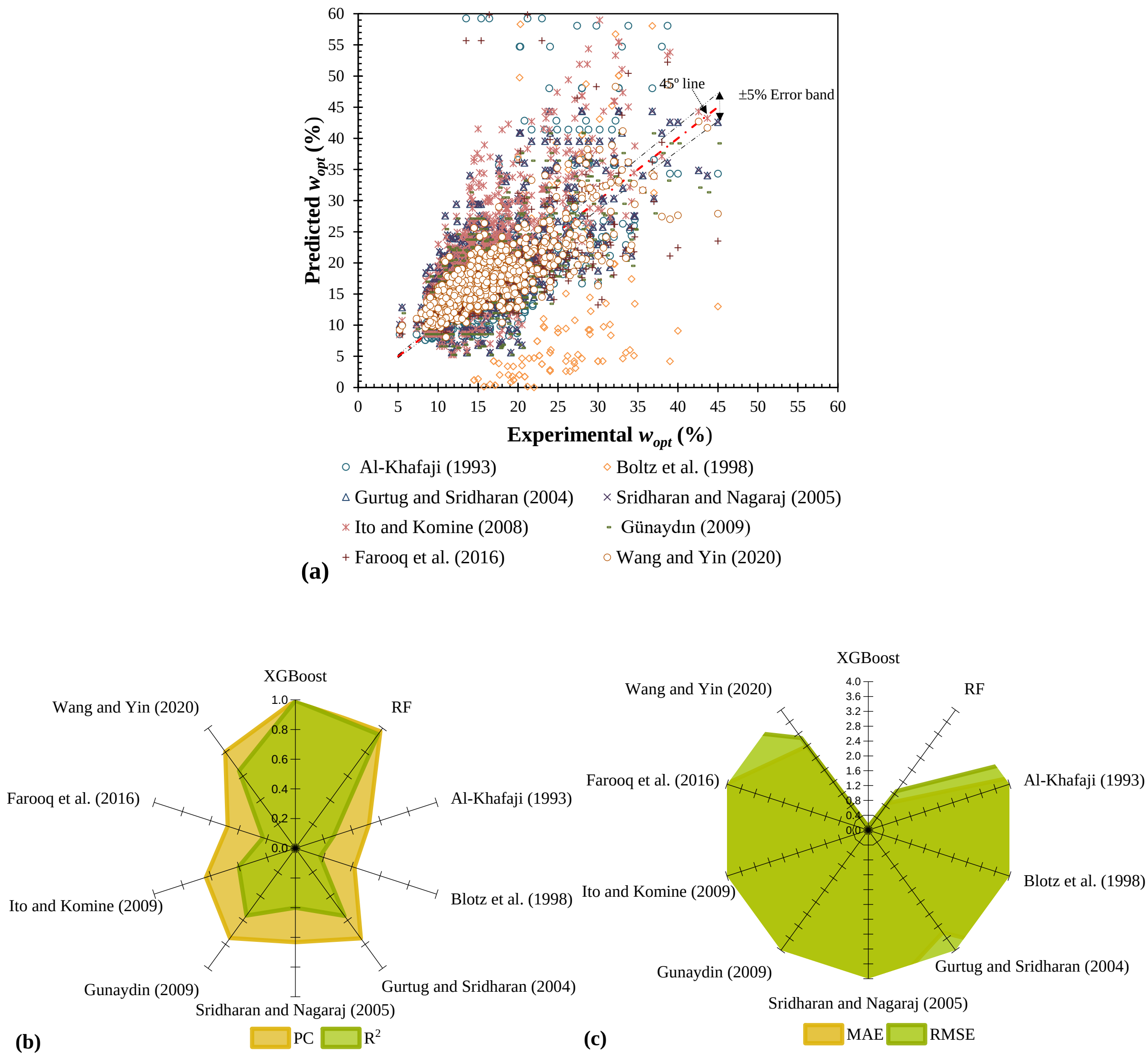


Fig. 15: Performance comparison of shortlisted models with existing models for w_{opt} (a) scatter plot; (b) correlation indices; (c) error indices

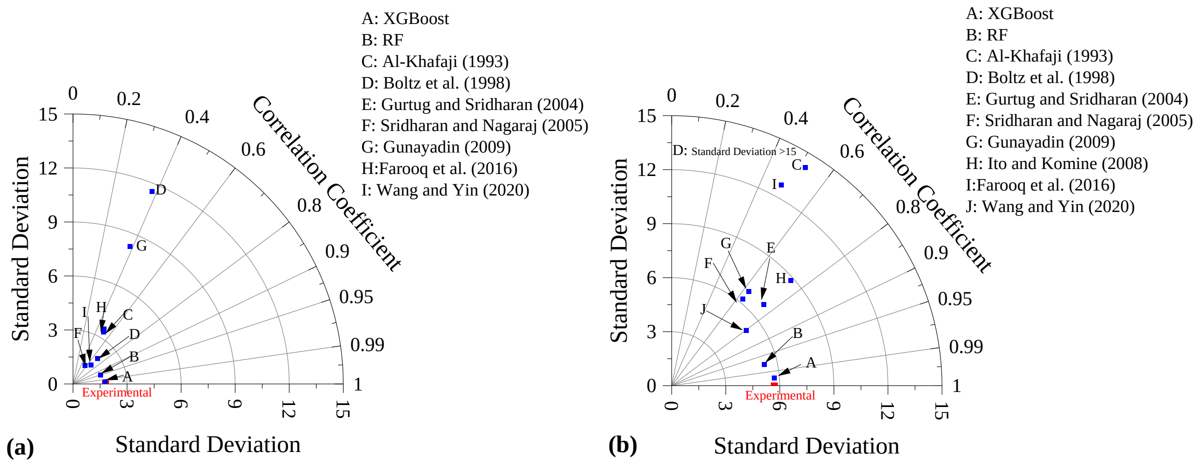


Fig. 16: Performance comparison of shortlisted models with the existing models through the Taylor diagram a) for γ_{dmax}; (b) for w_{opt}

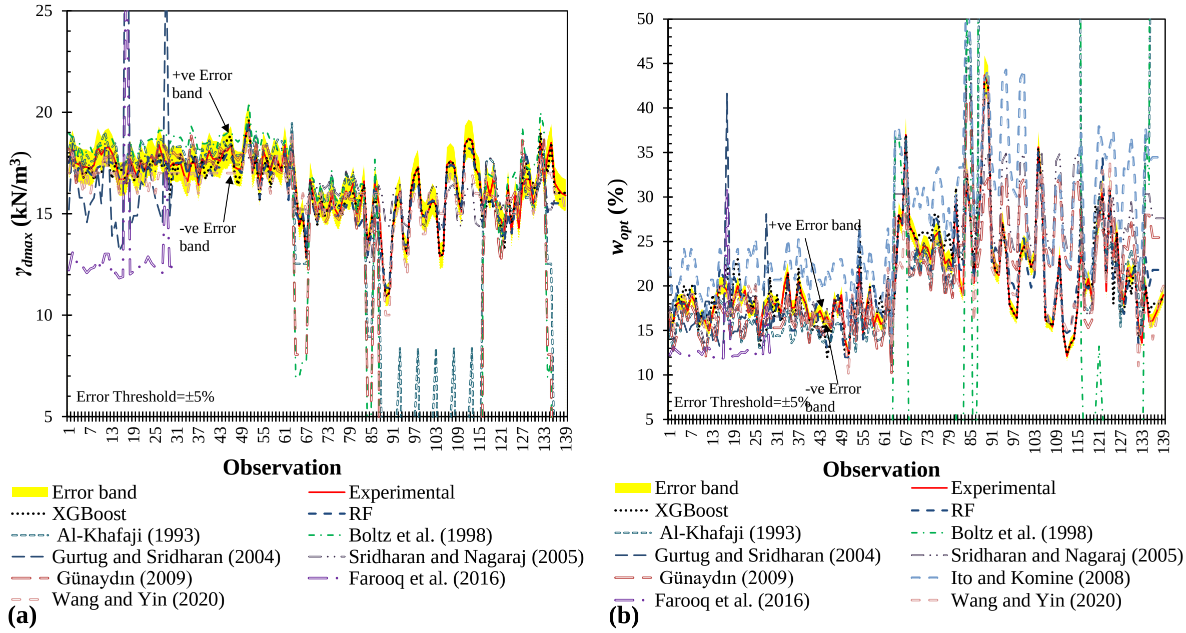


Fig. 17: External validation of shortlised models and comparison with existing models (a) for γ_{dmax}; (b) for w_{opt}

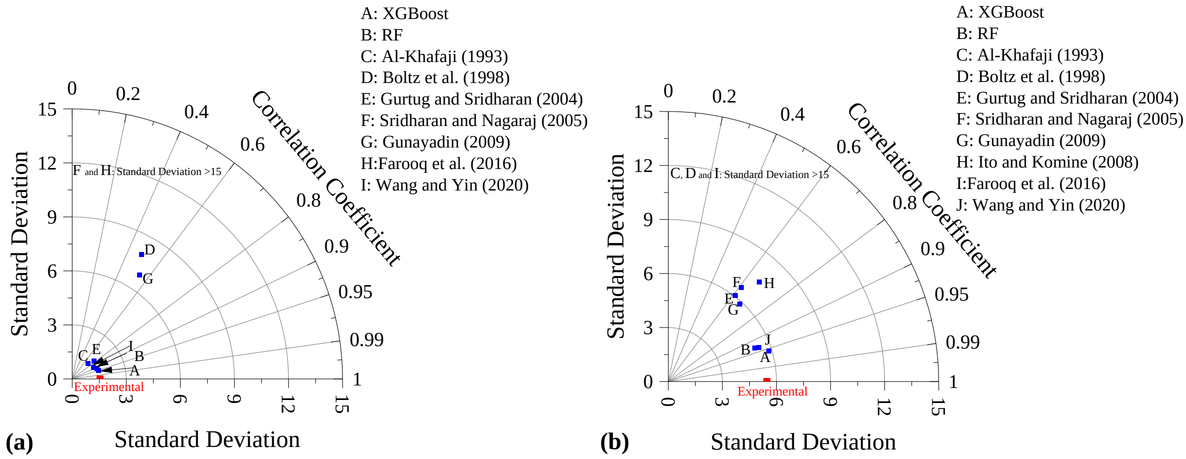


Fig. 18: External validation of shortlisted models and comparison with existing models through the Taylor diagram (a) for γ_{dmax}; (b) for w_{opt}

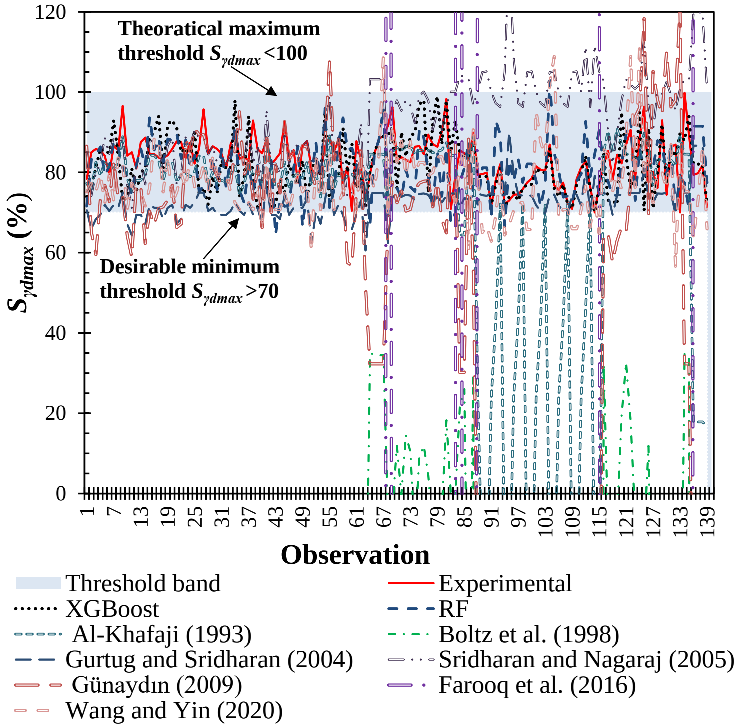


Fig. 19: External theoretical validation of shortlisted models and comparison with existing models through $S_{\gamma dmax}$ analysis

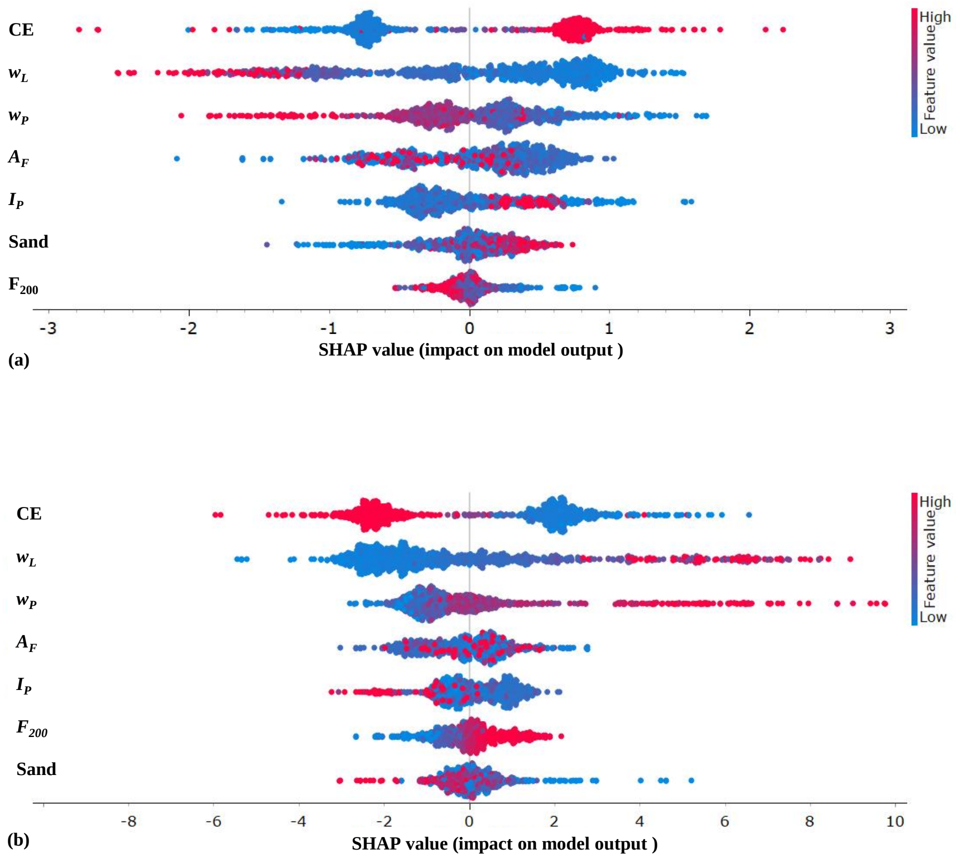


Fig. 20: XGBoost model explanation with SHAP analysis (a) for γ_{dmax}; (b) for w_{opt}

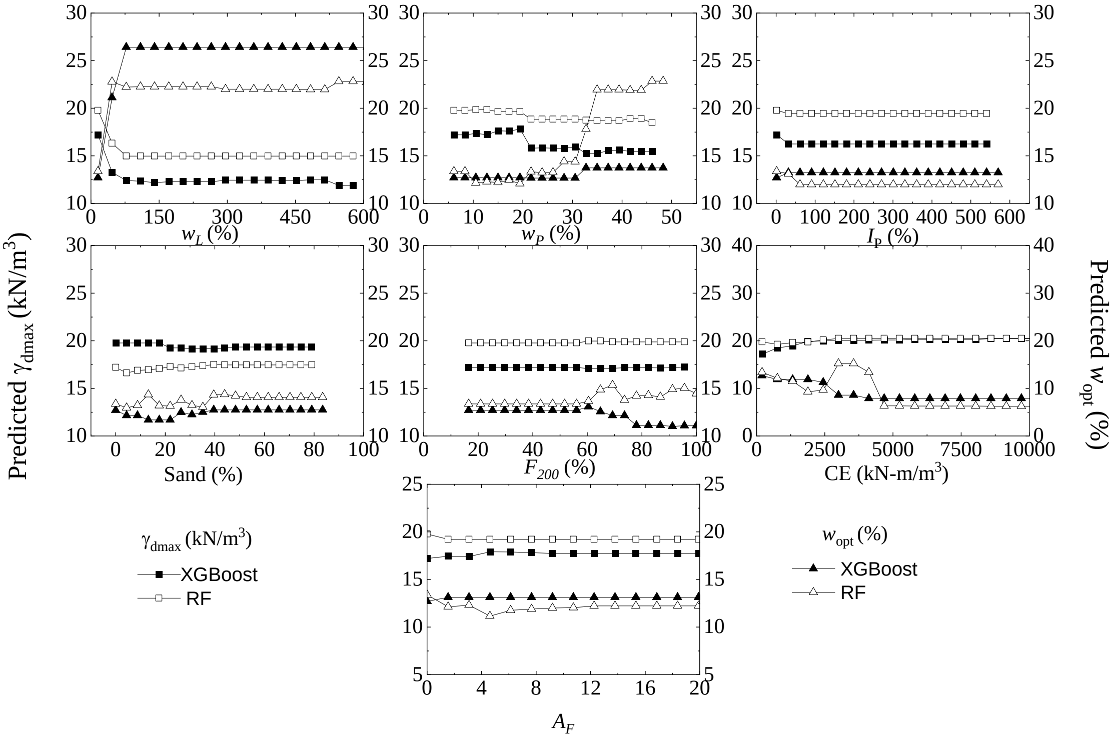


Fig. 21: Parametric analysis of XGBoost model and comparison with RF

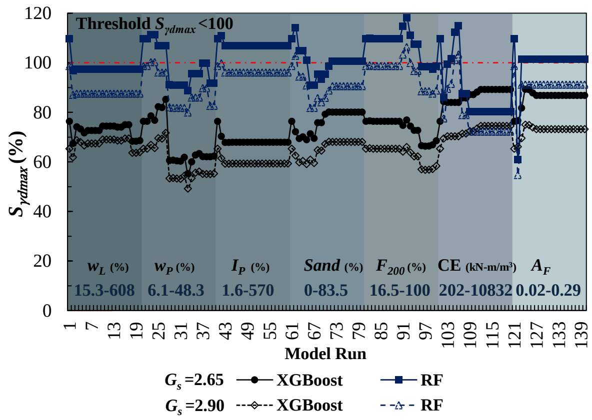


Fig. 22: Parametric analysis of XGBoost model and comparison with RF to analyze theoretical validation using $S_{\gamma dmax}$

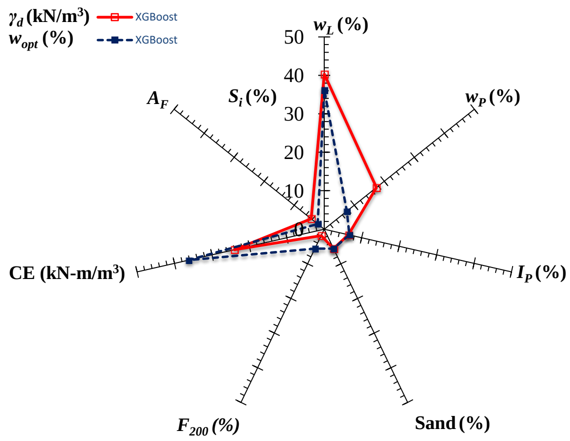


Fig. 23: Sensitivity analysis of input parameters for proposed models

Table 1: Overview of the dataset used for modeling and validation

		w_L	w_P	I_P	Sand	F_{200}	CE	A_F	w_{opt}	γ_d	$S_{\gamma dmax}$
		%	%	%	%	%	kN-m/m^3	-	%	kN/m^3	%
Model Training and Testing Data (Testing and literature)	Maximum	608.0	48.3	570.1	50.0	100.0	10832.1	11.4	45.0	22.5	97.1
	Minimum	15.0	6.1	1.6	0.0	50.0	202.0	0.001	5.2	10.9	50.2
	Count	1001	1001	1001	1001	1001	1001	1001	1001	1001	1001
	Mean	53.5	20.5	32.9	14.8	83.9	1614.5	0.4	16.1	17.5	76.1
	Median	32.0	20.0	14.0	10.0	89.0	594.0	0.2	14.8	17.6	75.0
	Standard Deviation	82.2	7.2	78.9	13.6	14.4	1280.6	1.2	5.7	1.8	8.5
	Variance	6758.1	52.2	6227.8	184.9	206.4	1639813.7	1.5	32.4	3.2	72.0
	Range	592.7	42.2	568.5	50.0	50.0	10630.1	11.4	39.8	11.6	49.5
	Excess kurtosis	28.6	2.0	30.2	-0.4	-0.6	11.6	43.4	3.2	0.9	0.0
	Skewness	5.2	1.0	5.4	0.9	-0.8	2.0	6.3	1.6	-0.8	0.1
External Validation Data (Literature)	Maximum	550	48.3	480	83.5	100.0	2693.0	29.1	43.7	37	98.9
	Minimum	23.0	12.0	3.4	0.0	16.5	551.0	0.03	12.1	12.3	70.0
	Count	139	139	139	139	139	139	139	139	139	139
	Mean	153.1	24.6	128.5	19.7	80.0	882.2	2.2	20.3	16.4	83.2
	Median	49.4	23.0	26.0	13.0	87.0	593.0	0.3	18.6	16.8	83.9
	Standard Deviation	196.8	7.2	192.2	19.7	20.0	726.4	4.2	5.4	2.0	6.0
	Variance	38733.3	51.4	36945.8	388.6	399.3	527597.7	17.7	28.9	4.0	35.7
	Range	527.0	36.3	476.7	83.5	83.5	2100	29.0	31.6	24.7	29.9
	Excess kurtosis	0.3	0.3	0.3	0.8	0.6	2.9	13.4	4.3	21.1	0.3
	Skewness	1.4	0.8	1.4	1.2	-1.2	2.2	3.1	1.8	-3.4	0.0

Table 2: Evaluation of developed models based on KPIs

Modeling	KPI	XGBoost		RF		GEP		MEP		ANN		MLR	
		Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
γ_{dmax} (kN/m^3)	R^2	0.99	0.99	0.93	0.83	0.77	0.71	0.65	0.60	0.70	0.70	0.55	0.54
	Adj R^2	0.99	0.97	0.93	0.83	0.77	0.70	0.65	0.60	0.70	0.70	0.55	0.55
	MSE	0.01	0.02	0.30	0.35	0.70	0.74	1.20	1.23	1.13	1.10	1.72	1.70
	RMSE	0.091	0.15	0.483	0.510	0.822	0.910	0.961	0.970	0.887	0.860	0.990	0.980
	NRMSE	0.0001	0.0003	0.0007	0.0016	0.0012	0.0010	0.0014	0.0031	0.0013	0.0027	0.0015	0.0031
	MAE	0.05	0.08	0.357	0.4	0.74	0.75	0.8	0.94	0.76	0.75	0.98	0.97
	MAPE	0.003	0.04	0.022	0.027	0.03	0.034	0.048	0.048	0.046	0.028	0.06	0.048
	NSE	0.99	0.97	0.9	0.81	0.76	0.7	0.64	0.6	0.68	0.68	0.52	0.5
	KGE	0.99	0.97	0.93	0.89	0.85	0.77	0.67	0.6	0.69	0.69	0.5	0.5
	PC	0.99	0.98	0.96	0.91	0.88	0.84	0.80	0.77	0.84	0.84	0.74	0.74
	NR	2.4	0.5	12.7	7.9	18.9	14.8	25.2	10.6	23.2	10.4	25.6	40.5
	RSS	5.6	0.3	160.4	63.2	359.1	219.5	634.5	113.0	540.2	109.0	655.3	1636.7
	F-value	316697	284159	8499	1538	1749	635	1213	461	1639	734	847	711
	Prob>F	0	0	0	0	0	0	0	0	0	0	0	0
w_{opt} (%)	R^2	0.98	0.98	0.95	0.92	0.78	0.71	0.71	0.66	0.77	0.66	0.63	0.71
	Adj R^2	0.98	0.97	0.95	0.92	0.78	0.71	0.71	0.66	0.77	0.66	0.63	0.71
	MSE	0.01	0.04	2.10	2.20	8.10	9.50	9.45	12.10	8.97	12.30	14.50	9.44
	RMSE	0.36	0.4	1.29	1.35	2.27	2.60	2.69	2.70	2.57	2.70	3.03	2.58
	NRMSE	0.0005	0.0006	0.0019	0.0043	0.0033	0.0083	0.0039	0.0087	0.0037	0.0087	0.0044	0.0083
	MAE	0.002	0.01	0.70	0.98	2.00	2.40	2.44	2.50	2.10	2.30	2.90	2.00
	MAPE	0.01	0.002	0.06	0.07	0.11	0.13	0.13	0.14	0.125	0.15	0.18	0.14
	NSE	0.99	0.97	0.94	0.92	0.77	0.7	0.7	0.6	0.75	0.6	0.55	0.62
	KGE	0.99	0.97	0.95	0.93	0.85	0.79	0.74	0.66	0.76	0.65	0.64	0.67
	PC	0.99	0.98	0.97	0.96	0.87	0.84	0.84	0.81	0.88	0.81	0.79	0.84
	NR	9.42	0.95	33.90	15.57	53.73	40.11	70.42	31.80	67.38	31.11	79.41	67.33
	RSS	88.77	0.90	1149.32	242.56	2887.06	1608.69	4959.37	1011.00	4540.05	967.60	6305.85	4533.24
	F-value	205364	1218313	13080	3426	1731	580	1657	590	2317	593	1156	1697
	Prob>F	0	0	0	0	0	0	0	0	0	0	0	0

Table 3: External validation analysis of proposed and existing models for γ_{dmax} (kN/m^3)

Dataset	Statistical Parameters								
	R^2	R_m	R_0^2	$R_0^{'2}$	k	$k^{'}$	MAE	$NAPE_{ol}$	$S_{\gamma dmax(ol)}$
Desirable range	>0.9	>0.5	around 1	around 1	0.85-1.15	0.85-1.15	<3	$\leq$5	0
A	0.91	0.65	0.99	0.99	0.99	0.84	0.30	4.3	0.0
B	0.84	0.51	0.99	0.99	0.98	0.80	0.40	10.1	0.7
C	0.29	0.06	0.90	0.90	0.82	0.10	3.86	37.4	8.6
D	0.21	0.04	0.90	0.90	0.20	2.32	15.54	56.7	84.9
E	0.59	0.26	0.94	0.94	0.94	0.85	1.23	51.1	0.0
F	0.51	0.19	0.93	0.93	1.00	0.69	0.81	28.1	30.2
G	0.21	0.04	0.90	0.90	0.39	1.53	11.11	35.6	25.8
H	0.23	0.04	0.90	0.90	1.71	1.17	12.57	94.9	100
I	0.77	0.42	0.98	0.98	0.98	0.93	0.70	25.9	3.6

A=XGBoost; B= RF; C= Al-Khafaji (1993); D= Boltz et al. (1998); E=Gurtug and Sridharan (2004); F= Sridharan and Nagaraj (2005); G= Günaydin (2009); H: Farooq et al. (2016); I: Wang and Yin (2020)

Table 4: External validation analysis of proposed and existing models for w_{opt} (%)

Dataset	Statistical Parameters								
	R^2	R_m	R_0^2	$R_0'^2$	k	k'	MAE	$NAPE_{ol}$	$S_{\gamma dmax(ol)}$
Desirable range	>0.9	>0.5	around 1	around 1	0.85-1.15	0.85-1.15	<3	≤5	0
A	0.92	0.70	0.99	0.99	1.00	0.85	1.2	3.1	0.0
B	0.88	0.60	0.99	0.99	1.00	0.87	1.5	5.0	0.7
C	0.11	0.01	0.9	0.9	1.75	8.88	18.0	44.0	8.6
D	0.08	0.01	0.8	0.8	3.03	117.72	71.4	97.3	84.9
E	0.45	0.15	0.9	0.9	1.01	1.81	3.4	31.6	0.0
F	0.37	0.10	0.9	0.9	1.09	1.95	4.0	30.9	30.2
G	0.37	0.10	0.9	0.9	1.00	2.51	3.9	35.9	25.8
H	0.45	0.15	0.9	0.9	1.30	0.33	6.8	58.7	NA
I	0.08	0.01	0.9	0.9	1.68	5.56	16.5	30.9	100
J	0.86	0.50	0.98	0.98	0.97	0.69	1.6	6.0	3.6

A=XGBoost; B= RF; C= Al-Khafaji (1993); D= Boltz et al. (1998); E=Gurtug and Sridharan (2004); F= Sridharan and Nagaraj (2005); G= Günaydin (2009); H: Ito and Komine (2020); I: Farooq et al. (2016); J: Wang and Yin (2020)