Estimation of load-set behavior of driven concrete piles using artificial neural network and cone penetration test

I V Ofrikhter\textsuperscript{1,2} and A B Ponomarev\textsuperscript{1}

\textsuperscript{1} Department of Construction operations and geotechnics, Perm National Research Polytechnic University, 29, Komsomolsky ave., Perm, 614990, Russia

\textsuperscript{2} ian.ofrikhter@gmail.com

Abstract. Pile foundations are among the most common types of deep foundations. Predicting their draft is a fairly standard engineering goal. There are many techniques for this. Many methods for determining the settlement of piles are based on various empirical relationships. In particular, the cone penetration test allows you to determine the settlement and bearing capacity of a pile using the developed empirical dependencies. However, some models must be used to develop such correlations. This study proposes to use a slightly different approach. Instead of using a pile model and a set of empirical dependencies, it is suggested to use machine learning methods. Artificial neural networks are one of the machine learning methods. The article presents a description of the development of a neural network that allows estimating pile settlement using data from the CPT test.

1. Introduction

The pile foundation is one of the most common types of deep foundations. Piles are generally used to transfer axial building loads through low-strength soil layers or bodies of water into more suitable bearing strata. The evaluation of the load-settlement performance of a single pile is one of the main aspects of the design of piled foundations. To predict the settlement of a pile, it is necessary to take into account many factors, such as soil characteristics, geometric dimensions of the pile, the method of pile installation, and others [1]. In analyzing pile settlement, Poulos and Davis [2] concluded that instant settlement accounts for a large proportion of total pile settlement. This also includes consolidation settlement for saturated clay soils and even for piles in clay as shown by Murthy [3]. With regards to piles in sand, immediate settlement accounts for almost the entire final settlement [3].

A large number of theoretical and experimental methods for calculating the settlement and bearing capacity of piles have been developed. As mentioned above, the calculation of pile settlement depends on many factors. Different calculation methods take into account different sets of factors and use different input data.

Theoretical calculation methods usually use several empirical dependences of pile settlement on the characteristics of the soil and pile material. Recently, the use of machine learning methods in various engineering calculations has become more and more relevant. Chan W. T., Chow Y. K. and Liu, L. F. offered to use neural networks as an alternative for pile driving formulas [4]. Some reanalyzing of driving formulas with the use of neural networks was done by Goh in 1996 [5].

The most common form of artificial neural network (ANN) that performs complex regression tasks is a fully connected feedforward neural network. This model consists of an input layer, a hidden layer, and an output layer (figure 1).
On figure 1 Xi values in the input layer are numbers that are the initial data of the predicted process. \(w_{ij}\) - link weights. The link weight \(w_{ij}\) is the number by which the output of the previous layer is multiplied. The neurons of the hidden and output layers are functions \(f(u)\) and \(g(u)\), where \(u\) is the sum of the input values multiplied by the corresponding connection weights \(w_{ij}\). The functions \(f(u)\) and \(g(u)\) are called activation functions. The parameters \(a_i\) and \(t_k\) are the outputs of the activation functions \(f(u)\) and \(g(u)\). In view of the above, \(a_i\) and \(t_k\) can be expressed in formulas 1 and 2.

\[
a_j = f \left( \sum_{i=1}^{N} w_{ij} x_i + b \right) \tag{1}
\]

\[
t_k = g \left( \sum_{j=1}^{N} w_{jk} a_j + b_k \right) \tag{2}
\]

The output layer can include any number of neurons. The hidden layer could include several neuron layers. The number of neurons and layers in the hidden layer is limited only by computational capabilities. In general, the more complex the process being modeled, the more neurons and layers are needed. In this study, it is proposed to develop a simple ANN for predicting the settlement of a driven pile based on the CPT test data.

2. Materials and methods
As stated above, in this paper, an artificial neural network is being developed to predict the load-settlement behavior of driven reinforced concrete piles. A compiled database is needed to develop a neural network. In total, the database contains 58 pairs of SLT and CPT tests. Databases from previously published works were used to analyze the results. For the development of ANN, it is necessary to define the input and output parameters of the model, the architecture of the model, stop criteria, the method for separating the data, and pre-preparation of the data. Some recommendations for the development of artificial neural networks are given in the work of Shahin M A, Maier H R, Jaksa M B [6]. Some of the approaches are taken from a study by Moayedi, Hossein Mosallanezhad, Mansour Rashid and others [7] are dedicated to basic approaches to solving geotechnical problems using neural networks.

2.1. Input and output parameters of the model
To get an artificial neural network capable of predicting certain parameters, it must be trained. The most common way to train a neural network is the backpropagation method. In this work, we use
predominantly multilayer direct propagation neural networks. The diagram of such a neural network is shown in the figure 1.

The neural network in figure 1 has an input layer, hidden layers, and an output layer. In this study, pile settlement is predicted as a function of load, pile geometry, and CPT test data. Accordingly, the load, pile characteristics, and CPT test parameters will be presented in the input layer and the determined settlement in the output layer. Thus, the output layer will consist of only one parameter. All data in the input layer must be represented as some limited set of numbers.

All piles collected in this study are square-section concrete driven piles, so these parameters are not taken into account in the ANN input layer. The only variable geometrical parameter is the pile length. To take into account the data of the CPT tests, the piles are divided into a certain number of segments of equal thickness. Since the tests in the existing CPT database were carried out without determining the pore pressure, these data are not taken into account in the model. The sleeve friction and cone resistance from the CPT test are averaged within these sections. F. Pooya Nejad, Mark B. Jaksa [8] proposes to break the pile into 5 sections. In this study, the piles are divided into 3 - 10 sections with subsequent analysis of the division density for the accuracy of the results. Besides, the characteristics in the area below the pile tip to a depth of up to half the pile length are taken into account. Hence the factors that are presented to the ANN as model input variables are: (1) embedded length of the pile \( L_{emb} \), (2) cone tip resistance at the end of the pile \( f_{tip} \), (3) applied load \( (P) \), (4-14) average sleeve friction from CPT test for layer \( i \) (\( q_{ci} \)), (14-24) average cone resistance from CPT test for layer \( i \) (\( f_{ci} \)), average sleeve friction from CPT under the pile tip (25), average cone resistance from CPT under the pile tip (26).

2.2. Data division

Artificial neural networks have an almost unlimited ability to reproduce various correlations. The amount of input and output data is essentially limited only by training time and computational capabilities. However, during training, two effects develop simultaneously: "generalization" and "memorization". The generalization effect is the ability of a neural network to correctly reproduce the dependencies between input and output parameters. The memory effect is the ability of a neural network to memorize training examples instead of finding dependencies between them. To train a neural network, it is necessary to minimize the "memorization" effect.

Another important feature of a neural network is the choice of its parameters. Weights and coefficients of activation functions can be found during training using algorithms. However, the choice of the network architecture, the minimized error function, the choice of input and output parameters, etc. are performed by a person. These parameters are called hyperparameters. As a result, when developing even simple neural networks, the entire existing database must be divided into several parts.

In this study, the Stone method is used to divide data [9]. SLT testing is performed in stages, each load stage is a separate case. There are 58 tests and 603 cases in the database. Loads and settlements at which the pile has lost its bearing capacity are not taken into account. All data is divided into three parts. 70% (422 cases) of the data is used to directly train the neural network using error backpropagation. 10% (60 cases) of the data is used to refine the hyperparameters. They do not participate in training directly, however, during training, the ANN is immediately tested on this sample. The remaining 20% (121 cases) of the data is used to test the obtained ANN. These data are used only after the completion of training to perform statistical analysis and compare the proposed methodology with existing ones.

2.3. Model training

After collecting all the necessary data, the neural network must be trained. The training process was carried out using the Tensorflow library. Tensorflow is a freely redistributable Python library. Training of artificial neural networks is usually performed using the backpropagation method. This method appeared in 1969 [10] and has become one of the main approaches for training neural networks. The main goal of training is to find such weights and values of activation functions at which the artificial neural network will correctly predict examples from the training sample.
F. Pooya Nejad, Mark B. Jaksa [8] suggested sorting the data to achieve an approximately equal ratio of key parameters in training and validation sets. In this study, only one type of pile is considered and more data is used, so random data comes across in the training and validation sets. To control the quality of the trained model, training was performed 10 times.

During training, the effect of overfitting is monitored. The effect of overfitting occurs when the neural network does not approximate the received data, but simply "memorizes" them. To maintain this effect, graphs are drawn up to reduce the error relative to the training sample, and the validation set. Since the validation dataset is not directly involved in training, it is assumed that in the case of overfitting, the error value on the training sample will be significantly lower than on the validation sample. Figure 1 shows a graph of error reduction. As it can be seen, in this case, the final value of the error of the training and validation samples is approximately the same. This suggests that no overfitting effect was observed with the adopted architecture. If the model was overfitted, these results were not taken into account. When overfitting (figure 2), the error on the validation set stops decreasing or increases after a certain epoch.

![Figure 2. Typical graph of a well-trained model.](image)

![Figure 3. Typical graph of an overfitted model.](image)

3. Results and discussion

A variety of ANN models have been tested to predict the settlement of driven piles. Models were trained with 2-8 layers of 10-200 neurons each (with a step of 10). To reduce the effect of overfitting, dropout layers [11] were additionally introduced. However, they did not have any substantive influence. The main goal was to find such an architecture in which the model is minimally retrained and shows the best results. The number of neurons in each layer largely depends on the number of pile segments. The pile was split into 3-10 segments. The correlation of the obtained accuracy with the number of segments is shown in table 1. The main statistical indicator and used loss function is the mean squared error (MSE).

| Number of segments | Best MSE      |
|--------------------|--------------|
| 3                  | 0.576023     |
| 4                  | 0.548246     |
| 5                  | 0.486842     |
| 6                  | 0.451754     |
| 7                  | 0.449877     |
| 8                  | 0.448099     |
| 9                  | 0.446491     |
| 10                 | 0.443275     |

Table 1. The correlation of the obtained accuracy with the number of segments.
As it can be seen from table 1, with more than 6 segments, the accuracy remains almost unchanged. F. Pooya Nejad, Mark B. Jaksa [8] suggested using 5 segments. One can agree with this technique. It is also worth noting that in this study, the database included piles no more than 15 meters in length. When analyzing longer piles, for example, drill piles, the required number of segments may be different.

As a result, an artificial neural network was obtained that allows assessing pile settlement based on CPT test data and vertical load. The final neural network consists of 4 hidden layers, layers of 200 neurons each. The quality of the results obtained is assessed on a test dataset. These data did not participate in the training of the neural network or the selection of hyperparameters. The correlation of the neural network prediction with the results of a static test is shown in figure 4.

![Figure 4. Correlation of ANN and SLT results.](image1)

![Figure 5. Correlation of SP and SLT results.](image2)

The mean average percentage error (MAPE) is 23.4%. To assess the effectiveness of the proposed approach. The results obtained are compared with the existing calculation method defined by the national standard SP 24.13330. The correlation of the SP 24.13330 prediction with the results of a static test is shown in figure 5. The mean average percentage error of the SP 24.13330 method is 43.40%.

As can be seen from the comparison of the methods, the use of a neural network allowed us to reduce the error. It should also be noted that the test dataset was collected in the same region as the training and testing ones. With enough data, an artificial neural network can significantly reduce error. In this study, it was not possible to achieve the indicators described in similar works [8]. However, in comparison with the existing technique, the neural network looks like a very promising method.

**Conclusion**

As a result of the study, a feed-forward artificial neural network was developed to predict the settlement of driven concrete piles. The error estimation showed that the ANN trained by the error propagation method allows predicting the pile settlement with greater accuracy than traditional methods. A comparison of the obtained data with existing methods shows that to obtain optimal results, it is necessary to collect a large amount of data for each type of pile. However, due to the ability to learn and high accuracy, neural networks look like a convenient and easy-to-use tool for assessing pile settlement.
Acknowledgments
The reported study was funded by RFBR according to the research project № 20-31-70001.

References
[1] Berardi R and Bovolenta R 2005 Pile-settlement evaluation using field stiffness non-linearity Geotechnical Engineering 158 35–44
[2] Rowe R K 1981 Pile foundation analysis and design: Book review Canadian Geotechnical Journal 18 472–3
[3] Murthy V N 2002 Geotechnical engineering: principles and practices of soil mechanics and foundation engineering Jurnal Ekonomi Malaysia 51 39–54
[4] Chan W T, Chow Y K, Liu L F 1995. Neural network: An alternative to pile driving formulas. Computers and Geotechnics 17 135–56
[5] Goh A 1996 Pile driving records reanalyzed using neural networks Journal of Geotechnical Engineering 122 492–5
[6] Shahin M A, Maier H R and Jaksa M B 2002 Predicting settlement of shallow foundations using neural networks Journal of Geotechnical and Geoenvironmental Engineering 128 783–5
[7] Moayedi H, Mosallanezhad M and Rashid A 2020 A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: theory and applications Neural Computing and Applications 32 495–518
[8] Pooya N F and Jaksa M B 2017 Load-settlement behavior modeling of single piles using artificial neural networks and CPT data Computers and Geotechnics 89 9–21
[9] Stone M 1976. Cross-validatory choice and assessment of statistical predictions (with discussion) Journal of the Royal Statistical Society: Series B (Methodological) 102
[10] Hecht-Nielsen R 1989 Theory of the backpropagation neural network 593–605
[11] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 Dropout: A simple way to prevent neural networks from overfitting Journal of Machine Learning Research 15 1929–58