Journal of Mining and Earth Sciences
Website: https://jmes.humg.edu.vn

Optimizing axial bearing capacity estimation of pre-bored grouted planted nodular (PGPN) pile: Enriched dataset and genetic algorithm approach

Tan Nguyen [1,*], Tuetakoun Aphistih [1], Minh The Quang Nguyen [1], Duy Khuong Ly [2, 3], Jim Shiau [4]

[1] *Ton Duc Thang University (Faculty of Civil Engineering), Hochiminh, Vietnam*
[2] *Van Lang University (Laboratory for Computational Mechanics - Institute for Computational Science and Artificial Intelligence), Hochiminh, Vietnam*
[3] *Van Lang University (Faculty of Civil Engineering - School of Technology), Hochiminh, Vietnam*
[4] *University of Southern Queensland (School of Engineering), Toowoomba, Australia*

ARTICLE INFO

Article history:
Received 23rd July 2024
Revised 29th Oct. 2024
Accepted 10th Nov. 2024

Keywords:
Axial bearing capacity,
Empirical Formula,
Genetic algorithm optimization,
Static Pile Load Test.

ABSTRACT

This study introduces a practical and efficient methodology to improve the estimation of the ultimate axial bearing capacity of Pre-bored Grouted Planted Nodular (PGPN) piles, addressing a critical need in geotechnical engineering. By enhancing the dataset with 98 additional case histories and pile load test data from various projects across Vietnam, we have developed a more reliable predictive tool for engineers. The use of genetic algorithms has refined existing empirical formulas, significantly improving their accuracy while remaining simple enough for hand calculations. The proposed formula achieves a correlation coefficient of 0.907, a 7.2% improvement over previous methods. This research offers a valuable solution for the challenges faced in predicting the load-bearing capacity of PGPN piles, providing a more dependable method that can streamline design and construction practices.

*Corresponding author
E - mail: nguyentan@tdtu.edu.vn
DOI: 10.46326/JMES.2024.65(6).10

1. Introduction

The Pre-bored Grouted Planted Nodular (PGPN) piles represent a groundbreaking innovation in pile foundation engineering. The use of PGPN is growing more popular in construction projects across Asia due to its practical advantages in ecofriendliness and efficiency, which have made it a preferred choice in many construction projects (Zhou et al., 2020, Nguyen et al., 2022). Its unique composition and load-bearing mechanisms have spurred researchers to develop methods for precisely estimating its ultimate axial bearing capacity while also addressing the challenge of minimizing environmental impact. This section revisits the evolution of bored PGPN piles, highlights the complexities inherent in their load-bearing mechanisms, reviews existing estimation methods, and underscores the quest for enhanced precision through innovative approaches like genetic algorithms.

The development of bored PGPN piles began in the 1990s, driven by continuous refinements aimed at optimizing load transfer mechanisms (Horiguchi & Karkee, 1995; Karkee et al., 1998). Japanese researchers pioneered the pre-bored precast piling method, introducing high-strength nodular concrete inserted into pre-excavated holes filled with cement milk. Subsequent innovations, such as the hyper-MEGA method, further enhanced axial load-bearing capacity by enlarging base excavations, making these piles highly effective in various Asian countries (Kobayashi & Ogura, 2007).

The distinctive composition of PGPN piles involves intricate interactions between the core Pre-tensioned spun High strength Concrete (PHC) pile, cemented-soil layer, and natural soils. With this unique composition, it would create novel load-bearing mechanisms (Yu et al., 2021; Huynh et al., 2022). Despite efforts to elucidate these mechanisms through analytical, experimental, and hybrid numerical methods, the complexities of cemented-soil interactions remain a challenge for conventional estimation techniques. Unlike traditional piles, the load transfer in PGPN piles is influenced by the cemented-soil layer, which could alter the interaction between the pile and surrounding soil, posing a significant challenge for the accurate prediction of load-bearing capacity.

Several methods have been proposed to estimate the axial load-bearing capacity of PGPN piles, each with its own advantages and limitations. Analytical approaches, such as those by Wang et al. (2019), used comprehensive equations derived from soil parameters. Experimental techniques, exemplified by Fang et al. (2014) and Zhou et al. (2020), involved physical testing to capture pile behavior but may not fully encapsulate the complexities of PGPN piles. Empirical formulas, derived from field data, offered simplicity but may lack applicability across diverse regions due to variations in soil behavior (Homma, 2014; Horiguchi & Karkee, 1995; Karkee et al., 1998; Kobayashi & Ogura, 2007; Yoshimi & Tokimatsu, 1983).

To address these challenges, this study explores the potential of genetic algorithms (GAs) to refine empirical formulas and improve accuracy. The decision to use GAs is driven by their ability to optimize complex, multi-objective problems, making them particularly suited for tackling the intricacies of predicting load-bearing capacities in PGPN piles. GAs, inspired by the principles of evolutionary theory, have been adapted to tackle complex engineering problems by optimizing multiple objectives simultaneously. Originating from the work of Holland and his research team in the 1960s and 1970s, GAs mimic the process of natural selection by representing potential solutions as chromosomes composed of discrete genes, each controlling specific aspects of the solution. Initially, these genes were conceptualized as binary digits, but subsequent developments introduced more diverse gene types (Lambora et al. 2019).

It is well known that engineering dilemmas are often met with multiple conflicting objectives, such as minimizing costs, maximizing performance, and enhancing reliability. These objectives pose formidable challenges, closely mirroring the complexities encountered in real-world scenarios. GA, as a widely embraced meta-heuristic, is adeptly suited to grapple with such multi-objective predicaments. It is adeptly modified to handle multiple objectives through the incorporation of specialized fitness functions and strategies geared towards preserving

solution diversity. The choice of GAs for this study is rooted in their success in balancing these objectives, especially in scenarios involving complex interactions, such as those found in PGPN piles (Huynh et al. 2022). Genetic algorithms are particularly well-suited to handle such multi-objective predicaments due to their ability to preserve solution diversity and adaptively search the solution space. This study employs genetic algorithm to optimize the coefficients of the empirical formula, significantly improving the prediction accuracy of axial load-bearing capacity for PGPN piles.

Building upon a dataset from Vietnam, Huynh et al. (2022) developed a direct Standard Penetration Test (SPT) method, where the subsurface condition is represented using in-situ SPT data. While this method is user-friendly and practical, its accuracy was somewhat limited, achieving a correlation coefficient of 0.846. Subsequently, Nguyen et al. (2022) leveraged the capabilities of artificial neural networks (ANN) to develop a closed-form solution with a significantly improved correlation coefficient of 0.98. Despite its accuracy, the ANN-based method's complexity made it challenging for practical use. Additionally, the dataset in this study was limited both quantitatively and qualitatively, with a scarcity of samples tested until failure.

Despite existing methodologies, accurately estimating axial load-bearing capacity for PGPN piles remains an open challenge. Accurate estimation of the axial load-bearing capacity of PGPN piles is vital for the safe and economical design of foundation systems in geotechnical engineering. The empirical formulas derived through genetic algorithm optimization provide a practical and efficient solution, incorporating the complexities of cemented-soil interactions. By enhancing the dataset with more comprehensive static pile load test data and employing advanced genetic algorithm techniques, this study aims to deliver a significantly improved predictive tool for practitioners in the field. Furthermore, the study seeks to expand the scope of this methodology, demonstrating its potential application across different regions and soil types beyond Vietnam.

In summary, this study builds upon previous work to develop a more accurate and practical method for estimating the axial load-bearing capacity of PGPN piles. By enriching the dataset and leveraging the power of genetic algorithms, the new empirical formulas offer a significant advancement over existing methods. This expanded scope enhances its relevance to geotechnical engineers globally, contributing to more efficient and reliable foundation design practices.

2. Static pile load test and dataset enrichment

The PGPN piles are composed of two main parts: the core PHC pile and a cemented soil layer. The core PHC pile combines nodular and cylindrical segments, as shown in Figure 1. It has a Young's modulus of about 40,000 MPa. The cemented soil layer, made of pure cement grout, has strengths of 10 MPa for the shaft and 20 MPa for the toe, yielding a minimum compressive strength of around 0.5 MPa. This data, provided by Phan Vu Investment Corporation, outlines the material properties used.

The dataset in this study builds upon the foundation established in the work of Huynh et al. (2022). Specifically, to maintain consistency with Huynh's study, the pile characteristic, material properties, and the procedure for determining the ultimate axial load-bearing capacity are based on the Chin-Kondner hyperbolic law (Chin, 1970). These aspects, having been comprehensively described in Huynh et al. (2022), will not be reiterated here. Instead, this study focuses on the significant enhancements made to the dataset to improve the accuracy and robustness of the empirical formulas.

To fortify the foundation of our study, we embarked on an extensive data enrichment process. Initially sourced from diverse case histories in Vietnam, our dataset was amplified to encapsulate a broader range of scenarios and geotechnical conditions. Specifically, the number of samples increased from 81 case histories in the study by Huynh et al. (2022) to a more robust dataset comprising 98 case histories (see Figure 2). Moreover, the number of case histories where piles were tested to the failure stage was significantly augmented, increasing from a limited number to 15 cases.

In this study, eight input variables were utilized to predict the pile head displacement of

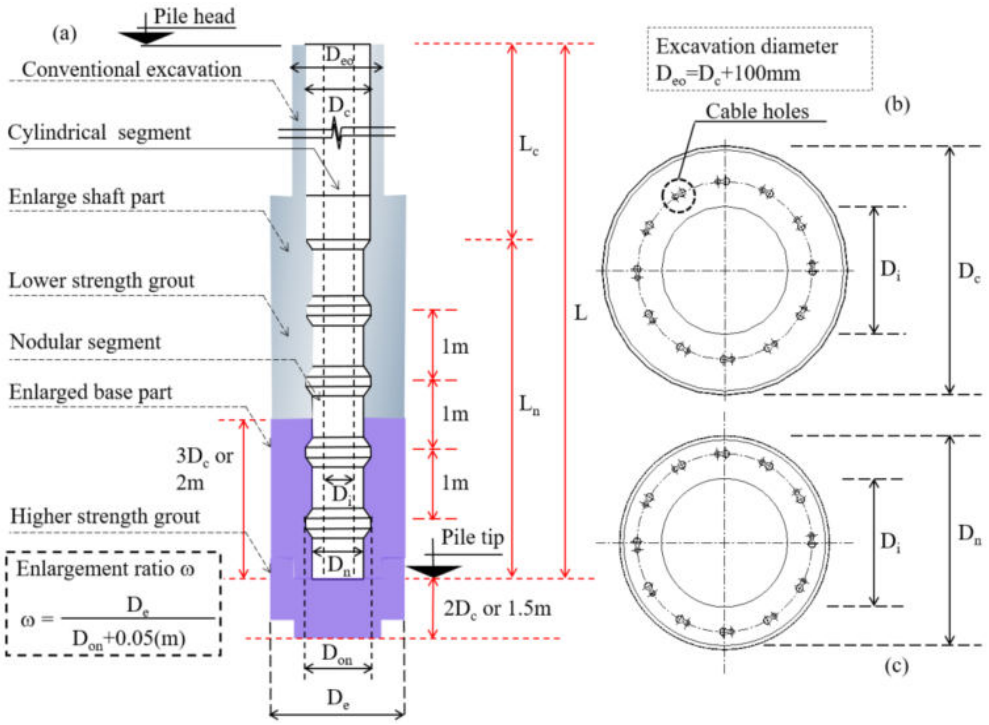


Figure 1 (a) Pile characteristics, (b) pile head, (c) pile tip (Nguyen et al., 2022; Huynh et al., 2022).

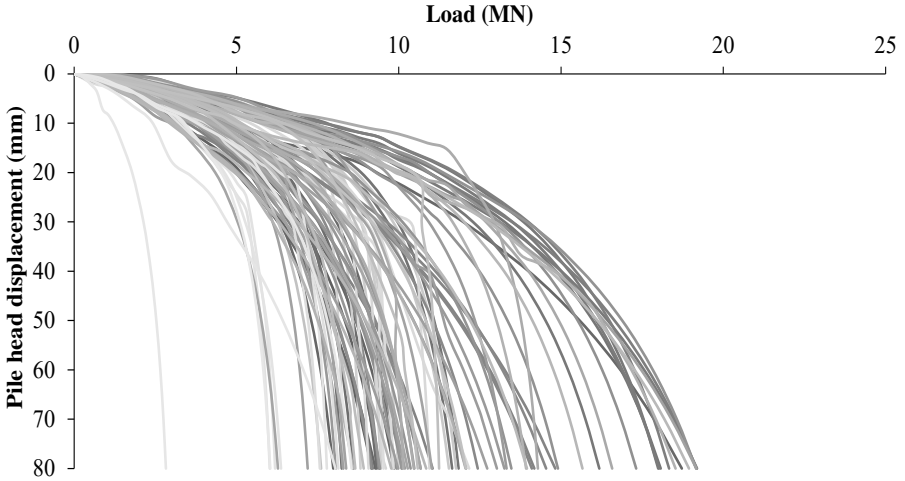


Figure 2. Load-displacement curves of all 98 case histories.

PGPN piles. These variables include the diameter of the nodular segment (D_{on}), the diameter of the cylindrical segment (D_c), the total pile length (L), the length of the nodular segment (L_n), the enlargement ratio (ω), the SPT index at the pile tip (SPT_{tip}), the average SPT index along the nodular segment (SPT_{nod}), and the average SPT index along the cylindrical segment (SPT_{cyl}). The statistical properties of the dataset are given in Table 1.

Important note: All the case histories used in this study involve friction piles, where shaft resistance is the predominant load-bearing mechanism. This focus ensures consistency across the dataset by concentrating on piles with similar geotechnical behavior, minimizing variations that could arise from different types of piles or load transfer mechanisms. This consistency is crucial in maintaining the reliability of the data used to develop and validate the proposed formula.

Table 1. Statistical properties of the dataset.

Variables	Unit	Min	Mean	Standard Deviation	Skewness	Max
D_{on}	mm	350	600	101.58	0.131	800
D_c	mm	350	600	94.11	0.245	800
L	m	22	40.50	8.00	-0.220	55
L_n	m	0	12	4.22	-1.230	12
ω	-	1	1.18	0.08	-1.354	1.23
SPT_{tip}	-	11.75	29.38	10.71	0.235	50
SPT_{nod}	-	0	19.67	13.74	0.353	50
SPT_{cyl}	-	2.45	13.05	8.13	1.714	40

Table 1 provides a comprehensive overview of the pile dimensions and soil conditions at the test sites. The pile diameters (D_{on} and D_c) have a mean of 600 mm (about 1.97 ft) with moderate variability of standard deviations of 101.58 mm (about 4 in) and 94.11 mm, respectively, and near-zero skewness values. This symmetry suggests a balanced dataset representative of typical pile sizes in geotechnical engineering. The pile lengths (L) range from 22 m to 55 m with a mean of 40.50 m, showing a slight skew towards shorter piles, while the nodular segment lengths (L_n) exhibit a higher negative skewness, indicating mostly shorter segments. The enlarged excavation ratio (ω as shown in Figure 1) shows low variability and a narrow range, reflecting high consistency across the dataset.

The soil properties, represented by SPT values, further highlight the dataset's diversity. SPT_{tip} values range from 11.75÷50, with a mean of 29.38 and a low skewness, indicating symmetrical tip resistance distribution. SPT_{nod} values, ranging from 0 to 50 with a mean of 19.67, show greater variability within the nodular segment. The SPT_{cyl} values, with a high skewness and a mean of 13.05, indicate lower resistance values in the cylindrical segment. These diverse soil conditions captured in the dataset are critical for validating the proposed formula's effectiveness.

In comparison to the previous work of Huynh et al. (2022), the 17 new cases expand the dataset's diversity by introducing a wider range of pile diameters, lengths, embedded lengths, and more extreme SPT values. This increased variability enhances the model's ability to generalize across different pile configurations and soil conditions, ultimately improving predictive accuracy. Omitting or replacing these cases with less diverse ones would reduce the dataset's variability, potentially decreasing the model's robustness. The inclusion of these cases plays a critical role in strengthening the model's predictive capacity. The comprehensive and diverse nature of the dataset ensures that the proposed formula for predicting the ultimate bearing capacity of PGPN piles is both robust and reliable, as evidenced by high reliability scores across different measurement methods, effectively accounting for the complexity and variability of geotechnical conditions.

Figure 2 shows the load-displacement curves of all 98 case histories. The value of the ultimate bearing capacity is determined by identifying the intersection of the Davisson offset line and the Chin extrapolated load-displacement curve.

Chin's method, based on Kondner's (1963) work, assumes a hyperbolic load-displacement curve near failure. Test data from the final loading cycle is replotted, with settlement on the x-axis and settlement-to-load ratio on the y-axis. A straight line is drawn from the last three points in the plastic phase to reduce errors from early elastic displacement, and the ultimate bearing capacity is calculated using Equation (1).

$$Q_u = \frac{n\sum_{i=1}^{n=3} S_i \frac{S_i}{P_i} - \sum_{i=1}^{n=3} S_i \sum_{i=1}^{n=3} \frac{S_i}{P_i}}{n\sum_{i=1}^{n=3} \left(\frac{S_i}{P_i}\right)^2 - \left(\sum_{i=1}^{n=3} \frac{S_i}{P_i}\right)^2} \quad (1)$$

Where: *i* - corresponds to the last 3 load steps; *n* - the number of load levels, where n = 3, S_i - the settlement at the i^{th} load step; P_i - the magnitude of the i^{th} load step.

The load-displacement curve has been extended according to Chin's method, as outlined in Equation (2).

$$S = a\frac{S}{P} + b \quad (2)$$

Where *a* and *b* are defined as follows:

$$a = Q_u = \frac{n\sum_{i=1}^{n=3} S_i \frac{S_i}{P_i} - \sum_{i=1}^{n=3} S_i \sum_{i=1}^{n=3} \frac{S_i}{P_i}}{n\sum_{i=1}^{n=3} \left(\frac{S_i}{P_i}\right)^2 - \left(\sum_{i=1}^{n=3} \frac{S_i}{P_i}\right)^2} \quad (2a)$$

$$b = \frac{1}{n}\left(\sum_{i=1}^{n=3} S_i - a\sum_{i=1}^{n=3}\frac{S_i}{P_i}\right) \quad (2b)$$

Davisson's method calculates the pile head's displacement (Equation 3). The ultimate bearing capacity is found where Davisson's offset line intersects the extrapolated Chin curve.

$$S = \frac{QL}{AE} + \frac{D}{120} + 4 \quad (3)$$

Where: S - the pile head deflection (mm); A - the cross-sectional area of the pile shaft; E - the Young's modulus of the pile material; L - the pile length; Q - the maximum applied load; D - the pile diameter or width (mm).

Similarly, ultimate bearing capacities defined by the total limited displacement criteria under Vietnamese standard TCVN 10304-2014 for a displacement of 40 mm and FHWA guidelines (Abu-Farsakh et al., 2010) for a displacement of 10%D are calculated by the intersection of Chin's extrapolated curve and horizontal lines at S = 40 mm and S = 10%D, respectively (Huynh et al., 2022). Considering that the pile diameters are either 600 mm or 800 mm, the 40 mm displacement corresponds to 5÷7% of the pile diameter.

The distribution characteristics of ultimate bearing capacity using the Chin method, Chin/Davisson method, Chin/10%D method, and Chin/40 mm method are illustrated in Figures 3÷7 respectively. It is evident that the ultimate bearing capacity obtained from the Chin method is the highest, followed by the Chin/Davisson and Chin/10%D methods, which are very comparable to each other.

The value from the Chin/40 mm method is the lowest among the methods. Crucially, the discrepancy in ultimate bearing capacity values derived from these four methods is less than 30% (see Figures 8 and 9). The minimal difference between the Chin/40 mm and Chin/10%D methods, where the Chin/10%D values are slightly higher, further validates the reliability of these approaches. This negligible variance highlights that the ultimate bearing capacity determined using the proposed approach, which integrates the Chin-Kondner extrapolation method with the total limited displacement criteria under Vietnamese standard TCVN 10304-2014, is both dependable and accurate. This consistency underscores the robustness of the methodologies employed.

3. Problem formulation

3.1. Problem statement

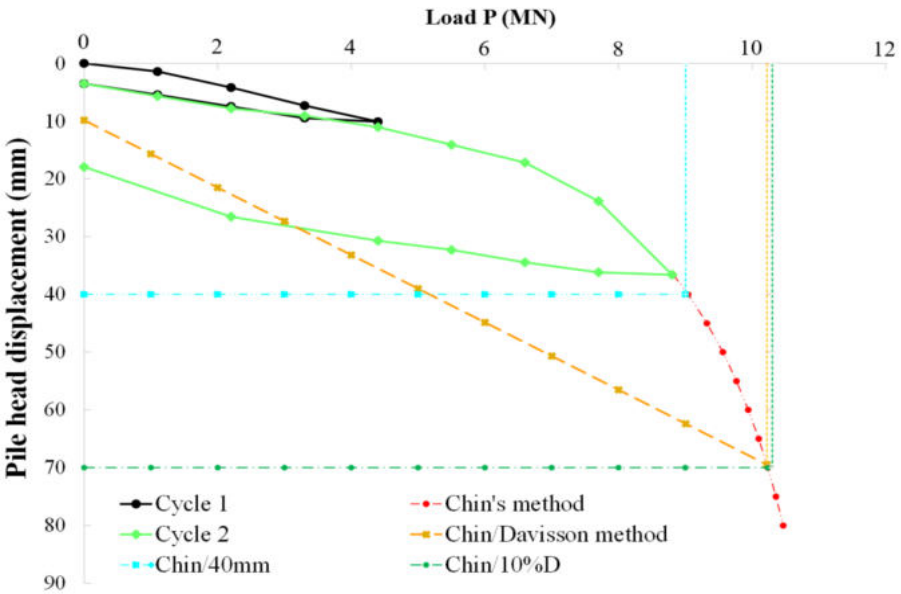


Figure 3. Chin's extrapolation method illustrating failure at various displacement levels

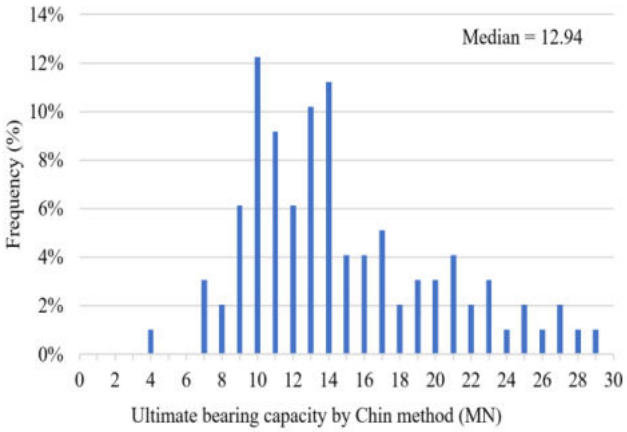


Figure 4. Distribution characteristics of ultimate bearing capacity of piles using the Chin method.

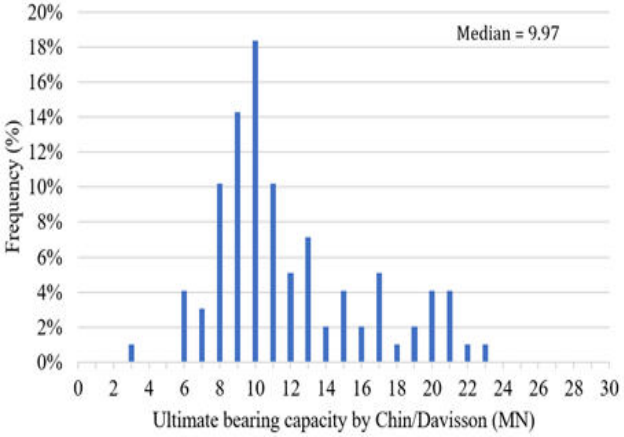


Figure 5. Distribution characteristics of ultimate bearing capacity of piles using the Chin/Davisson method.

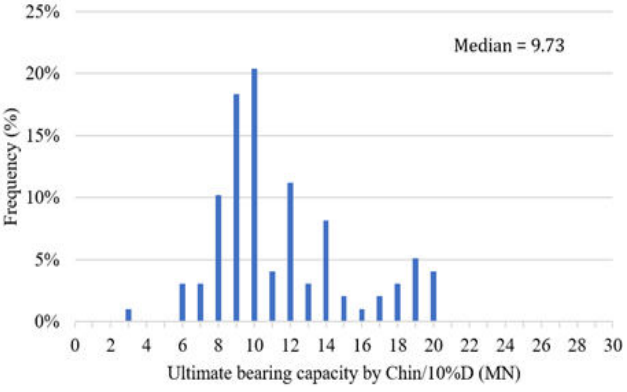


Figure 6. Distribution characteristics of ultimate bearing capacity of piles using the Chin/10%D method.

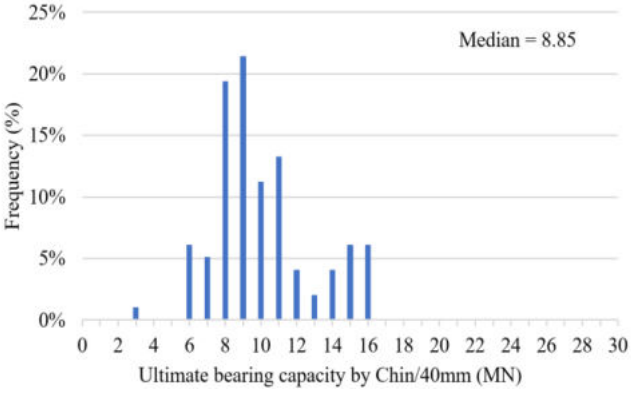


Figure 7. Distribution characteristics of ultimate bearing capacity of piles using the Chin/40 mm method.

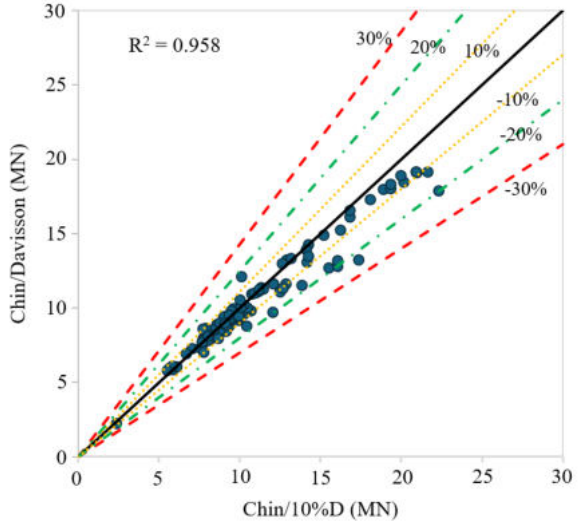


Figure 8. Comparison of measured load values from the Chin/Davisson method and Chin/10%D.

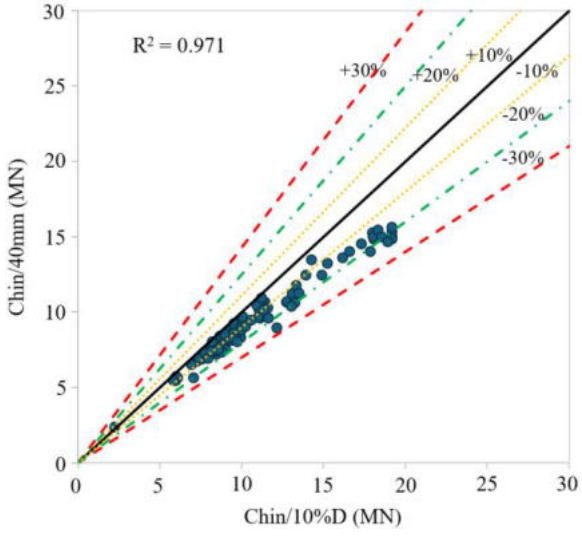


Figure 9. Comparison of measured load values from Chin/40 mm and Chin/10%D.

The total load capacity of the pile (Q_u) is equal to the sum of tip resistance (Q_p) and shaft resistance (Q_s), shaft resistance is divided into 2 components including the segment of PHC cylindrical pile (Q_s^c) and the nodular segment (Q_s^n).

$$Q_u = Q_p + Q_s = Q_p + (Q_s^c + Q_s^n) \quad (4)$$

Noting that the piles in the data set being surveyed are constructed through many layers of soil (such as sand, clay, clayey sand, sandy clay, and silt), the division of such a wide variety of soils is computational tedious nor is practical due to the limited number of survey piles. In general, the characteristics of survey soil fill are divided into two types, including sand and clay. So, the calculation formula is divided into six components as follows:

$$Q_u = Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6 \quad (5)$$

Where: Q_1 - the tip resistance in sandy layers; Q_2 - the tip resistance in clayey layers; Q_3 - shaft resistance along the cylindrical segment in sandy layers; Q_4 - shaft resistance along the cylindrical segment in clayey layers; Q_5 - shaft resistance along the nodular segment in sandy layers; Q_6 - shaft resistance along the nodular segment in clayey layers;

Note that when Q_1 has a non-zero value, Q_2 has a zero value and vice versa. Based on the tip resistance formula of the Hyper-MEGA JPC method $Q_p = \alpha \bar{N} A_p$, (where α - coefficient, $\bar{N}$ - the average SPT index of the soil beneath the pile tip, and A_p - the area of the pile's cross-section (details can be found in Homma 2014)). The tip resistance has a value equal to the product of a coefficient, the average SPT index of the soil under the tip of the pile and the cross section of the tip of the pile. Proceeding to transform the formula has the same meaning as follows:

$$Q_1 = X_1 \omega A_p N_1 \quad (6)$$

$$Q_2 = X_2 \omega A_p N_2 \quad (7)$$

Where: X_1 and X_2 - variables to be optimal; ω - the enlargement ratio, $\omega = D_e/D_s$ (D_e - the enlarged excavation diameter; $D_s = D_{on} + 0.05$(m), D_{on} is the diameter of PGPN pile); A_p - the area of pile tip; N_1 - the average N_{spt} of the sandy layers underneath the pile tip defined as:

$$N_1 = (N_u + 3N_L)/4 \le 20 \quad (8)$$

N_2 - the average N_{spt} of the clayey layers underneath the pile tip defined as:

$$N_2 = (N_u + 3N_L)/4 \le 20 \quad (9)$$

Where N_u - the average SPT index of the soil on the tip of the pile an interval of 2 m; N_L - the average SPT index of the soil under the tip of the pile an interval $[D_{on}+D_e]$.

Similar to the tip load capacity, the shaft load capacity has a value equal to the product of a coefficient, circumference, calculated pile segment length and average SPT index are presented as follows:

$$Q_3 = X_3 \psi L_3 N_3 \quad (10)$$

$$Q_4 = X_4 \psi L_4 N_4 \quad (11)$$

$$Q_5 = X_5 \psi L_5 N_5 \quad (12)$$

$$Q_6 = X_6 \psi L_6 N_6 \quad (13)$$

Where: X_3, X_4, X_5 and X_6 - variables to be optimal; Ψ - the circumference of the calculated segment pile; L_3 - the length of the cylindrical segment pile in the sandy layers; L_4 - the length of the cylindrical segment pile in the clayey layers; L_5 - the length of the nodular segment pile in the sandy layers; L_6 - the length of the nodular segment pile in the clayey layers; N_3 - the average SPT index of the cylindrical segment pile in the sandy layers, $N_3 \le 25$; N_4 - the average SPT index of the cylindrical segment pile in the clayey layers, $N_4 \le 15$; N_5 - the average SPT index of the nodular segment pile in the sandy layers, $N_5 \le 50$; N_6 - the average SPT index of the nodular segment pile in the sandy clayey, $N_6 \le 20$.

The formula is summarized as follows:

$$Q_u = X_1 Y_1 + X_2 Y_2 + X_3 Y_3 + X_4 Y_4 + X_5 Y_5 + X_6 Y_6 \quad (14)$$

In which X_1, X_2, X_3, X_4, X_5 and X_6 are variables to optimize and have values fluctuating in a range as follows:

$$150 \le X_1 \le 250;\ 150 \le X_2 \le 250;\ 5 \le X_3 \le 15;\ 5 \le X_4 \le 15;\ 5 \le X_5 \le 15;\ 5 \le X_6 \le 15$$

And Y_1, Y_2, Y_3, Y_4, Y_5 and Y_6 - a constant defined as follows:

$$Y_1 = \left(\omega A_p N_1\right)_{sand}^{tip} \quad (15)$$

$$Y_2 = \left(\omega A_p N_2\right)_{clay}^{tip} \quad (16)$$

$$Y_3 = \left(\psi L_3 N_3\right)_{sand}^{cylindrical} \quad (17)$$

$$Y_4 = \left(\psi L_4 N_4\right)_{clay}^{cylindrical} \quad (18)$$

$$Y_5 = \left(\psi L_5 N_5\right)_{sand}^{nodular} \quad (19)$$

$$Y_6 = \left(\psi L_6 N_6\right)_{clay}^{nodular} \quad (20)$$

Details of these constants are presented in Appendix A1.

3.2. Genetic algorithm approach for empirical formula development

In the pursuit of refining our empirical formula and achieving a heightened correlation coefficient, we turned to the power of a genetic algorithm a computational approach inspired by the mechanics of natural selection and genetics. This algorithmic framework enabled us to iteratively optimize the coefficients of our empirical formula, enhancing its accuracy and predictive capabilities.

In the context of genetic algorithms, the goal is to find the optimal solution to a problem by mimicking the process of natural evolution. This is achieved through a population of candidate solutions, which undergo processes akin to biological evolution, such as selection, crossover (recombination), and mutation.

Reproduction in genetic algorithms involves selecting chromosomes for the next generation based on their fitness. Various selection procedures exist, including proportional selection, ranking, and tournament selection, each employing different strategies for determining the probability of a chromosome's survival.

The genetic algorithm operates through a series of steps, mirroring the principles of evolution. The process can be summarized as follows:

3.2.1. Initialization

A population of potential solutions, known as "individuals," is generated. In our case, each individual corresponds to a unique set of coefficients for the empirical formula.

Formula: Let P={x_1, x_2,..., x_n}, represent the initial population, where x_i denotes an individual with a unique set of coefficients w=(w_1, w_2,..., w_m)

3.2.2. Evaluation

Each individual is evaluated based on a predefined fitness function. The fitness function quantifies how well the empirical formula with a particular set of coefficients aligns with the empirical measurements of axial load-bearing capacity. In our context, the fitness function captures the correlation between predicted and measured load-bearing capacities.

Fitness Function Formula:

$$f(x_1) = R^2 = 1 - \frac{\sum_{j=1}^{N}\left(y_j - \hat{y_j}\right)^2}{\sum_{j=1}^{N}\left(y_j - \bar{y_j}\right)^2} \quad (21)$$

where: y_j - the observed values; $\hat{y_j}$ - the predicted values from the empirical formula; $\bar{y}$ - the mean of the observed values; N - the number of data points.

3.2.3. Selection

Individuals with higher fitness values those that yield better correlations are more likely to be selected for the next generation. This mimics the concept of natural selection, favoring individuals with superior traits.

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^{N} f(x_j)} \quad (22)$$

where: $P(x_i)$ - the probability of selecting individual x_i ; $f(x_i)$ - is the fitness value of x_i.

3.2.4. Crossover and Mutation

Selected individuals undergo genetic operations, such as crossover and mutation, to create a new generation of solutions. Crossover involves combining traits from two parent individuals to create one or more offspring individuals. Mutation introduces small random changes in an individual's traits.

Crossover involves combining traits from two parent individuals to create one or more offspring individuals. This can be represented as:

$$Offspring_1=\alpha.Parent_1+(1-\alpha).Parent_2 \quad (23)$$

$$Offspring_2=(1-\alpha).Parent_1+\alpha.Parent_2 \quad (24)$$

where: α - a random crossover coefficient between 0 and 1.

Mutation introduces small random changes in an individual's traits. If x_i is a gene in an individual's chromosome, mutation can be represented:

$$X_i = X_i + \Delta x \quad (24a)$$

where: Δx - a small random perturbation.

Replacement:

The new generation replaces the previous generation, forming the basis for subsequent iterations. One common strategy is to replace the entire population with the new generation. Another strategy is elitism, where a certain number of the best individuals are carried over to the next generation to ensure the best solutions are retained.

Termination:

The algorithm iterates through multiple generations, continually improving the correlation coefficient of the empirical formula. Termination occurs when a predefined stopping criterion is met, such as a maximum number of iterations or a desired level of correlation.

The genetic algorithm's iterative nature refines the coefficients of the empirical formula over successive generations, converging towards an optimal solution coefficients that yield the highest correlation coefficient. This process harnesses the principles of natural selection and genetic variation to fine-tune the formula's accuracy and enhance its predictive capabilities.

Overall Process:

1. **Initialization**: Generate the initial population.
2. **Evaluation**: Calculate fitness for each individual.
3. **Selection**: Select individuals based on fitness.
4. **Crossover and Mutation**: Create new individuals.
5. **Replacement**: Form new generation.
6. **Termination**: Check if.

In summary, the genetic algorithm serves as a computational engine, dynamically optimizing the empirical formula's coefficients to achieve a higher correlation with empirical measurements of axial load-bearing capacity. This approach melds the power of mathematics and computation to elevate the precision of our method and its applicability to PGPN piles.

3.3. Mathematical representation of the new SPT direct method (Huynh et al. 2022)

To obtain the optimized variables, a genetic algorithm was applied to the dataset in Appendix A1, which includes constant values for $Y_1 \div Y_6$ and the target value Q_u for each pile, following the 40 mm method. The algorithm optimized variables $X_1 \div X_6$ to improve correlation coefficients and performance metrics for the new SPT direct method.

As above-mentioned, the variables $X_1 \div X_6$ were constrained as follows:

$$X_1, X_2: 150 \div 250$$

$$X_3 \div X_6: 5 \div 15$$

These constraints ensure practical and reliable results. Full data and constraints are provided in Appendix A1.

Table 2 presents the improved correlation coefficients and performance metrics achieved by the new SPT direct method.

Table 2. Results summary for variables X1 to X6 extracted via Genetic Algorithms.

Variables	X_1	X_2	X_3	X_4	X_5	X_6
Value	210	240	5.4	7.8	6.6	8.8

By substituting the values of variables X1 through X_6 into Equation (14), we derive the formula for estimating the ultimate bearing capacity of the nodular pile as follows:

$$Q_u=210Y_1+240Y_2+5.4Y_3+7.8Y_4+ 6.6Y_5+8.8Y_6 \quad (25)$$

The empirical formula is ultimately expressed in Equation (22) as follows:

$$Q_u = 210\ (\omega A_p N_1)_{sand}^{tip} + 240\ (\omega A_p N_2)_{clay}^{tip} + 5.4\ (\psi L_3 N_3)_{sand}^{cylindrical} + \quad (26)$$

$$7.8\,(\psi L_4 N_4)_{clay}^{cylindrical} + 6.6\,(\psi L_5 N_5)_{sand}^{no\ dular} + 8.8\,(\psi L_6 N_6)_{sand}^{cylindrical}$$

The strength of the proposed empirical formula given in Equation (22) lies in its simplicity, making it highly suitable for hand calculations. This ease of use ensures that practitioners can quickly and efficiently estimate the axial load-bearing capacity of PGPN piles without the need for complex computational tools. The streamlined nature of the formula minimizes the potential for errors during manual calculations and facilitates its adoption in field conditions where rapid decision-making is crucial. By providing a straightforward yet robust method for assessing pile capacity, the proposed formula enhances practical application in geotechnical engineering, streamlining the design process and improving overall efficiency. The reliability of this formula will be comprehensively assessed in the next section, further validating its utility and accuracy.

4. Performance assessment

Figure 10 shows a comparison between the proposed method and the measured value using the Chin/40 mm method. The scatter plot clearly demonstrates the superior performance of the proposed method in estimating the ultimate bearing capacity of PGPN piles. The correlation coefficient of 0.912 highlights the remarkable accuracy of the proposed method, even in cases where SPT data or soil conditions deviate from typical values. This high level of precision indicates that the proposed method can reliably predict the ultimate bearing capacity, closely matching the measured values. This accuracy is critical in ensuring safe and effective pile foundation design, particularly in non-typical geotechnical conditions. The robustness of the proposed method is demonstrated by its performance across a wide range of scenarios and geotechnical conditions included in the enriched dataset. This ensures that the method is versatile and adaptable, capable of providing accurate predictions under varying conditions, which is crucial for practical engineering applications.

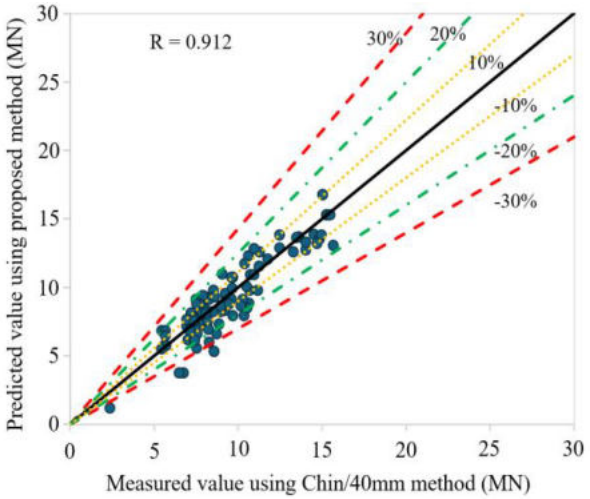


Figure 10. Comparison between the proposed method and the measured value using the Chin/40 mm method.

The comparison of the reliability of the proposed formula for calculating the ultimate axial load capacity against various established methods is summarized in Table 3. The reliability metrics span a range of methodologies, including Meyerhof (1976), Shioi & Fukui (1982), TCXD 205:1998, TCVN 10304:2014, Karkee et al. (1998), Horiguchi & Karkee (1995), JPC Basic (Homma 2014), JPC Hyper-Mega (Homma 2014), and the newly proposed formula. These methods were evaluated using ultimate bearing capacities determined through the Chin/40 mm, Chin/10%D, and Chin/Davisson approaches.

Table 3. Correlation coefficients comparing the reliability of the proposed formula with other methods.

Measured ultimate bearing capacity	Correlation coefficient (R)								
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Chin/ 40mm	0.533	0.541	0.551	0.722	0.856	0.872	0.810	0.835	0.912
Chin/ 10%D	0.518	0.521	0.536	0.698	0.847	0.860	0.795	0.820	0.903
Chin/ Davisson	0.515	0.521	0.536	0.675	0.815	0.820	0.743	0.764	0.871
[1] Meyerhof; [2] Shioi & Fukui; [3] TCXD 205:1998; [4] TCVN 10304:2014; [5] Karkee; [6] Horiguchi & Karkee; [7] JPC Basic; [8] JPC Hyper-Mega; [9] Proposed formula. Details shown in Appendix A2									

The comparison of the reliability of the proposed formula for calculating the ultimate axial load capacity against various established methods is summarized in Table 3. The reliability metrics span a range of methodologies, including

Meyerhof (1976), Shioi & Fukui (1982), TCXD 205:1998, TCVN 10304:2014, Karkee et al. (1998), Horiguchi & Karkee (1995), JPC Basic (Homma 2014), JPC Hyper-Mega (Homma 2014), and the newly proposed formula. These methods were evaluated using ultimate bearing capacities determined through the Chin/40 mm, Chin/10%D, and Chin/Davisson approaches.

The proposed formula consistently exhibits the highest reliability across all measurement methods (Chin/40 mm: 0.912, Chin/10%D: 0.903, Chin/Davisson: 0.871). This indicates superior performance in accurately predicting the ultimate axial load capacity of PGPN piles, reflecting the benefits of the enriched dataset and the advanced genetic algorithm approach. On the other hand, traditional methods such as Meyerhof, Shioi & Fukui, and TCXD 205:1998 show lower reliability scores across the board. This suggests that these methods may not capture the complex interactions in geotechnical conditions as effectively as the proposed approach.

4.1. Comparison of various measurement methods

The Chin/40 mm method shows slightly higher reliability values compared to Chin/10%D and Chin/Davisson, indicating it may be a more conservative measure for ultimate bearing capacity. The proposed formula significantly outperforms others here with a reliability of 0.912, highlighting its robustness. For Chin/10%D, reliability scores are marginally lower than Chin/40 mm but still showcase the proposed formula's high accuracy (0.903). The close values between Chin/40 mm and Chin/10%D suggest the proposed formula's adaptability and precision across different evaluation standards. Although the Chin/Davisson method generally yields the lowest reliability values, the proposed formula still leads with a reliability of 0.871, further underscoring its reliability and effectiveness, even under more stringent or varying conditions.

4.2. Robustness compared to JPC basic and JPC hyper-mega

The JPC Basic method, designed for PGPN piles, shows reliability scores of 0.810 (Chin/40 mm), 0.795 (Chin/10%D), and 0.743 (Chin/Davisson). These values are significantly lower than those of the proposed formula, indicating that the proposed method offers a more reliable and accurate prediction of the ultimate bearing capacity. This improvement can be attributed to the comprehensive dataset and the sophisticated genetic algorithm used. The JPC Hyper-Mega method, another advanced approach for PGPN piles, shows reliability scores of 0.835 (Chin/40 mm), 0.820 (Chin/10%D), and 0.764 (Chin/Davisson). While JPC Hyper-Mega outperforms JPC Basic, it still falls short of the proposed formula's reliability. The proposed formula's higher scores demonstrate its enhanced precision and robustness, likely due to the enriched dataset and optimization techniques employed. The substantial improvement over traditional methods, including the well-regarded JPC Basic and JPC Hyper-Mega, can be attributed to the enriched dataset, which captures a broader range of geotechnical scenarios. The use of a genetic algorithm refines empirical relationships with greater precision and allows for better adaptability across varying geotechnical conditions.

4.3. Discussion on the percentage of tip resistance obtained from the proposed formula

The histogram in Figure 11 illustrates the distribution of tip load capacity percentages obtained from the proposed formula, with a median value of 18.48%. This value is in line with the expected range for friction piles and reflects the typical contribution of tip resistance to the overall load-bearing capacity. In friction piles, the majority of the load is carried by skin friction along the shaft, rather than by end-bearing or tip resistance. The design of friction piles generally emphasizes the interaction between the pile surface and surrounding soil, which is well-represented by this moderate percentage of tip resistance.

The dataset used in this study encompasses various geotechnical conditions in Vietnam, where soils typically consist of soft to medium clays and loose to medium sands. These soils tend to promote significant shaft friction, which naturally results in a smaller contribution of tip resistance. The observed median of 18.48% aligns with this understanding. Furthermore, this value

is comparable to findings in geotechnical literature, which often report tip resistance ranging from 10÷30% of the total bearing capacity for friction piles in similar soil conditions. Thus, the median value supports the validity of the proposed formula.

From a distribution perspective, the histogram in Figure 11 shows a concentration of tip resistance percentages around the median, with a relatively symmetrical spread. This suggests that the proposed formula consistently estimates tip resistance across various case histories in the dataset. Appendix A3 further corroborates this by displaying the predicted tip resistance percentages for all case histories. The data show that the formula produces tip resistance values within a consistent range across different piles, and the absence of extreme values in both figures strengthens confidence in the formula's reliability. For practitioners, understanding the typical contribution of tip resistance is essential for designing efficient and safe pile foundations. The median value of 18.48% highlights that while tip resistance is a factor, it is not the dominant one-reaffirming the design principle that most of the load is carried by shaft friction. This insight aids in optimizing pile designs, ensuring that piles are neither overdesigned nor underdesigned.

In conclusion, the 18.48% median value for tip resistance obtained from the proposed formula is both reasonable and aligned with the expected behavior of friction piles, given the geotechnical conditions. This value not only corroborates findings from existing geotechnical literature but also reflects the typical load distribution for friction piles. The distribution shown in the histogram underscores the consistency and robustness of the proposed method in estimating ultimate bearing capacity, making it a valuable tool for geotechnical engineers.

The developed formula predicts that the minimal percentage of tip resistance, based on an analysis of 98 case histories, is around 13%. This minimal tip resistance typically occurs in short piles embedded in high bearing capacity soil layers, as exemplified by case history TP-2 of Project 4 (see Figure 12). In scenarios where piles are embedded in soils with moderate to good bearing capacities, characterized by SPT-N values ranging from 10÷20, the tip resistance is minimally mobilized. This is observed in case histories such as TP8-4 of Project 5 (Figure 13) and KT01 of Project 11 (Figure 14).

Conversely, the formula indicates that the maximal tip resistance is approximately 30%. This high level of tip resistance mobilization is observed in piles embedded in very soft soils, where the SPT-N value is less than 3, and the soft soil layer constitutes more than half of the entire embedment length. These conditions lead to a significant reliance on tip resistance for load-bearing capacity. Such scenarios are exemplified by the UTP1 pile of Project 7 (see Figure 15), the KT01 pile of Project 10 (see Figure 16), and the TP1 pile of Project 26 (see Figure 17).

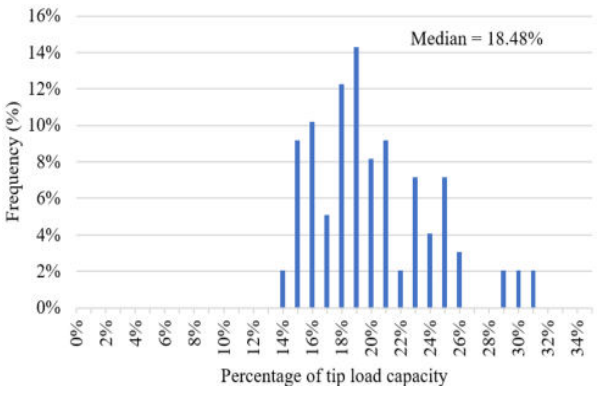


Figure 11. Distribution characteristics of pile tip load-bearing capacity.

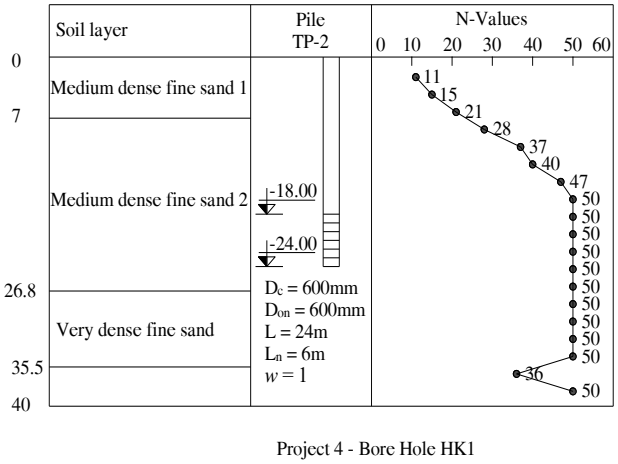
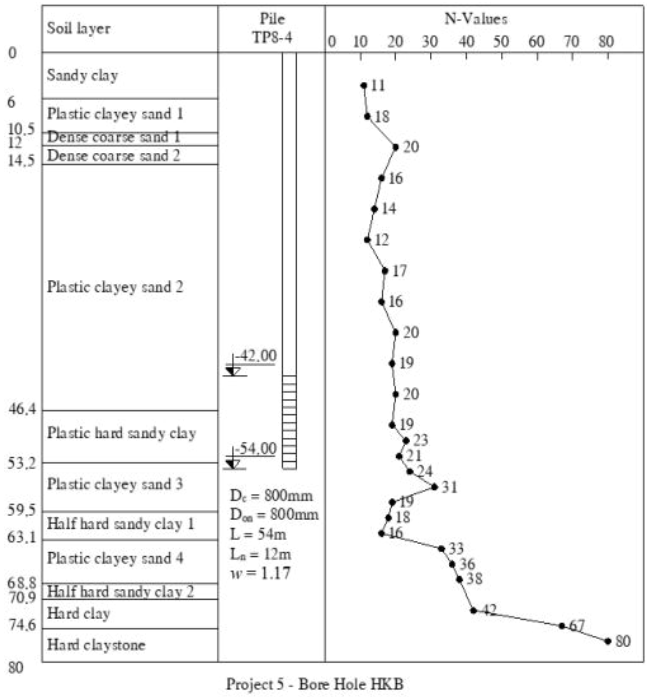
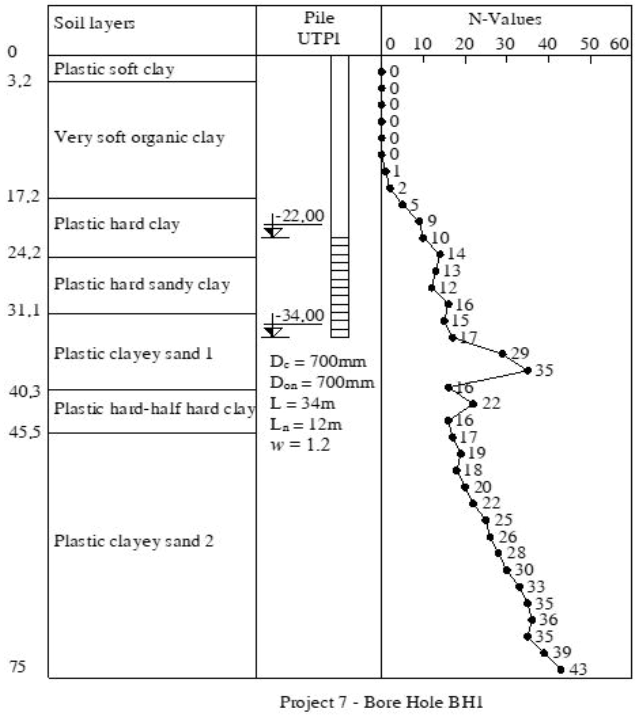


Figure 12. Characteristics of borehole 1 and characteristics of TP-2 pile of project 4.



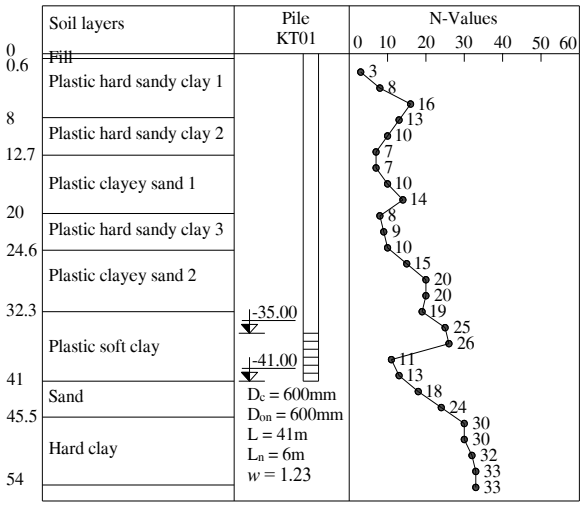
Project 5 - Bore Hole HKB

Figure 13. Characteristics of borehole HKB and characteristics of pile TP8-4 project 5.



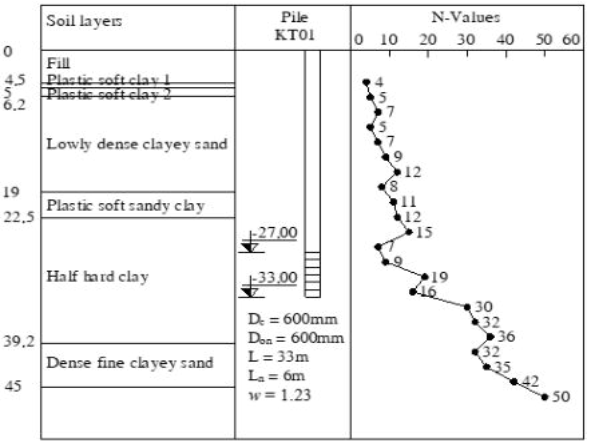
Project 7 - Bore Hole BH1

Figure 15. Characteristics of borehole BH1 and characteristics of UTP1 pile of project 7.



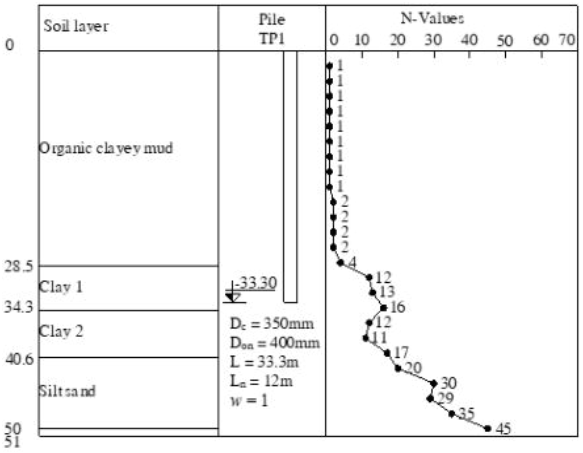
Project 11 - Bore Hole HK3

Figure 14. Characteristics of borehole HK3 and characteristics of KT01 pile of project 11.



Project 10 - Bore Hole HK2

Figure 16. Characteristics of borehole HK2 and characteristics of KT01 pile of project 10.



Project 26 - Bore Hole HK1

Figure 17. Characteristics of borehole HK1 and characteristics of TP1 pile of project 26.

The formula's ability to predict these variations in tip resistance, from minimal in strong soils to maximal in very soft soils, highlights its robustness and adaptability to different geotechnical environments.

The range of tip resistance percentages highlights the reliability and accuracy of the proposed formula. It effectively captures the interaction between subsurface conditions and the pile's shaft and tip resistance, reflecting the varying soil conditions encountered in the case histories. This variability demonstrates the formula's robustness in different geotechnical contexts, thereby supporting its practical applicability in predicting pile performance.

4.4. Discussion on the filed data of Zhou et al. (2021)

To further validate the reliability and robbusness of the proposed formula, we applied it to field data from recent studies. Notably, Zhou et al. (2021) provided valuable data from pile load tests on two instrumented piles, TP1 and TP2. This data serves as an excellent benchmark for assessing the accuracy of the proposed formula. Both piles, TP1 and TP2, exhibit similar physical characteristics, including dimensions and materials, as depicted in Figure 18. The soil profile for these piles consists of layers of clay and sand with varying densities and strengths. Standard Penetration Test (SPT) values indicate medium to dense sand and stiff clay layers, which are typical for friction piles where shaft friction is the primary mechanism of load resistance. The results of the static load tests, shown in Figure 19, highlight notable differences in performance despite the similarities in pile and soil characteristics. These differences become particularly pronounced when the displacement exceeds 40 mm, suggesting variations in load transfer and soil-pile interaction under higher load conditions. Specifically, the measured ultimate load capacities for TP1 and TP2 are 6204.9 kN and 6829.9 kN, respectively. However, when considering the pile head displacement of 40 mm, the measured ultimate load-bearing capacities of both piles are comparable, at approximately 5569 kN.

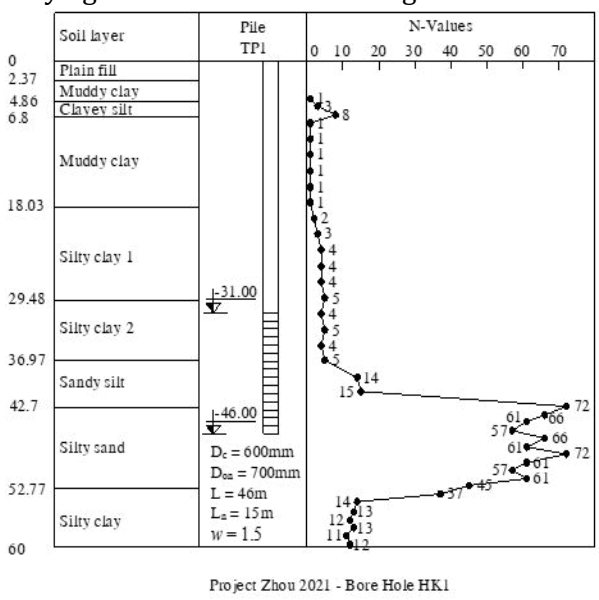


Figure 18. Characteristics of borehole HK1 and characteristics of TP1 pile of Zhou et al. (2021)'s study.

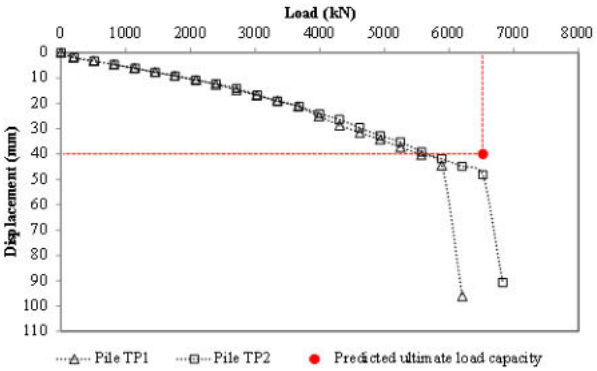


Figure 19. Results of static load tests on testing piles TP1 and TP2.

4.5. Application of proposed formula

The proposed formula's predicted ultimate load capacity, calculated as 6535 kN (see below), closely matches the measured value for pile TP2 (6829.9 kN) from Zhou et al. (2021), confirming its reliability and effectiveness.The total load capacity calculated as 6535 kN is shown below.

Tip load capacity:

$$Q_p=210Y_1+240Y_2=210x6.85+0=1438.5\ kN$$

Shaft load capacity:

$$Q_s= 5.4Y_3 + 7.8Y_4 + 6.6Y_5 + 8.8Y_6 = (5.4x0) + (7.8x115) + (6.6x336.94) + (8.8x224.52)$$

$$= 5096.5\ kN$$

Here, Y1, Y2, Y3, Y4, Y5, Y6 are derived from Table 4.

Table 4. Y1÷Y6 values of Zhou (2021) project calculated according to the proposed formula.

Y_1	Y_2	Y_3	Y_4	Y_5	Y_6
6.85	0	0	115	336.94	224.52

Total load capacity:

$$Q_u=Q_p+Q_s=1438.5+5096.5=6535\ kN$$

The proposed formula's predicted ultimate load capacity (6535 kN) closely matches the measured value for pile TP2 (6829.9 kN) from Zhou et al. (2021), confirming its reliability and effectiveness. Although there is a slight discrepancy in the predicted capacity for TP1, this is expected given the observed differences in test results for displacements over 40 mm. The formula's accurate prediction for TP2, despite significant differences in similar test conditions, demonstrates its precision, flexibility, and adaptability across various geotechnical environments and pile designs.

To sum up, the strong correlation with Zhou et al. (2021) field data validates the proposed method, suggesting it could become a new standard in geotechnical engineering. The formula's ability to accurately predict ultimate load capacities ensures safer and more efficient pile foundation designs. By utilizing enriched datasets and advanced genetic algorithms, the proposed method significantly improves upon traditional approaches, offering valuable insights for both practitioners and researchers.

5. Conclusion

This study presented a robust and reliable formula for predicting the ultimate bearing capacity of PGPN piles. The proposed formula, validated using an enriched dataset and advanced genetic algorithms, consistently exhibited higher reliability across various measurement methods. It has been successfully demonstrated with significant improvement over traditional methods. This enhanced performance indicated its superior capability in accurately predicting pile load capacities, thereby contributing to safer and more efficient pile foundation design.

Quantitatively, the proposed formula achieved impressive reliability scores across different measurement methods. For the Chin/40 mm method, it achieved a reliability score of 0.912, significantly higher than JPC Basic's 0.810 and JPC Hyper-Mega's 0.835. Similarly, for the Chin/10%D method, the proposed formula showed a reliability of 0.903, surpassing JPC Basic's 0.795 and JPC Hyper-Mega's 0.820. Even for the more stringent Chin/Davisson method, the proposed formula's reliability score of 0.871 outperformed JPC Basic's 0.743 and JPC Hyper-Mega's 0.764. These results highlight the robustness and consistency of the proposed method.

The reliability of the proposed formula was further confirmed through validation with field data from Zhou et al. (2021). In this case study, two instrumented piles (TP1 and TP2) were tested, with measured ultimate load capacities of 6204.9 kN and 6829.9 kN, respectively. The proposed formula accurately predicted the ultimate load capacity, especially for TP2, where the predicted value was consistent with the measured value. This agreement between the proposed method and field data underscores its flexibility and applicability to different locations and geotechnical conditions, demonstrating its practical utility and reliability.

The analysis of the percentage of tip resistance obtained using the proposed formula shows a median of 18.48%, which is reasonable for friction piles within the dataset. This median value aligns well with expected values for similar geotechnical conditions, further validating the proposed method's accuracy. The consistency of these results suggests that the proposed formula effectively captures the essential characteristics of pile-soil interaction, making it a reliable tool for predicting the ultimate bearing capacity.

In conclusion, the proposed formula, leveraging enriched datasets and advanced genetic algorithms, offers a significantly enhanced and reliable prediction method for the ultimate bearing capacity of PGPN piles. Its superior performance, validated by both quantitative evaluation and field data, positions it as a valuable tool for practitioners and researchers in geotechnical engineering. The proposed formula's ability to effectively address the complexities of

geotechnical conditions ensures more accurate and dependable predictions of ultimate bearing capacity, thereby contributing to safer and more efficient pile foundation design.

6. Limitations and Future Work

One limitation of this study is the lack of cross-validation to thoroughly assess the robustness of the genetic algorithm (GA) across different subsets of data. While we mitigated the risk of overfitting by using a diverse dataset and incorporating early stopping and fitness diversity mechanisms, the absence of formal cross-validation means that the model's ability to generalize to unseen data may not have been fully evaluated. Additionally, the formula's simplicity, though practical, may encounter limitations in more heterogeneous or challenging soil conditions, where complex interactions are not fully captured by an empirical formula.

For future work, implementing k-fold cross-validation or other validation techniques will be essential for rigorously testing the performance and robustness of the GA. Expanding the dataset to include more cases from different regions and geotechnical conditions would also enhance the model's applicability. Furthermore, we plan to explore comparative testing with computational models, such as finite element methods (FEM), to validate the formula's performance in more complex environments. Lastly, we aim to investigate hybrid methods that combine genetic algorithms with other machine learning techniques, which could offer improved accuracy while maintaining the formula's simplicity and practical usability.

Acknowledgements

This research is funded by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number NCUD.02-2022.08. This funding is gratefully acknowledged. The authors also thank the anonymous reviewers for their valuable and constructive comments, which significantly improved the manuscript. Finally, the kind handling of this manuscript by the editors is deeply appreciated.

Contributions of authors

Tan Nguyen - conceptualization, validation, investigation, writing original draft, writing review & editing, supervision, project administration, fund acquisition; Tuetakoun Aphistih - methodology, software, formal analysis, visualization, investigation, data curation; Minh The Quang Nguyen - data curation, formal analysis, writing review & editing; Duy Khuong Ly - methodology, writing review & editing, supervision; Jim Shiau - writing review & editing.

References

Abu-Farsakh, M. Y., Yu, X., Yoon, S., & Tsai, C. (2010). *Calibration of resistance factors needed in the LRFD design of drilled shafts* (No. FHWA/LA. 10/470). Louisiana Transportation Research Center.

Chin, F. K. (1970). Estimation of the ultimate load of piles from tests not carried to failure. In *Proc. 2nd Southeast Asian Conference on Soil Engineering, Singapore, 1970*.

Fang, P., Xie, X., & Qi, J. (2014). Engineering character of a new-style pretensioned spun concrete nodular pile. In *Advances in Soil Dynamics and Foundation Engineering* (pp. 404-413).

Homma, Y. (2014). Introduction of base enlarged pre-boring method with nodular pile. In *The 1st workshop on new pile foundation technologies in Vietnam, Ho Chi Minh University of Technology, Vietnam National University-Ho Chi Minh city.*

Horiguchi, T., and Karkee, M. B. (1995). Load Tests on Bored Phc Nodular Piles in Different Ground Conditions and the Bearing Capacity Based on Simple Soil Parameters. *AIJ Journal of Technology and Design*, 1(1). 89-94. doi:10. 3130/aijt. 1.89.

Huynh, V. H., Nguyen, T., Nguyen, D. P., Nguyen, T. S. & Nguyen, T. C., (2022). A novel direct SPT method to accurately estimate ultimate axial bearing capacity of bored PHC nodular piles with 81 case studies in Vietnam. *Soils and Foundations,* 62(4). p.101163. doi.org/10. 1016/j.sandf.2022. 101163

Karkee, M. B., Kanai, S. & Horiguchi, T. (1998). Quality assurance in bored PHC nodular piles through control of design capacity based on loading test data. In *Proceedings of the 7th International Conference and Exhibition, Piling and Deep Foundations* (Vol. 1, No. 24, pp. 1-9).

Kobayashi, K. & Ogura, H. (2007). Vertical bearing capacity of bored pre-cast pile with enlarged base considering diameter of the enlarged excavation around pile toe. In *Advances in Deep Foundations* (pp. 289-296). CRC Press.

Lambora, A., Gupta, K. & Chopra, K. (2019). Genetic algorithm-A literature review. In *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)*. (pp. 380-384). IEEE.

Meyerhof, G. G. (1976). Bearing capacity and settlement of pile foundations. *Journal of the Geotechnical Engineering Division*, 102(3). 197-228.

Nguyen, T., Ly, K.-D., Nguyen-Thoi, T., Nguyen, B.-P. & Doan, N.-P. (2022). Prediction of axial load bearing capacity of PHC nodular pile using Bayesian regularization artificial neural network. *Soils and Foundations*, 62(5). 101203. doi:10.1016/j.sandf.2022.101203.

Shioi, Y. & Fukui, J. (2021). Application of N-value to design of foundations in Japan. In *Penetration Testing, volume 1* (pp. 159-164). Routledge.

TCVN 10304:2014. Pile foundation - Design Standard. Vietnam national standard.

TCVN 205:1998. Pile foundation - Specifications for design. Vietnam national standard.

Wang, Z. J., Zhang, R. H., Xie, X. Y., Fang, P. F., Zheng, L. W., Li, J. Z., & Zhu, D. Y. (2019). Field tests and simplified calculation method for static drill rooted nodular pile. *Advances in Civil Engineering,* 2019(1), 1-13. doi:10.1155/2019/5841840.

Yoshimi, Y. & Tokimatsu, K. (1983). SPT Practice Survey and Comparative Tests. *Soils and Foundations,* 23(3). 105-111. doi:10.3208/sandf 1972.23.3_105.

Yu, J. L., Zhou, J. J., Gong, X. N., & Zhang, R. H. (2021). Shaft capacity of prestressed high strength concrete (PHC) pile-cemented soil column embedded in clayey soil. *Soils and Foundations,* 61(4), 1086-1098. doi:10.1016 /j.sandf.2021.05.006.

Zhou, J. J., Yu, J. L., Gong, X. N., El Naggar, M. H., & Zhang, R. H. (2021). Field study on the behavior of pre-bored grouted planted pile with enlarged grout base. *Acta Geotechnica, 16*, 3327-3338. doi.org/10. 1007/s11440-021 - 01208-7.

Zhou, J., Yu, J., Gong, X., El Naggar, M. H., & Zhang, R. (2020). The effect of cemented soil strength on the frictional capacity of precast concrete pile-cemented soil interface. *Acta Geotechnica*, 15(11). 3271-3282. doi:10.1007/s11440-020-00915-x.

Appendix

Appendix A1.

Table 5. Summary of constant values Y_1, Y_2, Y_3, Y_4, Y_5, Y_6 and target value Q_u according to the 40mm method.

Pile	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Q_u
	kN	kN	kN	kN	kN	kN	kN
1	6.96	0.00	392.82	139.20	505.17	0.00	7743.62
2	6.96	0.00	392.82	139.20	505.17	0.00	7267.83
3	6.18	0.00	508.94	127.11	331.75	0.00	8030.19
4	6.18	0.00	508.94	127.11	331.75	0.00	6906.47
5	6.18	0.00	508.94	127.11	331.75	0.00	7203.66
6	0.00	11.83	761.38	111.72	530.80	160.85	13998.55
7	0.00	6.96	49.20	342.12	73.32	177.47	8698.68
8	0.00	6.96	62.20	339.29	135.15	213.00	8150.59
9	0.00	6.96	62.20	339.29	135.15	213.00	7613.09
10	0.00	6.96	49.20	342.12	73.32	177.47	7590.00
11	5.65	0.00	844.70	0.00	334.27	0.00	10356.12
12	5.65	0.00	725.08	0.00	565.49	0.00	10638.89
13	0.00	5.65	938.71	0.00	47.12	207.35	10517.43
14	0.00	5.65	844.46	0.00	235.62	131.95	10202.95
15	5.65	0.00	635.23	0.00	565.49	0.00	8869.55
16	5.65	0.00	635.23	0.00	565.49	0.00	9195.62
17	0.00	5.65	962.27	98.96	0.00	226.19	9263.47
18	0.00	5.65	962.27	98.96	0.00	226.19	8421.83
19	0.00	5.65	962.27	98.96	0.00	226.19	8038.90
20	11.83	0.00	701.20	524.60	237.50	197.79	14838.09
21	11.83	0.00	701.20	524.60	237.50	197.79	13590.86
22	0.00	11.83	1308.16	165.88	445.98	138.48	15249.85
23	11.83	0.00	1526.45	165.88	266.41	341.81	15003.20
24	11.83	0.00	1526.45	165.88	266.41	341.81	15007.42
25	0.00	11.83	1308.16	165.88	445.98	138.48	15303.19
26	0.00	11.83	1308.16	165.88	445.98	138.48	15451.34
27	0.00	6.96	404.70	275.55	199.33	282.74	10676.56
28	0.00	6.96	404.70	275.55	199.33	282.74	8532.55
29	0.00	6.96	580.57	275.55	27.80	433.54	10725.66
30	0.00	6.96	404.70	275.55	199.33	282.74	11126.47
31	0.00	6.96	580.57	275.55	27.80	433.54	10952.87
32	0.00	6.96	409.15	171.91	211.49	286.51	8127.15
33	0.00	6.96	409.15	171.91	211.49	286.51	7508.28
34	0.00	6.96	409.15	171.91	211.49	286.51	9596.22
35	0.00	6.96	409.15	171.91	211.49	286.51	8298.26
36	0.00	6.96	409.15	171.91	211.49	286.51	10274.60
37	0.00	5.65	599.42	334.93	0.00	0.00	7447.50
38	5.65	0.00	486.96	275.55	0.00	0.00	8291.94
39	0.00	6.96	239.88	171.91	365.00	98.02	8178.13
40	7.62	0.00	0.00	97.64	102.04	264.70	8577.11
41	11.83	0.00	15.08	344.53	524.27	153.31	10464.51
42	11.83	0.00	0.00	239.26	352.61	175.18	8937.89
43	7.62	0.00	0.00	97.64	102.04	264.70	8561.67
44	11.83	0.00	0.00	376.72	197.04	323.71	9252.03
45	11.83	0.00	10.05	282.20	161.85	332.96	9221.47
46	11.83	0.00	50.27	703.57	665.85	140.74	14506.31
47	11.83	0.00	114.10	623.71	602.05	109.33	12479.04
48	11.83	0.00	517.94	481.21	395.21	167.13	15643.65
49	10.79	0.00	364.17	364.98	350.98	248.31	11229.58

Pile	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Q_u
	kN	kN	kN	kN	kN	kN	kN
50	0.00	6.96	582.11	189.44	18.10	203.58	8489.10
51	0.00	6.96	755.68	220.54	0.00	226.19	7918.28
52	0.00	6.96	685.78	254.47	45.24	188.50	8604.25
53	0.00	6.96	605.48	243.16	45.24	169.65	8971.23
54	0.00	6.96	493.44	254.47	232.48	37.70	7213.33
55	11.83	0.00	0.00	654.16	381.52	248.31	10597.50
56	11.83	0.00	0.00	665.14	339.29	249.76	11770.90
57	11.83	0.00	0.00	613.24	315.16	278.22	10380.33
58	0.00	6.96	182.68	184.25	0.00	165.88	7535.00
59	0.00	6.96	649.24	146.46	0.00	226.19	8128.73
60	5.83	0.00	410.70	368.51	0.00	188.50	9683.37
61	0.00	6.96	220.16	361.72	149.29	21.49	7117.66
62	4.24	0.00	351.48	122.05	0.00	0.00	6745.42
63	4.24	0.00	351.48	122.05	0.00	0.00	6502.58
64	0.00	9.31	143.63	814.30	0.00	482.55	14835.75
65	0.00	7.24	0.00	679.84	0.00	331.75	9644.61
66	0.00	4.26	0.00	422.80	0.00	211.12	6982.86
67	0.00	4.09	0.00	422.80	0.00	211.12	7368.23
68	0.00	4.09	0.00	422.80	0.00	211.12	7402.75
69	0.00	7.84	0.00	355.33	0.00	351.86	7397.61
70	0.00	4.61	0.00	304.56	0.00	263.89	5570.70
71	0.00	4.61	0.00	304.56	0.00	263.89	5717.07
72	0.00	7.84	0.00	355.33	0.00	351.86	7331.17
73	5.65	0.00	777.85	183.41	0.00	0.00	5466.95
74	5.65	0.00	777.85	183.41	0.00	0.00	5659.98
75	4.59	0.00	866.11	167.20	0.00	0.00	8249.88
76	5.65	0.00	743.25	275.02	0.00	0.00	8839.41
77	0.00	11.83	798.97	189.50	291.54	311.65	14000.49
78	0.00	11.83	953.03	107.32	269.17	346.83	13446.41
79	11.83	0.00	151.59	450.97	633.35	0.00	9041.33
80	11.83	0.00	129.20	450.97	613.24	0.00	9679.02
81	11.83	0.00	129.20	450.97	613.24	0.00	9585.64
82	0.00	11.83	0.00	730.36	0.00	603.19	14968.30
83	0.00	11.83	0.00	730.36	0.00	603.19	12468.41
84	5.65	0.00	376.99	575.23	0.00	0.00	6915.70
85	5.65	0.00	447.68	503.21	0.00	0.00	7353.29
86	5.65	0.00	637.11	456.29	0.00	0.00	8057.55
87	10.05	0.00	190.66	393.88	969.47	0.00	13242.14
88	10.05	0.00	217.90	450.15	969.47	0.00	14660.26
89	10.05	0.00	190.66	393.88	969.47	0.00	11258.32
90	10.05	0.00	376.07	124.53	186.99	277.72	7714.56
91	10.05	0.00	376.07	124.53	186.99	277.72	7539.17
92	10.05	0.00	376.07	124.53	186.99	277.72	9460.89
93	0.00	11.83	799.10	159.09	183.47	366.94	10933.02
94	0.00	6.96	626.89	99.15	136.85	218.65	10139.13
95	0.00	5.65	81.43	600.83	0.00	0.00	5662.78
96	0.00	1.47	0.00	111.53	0.00	0.00	2364.78
97	0.00	6.96	368.70	174.36	312.71	56.55	7371.36
98	5.23	0.00	673.56	97.22	0.00	0.00	5439.15

Appendix A2. Regression Plot Demonstrating the Performance of Other Empirical Methods

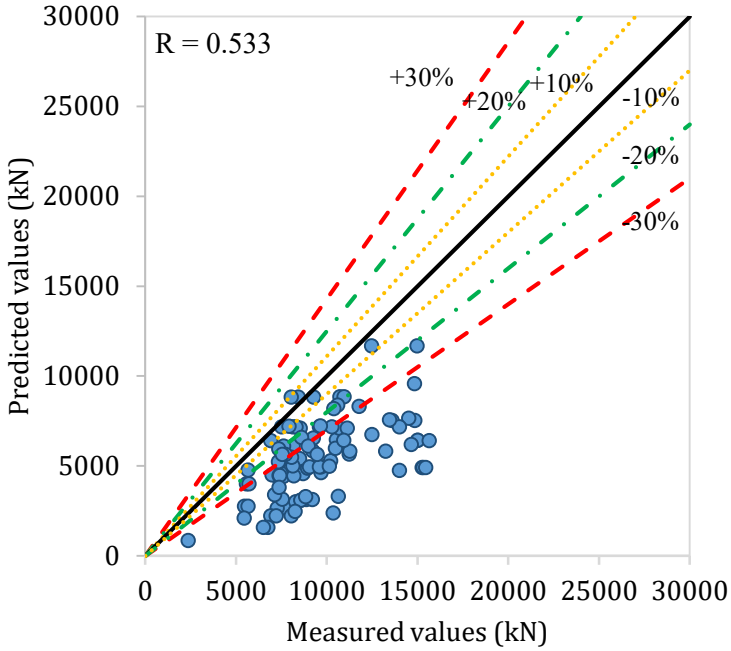


Figure 20. Comparison between the predicted values using Meyerhof (1976)'s method and measured values using the Chin/40 mm method.

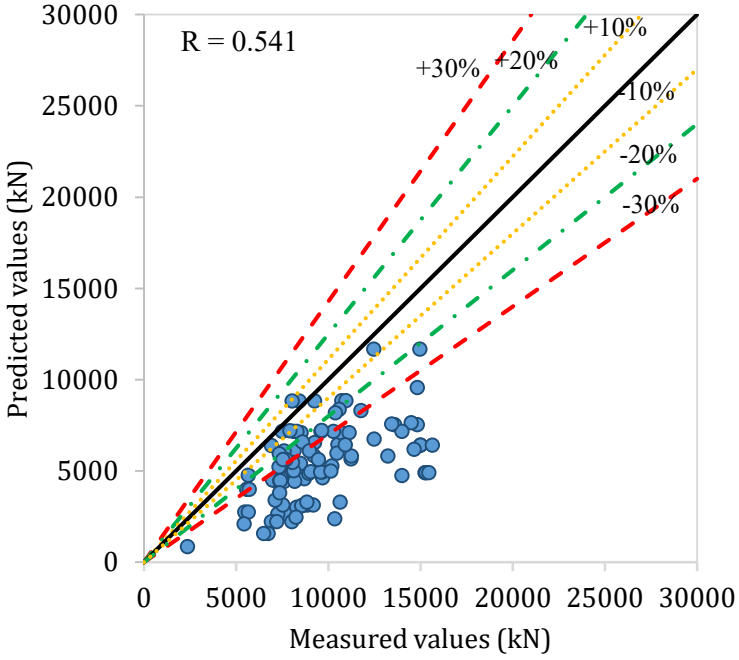


Figure 21. Comparison between the predicted values using Shioi & Fukui (1982)'s method and measured values using the Chin/40 mm method.

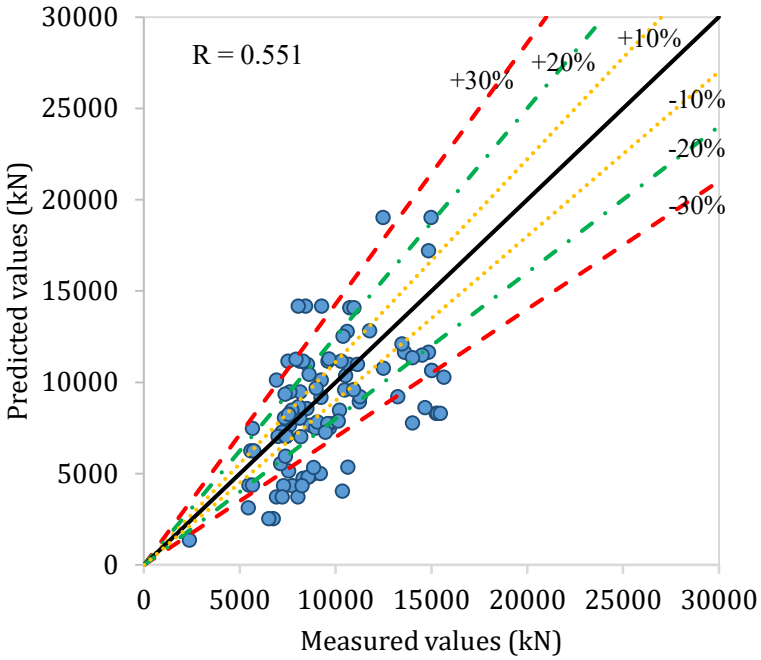


Figure 22. Comparison between the predicted values using TCVN 205:1998 and measured values using the Chin/40 mm method.

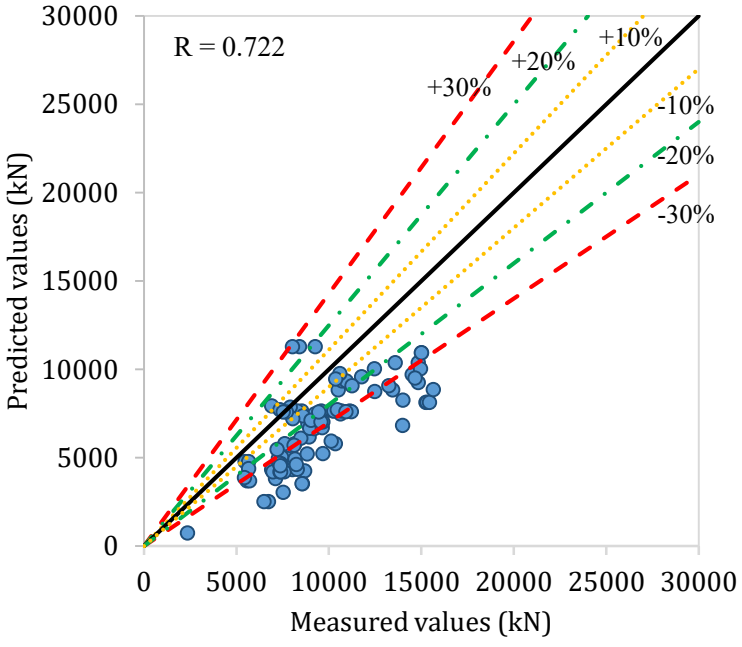


Figure 23. Comparison between the predicted values using TCVN 10304:2014 and measured values using the Chin/40 mm method.

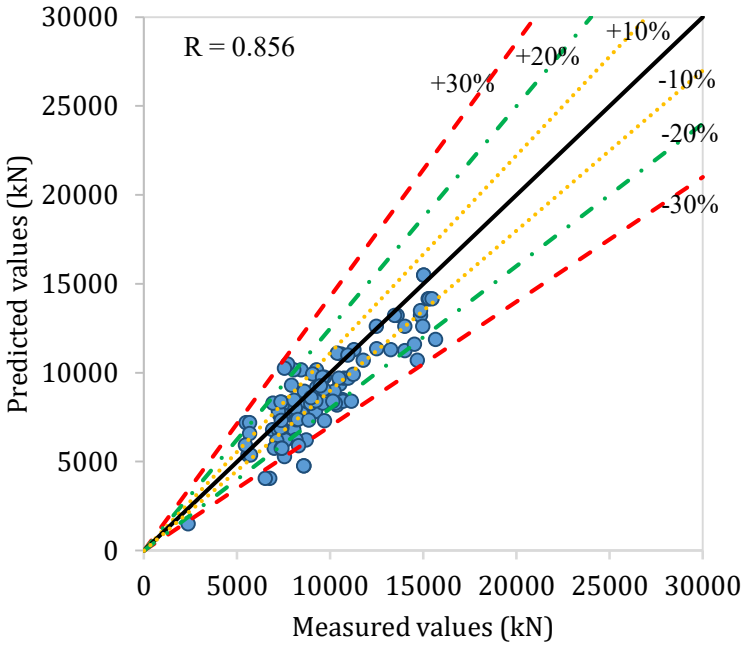


Figure 24. Comparison between the predicted values using Horiguchi and Karkee (1995)'s method and measured values using the Chin/40 mm method.

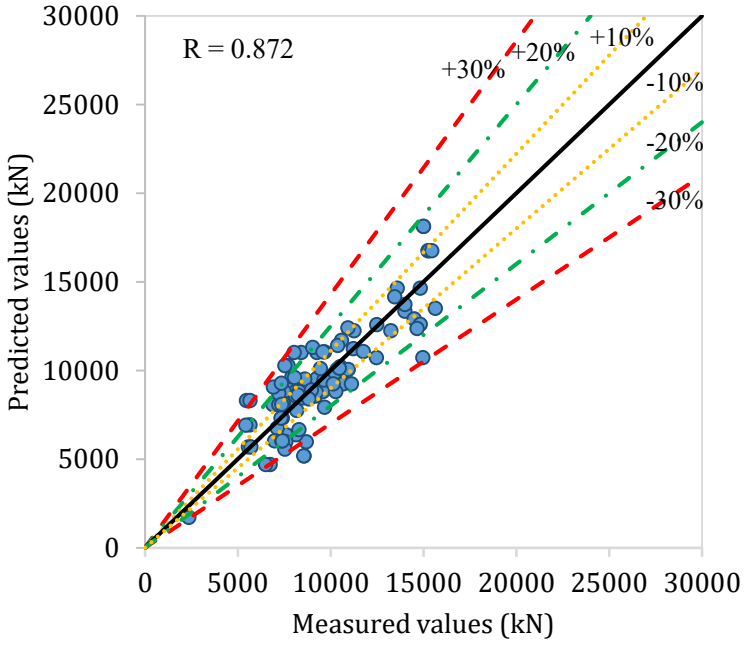


Figure 25. Comparison between the predicted values using Karkee et al. (1998) and measured values using the Chin/40 mm method.

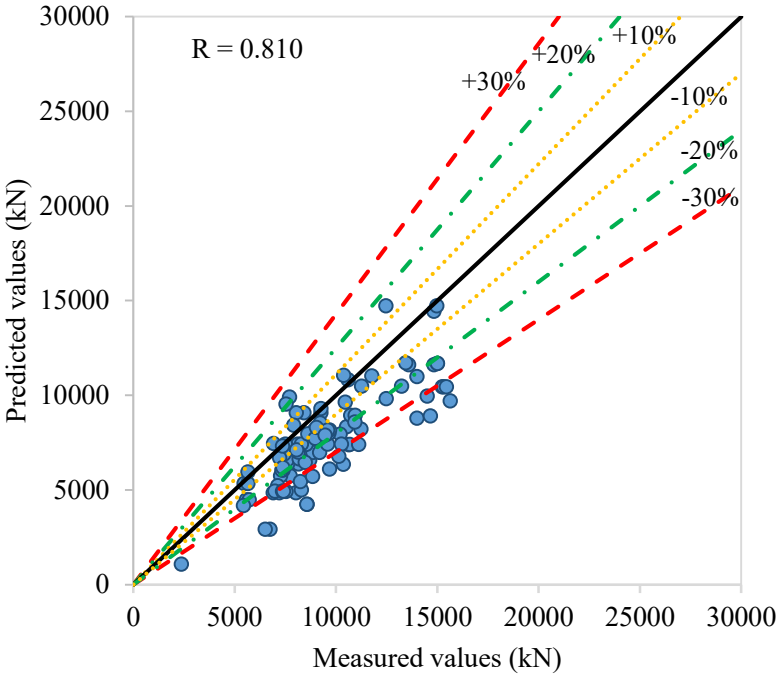


Figure 26. Comparison between the predicted values using JPC basic formula (Homma, 2014) and measured values using the Chin/40 mm method.

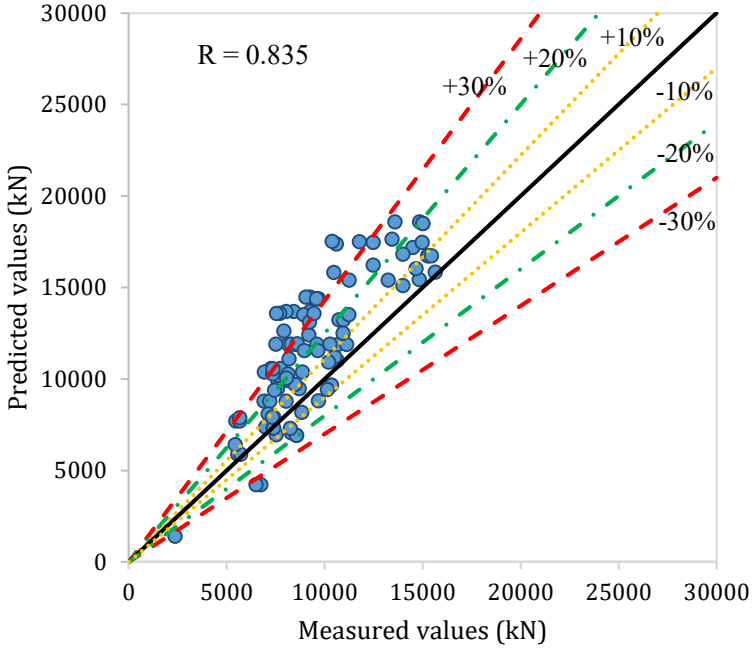


Figure 27. Comparison between the predicted values using JPC hyper-MEGA (Homma, 2014) and measured values using the Chin/40 mm method.

Appendix A3

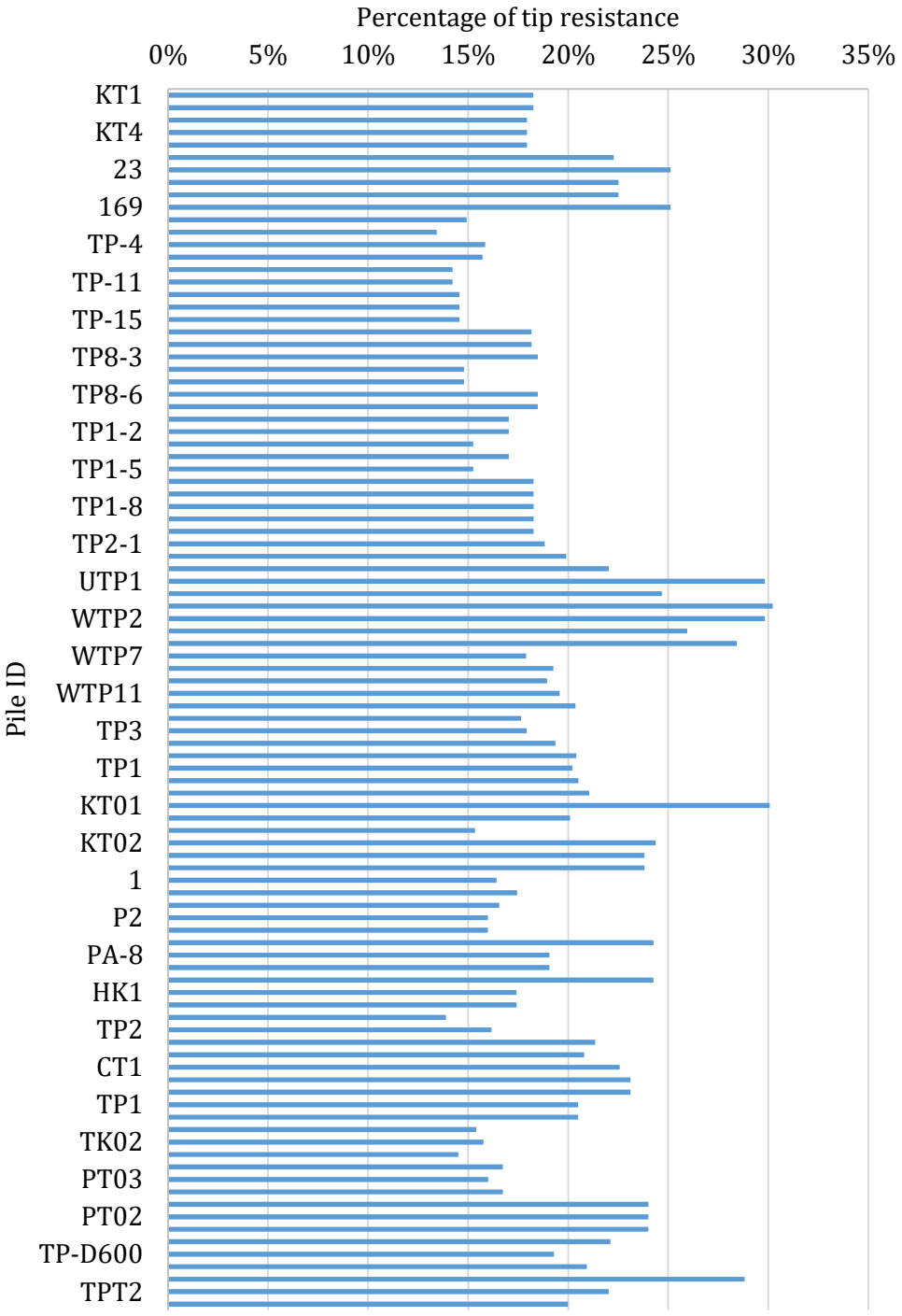


Figure 28. Percentage of Tip Resistance in All Case Histories Predicted by the Proposed Formula.