Geomechanics and Engineering, *Vol. 28, No. 4 (2022) 397-404*
https://doi.org/10.12989/gae.2022.28.4.397

Development of an optimized model to compute the undrained shaft friction adhesion factor of bored piles

Saif Alzabeebee*[1], Ali Adel Zuhaira[2] and Rwayda Kh. S. Al-Hamd[3]

[1]*Department of Roads and Transport Engineering, College of Engineering, University of Al-Qadisiyah, Al-Qadisiyah, Iraq*
[2]*Technical Institute of Al-Najaf, Al-Furat Al-Awsat Technical University, Iraq*
[3]*School of Applied Sciences, Abertay University, Dundee, UK*

(Received September 6, 2021, Revised December 16, 2021, Accepted January 9, 2022)

Abstract. Accurate prediction of the undrained shaft resistance is essential for robust design of bored piles in undrained condition. The undrained shaft resistance is calculated using the undrained adhesion factor multiplied by the undrained cohesion of the soil. However, the available correlations to predict the undrained adhesion factor have been developed using simple regression techniques and the accuracy of these correlations has not been thoroughly assessed in previous studies. The lack of the assessment of these correlations made it difficult for geotechnical engineers to select the most accurate correlation in routine designs. Furthermore, limited attempts have been made in previous studies to use advanced data mining techniques to develop simple and accurate correlation to predict the undrained adhesion factor. This research, therefore, has been conducted to fill these gaps in knowledge by developing novel and robust correlation to predict the undrained adhesion factor. The development of the new correlation has been conducted using the multi-objective evolutionary polynomial regression analysis. The new correlation outperformed the available empirical correlations, where the new correlation scored lower mean absolute error, mean square error, root mean square error and standard deviation of measured to predicted adhesion factor, and higher mean, $\boldsymbol{a20-index}$ and coefficient of correlation. The correlation also successfully showed the influence of the undrained cohesion and the effective stress on the adhesion factor. Hence, the new correlation enhances the design accuracy and can be used by practitioner geotechnical engineers to ensure optimized designs of bored piles in undrained conditions.

Keywords: evolutionary computing; statistical assessment; undrained adhesion factors; undrained shaft capacity

1. Introduction

Piled foundations are routinely used to support buildings, bridges, railways, and railway stations when the soil beneath these structures is weak. However, the correlations available in the literature to predict the capacity of bored piles are of empirical nature. These correlations have also been proposed using simple regression analysis (Dias and Grippon 2017, Shin *et al.* 2017).

Undrained shaft resistance of bored piles (fs) is routinely estimated as a percentage of the undrained cohesion of the soil (Su) using Eq. (1). The factor employed to judge the percentage of the undrained cohesion which contributes to the shaft friction is called the undrained adhesion factor (α) (Robert 1997). However, there are different correlations proposed in the literature to estimate the undrained adhesion factor; these correlations proposed by Weltman and Healy (1978), Sladen (1992), Chen and Kulhawy (1994), Coduto (1994), and Cherubini (1999). Weltman and Healy (1978) collected data of bearing capacity of piles embedded in glacial till and boulder clay and proposed simplified correlations using one dimensional regression analysis to correlate the adhesion factor with the undrained cohesion as shown in Eq. (2). On the other hand, Sladen (1992) correlated the adhesion factor to the ratio of the average effective stress to the undrained cohesion as shown in Eq. (3). Furthermore, Chen and Kulhawy (1994) employed regression analysis and proposed a correlation between the undrained adhesion factor and the undrained cohesion of the soil as presented in Eq. (4). Coduto (1994) also proposed correlations to estimate the adhesion factor of bored pile based on the undrained cohesion as shown in Eq. (5). Finally, Cherubini (1999) suggested that the correlation of Sladen (1992) produces better results if it is slightly amended; this amendment is shown in Eq. (6).

$$fs = \alpha \times Su \tag{1}$$

$$\begin{cases} \alpha = 1.00 & \text{for } Su \le 30 \text{ kPa} \\ \alpha = 1.16 - \left(\frac{Su}{185}\right) & \text{for } 30 \text{ kP} \le Su \le 150 \text{ kPa} \\ \alpha = 0.35 & \text{for } Su \ge 150 \text{ kPa} \end{cases} \tag{2}$$

$$\alpha = 0.5 \times \left(\frac{\sigma_o'}{Su}\right)^{0.45} \tag{3}$$

$$\alpha = 0.31 + \frac{0.17}{\frac{Su}{P_a}} \tag{4}$$

$$\begin{cases} \alpha = 1.00 & \text{for } Su \le 51 \text{ kPa} \\ \alpha = 0.32 + 250 \times Su^{-1.5} & \text{for } Su > 51 \text{ kPa} \end{cases} \tag{5}$$

$$\alpha = 0.5 \times \left(\frac{\sigma_o'}{Su}\right)^{0.36} \tag{6}$$

*Corresponding author, Ph.D.
E-mail: Saif.Alzabeebee@qu.edu.iq

ISSN: 2005-307X (Print), 2092-6219 (Online)

Where, σ_o' is the average vertical effective stress and P_a is the atmospheric pressure.

It is evident based on Eqs. (2) to (6) that there are many available correlations to calculate the undrained adhesion factor. However, no attempt has been paid in past studies to statistically assess all of the available correlations. In addition, few attempts have been conducted in previous research papers to develop accurate correlation to predict the undrained adhesion factor for bored piles, where only one attempt have been made in the past to improve the accuracy of the prediction of the adhesion factor for bored piles. This attempt has been made by Goh *et al*. (2005), who employed the Bayesian neural network to propose new correlation to estimate the undrained adhesion factor. However, Goh *et al*. (2005) correlation is complicated and hence, it is difficult to use it in routine designs. This is indeed one of the limitations of the Bayesian neural network (Alzabeebee and Chapman 2020). Therefore, it is very necessary to assess the accuracy of the available empirical correlations and propose new correlation to improve the accuracy of the predictions using advanced data mining technique. Thus, this research aims to address these gaps in knowledge. Accordingly, the objectives of this study are:

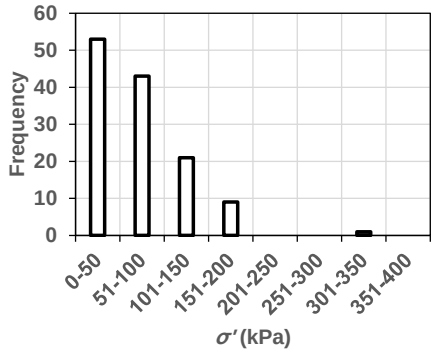
1- Assessing the accuracy of the available empirical correlations to address the gap regarding lack of studies on the accuracy of the available empirical correlations of the undrained adhesion factor.
2- Developing a simple and accurate hybrid intelligent correlation that can be used by practitioner geotechnical engineers to optimize designs of bored piles in undrained conditions. This objective aims to address the gap of lack of robust and simple correlation to predict the undrained adhesion factor. The multi-objective evolutionary polynomial regression analysis (EPR-MOGA) has been used in the development of the new correlation as this method is superior compared to other soft computing techniques (such as the Bayesian neural network) as it provides simple design correlations (Alzabeebee 2020, 2022).

2. Database used in the study

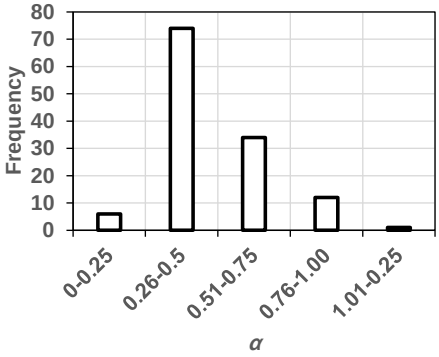
The database used in this research study is adopted from Chen and Kulhawy (1994). It includes 127 data collected from field tests of bored piles embedded in different profiles of cohesive soils. The field tests were for bored piles with a diameter range of 0.18 m to 1.80 m and a length range of 1.62 m to 77.00 m (Chen and Kulhawy 1994). The database consists of the average vertical effective stress (σ_o'), undrained cohesion (Su), and the undrained adhesion factor (α). It is important to mention that only the average vertical effective stress and the undrained cohesion have been considered as the independent parameters that affect the adhesion factor based on the observations of Chen and Kulhawy (1994) and Goh *et al*. (2005). Table 1 presents the statistics of the collected data. The frequency of each variable is also illustrated in Fig. 1. It is obvious from Table 1 and Fig. 1 that the data used in the modelling considers wide range of scenarios. The average vertical effective stress ranges between 11.00 kPa to 343.00 kPa and the undrained cohesion ranges between 21.00 kPa and 483.00 kPa.

Table 1 Statistical indictors of the data employed in this research

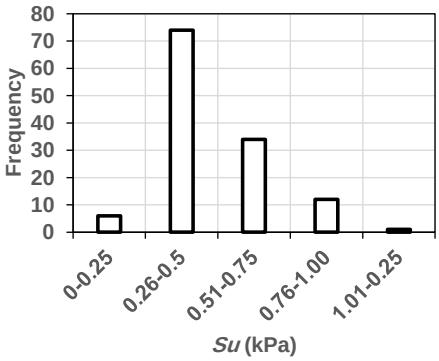
Statistical indicator	σ_o' (kPa)	Su (kPa)	α
Minimum	11.00	21.00	0.24
Maximum	343.00	483.00	1.03
Average	72.43	143.70	0.49
Standard deviation	49.53	92.81	0.18



(a) σ_o'



(b) Su



(c) α

Fig. 1 Frequency of the data used in this study

3. Methodology of the study

The following subsections discuss the methodology of the statistical assessment and the EPR-MOGA, which has been employed in development of the new correlation.

3.1 Statistical assessment

The accuracy of the available and the developed correlations have been tested by calculating the mean absolute error (MAE), the mean square error (MSE), the root mean square error ($RMSE$), the mean (μ), the standard deviation of the ratio of the predicted to measured adhesion factor (σ), the percentage of predictions within error range of ±20% (a20-index), and the coefficient of determination (R^2). These statistical performance measures have been calculated using Eqs. (7) to (13) (Zhang and Goh 2013, 2016, Alkroosh *et al*. 2020, Wang *et al*. 2020, Moayedi *et al*. 2020, Liu *et al*. 2020, Goh *et al*. 2020, Nguyen *et al.* 2020a, b, Zhang *et al*. 2020, Luat *et al*. 2020a, b, Armaghani *et al*. 2021, Bai *et al*. 2021, Tang *et al*. 2021, Zhang *et al*. 2021a, b, c).

$$MAE = \frac{1}{n}\sum_{1}^{n}\left|\alpha_{(p)} - \alpha_{(m)}\right| \tag{7}$$

$$MSE = \frac{1}{n}\sum_{1}^{n}\left(\alpha_{(p)} - \alpha_{(m)}\right)^2 \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}\left(\alpha_{(p)} - \alpha_{(m)}\right)^2} \tag{9}$$

$$\mu = \frac{1}{n}\sum_{1}^{n}\left(\frac{\alpha_{(p)}}{\alpha_{(m)}}\right) \tag{10}$$

$$\sigma = \sqrt{\frac{\sum_{1}^{n}\left(\frac{\alpha_{(p)}}{\alpha_{(m)}} - \mu\right)^2}{n-1}} \tag{11}$$

$$R^2 = \left(\frac{\sum_{i=1}^{n}(\alpha_{(p)_i} - \alpha_{(p)_{average}})(\alpha_{(m)_i} - \alpha_{(m)_{average}})}{\sqrt{\sum_{i=1}^{n}(\alpha_{(p)_i} - \alpha_{(p)_{average}})^2 \sum_{i=1}^{n}(\alpha_{(m)_i} - \alpha_{(m)_{average}})^2}}\right)^2 \tag{12}$$

$$a20 - index = \frac{mp20}{n} \tag{13}$$

Where, $\alpha_{(p)}$ is the predicted undrained adhesion coefficient, $\alpha_{(m)}$ is the measured undrained adhesion coefficient, and $mp20$ is the number of predictions within an error range of ±20%.

The evaluation of the accuracy has been achieved by comparing the values of the statistical performance measures obtained from Eqs. (7) to (13) against the optimum values. The optimum value for each statistical performance measure along with a simplified discussion on the interpretation of each measure is shown in Table 2. Table 2 explicitly shows how these statistical measures (indicators) are utilized to assess the accuracy.

Table 2 Performance indicators clarification (Alzabeebee 2020a)

Performance measure	Optimum value	Indicator descriptor	Attained value clarification
MAE	0.0	Measuring the overall error of prediction.	When the attained value is close to zero that denotes a higher accuracy for the developed correlation
MSE	0.0	Measuring the overall square error of prediction.	When the attained value is close to zero that denotes a higher accuracy for the developed correlation.
$RMSE$	0.0	Measuring the overall error of prediction and is more sensitive to large error in prediction.	When the attained value is close to zero that denotes a higher accuracy for the developed correlation.
μ	1.0	Assessing the overall accuracy of the prediction. Especially, overestimation and underestimation of the predication.	When the attained value is close to 1.0 that denotes a better performance. If the value of this indicator is lower than 1.0 that denotes underestimation of the correlation and vice versa.
σ	0.0	Analyzing the predictions distribution around the mean.	When the attained value is close to zero that denotes a low distribution of the attained values around the mean.
R^2	1.0	Measuring if the predicted values are close to the no-error line. The no-error line can be defined as a hypothetical line where the predicated values are equal to the measured values.	When the attained value is close to 1.0 that means most of the predicted values are close to the measured values.
$a20-\text{index}$	1.0	Measuring the prediction within error range of ±20%.	When the attained value is close to 1.0 that denotes a better performance. This explicitly demonstrates the proportion of predicted values within a range of error equal or less than ±20% (i.e., a20-index of 0.75 denotes that 75% of predicted values are within a range of error of ±20%).

3.2 Evolutionary polynomial regression analysis

The multi-objective genetic algorithm evolutionary polynomial regression (EPR-MOGA) is a regression analysis driven by genetic algorithm (GA). The approach of EPR-MOGA is based on regression analysis that evolves automatically utilizing genetic algorithm (GA) (Ahangar-Asr *et al*. 2014, Alzabeebee 2020a, b). The EPR-MOGA generates a number of selected correlations between the dependent and independent variables with the help of the genetic algorithm. The developed correlations are controlled by the suggested correlation type (based on the

user preference), complexity of the interrelationship between the variables, the number of the data incorporated in the analysis, the range of the exponents (based on the user preference), and finally the number of terms of the correlation. The typical formation of the EPR-MOGA is shown in Eq. (14) (Giustolisi and Savic 2006).

$$Y = \sum_{j=1}^{m} F\left(\mathbf{X}, f(\mathbf{X}), a_j\right) + a_0 \quad (14)$$

Where, Y is the dependent input value, a_j is a constant value, F is a function advances with the aid of the genetic algorithm depending on the dependent and independent variables, $\mathbf{X}$ is the matrix of the input independent variables, $f(\mathbf{X})$ is the proposed general form of the correlation that advances during the analysis, m is the number of terms of the suggested correlation and a_0 is the bias.

According to Alzabeebee *et al*. (2020a, b) both the proposed correlation $\left(f(\mathbf{X})\right)$ and the number of terms (m) are defined as preferred by the user. Although the general form of the correlation is suggested by the user, the number of terms, order of exponents, and the interrelationship between the dependent variables within the correlation are all determined using the genetic algorithm (GA).

As mentioned before, the genetic algorithm is adopted in the EPR-MOGA to suggest the best mathematical correlation. This is done by a random creation induced by Darwinian evolution (from the GA) to generate the initial population of solutions, where each parameter set in the population indicates chromosomes of the individuals. Each individual is assigned with its own fitness which correlates with the performance of the individual in its environment. Then, crossover and mutation operations are conducted to conclude the best equation for each specific number of terms (Giustolisi and Savic 2009, Ahangar-Asr *et al*. 2014). The process continues until the desired number of terms of the established correlation is reached.

The coefficient of determination (CD) is computed for each candidate relationship to evaluate the accuracy of each modelling round. The correlation that employed in the EPR-MOGA in order to compute the CD is presented in Eq. (15) (Giustolisi and Savic 2006).

$$CD = 1 - \frac{\sum_N (\alpha_{(m)} - \alpha_{(p)})^2}{\sum_N (\alpha_{(m)} - \frac{1}{N}\sum_N \alpha_{(p)})^2} \quad (15)$$

4. Results

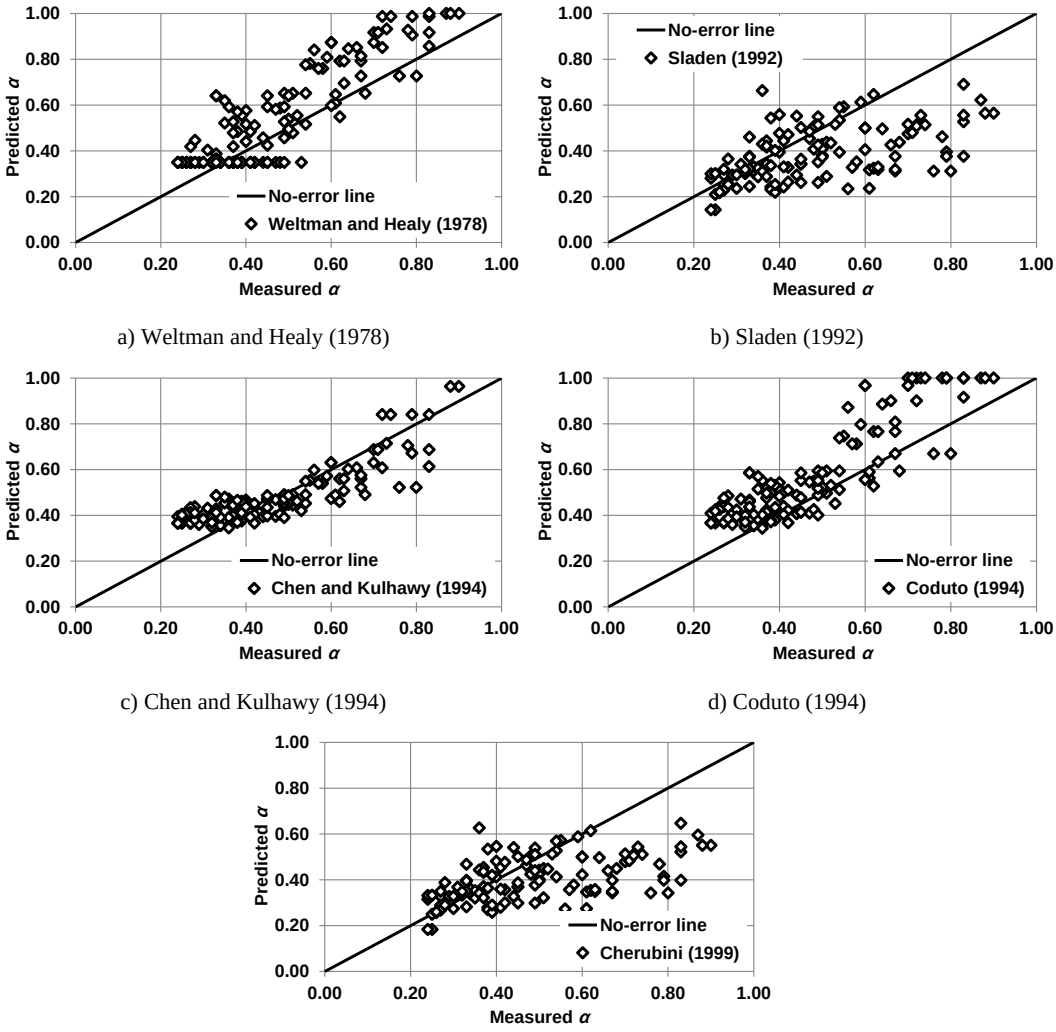
4.1 Assessment of the available correlations

One of the aims of this research is to assess the accuracy of the available correlations to find out which correlation provides the best prediction and to consider the assessment results for comparison with the correlation developed in this paper (which will be discussed further in the next subsection). Therefore, the correlations discussed in the introduction (Eq. (2) to (6)) have been used to calculate the undrained adhesion factor.

Fig. 2 shows the relationship between the measured and predicted undrained adhesion factor in comparison with the no-error line for the aforementioned correlations. It is clear based on Fig. 2 that the predictions of all of the available correlations have a remarkable scatter around the mean. The obtained results from the aforementioned empirical correlations have also been used to do the statistical assessment. Table 3 shows the results of this assessment using the statistical performance indicators. It is clear from Table 3 that Chen and Kulhawy (1994) correlation produces the lowest error with MAE, MSE and $RMSE$ of 0.08, 0.010, and 0.100. Table 3 also shows that Chen and Kulhawy (1994) correlation scored the closest μ value to 1.00 (1.07), indicating an average overestimation in the prediction of about 7%. Furthermore, the same correlation scored $a20-index$ of 0.66 indicating that 66% of the predictions of this correlation are within error range of ±20%. However, the results show that all of the correlations scored similar standard deviation of the predicted to measured values. This means that the scatter of the predictions of all of the correlations is similar. Also, Chen and Kulhawy (1994) correlation did not score the best R^2, where it scored a value of 0.68.

It is also evident from Table 3 that Weltman and Healy (1978) and Coduto (1994) correlations scored almost similar values for all the statistical indicators. Both correlations scored higher average error than Chen and Kulhawy (1994) correlation and slightly higher μ compared with Chen and Kulhawy (1994) correlation. Also, these correlations scored lower $a20-index$, which is equal to 0.47 and 0.46, respectively. This means that only 47% and 46% of the values of the adhesion factor can be predicted within error range of 20% by these correlations. However, it is interesting to note that both correlations scored the highest R^2 compared with other empirical correlations. The scored R^2 is equal to 0.77 and 0.76 for both correlations, respectively. Finally, it is worthy to mention that Sladen (1992) and Cherubini (1999) correlations achieved the poorest predictions compared with other empirical correlations. These correlations scored the highest error, the lowest mean, and the lowest R^2 values. However, these correlations also scored $a20-index$ of 0.49 and 0.48, respectively, which is almost similar to Weltman and Healy (1978) and Coduto (1994) correlations. Finally, it is worthy to mention that the results of the statistical assessment are strongly influenced by the collected database, where Chen and Kulhawy (1994) scored the best, in general, because the database employed in the present assessment is the same database used by Chen and Kulhawy (1994) in the development of the correlation. This also indicates that other correlations which has scored low values might be developed using limited database which did not enable good predictions for these correlations when examined against database with more different scenarios. However, this limitation is well known, and it is one of the major shortcomings of the machine learning based models.

In general, it can be said based on the results and discussions in this section that while some of the correlations produced predictions with low error on average, these correlations failed to show low scatter



a) Weltman and Healy (1978)

b) Sladen (1992)

c) Chen and Kulhawy (1994)

d) Coduto (1994)

e) Cherubini (1999)

Fig. 2 Predicted and measured alpha values for the existed empirical correlations

Table 3 Performance indicators of the existed empirical correlations

Performance indicator	Weltman and Healy (1978)	Sladen (1992)	Chen and Kulhawy (1994)	Coduto (1994)	Cherubini (1999)
MAE	0.108	0.134	0.080	0.115	0.122
MSE	0.017	0.032	0.010	0.020	0.028
$RMSE$	0.140	0.180	0.100	0.147	0.170
μ	1.19	0.84	1.07	1.21	0.90
σ	0.24	0.25	0.23	0.24	0.25
R^2	0.77	0.32	0.68	0.76	0.32
a20-index	0.47	0.49	0.66	0.46	0.48

around the mean. On the other hand, while other correlations produced predictions with low scatter around the mean, these correlations performed poorly in the average error and in the percentage of predictions within error range of ±20%. Hence, further improvement in the predictions is necessary for optimized designs.

4.2 Development of the hybrid intelligent correlation

The data collected from the literature is divided into two groups: training data and testing data. The training data group consisted of 80% of the whole data; however, the remaining 20% of the data is used as testing data. The training data has been employed in the EPR-MOGA to enable the EPR-MOGA to learn the best correlation, while the testing data has been employed to test the accuracy of the established correlation in computing the adhesion factor using data that have not been employed in the correlation development (Alzabeebee 2021a, b). This process is quite common in soft computing to ensure the robustness of the developed correlation (Ahangar-Asr *et al*. 2014, Alzabeebee 2020a, Alzabeebee and Chapman 2020, Armaghani *et al*. 2021). Importantly, considerable efforts have been paid in selecting the training and testing data to guarantee that the testing data is within the range of the training data. This approach has been deemed necessary to prevent extrapolation in the predictions of the developed correlation as recommended by many previous studies (Ahangar Asr *et*

al. 2018; Alzabeebee *et al*. 2020a, b, Luat *et al*. 2020a, b, Shams *et al*. 2020). Tables 4 and 5 show the statistical indicators of the training and testing data. It is clear from both tables that the testing data is within the range of the training data as recommended in previous studies (Ahangar-Asr *et al*. 2014, Alzabeebee, 2021a, b).

Trails have been conducted employing the EPR-MOGA and the training and testing data to find the best suitable mathematical correlation that correlates the undrained adhesion factor with average vertical effective stress and undrained cohesion. For each trial, the statistical performance indicators and the ability of the correlation to predict the trend of the relationship between the independent and dependent variables have been carefully examined. Considering all of the aforementioned criteria, the best correlation obtained from the EPR-MOGA is shown in Eq. (16).

$$\alpha = \frac{-618.1}{Su^2} + \frac{44.25}{Su} + 0.00073 \times \sqrt{\sigma_o'} \times \sqrt{Su} + 0.053 \quad (16)$$

Figs. 3(a) and 3(b) show the relationship between the measured and predicted adhesion factor for training and testing data, respectively. The figures clearly demonstrate the accuracy of the correlation as most of the points are near to the no-error line. In addition, Table 6 shows the statistical performance indicators of the developed correlation for both training and testing data. The table shows that the correlation achieved almost the same statistical performance indicators for both training and testing data. This is a testament of the ability of the developed correlation to predict unseen data with accuracy similar to the predictions of the data used in the training of the correlation. The developed correlation scored MAE, MSE, and $RMSE$ of 0.060, 0.010, and 0.080 for training data, and 0.07, 0.01, and 0.09 for testing data indicating a marginal error of prediction. The correlation also scored a mean value near to the optimum value (1.03 for training data and 1.04 for testing data). Also, the standard deviation of the predicted to measured values is very close to zero (0.18 for training data and 0.20 for testing data), again indicating an excellent prediction capability. Importantly, the results of the $a20-index$ analyses indicates that the correlation is capable of predicting 78% (on average) of the data with error equal to or less than ±20%, where the scored $a20-index$ for training and testing data is equal to 0.78 and 0.79, respectively. Finally, the results of the R^2 also show very good prediction ability with scored values of 0.78 and 0.79 for training and testing data, respectively.

Comparing the statistical performance indicators of the developed correlation (Table 6) with the available correlations (Table 3) clearly demonstrated the better accuracy of the new correlation, where it scored lower MAE, MSE, $RMSE$, μ value closer to 1.0, lower σ, higher $a20-index$ and higher R^2 for both training and testing data. This is because these previous correlations have been developed using simple regression analysis while the present correlation have been developed using more advanced data mining technique with the aid of the genetric algirthm which enabled a robust serach for the best mathmatical correlation. In addition, this comparisoin provide additonal evidence of the capabilities of the new correlation, and hence, this corelation can be recommended to be used by geotechnical engineers for optimized design of shaft capacity of bored piles in undrained conditions.

Table 4 Statistical indicators of the training data

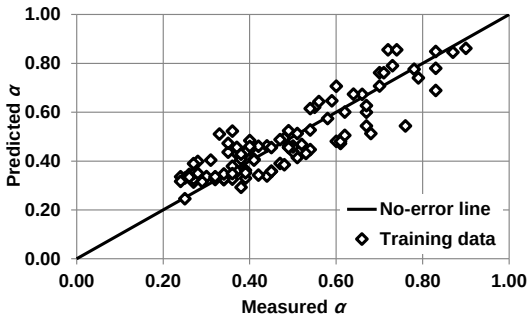
Statistical indicator	σ_o' (kPa)	Su (kPa)	α
Minimum	11.00	21.00	0.24
Maximum	343.00	483.00	1.03
Average	76.83	145.67	0.49
Standard deviation	52.95	95.63	0.17

Table 5 Statistical indicators of the testing data

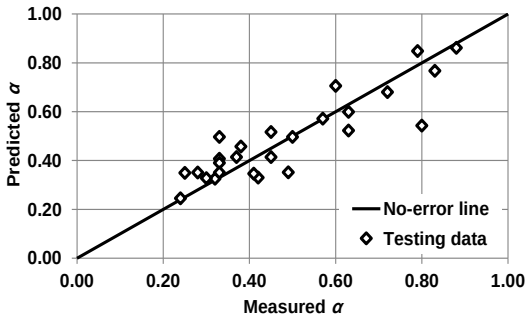
Statistical indicator	σ_o' (kPa)	Su (kPa)	α
Minimum	17.00	26.00	0.24
Maximum	119.00	307.00	0.88
Average	55.35	136.04	0.49
Standard deviation	27.67	82.19	0.19

Table 6 Performance indicators of the developed hybrid intelligent correlation for training and testing data

Performance indicator	Training data	Testing data
MAE	0.060	0.070
MSE	0.010	0.010
$RMSE$	0.080	0.090
μ	1.03	1.04
σ	0.18	0.20
R^2	0.78	0.79
a20-index	0.78	0.78



(a) Training data



(b) Testing data

Fig. 3 Predicted and measured alpha values for the EPR correlation using: (a) training data and (b) testing data

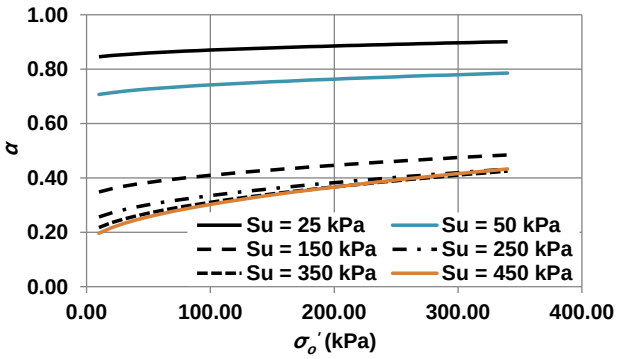


Fig. 4 Results of the parametric study using the developed correlation

5. Sensitivity analysis

Finally, the results of the developed correlation (Eq. (16)) have been employed in a parametric study to illustrate the effect of the undrained shear strength and the effective stress on the adhesion factor. An undrained cohesion range of 25 to 450 kPa is considered in the parametric study. In addition, the considered mean vertical effective stress range is 10 to 340 kPa. The results of the parametric study are shown in Fig. 4. It is clear from the presented results that increasing the undrained shear strength from 25 to 150 kPa remarkably decreases the adhesion factor; however, the percentage decrease remarkably reduces as the undrained cohesion increases above 150 kPa. The reduction of the adhesion factor becomes negligible as the undrained cohesion increases from 250 to 350 kPa for mean vertical effective stress equal to or greater than 200 kPa. Moreover, the figure reveals that increasing the undrained cohesion from 350 to 450 kPa did not noticeably affect the adhesion factor. It can also observe from the figure that increasing the stress level rises the adhesion factor in a non-linear fashion. However, the rate of non-linearity depends on the undrained cohesion, and it increases as the undrained cohesion rises. Hence, it is clear from the results that it is necessary to consider the mean vertical effective stress in the prediction of the adhesion factor.

It is useful to state that the general trend of the obtained adhesion factors agrees with the experimental observation of Kulhawy and Jackson (1989). Thus, this agreement provides additional confidence in the performance of the developed correlation.

5. Conclusions

This research provided a comprehensive assessment of the available correlations to predict the undrained adhesion factor of bored piles and proposed novel and more accurate correlation using multi-objective evolutionary polynomial regression analysis to address the gaps in knowledge regarding the prediction of the undrained adhesion factor of bored piles. The following points summarize the main findings of this research:

1- Chen and Kulhawy (1994) correlation scored, on average, the lowest error and the best mean compared with other available correlations. However, this method could not score the best R^2 compared with other available correlations.
2- Weltman and Healy (1978) and Coduto (1994) correlations scored the best R^2 compared with other available correlations. However, these correlations also scored average error higher than Chen and Kulhawy (1994) correlations.
3- Sladen (1992) and Cherubini (1999) correlations scored poorest predictions (highest error, lowest mean, and lowest R^2 values) among the other available correlations. Additionally, these correlations scored $a20-index$ of 0.49 and 0.48, respectively, which is almost similar to the $a20-index$ of Weltman and Healy (1978) and Coduto (1994) correlations.
4- A novel surrogate correlation has been developed in this research using the multi objective artificial intelligence evolutionary polynomial regression analysis. The developed correlation achieved better predictions than the available correlations.
5- The new correlation can be used by researchers and geotechnical design engineers to optimize designs of bored piles in undrained conditions as it scored accuracy higher than the available correlations.

References

Ahangar-Asr, A., Javadi, A.A., Johari, A. and Chen, Y. (2014), "Lateral load bearing capacity correlationling of piles in cohesive soils in undrained conditions: an intelligent evolutionary approach", *Appl. Soft Comput.*, **24**, 822-828. https://doi.org/10.1016/j.asoc.2014.07.027.

Alani, A.M. and Faramarzi, A. (2014), "An evolutionary approach to correlationling concrete degradation due to sulphuric acid attack", *Appl. Soft Comput.*, **24**, 985-993. https://doi.org/10.1016/j.asoc.2014.08.044.

Alkroosh, I., Alzabeebee, S. and Al-Taie, A.J. (2020), "Evaluation of the accuracy of commonly used empirical correlations in predicting the compression index of Iraqi fine-grained soils", *Innov. Infrastruct. Solut.*, **5**(3), 1-10. https://doi.org/10.1007/s41062-020-00321-y.

Alzabeebee, S. (2020a), "Application of EPR-MOGA in computing the liquefaction-induced settlement of a building subjected to seismic shake", *Eng. Comput.*, https://doi.org/10.1007/s00366-020-01159-9.

Alzabeebee, S. (2020b), "Dynamic response and design of a skirted strip foundation subjected to vertical vibration", *Geomech. Eng.*, **20**(4), 345-358. https://doi.org/10.12989/gae.2020.20.4.345.

Alzabeebee, S. and Chapman, D.N. (2020), "Evolutionary computing to determine the skin friction capacity of piles embedded in clay and evaluation of the available analytical methods", *Transp. Geotechn.*, 100372. https://doi.org/10.1016/j.trgeo.2020.100372

Alzabeebee, S., Alshkane, Y.M., Al-Taie, A.J. and Rashed, K.A. (2021a), "Soft computing of the recompression index of fine-grained soils", *Soft Comput.*, **25**(24), 15297-15312. https://doi.org/10.1007/s00500-021-06123-3.

Alzabeebee, S., Mohamad, S.A. and Al-Hamd, R.K.S. (2021b), "Surrogate models to predict maximum dry unit weight, optimum moisture content and California bearing ratio form grain size distribution curve", *Road Mater. Pav. Design*, https://doi.org/10.1080/14680629.2021.1995471.

Alzabeebee, S. (2022), "Explicit soft computing model to predict the undrained bearing capacity of footing resting on aggregate pier reinforced cohesive ground", *Innov. Infrastruct. Solut*., 7(1), 1-10. https://doi.org/10.1007/s41062-021-00706-7.

Armaghani, D.J., Mamou, A., Maraveas, C., Roussis, P.C., Siorikis, V.G., Skentou, A.D. and Asteris, P.G. (2021), "Predicting the unconfined compressive strength of granite using only two non-destructive test indexes", *Geomech. Eng*, **25**(4), 317-330. https://doi.org/10.12989/gae.2021.25.4.317.

Bai, X.D., Cheng, W.C., Ong, D.E. and Li, G. (2021), "Evaluation of geological conditions and clogging of tunneling using machine learning", *Geomech. Eng.*, **25**(1), 59-73. https://doi.org/10.12989/gae.2021.25.1.059.

Chen, Y.J. and Kulhawy, F.H. (1994), "Case history evaluation of the behavior of drilled shafts under axial and lateral loading", Rep. No. TR-104601, Electric Power Research Institute, Palo Alto, Calif.

Cherubini, C. and Yves, R. (1998), "A few comments on pile design: Discussion/Reply", *Can. Geotech. J.*, **35**(5), 905. https://doi.org/10.1139/t98-036.

Coduto, D.P. (1994), Foundation design, principles and practices. Prentice Hall Inc., Englewood Cliffs, N.Y.

Dias, D. and Grippon, J. (2017), "Numerical modelling of a pile-supported embankment using variable inertia piles", *Struct, Eng, Mech.*, **61**(2), 245-253. https://doi.org/10.12989/sem.2017.61.2.245.

Giustolisi, O. and Savic, D.A. (2006), "A symbolic data-driven technique based on evolutionary polynomial regression", *J. Hydroinform.*, **8**(3), 207-222. https://doi.org/10.2166/hydro.2006.020b.

Giustolisi, O. and Savic, D.A. (2009), "Advances in data-driven analyses and correlationling using EPR-MOGA", *J. Hydroinform.*, **11**(3-4), 225-236. https://doi.org/10.2166/hydro.2009.017.

Goh, A.T., Kulhawy, F.H. and Chua C.G. (2005), "Bayesian neural network analysis of undrained side resistance of drilled shafts", *J. Geotech. Geoenv. Eng*., **131**(1), 84-93. https://doi.org/10.1061/(ASCE)1090-0241(2005)131:1(84).

Goh, A.T.C., Zhang, R.H., Wang, W., Wang, L., Liu, H.L. and Zhang, W.G. (2020), "Numerical study of the effects of groundwater drawdown on ground settlement for excavation in residual soils", *Acta Geotech*., **15**(5), 1259-1272. https://doi.org/10.1007/s11440-019-00843-5.

Hwang, T.H., Kim, K.H. and Shin, J.H. (2017), "Bearing capacity of micropiled-raft system", *Struc, Eng, Mech*., **63**(3), 417-428. https://doi.org/10.12989/sem.2017.63.3.417.

Kulhawy, F.H. and Jackson, C. (1989), "Some observations of undrained side resistance of drilled shafts", *Foundation Engineering: Current principles and practices*, 2, 1011-1025.

Liu, L., Moayedi, H., Rashid, A.S.A., Rahman, S.S.A. and Nguyen, H. (2020), "Optimizing an ANN correlation with genetic algorithm (GA) predicting load-settlement behaviours of eco-friendly raft-pile foundation (ERP) system", *Eng. Comput*., **36**(1), 421-433. https://doi.org/10.1007/s00366-019-00767-4.

Luat, N.V., Lee, K. and Thai, D.K. (2020a), "Application of artificial neural networks in settlement prediction of shallow foundations on sandy soils", *Geomech. Eng.*, **20**(5), 385-397. https://doi.org/10.12989/gae.2020.20.5.385.

Luat, N.V., Nguyen, V.Q., Lee, S., Woo, S. and Lee, K. (2020b), "An evolutionary hybrid optimization of MARS model in predicting settlement of shallow foundations on sandy soils", *Geomech. Eng.*, **21**(6), 583-598. https://doi.org/10.12989/gae.2020.21.6.583.

Moayedi, H., Moatamediyan, A., Nguyen, H., Bui, X.N., Bui, D.T. and Rashid, A.S.A. (2020), "Prediction of ultimate bearing capacity through various novel evolutionary and neural network correlations", *Eng. Comput.*, **36**, 671-687. https://doi.org/10.1007/s00366-019-00723-2.

Nguyen, H., Moayedi, H., Foong, L.K., Al Najjar, H.A.H., Jusoh, W.A.W., Rashid, A.S.A. and Jamali, J. (2020b), "Optimizing ANN correlations with PSO for predicting short building seismic response", *Eng. Comput.*, **36**, 823-837. https://doi.org/10.1007/s00366-019-00733-0.

Nguyen, H., Moayedi, H., Jusoh, W.A.W. and Sharifi, A. (2020a), "Proposing a novel predictive technique using M5Rules-PSO correlation estimating cooling load in energy-efficient building system", *Eng. Comput.*, **36**, 857-7866. https://doi.org/10.1007/s00366-019-00735-y.

Robert, Y. (1997), "A few comments on pile design", *Can. Geotech. J*., **34**(4), 560-567. https://doi.org/10.1139/t97-024.

Sladen, J.A. (1992), "The adhesion factor: applications and limitations", *Can. Geotech. J*., **29**(2), 322-326. https://doi.org/10.1139/t92-036.

Tang, L. and Na, S. (2021), "Comparison of machine learning methods for ground settlement prediction with different tunneling datasets", *J. Rock Mech. Geotech. Eng*., https://doi.org/10.1016/j.jrmge.2021.08.006.

Wang, H., Moayedi, H. and Foong, L.K. (2020), "Genetic algorithm hybridized with multilayer perceptron to have an economical slope stability design", *Eng. Comput*., https://doi.org/10.1007/s00366-020-00957-5.

Weltman, A.J. and Healy, P.R. (1978), "Piling in 'Boulder Clay' and Other Glacial Tills", Construction Industry Research and Information Association, Report PG5.

Zhang, R.H., Li, Y.Q., Goh, A.T.C. and Zhang, W.G. (2021c), "Analysis of ground surface settlement in anisotropic clays using XGBoost and RFR models", *J. Rock Mech. Geotech. Eng*., https://doi.org/10.1016/j.jrmge.2021.08.001.

Zhang, W. and Goh, A.T. (2016), "Multivariate adaptive regression splines and neural network models for prediction of pile drivability", *Geosci. Front*, **7**(1), 45-52. https://doi.org/10.1016/j.gsf.2014.10.003.

Zhang, W., Li, H., Li, Y., Liu, H., Chen, Y. and Ding, X. (2021b), "Application of deep learning algorithms in geotechnical engineering: a short critical review", *Artif. Intell. Rev.*, **54**, 5633-5673. https://doi.org/10.1007/s10462-021-09967-1.

Zhang, W., Wu, C., Zhong, H., Li, Y. and Wang, L. (2021a), "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization", *Geosci. Front.*, 12(1), pp. 469–477. https://doi.org/10.1016/j.gsf.2020.03.007

Zhang, W., Zhang, R., Wu, C., Goh, A.T.C., Lacasse, S., Liu, Z. and Liu, H. (2020), "State-of-the-art review of soft computing applications in underground excavations", *Geosci. Front*., **11**(4), 1095-1106. https://doi.org/10.1016/j.gsf.2019.12.003.

Zhang, W.G. and Goh, A.T.C. (2013), "Multivariate adaptive regression splines for analysis of geotechnical engineering systems", *Comput. Geotech*., **48**, 82-95. https://doi.org/10.1016/j.compgeo.2012.09.016.

GC