Non-Linear Multivariate Analysis with Artificial Neural Network in Estimating Compression Index for Cohesive Soils of Northern Jakarta Coast

Andri Mulia[1], Elyada Eben Ezer[1,*],M. Addifa Yulman[1], Shandy Yudha Isa[1], Kenandio Darmadi[1], Miftahurrohman[2], Aditya Hadyan Putra[2]

[1]PT Haskoning Indonesia, Jakarta, Indonesia, 12520; Andri.mulia@rhdhv.com
[2]PT Pelindo Solusi Logistik, Jakarta, Indonesia, 14230
*Correspondence: elyada.eben.ezer@rhdhv.com

ABSTRACT This study presents a novel application of artificial neural network (ANN) to develop a model for predicting compression index (C_c) of cohesive soils from their index properties. The model was trained using data from 347 undisturbed samples on a variety of cohesive soils from Northern Jakarta. It takes up to three variables as inputs: specific gravity (G_s), liquid limit (LL), and plastic limit (PL). The model was tested on a separate dataset of 117 samples and found to have a strong capability to predict C_c values when compared to some reference correlations. The ANN model has demonstrated good performance for each set by producing overall error of 29.6%, compared to 38.1% and 30.5% for the empirical formulas. This study shows that the application of ANN offers an essential advancement in this area, helping to overcome the limitation of conventional statistical correlation.

KEYWORDS Artificial neural network; Non-linear Multivariate Analysis; Compression Index; Northern Jakarta Cohesive Soil.

1 INTRODUCTION

Soils are naturally compressible, be subjected to volume changes in response to applied stresses (Balasubramaniam & Brenner, 1981). This phenomenon consists of three phases: immediate, primary consolidation, and secondary consolidation settlements (Das & Sobhan, 2018). Soils with high plasticity are predominantly encountered in North Jakarta Bay's, and the soil often experiences high consolidation settlement.

Assessing soil consolidation properties through testing often require a lot of time and geotechnical drilling operations yield only a limited number of undisturbed samples (UDS). Empirical equations offer a solution to reduce the cost, but the application of empirical equations to different sites can be questionable (Al-Taie et al., 2017). This paper explores a modern alternative named artificial neural networks (ANNs), and the utilization of ANNs is significantly increased in geotechnical engineering with success (Shahin et al. 2001). ANNs are superior in modeling complex relationships, making it an ideal tool for conditions which the variables' connections are elusive (Hubick, 1992).

This study aims to create a model to predict compression index (C_c) using soil properties gathered from 347 undisturbed soil samples in Northern Jakarta. The model's performance is subsequently assessed using an additional 117 undisturbed soil samples, allowing for a comparative evaluation against other empirical equations.

2 COMPRESSION INDEX

The soil settlement calculation under structural load is the most important aspect besides checking the integrity of the structure. Settlement is defined as the decrease in soil volume due to water flowing

out of soil and particle re-arrangement under the effect of applied pressure. Compression index (C_c) is one of the parameters that is used to calculate settlement. High C_c value means the soil is more compressive (Dwivedi et al., 2016a).

Oedometer test takes a lot of time, requires precision, precautions, and expertise in the process. Therefore, it is very tough to get ideal value of C_c. Furthermore, even a very small disturbance in the process can lead to overestimation or underestimation C_c value. Therefore, correlations are developed to limit time disadvantages of getting oedometer test results. Two popular correlations will also be compared for this study:

$$C_c = 0.009(LL - 10) \text{ (Terzaghi \& Peck, 1967)} \tag{1}$$

$$C_c = 0.2343 \left(\frac{LL}{10}\right) \times Gs \text{ (Nagaraj \& Murty, 1985)} \tag{2}$$

Although empirical correlations can provide a quick and inexpensive way to estimate soil parameters with simple tests, most of these correlations are derived from fitting data measurement made under specific site condition. This may cause large deviations when used at other sites (Dehghanian & Ipek, 2022)

3 ARTIFICIAL NEURAL NETWORK MODEL

3.1 Artificial Neural Networks (ANNs)

ANNs are artificial adaptive systems that are inspired by the functioning processes of human brain and nervous system (Grossi & Buscema, 2008). ANNs provide strong solutions to problems in several areas, including classification, prediction, filtering, optimization, pattern recognition, and function approximation.

The biological nervous system is extremely complicated; artificial neural networks algorithms seek to simplify this complexity and focus on what may theoretically matter most from an information-processing standpoint (Thakur & Konde, 2021). A comprehensive description of ANNs is beyond the scope of this paper. Many authors have described the structure and operation of ANNs (e.g., Hecht-Nielsen 1990; Maren et al. 1990; Zurada 1992; Fausett 1994; Ripley 1996).

The fundamental of artificial neural networks (ANNs) lies on their constituent elements which is the artificial neurons or processing elements (PEs). These PEs operate on a simple mathematical model defined by three basic rules consist of multiplication, summation, and activation. Within the artificial neuron, the input values are subject to weighting, whereas each input is multiplied by an associated weight. In the core of the artificial neuron resides a summation function that aggregates all the weighted inputs along with a bias term. Finally, the cumulative result of the weighted inputs and bias undergoes an activation process, often referred to as a transfer function (Andrej et al., 2011) at the output of the artificial neuron. Figure 1 illustrates the operational principle of an artificial neuron.
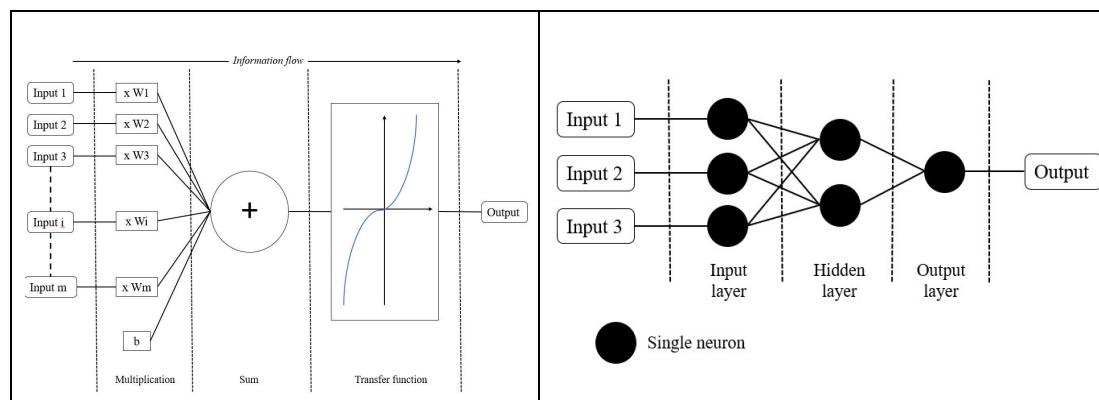

Figure 1. Working principles of an artificial neuron (left); Example of simple ANN network (right).

While the foundational principles governing this operation may appear deceptively simple, the true power and computational prowess of these models emerge when we begin to interconnect them within an artificial neural network (ANNs), as depicted in **Error! Reference source not found.**.

This interconnectedness generates the ability of ANNs to modify their own connections over time, thereby initiating a learning process that characterizes the entire ANN (Hebb, 1949). This process of connection modification is often referred to as the 'Law of Learning.' Notably, the dynamism of an ANN is essentially linked to time. To facilitate the modification of connections, the ANN requires continuous interaction with its environment, typically represented by data, over extended periods (Rosenblatt, 1958). This learning process is a key mechanism that defines ANNs as adaptive processing systems.

Neurons can be organized in any topological manner (e.g., one- or two-dimensional layers, three-dimensional blocks, or more-dimensional structures), depending on the quality and amount of input data. The most common ANNs are composed in a so-called forward topology (Wasserman, 1989; Aleksandar & Morton, 1990). Therefore, type of ANNs also adopted in this study. A certain number of PEs is combined to an input layer, normally depending on the amount of input variables. The information is forwarded to one or more hidden layers working within the ANN. The output layer, as the last element of this structure, provides the result. The output layer contains one PE only, regardless the result is a binary value or a single number. Figure 2 represents the most popular architecture of neural networks; forward propagation (Fahlman, 1988; Le Cun, 1989).
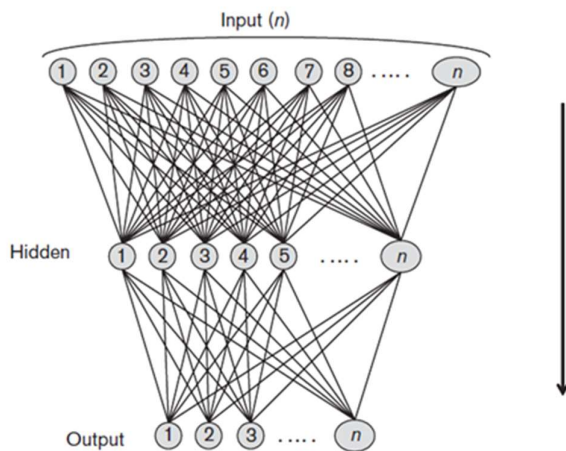

Figure 2 Forward Propagation neural network architecture.

3.2 Activation Function

Activation functions, also known as transfer functions, are instrumental in artificial neural networks. They convert input signals into outputs, which then serve as inputs for subsequent layers. Net inputs, central to the network's structure, undergo transformation into unit activations through these functions, constituting a scalar-to-scalar conversion (Sharma et al., 2020).

The Leaky Rectified Linear Unit is employed as activation function for this study. Leaky ReLU is defined as:

$$f(x_i) = \begin{cases} x_i, x_i > 0 \\ \alpha_i x_i, x_i < 0 \end{cases} \tag{3}$$

$$f'(x_i) = \begin{cases} 1, x_i > 0 \\ \alpha_i, x_i < 0 \end{cases} \tag{4}$$

Where α_i is the coefficient of i-th channel for negative inputs, and it is a small constant, typically 0.01 or so. Its output is not 0 for negative inputs, so it is the improvement of ReLU function for the problem "Dying ReLU" which encountered when ReLU is employed. It is straightforward for us to get its derivative. Figure 3 presents the function curve and its derivative curve. Because it can produce a constant times input value for negative inputs, compare to ReLU, where ReLU will not saturate for both directions (Feng & Lu, 2019).
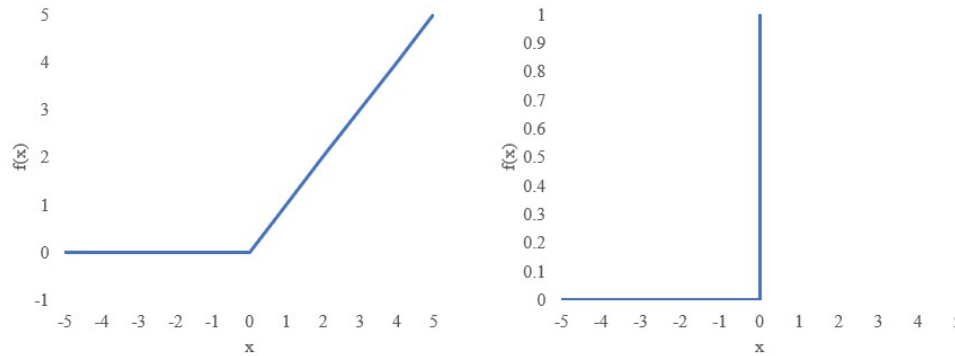

Figure 3 Leaky ReLU function curve (left); Leaky ReLU derivative curve (right)

3.3 Back Propagation

Back propagation is a fundamental technique employed in supervised learning and multi-layered training programs. It operates by leveraging errors both in the forward and backward propagation phases of neural network training (Izhari et al., 2020). This method involves adjusting the weights of a neural network based on error values obtained in previous epochs, emphasizing iterative refinement. Precisely tuned weights facilitate the reduction of error rates and enhance the model's reliability by expanding its predictive capabilities.

The term "back propagation" is a concise way of denoting "backward propagation of errors," and it serves as a standard procedure for training artificial neural networks (Soemartono et al. 2018). The back propagation process involves several key steps:

a) Error Rate Computation: This initial step requires calculating the discrepancy between the model's output and the actual target output.
b) Error Minimization: Subsequently, the process verifies whether the error has been minimized effectively.
c) Weight and Bias Updates: If the error exceeds an acceptable threshold, the weights and biases are updated accordingly. This cycle continues until the error converges to a satisfactory level.
d) Neural Network Model Finalization: Once the error rate falls within an acceptable range, the neural network model becomes ready for deployment, enabling it to effectively forecast data.

Figure 4 presents the workflows of back propagation mechanism (Sekhar & Meghana, 2020).

In the context of facilitating the back propagation process, the gradient descent (GD) algorithm assumes a pivotal role. This iterative optimization technique, widely employed in ANNs, enables the gradual adjustment of parameters represented by θ, with the overarching objective of minimizing the error function J(θ) (Mustapha et al. 2020). Notably, the GD algorithm employs the entire dataset for each parameter update, a precision-focused approach that, while accurate, demands a substantial computational workload.

In our study, we turn to the Adaptive Moment Algorithm (Adam) to efficiently navigate the optimization landscape. Adam is a sophisticated amalgamation of two prominent optimization methods, Momentum and RMSProp (Kingma & Ba, 2015) This algorithm excels in its ability to calculate adaptive learning rates tailored to each parameter. Optimization continues until the

objective function reaches its minimum, which corresponds to the smallest deviation between the predicted values and the test set values.

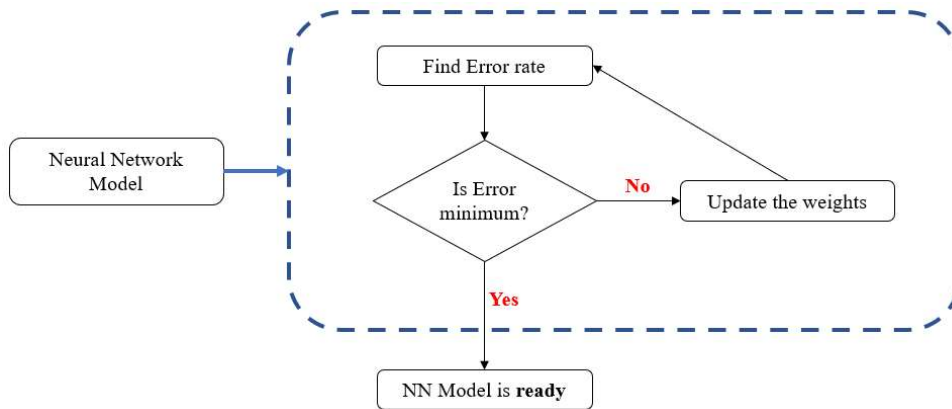$$\text{Overall error } (\delta) = \frac{\sum_{i=1}^{n}(x_{set,i} - x_{predicted,i})}{n} \tag{5}$$

Figure 4 Flow diagram of back propagation mechanism.

3.4 ANN Model Design

The dataset utilized for the training phase in this study comprised 347 undisturbed cohesive soil samples collected from multiple projects in northern bay area of Jakarta. This dataset incorporated three essential index properties as variables: specific gravity (G_s), liquid limit (LL), and plastic limit (PL). The objective is to employ these variables as inputs to predict the compression index (C_c) using Artificial Neural Networks (ANNs). ANN model was developed using Visual Basic for Application (VBA) in Microsoft Excel program. The VBA comprised algorithms as described in Chapter 2 using 8 hidden layers with 7 neurons in each hidden layer and Adam Optimizer. To assess the performance of the model, a separate test set comprising 117 samples contained similar properties as dataset were employed. Figure 5 provides a snapshot of compression index (C_c) from the dataset and test set in this study.
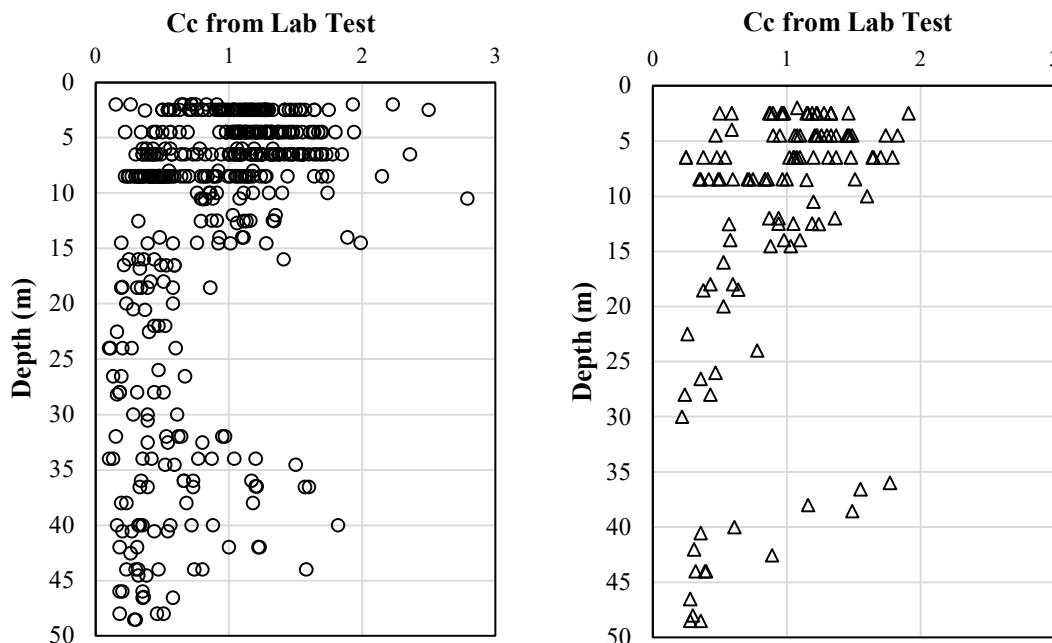

Figure 5 Compression index (C_c) from dataset (left); Compression index (C_c) from test set (right).

4 RESULT AND DISCUSSION

4.1 Training Phase

Training phase aimed to check the suitability of ANN model to predict C_c values. This was measured by comparing the overall error (δ) calculated using formula (5) between ANN model and reference equations. Figure 6 presents the calculated C_c by employing those equations compared with C_c from laboratory tests (data set). Whereas Figure 7 depicted similar features but for the created ANN model.

Notably, the bold black line represents the calculated or predicted C_c values that exactly align with the laboratory test (y = x), while the dotted line represented the trend line of results, originating from the intercept point at (0,0). It shall be noted that the features of a trend line, gradient and R-squared (R^2), could not be a suitable indicator. The reason is that these values were measured relative to its own trendline and not to the y = x line; hence, it is not the best performance reflection of a model. This is proven from Figure 6 which shows that although the trendline from Nagaraj's equation model has lower gradient value and R^2 than Terzaghi's, it has a lower overall error and becomes more suitable in this case.

Therefore, overall error (δ) of a model, which considered all the deviation for every single data point, was adopted as the indicator to determine the suitability of a model. The lower the value means the better the accuracy of a model.

ANN model resulted δ of 38.9%. This value was lower than that of the equations, Terzaghi's and Nagaraj's, with numbers 55.9% and 43.2%, respectively. This implies that ANN model produced a better accuracy than that of equations. In addition to that, for higher Cc values (larger than 1.5), ANN model can predict better than the empirical formulas. These findings provided a solid basis to verify that ANN model is suitable to predict the C_c values.
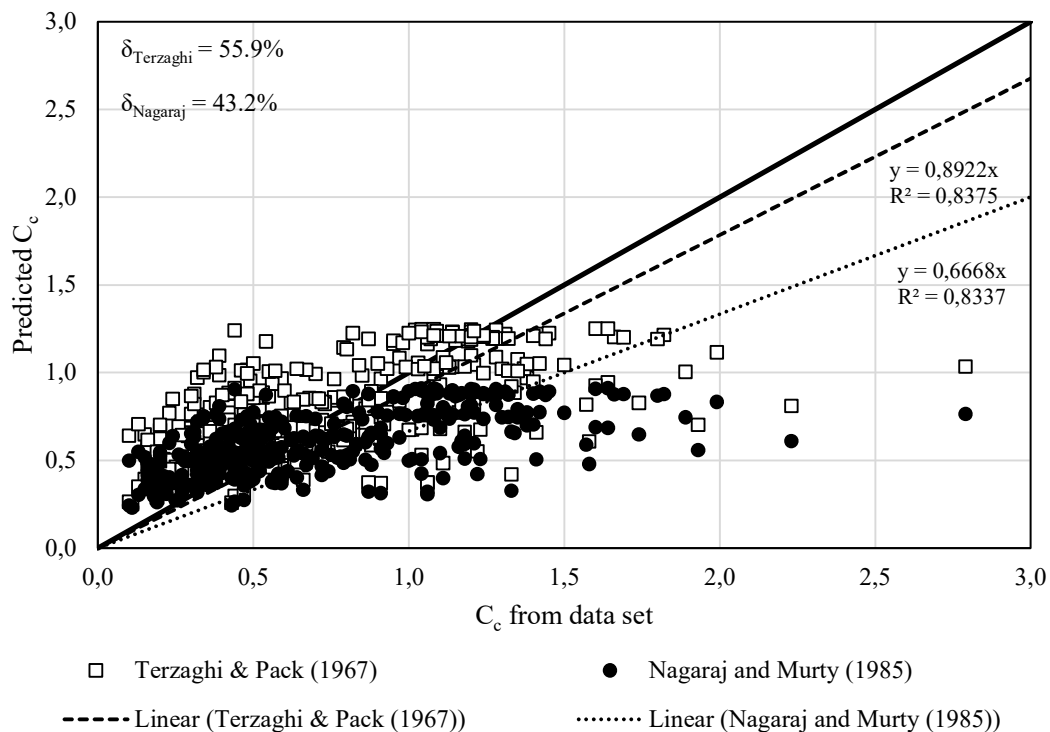

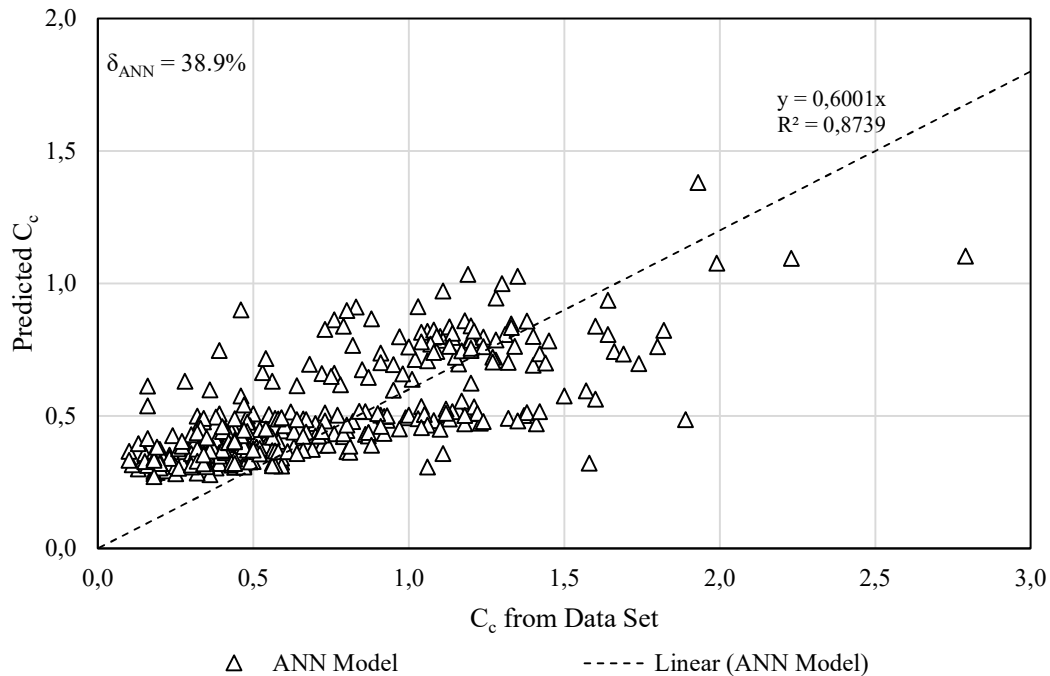Figure 6 Calculated Cc from equations based on data set.

Figure 7 Predicted C_c from ANN model based on data set.

4.2 Testing Phase

After the training phase, the trained ANN model was employed to predict the C_c values. These values then compared with C_c from the test set (117 data). Similar method was also conducted for models based on the equations. Figure 8 displays the results distribution for the ANN model and equations. The reference model, based on equations, resulted δ of 38.1% for Terzaghi's equation and δ of 30.5% for Nagaraj's. While the trained ANN model produced δ of 29.6%, smaller compared to the equations. This showed that the ANN model aligned with result during training phase, therefore applicable for determining C_c.
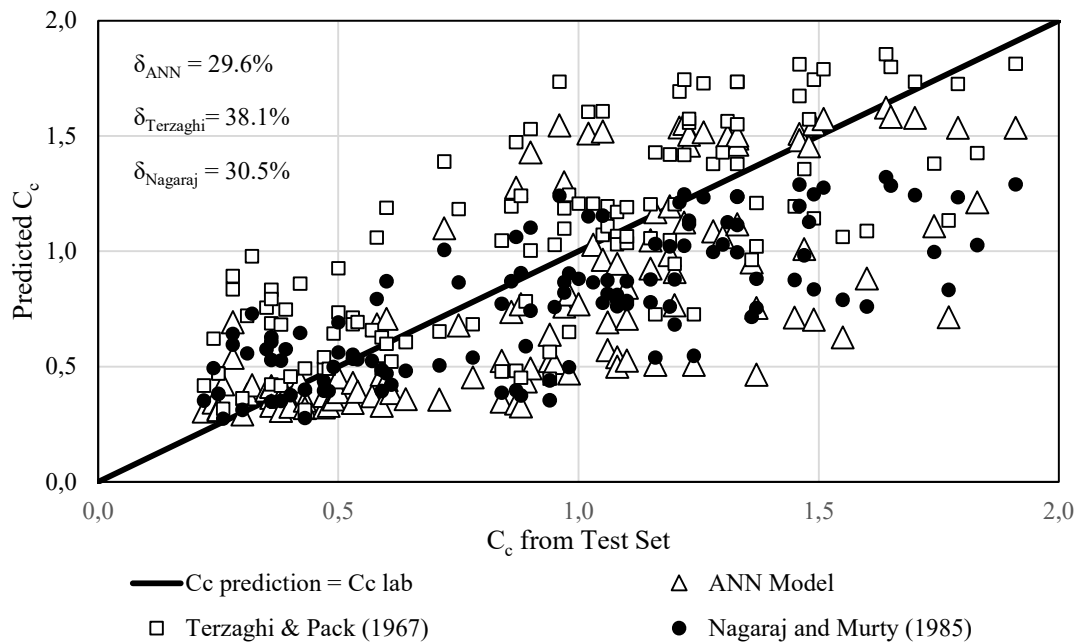

Figure 8 Calculated Cc compared to testing set.

5   CONCLUSION AND RECOMMENDATION

ANN model is proven to give a promising alternative solution to predict the value of compression index (C_c). ANN examination provides good prediction compared to existing empirical formulas. In addition, the ANN model is developed by considering the typical soil condition in particular area, in this case North Jakarta's Bay whereas the empirical formula is not suited to North Jakarta's condition .

Further studies could improve this model's reliability by exploring other activation functions or other gradient descent algorithms.

6   REFERENCES

Aleksandar, I. & Morton, H., 1990. *An introduction to neural computing.* London: Chapman & Hall.

Al-Khafaji, A. & Andersland, O., 1992. Equations for compression index. *Journal of Geotechnical Engineering,* 118(1), pp. 148-153.

Al-Taie, A. J., Al-Bayati, A. F. & Taki, Z. N. M., 2017. Compression index and compression ratio prediction by artificial neural networks. *Journal of Engineering,* 23(12), pp. 96-97.

Balasubramaniam, A. & Brenner, R., 1981. *Consolidation and settlement of soft clay.* Amsterman - Oxford - New York, Elsevire Scientific Publishing Company.

Carter, M. & Bentley, S., 1991. *Correlations of soil properties.* London: Pentech Press.

Das, B. & Sobhan, K., 2018. *Principles of geotechnical engineering.* Ninth ed. Boston: Cengage Learning.

Dehghanian, K. & Ipek, S. Ö., 2022. A survey on the relationships between compression index, coefficient of consolidation, and atterberg limits. *Journal of Sustainable Construction Materials and Technologies,* 7(4), pp. 302-315.

Dwivedi, P., Kumar, R. & Jain, P. K., 2016. Prediction of compression index (Cc) of fine grained remolded soils. *International Journal of Applied Engineering Research,* 11(1), pp. 592-598.

Fahlman, S., 1988. *An empirical study of learning speed in back-propagation networks,* Pittsburg: Carnegie-Mellon University.

Fausett, L., 1994. *Fundamentals neural networks: architecture, algorithms, and applications.* New Jersey: Prentice-Hall, Inc..

Feng, J. & Lu, S., 2019. Performance analysis of various activation. *Journal of Physics: Conference Series.*

Figueiredo, D., Junior, S. J. & Rocha, E. C., 2011. What is r2 all about?. pp. 60-68.

Grossi, E. & Buscema, M., 2008. Introduction to artificial neural networks. *European Journal of Gastroenterology & Hepatology,* pp. 1046-1054.

Hebb, D., 1949. *The organization of behavior.* New York: Wiley.

Hecht-Nielsen, R., 1990. *Neurocomputing.* s.l.:Addison-Wesely Publishing Company.

Hubick, K., 1992. *Artificial neural networks in australia,* Canberra: Department of Industry, Technology and Commerce,.

Hutabarat, L. E., Rainaldo, D., Ilyas, T. & Prakoso, W. A., 2019. Site characterization of marine clay consolidation ratio on kamal. *IOP Conference Series: Earth and Environmental Science,* 258(012020).

Izhari, F., Zarlis, M. & Sutarman, 2020. Analysis of backpropagation neural neural network algorithm on student ability based cognitive aspects. *IOP Conference Series: Materials Science and Engineering.*

Kingma, D. & Ba, J., 2015. Adam: a method for stochastic optimization. *3rd International Conference on Learning Representations,* pp. 1-15.

Krenker, A., Bester, J. & Kos, A., 2011. *Introduction to the artificial neural networks, artificial neural networks - methodological advances and biomedical applications.* s.l.:InTech.

Le Cun, Y., 1989. Generalization and network design strategie. In: R. Pfeifer et al, ed. *Connectionism in perspective.* North Holland: Amsterdam, pp. 143-156.

Mustapha, A., Mohamed, L. & Ali, K., 2020. An overview of gradient descent algorithm optimization in machine learning: Application in the ophthalmology field. *International Conference on Smart Applications and Data Analysis,* pp. 349-359.

Mustapha, A., Mohamed, L. & Ali, K., 2020. Comparative study of optimization techniques in deep learning: Application in the ophthalmology field. *Journal of Physics: Conference Series.*

Nagaraj, T.S. & Murthy, B.R.S., 1985. A critical reappraisal of compression index equations. Géotechnique, 36(1), pp. 27-32.

Ranjan, G. & Rao, A., 2000. *Basic and applied soil mechanics.* 2nd ed. New Delhi: New Age International.

Sekhar, C. & Meghana, P., 2020. A study on backpropagation in artificial neural networks. *Asia-Pacific Journal of Neural Networks and Its Applications,* Volume 4, pp. 21-28.

Shahin, M., Jaksa, M. B. & Maier, H. R., 2001. Artifical neural network applications in geotechnical. *Australian Geomechanics Jorunal,* p. 49.

Sharma, S., Sharma, S. & Athaiya, A., 2020. Activation functions in neural networks. *International Journal of Engineering Applied Sciences and Technology,* 4(12), pp. 310-316.

Soemartono, T., Windarto, A. P., Herawan, T. & Budiharjo, 2018. Predicting tuition fee payment problem using backpropagation. *International Journal of Advanced Science and Technology,* Volume 120, pp. 85-96.

Terzaghi, K. & Peck, R.B., 1967. *Soil mechanics in engineering application (3rd edition).* The Macmillan Company, London, England.

Thakur, A. & Konde, A., 2021. Fundamentals of neural networks. *International Journal for Research in Applied Science & Engineering Technology (IJRASET),* 9(VIII), pp. 407-426.

Wasserman, P., 1989. *Neural computing: theory and practice.* New York: Van Nostrand.

- This page is intentionally left blank -