Enhancing Ultimate Bearing Capacity Assessment of Rock Foundations using a Hybrid Decision Tree Approach

Mei Guo, Ren-an Jiang
College of Road and Bridge Engineering, Jilin Communications Polytechnic, Changchun Jilin, 130012, China

Abstract—Accurately estimating the ultimate bearing capacity of piles embedded in rock is of paramount importance in the domains of civil engineering, construction, and foundation design. This research introduces an innovative solution to tackle this issue, leveraging a fusion of the Decision Tree method with two state-of-the-art optimization algorithms: the Zebra Optimization Algorithm and the Coronavirus Herd Immunity Optimizer. The research approach encompassed the creation of a hybridized model, unifying the DT with the Zebra Optimization Algorithm and Coronavirus Herd Immunity Optimizer. The primary objective was to augment the precision of the ultimate bearing capacity of prediction for piles embedded in rock. This hybridization strategy harnesses the capabilities of DT along with the two pioneering optimizers to address the inherent uncertainty stemming from diverse factors impacting bearing capacity. The Zebra Optimization Algorithm and Coronavirus Herd Immunity Optimizer showcased their efficacy in refining the base model, leading to substantial enhancements in predictive performance.

I. INTRODUCTION

A. Background

Pile foundations are essential for the structural load transmission into the ground, guaranteeing the stability of the structure. Precisely assessing the load-bearing capacity of piles holds paramount importance in planning geotechnical structures [1]. Numerous experimental and theoretical approaches established in the past for predicting pile capacity rely on assumptions related to the factors governing the ultimate bearing capacity (\(Q_u\)) [2]. Nevertheless, owing to the intricate behavior of piles, nearly all of the existing methods and models fall short of delivering precise predictions, and a considerable number of them are tailored to specific construction sites [3].

While the static load test (SLT) remains the most reliable approach for evaluating pile-bearing capacity, its implementation can be a time-consuming, expensive, and demanding process [4]. An advanced method for forecasting the bearing capacity of piles is high-strain dynamic testing (HSDT). This method relies on wave propagation theory and is executed by utilizing a pile-driving analyzer (PDA). The HSDT process is usually standardized according to American Standard Test Methods [5]. Previous studies have revealed a strong correlation between the bearing capacity projected through PDA and values anticipated through SLT.

Moreover, PDA (HSDT) offers the advantages of enhanced speed and cost-effectiveness when compared to SLT. However, the attainment of trustworthy results mandates the execution of multiple PDA tests for every construction project. Hence, the aim to reduce the necessary number of PDA tests is of great significance, as it would reduce overall project costs. In pursuit of this goal, novel techniques such as artificial intelligence (AI) have surfaced, demonstrating the ability to provide more accurate forecasts of pile-bearing capacity and expedite solutions for complex engineering problems compared to traditional methods [6].

B. Literature Review

Machine learning (ML) methods have recently become pivotal in civil engineering and geotechnical projects. They streamline labor-intensive engineering procedures and substantially contribute to these projects’ cost efficiency. Additionally, numerous research studies utilize ML techniques to estimate the \(Q_u\) (ultimate bearing capacity) of rocks [7,8]. Yagiz et al. [9] presented ANN models, whereas Jahed Armaghani et al. [10] introduced adaptive neuro-fuzzy inference system (ANFIS) models to forecast rock strength. Concurrently, Singh et al. [11] recorded the effective application of an ANFIS estimation model to estimate the Young’s modulus of rock. In another study, Shirani Faradonbeh et al. [12] conducted research focused on developing a genetic programming (GP) technique for predicting backbreak caused by blasting. Monjezi et al. [13] applied ANN models to forecast ground vibration, whereas Marto et al. [14] utilized imperial competitive algorithm (ICA)-ANN models to predict fly rock incidents in their studies. Azimi et al. [15] also proposed a new control method, called Swarm-Based Parallel Control (SPC), inspired by the intelligence of swarms in
nature. In this study, sharing the response data among adjacent buildings using a wireless sensor network (WSN) at each floor is proposed to improve the seismic performance and minimize the risks of knocking.

In foundation design and analysis, these methodologies have garnered extensive usage. For example, Chan et al. [13] introduced an ANN model for predicting pile driving forces. This ANN was trained to employ data associated with the pile configuration, energy released during pile driving, and the actual bearing capacity (Q_u) of the pile as the network's output. In another investigation by Pal and Deswal [16], they developed two soft computing approaches, namely ANN and support vector machine (SVM), to evaluate the Q_u of concrete spun pile pipes. Their results demonstrated that among the models examined, ANN delivered the highest level of predictive accuracy. Shahin et al. [17] designed an ANN predictive model for forecasting the bearing capacity of drilled shafts. In a distinct study, Jianbin et al. [18] evidenced the effectiveness of ANN in forecasting the axial bearing capacity (ABC) of pile pipes in sandy soil. They included significant factors like the effective length and diameter of the pile, soil cohesion, unit weight, internal friction angle, and results from the standard penetration test (SPT) in the development of the network. In another research endeavor, Yu Lei et al. [19] conducted a comparative study involving six hybrid models that incorporated neural networks in conjunction with six different swarm intelligence optimization algorithms, including the innovative Seagull Optimization Algorithm (SOA). These hybrid models were evaluated against the predictive capabilities of two single models without any optimization techniques in forecasting Uniaxial Compressive Strength. Moreover, other researchers have contributed to the advancement of meta-heuristic algorithms. For instance, Agushaka [20] conducted studies aimed at enhancing arithmetic optimization, while Gaurav Dhiman et al. [21] focused on refining the Seagull Optimization Algorithm. These efforts collectively contribute to developing and improving hybrid models that leverage both ML techniques and advanced optimization methods for enhanced predictive accuracy.

C. Objective

This research introduces a novel ML approach to attain precise and optimal predictions. The hybridization method employed in this study is custom-tailored to enhance the performance of DT models, guaranteeing reliable outcomes. Through the combination of 2 cutting-edge and effective optimization techniques, namely the Zebra Optimization Algorithm (ZOJA) and the Coronavirus Herd Immunity Optimizer (CHIO), the creation of these innovative hybrid models exceeded the capabilities of traditional methods, marking a substantial leap forward. A thorough assessment was carried out on these models, individually and in hybrid setups, to guarantee a fair and unbiased evaluation of their performance. To ensure the trustworthiness of the outcomes, the evaluation of model outputs included well-recognized performance metrics such as R² and RMSE. This approach played a crucial role in mitigating any possible bias in the results, offering a more precise understanding of the models’ effectiveness. In addition to the technical aspects, this study acknowledged the practical importance of these discoveries. The increased precision achieved with the hybrid models holds the promise of improving decision-making in real-world geotechnical engineering projects, thus reducing the risks associated with inaccurate Q_u estimations. The ability of the DT model to provide dependable predictions, whether applied individually or in hybrid configurations, highlights their flexibility and appropriateness for various project needs.

II. MATERIALS AND METHODOLOGY

A. Data Gathering

In ML projects, the careful choice of input variables and the precise definition of desired outputs are crucial for attaining optimal model performance. An extensive set of input variables is systematically gathered from published literature [22]. It is important to reiterate, as highlighted in multiple references, including [23], that the pile’s dimensions, particularly its length and diameter, play a paramount role in determining the Q_u. Therefore, to consider the impact of rock and soil layers, 2 ratios associated with pile geometry were chosen: the ratio of length within the soil layer (L_s) to socket length (L_r), and the ratio of total length (L_t) to diameter (D).

In summary, the model inputs for estimating the Q_u of rock-socketed piles consisted of L_s/L_r, L_t/D, UCS, and SPT N –value. In addition, H_r is the height of the layer. These inputs were chosen to simplify the predictive model using smaller parameters. Table I indicates the statistical properties of these inputs and Q_u. Fig. 1 presents a column plot for determining the input frequency in correlation with Q_u.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Category</th>
<th>Min</th>
<th>Max</th>
<th>Avg</th>
<th>St. Div</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lp/D</td>
<td>Input</td>
<td>4.331</td>
<td>96.30</td>
<td>31.39</td>
<td>22.57</td>
</tr>
<tr>
<td>Ls/Lr</td>
<td>Input</td>
<td>0.29</td>
<td>31.71</td>
<td>4.86</td>
<td>5.66</td>
</tr>
<tr>
<td>N_SPT</td>
<td>Input</td>
<td>0.00</td>
<td>166.42</td>
<td>44.67</td>
<td>59.57</td>
</tr>
<tr>
<td>UCS</td>
<td>Input</td>
<td>0.00</td>
<td>68.49</td>
<td>24.23</td>
<td>23.55</td>
</tr>
<tr>
<td>H_r</td>
<td>Input</td>
<td>0.00</td>
<td>8.36</td>
<td>0.79</td>
<td>1.50</td>
</tr>
<tr>
<td>Q_u</td>
<td>Output</td>
<td>1449.00</td>
<td>42700.73</td>
<td>17421.79</td>
<td>10230.86</td>
</tr>
</tbody>
</table>
Fig. 1. The column plot between input and output.

B. DT

Derived from ML theory, a DT functions as a potent instrument for proficiently tackling regression and classification challenges. In contrast to other classification techniques that rely on a combined set of features for one-step classification, the DT employs a multi-stage or hierarchical decision approach, displaying a structure resembling a tree [24]. In contrast to other classification methods that depend on a unified set of features for immediate classification, the DT employs a multi-stage or hierarchical decision approach, showcasing a structure that resembles a tree. This tree includes a root node housing all the data, a sequence of splits (internal nodes), and an assortment of leaves (terminal nodes) [25]. Within the DT structure, each node performs a binary decision, separating either a singular class or a subset of classes from the remaining ones. Typically, the process entails traversing the tree from the top to the bottom, following a top-down methodology [26].

1) Decision tree regression (DTR): As shown in Fig. 2, a distinct regression tree is built for every class to enable soft classification. Within the context of regression trees, the target vector is the defined class proportions of a pixel, also known as soft reference data, whilst the pixel intensity values from
various bands function as estimator feature vectors or variables. The technique generates predicted class proportions as an output by using the intensity values as input for each individual regression tree [27]. The steps involved in creating regression trees using the training dataset are as follows:

a) Employ the pixel intensity values from different bands as the variables that predict.

b) Employ the known class fraction of class $i$ within the target variable, a pixel.

c) Create the regression tree specific to class $i$.

d) Iterate through the process for class $i$, varying from 1 to $M$.

The DT regression technique facilitates the soft classification of remote sensing data.

The regression tree algorithm for soft categorization is as follows:

1) Input the pixel intensity values from various bands.

2) Execute the regression tree for class $i$.

3) Retrieve the output of regression tree $i$, which represents the fraction of class $i$ in a pixel.

4) Repeat the process for class $i$, ranging from 1 to $M$.

It is common practice to rescale the soft classification outputs to a range of 0 to 1 for each pixel, indicating the class fractions within the ground pixel area. Thus, the estimated class proportions from each tree, denoted as $DT(i)$ for $i = 1,...,M$, undergo normalization through the following process:

$$P(i) = \frac{DT(i)}{\sum DT(i)}, i = 1, ..., M$$  \hspace{1cm} (1)

C. Zebra Optimization Algorithm (ZOA)

Zebras are included in the cohort population of ZOA, a population-based optimization method. It updates its members by utilizing 2 innate behaviors seen in wild zebras. These include hunting and gathering food as well as protecting oneself from predators. Therefore, the ZOA population members receive updates in 2 distinct stages for each iteration [28].

1) Foraging behavior: Zebras predominantly graze on grasses and sedges; however, they might turn to buds, bark, fruits, roots, and leaves when their preferred food is scarce. During the initial stage, updates to the population members are executed by simulating zebra behavior while searching for food. The amount of time zebras dedicate to eating can vary between 60-80 percent, contingent on the quality and accessibility of vegetation. Among zebras, the plains zebra stands out as a primary grazer. It consumes the upper canopy of grass, which is often less nutritious, creating an environment conducive for other species that rely on shorter and more nutrient-rich grasses. In ZOA [29], the lead zebra is regarded as the top-performing member of the population and directs other members toward its location in the search area. Consequently, the adjustment of zebra positions during the foraging stage can be expressed mathematically utilizing Eq. (2) and (3).
In this context, $x_{ij}^{new,p1}$ represents the updated state of the $ith$ zebra, with $x_{ij}^{new,p1}$ denoting its value in the $jth$ dimension and $F_{ij}^{new,p1}$ indicating its updated objective function value. The lead zebra, referred to as $BM_j$, stands as the top-performing member, and $BM_j$ signifies its value in the $jth$ dimension [30]. The value of $i$ is determined using a random number $e$, which can be any value between 0 and 1. The round function $(1 + rand)$ is then used to round the value of $i$ to the closest integer. The parameter $i$ has 2 possible values: 1 and 2. Setting $i$ to 2 results in more significant changes in population mobility.

2) Defense strategies against predators: Updates to the positions of ZOA residents in the search area are made in the second stage by using simulations of a zebra’s defensive maneuvers against predator attacks. Not only do lions pose a threat to zebras, but wild dogs, leopards, brown hyenas, and cheetahs also pose a hazard, each of which calls for a different approach to defense. Zebras defend themselves against lion attacks by using zigzag patterns and occasional sideways spins as evasive maneuvers. Zebras usually react more aggressively to smaller predators such as hyenas and dogs. They regroup and cause chaos to make their enemies more difficult to deal with. In the context of ZOA, the assumption is that lions or other predators may initiate an attack, prompting zebras to decide between an escape strategy when facing lions or an offensive approach against other predators [31]. The escape strategy is represented by mode S1, while the defensive strategy, where the herd forms a protective structure, is represented by mode S2. A zebra’s new position is valid if it results in an improved objective function value.

$$ x_{ij}^{new,p2} = \begin{cases} 
S1: x_{ij} + e \times (2e - 1) \times (1 - \frac{1}{n} + x_{ij}, P_i \leq 0.5 \\
S2: x_{ij} + e \times (BM_j - l \times x_{ij}).
\end{cases} \quad (4)$$

$$ x_i = \begin{cases} 
x_{i}^{new,p2}, F_{ij}^{new,p2} < f_i, \text{otherwise}
\end{cases} \quad (5)$$

In this context, the constant $E$ is fixed at 0.01, and $P_i$ represents the likelihood of selecting one of two methods at random from the range [0,1]. $BM_j$ refers to the condition of the zebra under attack, where $BM_j$ signifies its value in the $jth$ dimension.

D. Coronavirus Herd Immunity Optimizer (CHIO)

The algorithm for optimization introduced in this study integrates the concept of herd immunity, drawing parallels between the principles of COVID-19 and the optimization process [32]. The steps that make up the CHIO are listed below, with a detailed explanation of each. The algorithm explores each of the 6 primary phases in the parts that follow [33, 34].

Phase 1: To fit the optimization difficulty's context, the objective function is formulated as follows during the CHIO and optimization problem's initialization phase:

$$\min O(x) \quad x \in [lb, ub] \quad (6)$$

An individual’s or case $x$’s immunity rate is reflected by the objective function $O(x)$, which is denoted by the gene values $x_1, x_2, \ldots, x_n$, where each $x_i$ corresponds to a gene or choice variable. Each person’s total gene count is denoted by the index $n$. It is crucial to emphasize that each gene $x_i$ falls within the value range of $[lb_i, ub_i]$, with $ub_i$ and $lb_i$ denoting the upper and lower boundaries, respectively, of gene $x_i$.

Through this phase, CHIO initializes its two primary control parameters [34]: $BR_r$. It governs the propagation of the virus pandemic among people and acts as the fundamental reproduction rate governing the CHIO operators. $Max_{age}$ This denotes the oldest age at which infected cases can be documented and describes how those instances will end. Cases reaching $Max_{age}$ either recover or pass away.

Two control parameters and four algorithmic parameters make up CHIO. $C_0$: This is the initial number of infected instances, which in this particular case is one. $Max_{it}$: the highest quantity of repetitions. $HIS$: The number of people. $n$: The problem’s dimensionality.

Phase 2: The examples that are generated are then saved in HIP in the form of a 2-dimensional matrix with dimensions of $n \times HIS$. Here, $n$ stands for each person’s size, which is indicated as follows:

$$HIP = \begin{bmatrix}
x_1^1 & x_2^1 & \cdots & x_n^1 \\
x_1^2 & x_2^2 & \cdots & x_n^2 \\
\vdots & \vdots & \ddots & \vdots \\
x_1^{HIS} & x_2^{HIS} & \cdots & x_n^{HIS}
\end{bmatrix} \quad (7)$$

Each row in HIP denoted as $x_i$, fits a case that is produced by the following formula: $x_i = lb_1 + (ub_1 - lb_1) \times U(0,1)$, where $i$ series from 1 to $n$, and where $ub_l$ and $lb_l$ represent the upper and lower limitations of the gene $x_i$. Eq. (6) is utilized to calculate the immunity rate, or objective function, for every instance. Every instance in HIP also has a status vector $(S)$ constructed for it, with a length of $HIS$. In this vector, a vulnerable case is represented by 0, and an infected case is represented by 1 value.

Phase 3: The progression of coronavirus herd protection constitutes the central iterative process of CHIO. This process encompasses three rules governing the change of genes $x_i$ within case $x_j$. These genes can either remain unchanged or undergo alterations due to the implementation of social estrangement procedures. The specific rules are streamlined according to the percentage of $BR_r$ and are detailed as follows:

$$x_i' (t + 1) \left\{ \begin{array}{ll}
x_i' (t) & r \geq BR_r \\
C \left( x_i' (t) \right) & r < \frac{1}{2} \times BR_r \text{ infected} \\
N \left( x_i' (t) \right) & r < \frac{2}{3} \times BR_r \text{ suspected} \\
R \left( x_i' (t) \right) & r < BR_r \text{ immuned}
\end{array} \right. \quad (8)$$
Phase 4: The immunity rate, denoted as \( O(x^i(t+1)) \), is computed for each newly generated case, \( x^i(t+1) \). If it outperforms the current case, \( x^i(t) \), as indicated by \( (x^i(t+1)) > 0(x^i(t)) \), then \( x^i(t) \) is substituted with \( x^i(t+1) \). Furthermore, if \( E_j = 1 \), the age vector \( A^j \) is incremented by one.

The status vector, \( E_j \), for each case, \( x^j \), is adjusted based on the herd immunity threshold, calculated as follows:

\[
E_j \left\{ \begin{array}{l}
0 \left( x^j(t+1) \right) < \frac{\delta(x^j(t+1))}{\Delta O(x)} , E_j = 0, \text{is.corona}(x^j(t+1)) \\
2 \left( x^j(t+1) \right) < \frac{\delta(x^j(t+1))}{\Delta O(x)} , E_j = 1
\end{array} \right.
\]

(9)

The equation is formulated based on the following variables:
- \( \text{is.corona}(x^j(t+1)) \): A dual value set to 1 if the newly created case, \( x^j(t+1) \) inherits an amount from each affected instance.
- \( \Delta O(x) \): This denotes the mean immunity rate of the population, which is calculated as the sum of all immunity rates divided by the population size.

It is crucial to stress that determined social distancing measures may have an impact on the population’s immunity rate among persons [35], considering that the freshly formed individual’s immunity rate is higher than the population’s average immunity rate.

E. Performance Evaluator

This section delineates a set of metrics designed to assess hybrid models. These metrics gauge error and correlation, providing respected insights into the models’ performance. Table II displays the equations for the metrics employed in this paper [36].

Respectively, the variables can be expressed as:
- The symbol \( n \) represents the sample size.
- Predicted values are indicated as \( b \).
- \( \bar{m} \) and \( \bar{b} \) stand for the mean of the evaluated and anticipated values, respectively.
- The measured value is denoted as \( m \).
- The mean of the predictor variable in the dataset is symbolized as \( \bar{x} \).

| Coefficient Correlation (R²): | \( R^2 = \left( \frac{\sum_{i=1}^{n}(b_i - \bar{b})(m_i - \bar{m})}{\sqrt{\sum_{i=1}^{n}(b_i - \bar{b})^2 \sum_{i=1}^{n}(m_i - \bar{m})^2}} \right)^{2} \) | (10) |
|-----------------------------|-----------------------------------------------------------------|
| Root Mean Square Error (RMSE): | \( RMSE = \frac{1}{n} \sum_{i=1}^{n} (m_i - b_i)^2 \) | (11) |
| Mean Absolute Error (MAE): | \( MAE = \frac{1}{n} \sum_{i=1}^{n} |b_i - m_i| \) | (12) |
| Mean Absolute Percentage Error (MAPE): | \( MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{b_i - m_i}{m_i} \right| \) | (13) |
| weight absolute percentage error (WAPE): | \( WAPE = \max \left( \frac{|b_i - m_i|}{b_i} \right) \) | (14) |

III. RESULT

In this section, a comparative assessment of the results was carried out to obtain the proposed model using single and hybrid frameworks. The hybrid variants developed to achieve optimal outcomes include DT+ZOA (DTZ0) and DT+CHIO (DTCH). To train these proposed models, 70% of the input data was dedicated to the training process, while the remaining 30% was further divided, with 15% allocated for validation and 15% for testing purposes. Various metrics, including \( R^2 \), RMSE, MAE, MAPE, and WAPE, were employed to conduct a thorough assessment of the acquired results and ensure unbiased findings. In the case of the \( R^2 \) metric, values close to 1 suggest excellent results. Conversely, values approaching 0 indicate accurate outcomes when considering the error indicators.

Table III presents the results of predicting the \( Q_u \) of rocks for each framework. The performance of the traditional DT models within a single framework does not show favorable outcomes according to the mentioned metrics. While the \( R^2 \) value for DT is 0.9817 during the training phase, it notably decreases in both the validation and testing phases. On the other hand, DTZA exhibits steady performance, especially in the training phase, with an \( R^2 \) value of 0.9962 and an RMSE value of 629.7812. Among the models, the DTCH model showed moderate results, achieving higher accuracy compared to DT but slightly lower accuracy than the DTZA model. Overall, it can be concluded that the accuracy of DT models in predicting the \( Q_u \) of rocks is successfully enhanced by the incorporation of ZOA and CHIO optimizers, which contributes to an overall improvement in the reliability of the results.

Fig. 3 displays a scatter plot depicting the models’ performance, evaluated through their \( R^2 \) and RMSE values. The X-axis correlates with the measured values, while the Y-axis correlates with the predicted values generated by the models. Triangular shapes in various colors are incorporated into the scatter plot to distinguish between the training, testing, and validation phases. These shapes are dispersed around a
diagonal line, representing an ideal scenario with an $R^2$ value of 1. The limited accuracy of the DT model becomes apparent due to the significant spread of data points.

Conversely, the DTZA and DTCH models demonstrate enhanced performance compared to the standalone DT approach. Data points for DTZA are tightly clustered near the central line, indicating a more favorable outcome. However, some broader dispersions are observable in the case of DTCH.

TABLE III. THE RESULT OF DEVELOPED MODELS FOR DT

<table>
<thead>
<tr>
<th>Model</th>
<th>Phase</th>
<th>RMSE</th>
<th>$R^2$</th>
<th>MAE</th>
<th>MAPE</th>
<th>WAPE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>1440.54</td>
<td>0.9817</td>
<td>1235.19</td>
<td>11.3086</td>
<td>0.0713</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>1829.53</td>
<td>0.9683</td>
<td>1471.95</td>
<td>7.6086</td>
<td>0.0726</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>1695.50</td>
<td>0.9746</td>
<td>1121.61</td>
<td>7.4247</td>
<td>0.0745</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>1545.41</td>
<td>0.9774</td>
<td>1253.81</td>
<td>10.1622</td>
<td>0.0720</td>
</tr>
<tr>
<td>DTZA</td>
<td>Train</td>
<td>629.78</td>
<td>0.9962</td>
<td>446.47</td>
<td>2.6926</td>
<td>0.0258</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>1221.55</td>
<td>0.9873</td>
<td>926.96</td>
<td>4.6943</td>
<td>0.0457</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>1229.66</td>
<td>0.9893</td>
<td>814.87</td>
<td>5.2731</td>
<td>0.0541</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>854.89</td>
<td>0.9934</td>
<td>574.80</td>
<td>3.8535</td>
<td>0.0330</td>
</tr>
<tr>
<td>DTCH</td>
<td>Train</td>
<td>1001.70</td>
<td>0.9903</td>
<td>742.39</td>
<td>4.6906</td>
<td>0.0429</td>
</tr>
<tr>
<td></td>
<td>Validation</td>
<td>1503.80</td>
<td>0.9806</td>
<td>1145.34</td>
<td>5.7609</td>
<td>0.0565</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>1250.77</td>
<td>0.9834</td>
<td>850.54</td>
<td>5.7691</td>
<td>0.0604</td>
</tr>
<tr>
<td></td>
<td>All</td>
<td>1178.30</td>
<td>0.9872</td>
<td>841.02</td>
<td>5.0565</td>
<td>0.0483</td>
</tr>
</tbody>
</table>

Fig. 3. The correlation between the predicted and measured values.
Fig. 4 presents the relationship between predicted and measured values of the DT base models. In this figure, the measured values are represented by the black box. When the predicted values closely align with these measured values, it signifies the model's accuracy. As shown in Fig. 4, the DT model exhibits a significant deviation of its data from the measured values, particularly during the testing phase. Conversely, the DTZA model showcases a precise outcome, with an almost perfect match between its predicted and measured values, particularly within the sample numbers from 40 to 80. In contrast, the DTCH model demonstrates a noticeable lack of accuracy within the sample numbers ranging from 80 to 120, rendering it less precise than the DTZA model.

It is crucial to conduct an error assessment to achieve a deeper understanding of the uniqueness and precision of the models. Fig. 5 emphasizes that the DT model displayed a significant error rate, especially during the testing phase, peaking at a maximum error of 35% within the sample range of 0 to 50. On the other hand, most data points in the DTZA model exhibited errors that were nearly 0%, in contrast to the DTCH model, where the maximum error reached 20%.

Fig. 6 presents a box plot that simplifies the comparison of errors among the models within a single visual representation. The precision of the DTZA model becomes evident as the data points are closely grouped, primarily concentrated within the error range of -5% to 5%. This pattern contrasts with the
spread seen in the other models, underscoring the dependability of the DTZA model’s predictions. The concentration of data points within a relatively narrow error range signifies its consistent and accurate performance. In contrast, the dispersion observed in the error distribution of the other models indicates a broader variability in their predictions.

Fig. 5. The error percentage for the hybrid models is based on a line-symbol plot.
IV. DISCUSSION

A. Comparison between the Results of Previous Articles and the Present Study

Table IV presents the results of previous research efforts on Qu prediction, providing a comprehensive benchmark for comparison with the current study. Among the five models discussed in this table from earlier research, the GA-DLNN model, described in Pham et al.’s research [37], demonstrated the most impressive performance, achieving an \( R^2 \) of 0.882 and an RMSE of 109.965. As detailed in the previous Section, the current investigation emphasizes the superior performance of the DTZA model during the training phase, producing commendable metric scores with an \( R^2 \) value of 0.9962 and an RMSE of 629.78. This exceptional performance in critical metrics decisively positions the DTZA model in this study as outperforming its counterparts, confirming its effectiveness in Qu prediction.

<table>
<thead>
<tr>
<th>Articles</th>
<th>Models</th>
<th>Models' performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
</tr>
<tr>
<td>Armaghani et al. [38]</td>
<td>ANN</td>
<td>0.135</td>
</tr>
<tr>
<td>Pham et al. [37]</td>
<td>GA-DLNN</td>
<td>109.965</td>
</tr>
<tr>
<td>Momeni et al. [39]</td>
<td>ANN</td>
<td>0.529</td>
</tr>
<tr>
<td>Momeni et al. [39]</td>
<td>ANFIS</td>
<td>0.048</td>
</tr>
<tr>
<td>Kulkarni et al. [40]</td>
<td>GA-ANN</td>
<td>0.0093</td>
</tr>
<tr>
<td>Present Study</td>
<td>DTZA (DT+ZAO)</td>
<td>629.78</td>
</tr>
</tbody>
</table>

B. Results of an Ablation Study

In this study, an ablation study was conducted as an alternative assessment approach. Two input values, as detailed in previous sections, were deliberately removed from the dataset. Subsequently, the prediction process was reiterated utilizing the best hybrid model. The outcomes of this ablation study were then compared with those of the main prediction process, with the findings summarized in Table V. Analysis of the table reveals that the results of the Main Study outperformed those of the Ablation Study. Specifically, the Root Mean Square Error (RMSE) of the Main Study was 17.22% lower than that of the Ablation Study, indicating a superior predictive performance in the former.

<table>
<thead>
<tr>
<th>phase</th>
<th>Main Study</th>
<th>Ablation Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( R^2 )</td>
<td>RMSE</td>
</tr>
<tr>
<td>Train</td>
<td>0.9962</td>
<td>629.78</td>
</tr>
<tr>
<td>Validation</td>
<td>0.9873</td>
<td>1221.55</td>
</tr>
<tr>
<td>Test</td>
<td>0.9893</td>
<td>1229.66</td>
</tr>
<tr>
<td>All</td>
<td>0.9934</td>
<td>854.89</td>
</tr>
</tbody>
</table>

C. Limitations of the Study

The study’s limitations are acknowledged, including the potential restriction of results’ generalizability by the specific dataset and experimental setup utilized. Variations in geological conditions, pile types, and other site-specific factors could affect the performance of the proposed hybrid models in different contexts. Additionally, the focus on predictive accuracy in the analysis may neglect considerations of computational efficiency or scalability, which could be crucial in real-world applications. Furthermore, biases or uncertainties may be introduced by the assumptions and parameters chosen for the optimization algorithms. Lastly, while efforts were made to provide insightful analysis of the experimental results, there may be aspects that require further investigation or validation in future studies. These limitations underscore the need for a cautious interpretation of the findings and highlight avenues for future research to address these constraints and refine the proposed methodologies.

V. CONCLUSION

The estimation of the ultimate bearing capacity (\( Q_u \)) using ML methods, specifically the DT model, coupled with advanced optimization algorithms, including the Zebra Optimization Algorithm (ZOA) and the Coronavirus Herd Immunity Optimizer (CHIO), has demonstrated significant promise and yielded valuable insights into the field of civil engineering. The incorporation of ML techniques, particularly the DT model, has proven to be a powerful tool for accurately estimating the ultimate bearing capacity of soils and rocks. The
utilization of this model allows for the efficient handling of complex datasets and the extraction of meaningful patterns and relationships within the data. However, as revealed in the analysis, the standalone DT model exhibited limitations in accuracy, especially during the testing phase, highlighting the need for further refinement. The fusion of the DT model with optimization algorithms, such as ZOA and CHIO, has been a pivotal advancement in enhancing predictive accuracy. The ZOA and CHIO optimizers have contributed to refining the model’s performance during the training, validation, and testing phases. The analysis of the prediction results, as depicted in correlation and error assessments, showcased the distinctiveness of these models. The DTZA model demonstrated remarkable accuracy, with a close alignment between predicted and measured values. Its ability to maintain errors close to 0% over a wide range of samples is a testament to its consistent and precise performance. On the other hand, the DTCH model, while showing improved accuracy compared to the standalone DT model, exhibited some variability in its predictions, particularly in a specific sample range. The integration of ML methods, DT models, and advanced optimization algorithms like ZOA and CHIO has proven to be a valuable approach for predicting the ultimate bearing capacity in civil engineering applications. These hybrid models have shown substantial improvements in accuracy and reliability, which are crucial for making informed decisions in geotechnical engineering projects. Further research and fine-tuning of these models can advance the understanding and predictive capabilities in civil engineering, particularly in earthquake analysis and related areas.

ACKNOWLEDGMENTS

2023 Jilin Provincial Department of Education Science Research Project, Project Name: "Studying on the Shrinkage Behavior of Semi-flexible Grading Gravel Base Material in Cold Regions" (Project Number: JJKH20231038KJ)

REFERENCES


