Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-2385

Proposal of estimation model for clay consolidation behavior using artificial neural networks and Monte Carlo cross-validation

Kazuhiro Oda
Department of Architecture and Environment design, Osaka Sangyo University, Daito, Japan, oda@aed.osaka-sandai.ac.jp

ABSTRACT: This paper presents an estimation model for the consolidation behavior of spatially variable clay soils. The proposed framework uses an Artificial Neural Network (ANN) integrated with Monte Carlo cross-validation (MCCV) to estimate the *e*-log *p*' relationship and relationship between coefficient of permeability and void ratio from sparse geotechnical investigation data. The ANN architecture utilizes geospatial coordinates and soil state variables as inputs to continuously estimate the soil's response. MCCV is employed to generate a robust ensemble of models, mitigating biases associated with specific training data configurations and enhancing predictive reliability. The model's performance was rigorously validated against oedometer test results from locations not included in the training set. The estimations demonstrated excellent correlation with experimental data, successfully capturing complex non-linear behaviors. Furthermore, the framework was applied to a practical case by incorporating it into a soil-water coupled finite element analysis. The simulation accurately reproduced the observed time-settlement history of the reclaimed island, verifying the model's practical applicability. This study concludes that the proposed ANN-MCCV framework offers a promising approach for geotechnical design and analysis, with the potential to enhance the characterization of heterogeneous ground conditions.

KEYWORDS: Artificial neural network, Monte Carlo cross-validation, Consolidation, Numerical simulation.

1 INTRODUCTION

Predicting ground deformation during and after construction of structures on soil remains one of the most challenging problems in geotechnical engineering. It is necessary to determine the ground model at construction sites based on limited geotechnical investigation results. Kriging is a representative approach for this purpose. In recent years, the application of Gaussian processes has become increasingly applied (Ching & Yoshida 2023). Meanwhile, reproducing the mechanical behavior of geomaterials is also important, and research on constitutive models and their application to practical problems has been actively conducted. Currently, artificial intelligence technologies are being applied to constitutive models (Zhang et al. 2023).

In this study, we propose an estimation model for the spatial distribution of both *e*-log *p*' relationship and the relationship between coefficient of permeability and void ratio using an artificial neural network (ANN), which are representative artificial intelligence technologies. Furthermore, we reproduce the consolidation settlement of seabed ground associated with the construction of an actual reclaimed island through consolidation settlement simulation applying the estimation models derived by the proposed method and examine the effectiveness of the proposed approach.

2 ARTIFICIAL NEURAL NETWORK-BASED ESTIMATION MODEL FOR CONSOLIDATION BEHAVIOR

2.1 *Artificial neural network*

The two estimation models developed in this study estimate the *e*-log *p*' relationship and the relationship between coefficient of permeability and void ratio at arbitrary locations, respectively. Specifically, the estimation model for the *e*-log *p*' relationship employs latitude, longitude, depth (as geographical coordinates), and consolidation pressure as explanatory variables, with void ratio as the target variables. On the other hand, the estimation model for the relationship between coefficient of permeability and void ratio uses latitude, longitude, depth, and void ratio as explanatory variables, with coefficient of permeability as the target variables.

Figure 1 schematically illustrates the structure of the ANN used in the estimation model for the *e*-log *p*' relationship. An ANN is a representative form of artificial intelligence that mimics the neural circuits of the human brain. This ANN consists of an input layer, an output layer, and hidden layers that connect these components. Explanatory variables are provided to the network through the input layer. After computational processing within the ANN, the estimation results are delivered through the output layer. In Figure 1, four lines extend from each blue solid circle to connect with hollow circles. The hollow circles represent neurons. Neurons transmit signals from the preceding layer to the subsequent layer. The signals transmitted within neurons are mathematically manipulated using the following equations:

$$\theta = \sum_{i=1}^{n} w_i^k x_i^k + \beta_j^k \tag{1}$$

$$x_j^{k+1} = f(\theta) \tag{2}$$

where k is the level of the hidden layer, n is the number of neurons in the k-th hidden layer, x^k_i is the signal transmitted from the i-th neuron of the (k-1)-th hidden layer to the j-th neuron of the k-th hidden layer, w^k_i is the weight of the i-th signal at the k-th level, and β^k is the bias at the k-th level. The signal transmitted from the j-th neuron at the k-th level to the neurons in the (k+1)-th hidden layer is calculated by the activation function f. Hidden layers forcibly associate the explanatory variables from the input layer with the target variables in the output layer, even without theoretical relationships. Furthermore, by making the hidden layers more complex, i.e., increasing the number of neurons and adding more hidden layers, the expressive ability for the target variables in the output layer is enhanced. However, when implementing deeper hidden layers, the vanishing gradient problem may occur during the backpropagation process depending on the choice of activation function. Compared to sigmoid and tanh functions, the ReLU function is less prone to vanishing gradients because its first derivative is always unit when θ is positive. Therefore, this study also employs the ReLU function as the activation function. In this research, a fully connected model with 16 hidden layers and 250 neurons was utilized.

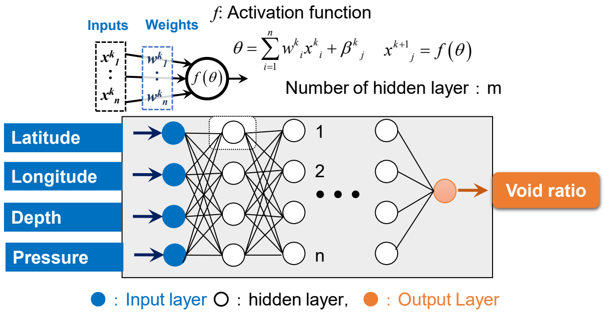

Figure 1. Structure of artificial neural network used in this study.

2.2 *Monte Carlo cross-validation*

To develop an ANN as an estimation model, a "training" process is essential. In this process, the unknown parameters w_i and β are optimized. During the training phase, a dataset of predefined explanatory variables and actual target variables is utilized, and w_i and β are iteratively adjusted through the backpropagation algorithm to minimize the discrepancy between the estimated target variables and the actual target variables. In this study, 75% of the total data was used for training. Subsequently, the generalization performance of the model was validated using the remaining 25% of the data that was not used during training. By comparing the training error with the validation error during generalization performance assessment and applying early stopping to terminate the iterative computation before the difference between them becomes significant, overfitting was prevented while achieving computational efficiency.

In constructing the estimation models, training data is randomly extracted from the entire dataset. Different training data results in different estimation models, which consequently yield different outputs for the same input values. To minimize such dependency of estimated values on the training data, Monte Carlo cross-validation (MCCV) was applied (Xu & Liang 2001). Figure 2 conceptually illustrates the MCCV method. MCCV is a cross-validation technique that evaluates model prediction performance by repeatedly and randomly partitioning the entire dataset into training data and validation data. In this study, 10 iterations were performed.

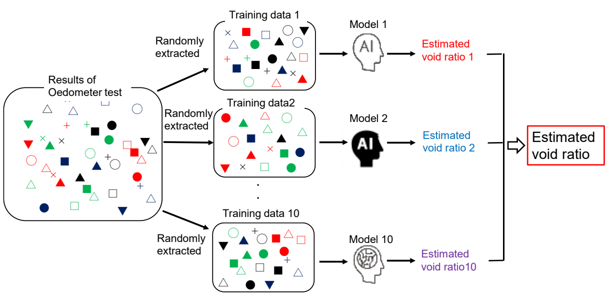

Figure 2. Conceptual view of Monte Carlo cross-validation.

3 STUDY AREA

3.1 *Target reclamation site*

Figure 3 shows the reclaimed island examined in this study. The blue solid circles (●) in the figure indicate the locations where geotechnical investigations were conducted and used for training the estimation models for both the *e*-log *p*' relationship and the relationship between coefficient of permeability and void ratio. Points A and B (both indicated by ●) represent geotechnical investigation sites used to validate the performance of the estimation models. KC-1 (●) indicates the location where ground surface settlement has been measured. Since the geotechnical investigation results from Points A and B were not used for training the estimation models, they are independent of the estimation models. At KC-1, settlement monitoring for runway maintenance has been conducted since the reclamation construction began. However, since no geotechnical investigation has been performed at this location, the consolidation characteristics remain unknown.

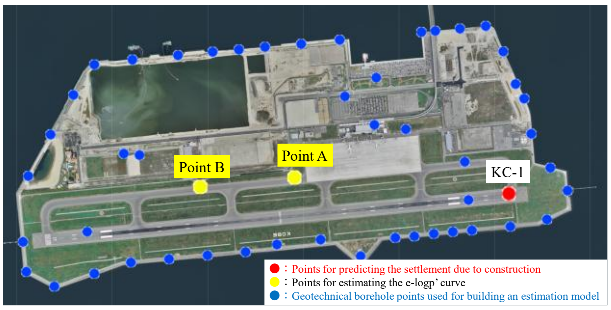

Figure 3. Reclaimed artificial island examined in this study.

3.2 *Geological profile of the study area*

Figure 4 shows the geological profile of the targeted reclamation site. Ma13 represents a Holocene clay layer, while Ma12 to Ma9 are Pleistocene clay layers, respectively. Ds1 to Ds10 are Pleistocene formations composed primarily of gravel and sand. Since Ma13 is extremely soft, ground improvement using sand drains has been implemented around the runway area. Consequently, the consolidation behavior of Ma13 differs significantly from that of Ma12 and deeper layers where no ground improvement has been performed. Additionally, detailed geotechnical investigations have been conducted primarily for Ma13. Geotechnical investigations extending to Ma12 and deeper layers are limited to only a few boreholes. Therefore, this study focuses exclusively on Ma13.

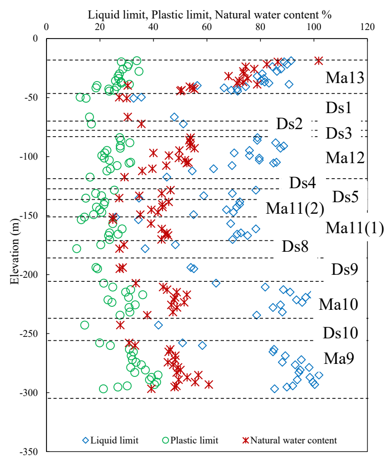

Figure 4. Geological profile of the targeted reclamation site.

4 VALIDATION OF ESTIMATION MODEL

4.1 *e-logp' relationship*

Figures 5 and 6 show the estimated *e*-log *p*' relationships along with the results from oedometer tests. The estimated values are in good agreement with the results from the oedometer tests. In particular, the estimation models successfully capture the consolidation characteristics that differ due to aging effects, such as the behavior shown in Figure 5 where compressibility increases rapidly after yielding but gradually decreases as pressure increases, and the behavior illustrated in Figure 6 where compressibility after yielding remains relatively constant regardless of pressure magnitude. Therefore, by combining ANN with MCCV, appropriate estimation models for e-log p' relationships at arbitrary locations can be constructed.

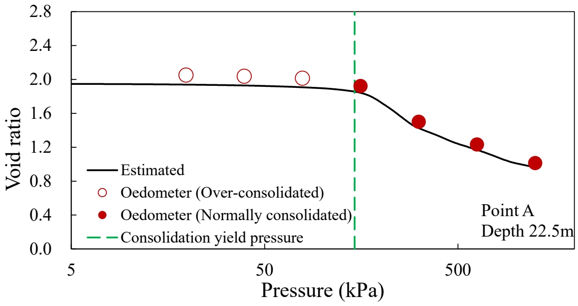

Figure 5. Estimated e-log p' relationship (Point A, depth 22.5m).

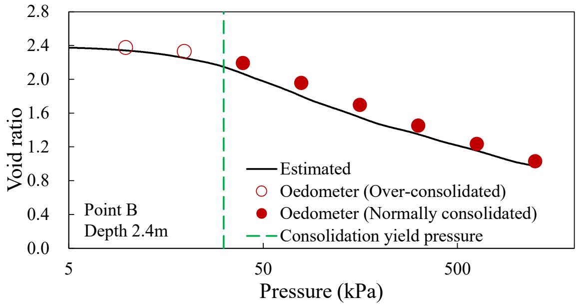

Figure 6. Estimated e-log p' relationship (Point B, depth 2.4m).

4.2 *Relationship between coefficient of permeability and void ratio*

Figures 7 and 8 show the estimated relationships between coefficient of permeability and void ratio along with the results from oedometer tests. At Point A (Figure 7), under normally consolidated conditions, the estimated relationship between coefficient of permeability and void ratio can approximately follow the results from the oedometer test. At Point B (Figure 8), the estimated coefficient of permeability is somewhat larger than the results from the oedometer test, but the trend of coefficient of permeability with decreasing void ratio. Therefore, likewise the e-log p' relationships, for the relationship between coefficient of permeability and void ratio, estimation model can be generally constructed by combining ANN with MCCV.

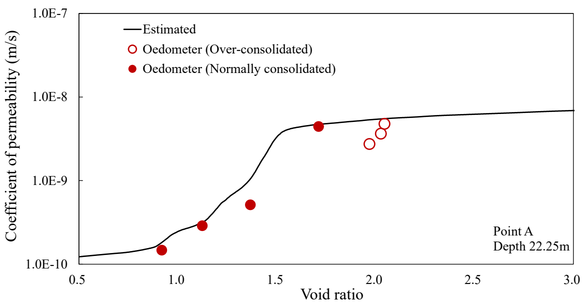

Figure 7. Estimated relationship between coefficient of permeability and void ratio (Point A, depth 22.5m).

To validate the capability of the proposed models to reproduce consolidation settlement behavior, a simulation of the settlement at the KC-1, as shown in Figure 3, due to land reclamation was performed. Figure 9 shows the analysis model. The sand drains have been implemented to accelerate the consolidation due to the reclamation. Therefore, pore water flows primarily in the horizontal direction and drains through the sand drain elements.

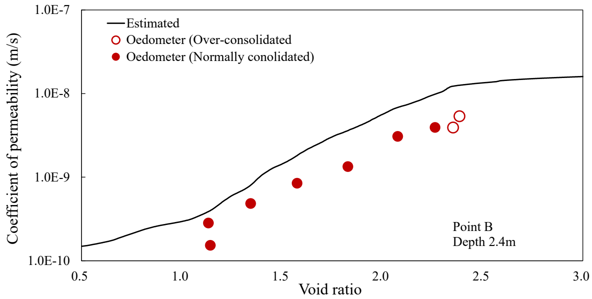

Figure 8. Estimated relationship between coefficient of permeability and void ratio (Point B, depth 2.4m).

To simulate this effect, the ground was modeled under axisymmetric conditions, with a drained boundary condition applied to the elements corresponding to the sand drains. Furthermore, because the sand drains are partially penetrating and do not reach the Ds-1 stratum (see Figure 4), an undrained boundary condition was applied to the non-penetrating base. The model was discretized into a mesh of 144 elements, arranged in 16 layers vertically and 9 columns horizontally. The consolidation settlement simulation was conducted using a coupled soil-water finite element method. The respective ANN estimation models for the e–logp' relationship and the relationship between coefficient of permeability and void ratio, as determined in the chapter 4, were incorporated directly into the analysis code. This approach allows the simulation to account for the changes in compressibility and permeability for each element as the consolidation pressure and void ratio evolve during the consolidation process. It should be noted that while the input variables for the estimation models include latitude, longitude, depth, and either consolidation pressure or void ratio (as shown in Figure 1), this simulation used the specific coordinates of KC-1 for latitude and longitude, and the initial state values for depth.

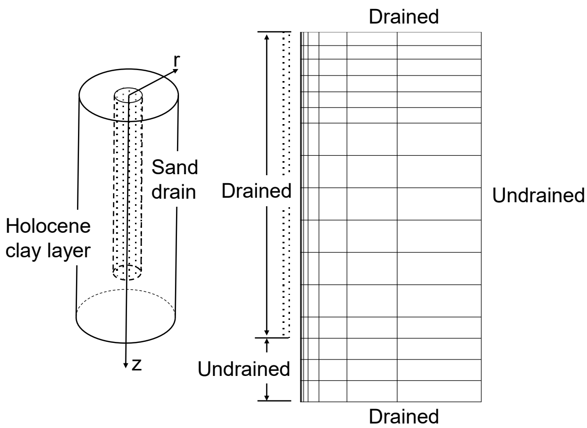

Figure 9. Analytical model.

Figure 10 shows examples of e–logp' relationship used in the consolidation settlement simulation. In the figure legend, for a label such as "KC-1_02," "KC-1" denotes the location of settlement measurement, and "02" indicates the second layer from the top of the analytical model. For layers 2 to 7, the e–logp' relationships do not show a significant jump in consolidation yield stress (p_c'), and their post-yield behavior nearly overlaps with the e–logp' relationship of the 2nd layer. In contrast, the relationships for the 9th and 10th layers exhibit a remarkable jump in consolidation yield stress, forming a so-called inverse S-shaped curve. For the 12th layer and below, this sharp increase in yield stress is no longer apparent.

Furthermore, in the 13th and 16th layers, the post-yield *e*–log*p*' relationships plot below that of the 2nd layer. This collective behavior accurately reproduces the characteristic consolidation properties of the ground at this site (Hasegawa et al. 2007).

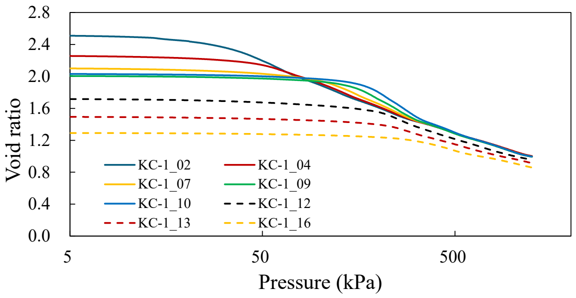

Figure 10. Examples of *e*–log*p*' relationship used for the consolidation settlement simulation.

Figure 11 shows examples of the relationship between coefficient of permeability (*k*) and the void ratio (*e*) used for the consolidation settlement simulation. At this location, the initial void ratio is approximately 3.0 near the seabed and decreases to about 1.3 in the vicinity of the Ds-1 stratum at a depth of about 30 m. Furthermore, considering the change in void ratio based on the *e*-log*p*' relationship shown in Figure 10, the coefficient of permeability for any given layer can be inferred to be in the range of approximately 10^{-8} to 10^{-9}m/s.

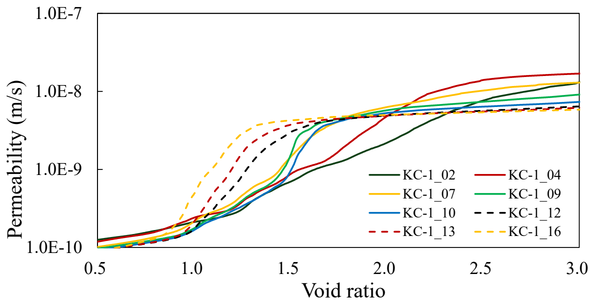

Figure 11. Examples of relationship between void ratio (*e*) and coefficient of permeability (*k*) used for the consolidation settlement simulation.

Figure 12 shows the time history of the reclamation load. The land reclamation work commenced in August 2000 and was completed in April 2004. The maximum fill load applied was approximately 400 kPa.

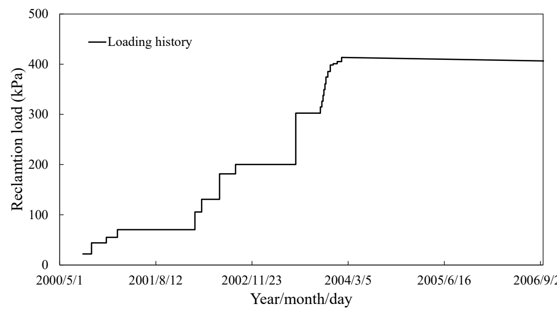

Figure 12. Time history of the reclamation load.

Figure 13 shows the time history of settlement due to land reclamation. Until approximately July 2003, the simulated results are in excellent agreement with the measured values. After this point, the measured settlement becomes larger than the simulated settlement. The measurements were conducted for the purpose of runway maintenance. Therefore, they include not only the settlement of the Holocene clay layer (Ma13), which is the target of the current analysis, but also the settlement of the underlying Pleistocene strata (the Ma12 layer and deeper). The significant discrepancy between the simulated and measured values becomes apparent because the influence of settlement in these deep Pleistocene strata has become more significant (Hasegawa et al. 2006). This settlement within the Pleistocene strata is a common phenomenon for reclaimed islands in Osaka Bay, with the settlement issues at Kansai International Airport (KIX) being a well-known example (Furudoi 2010).

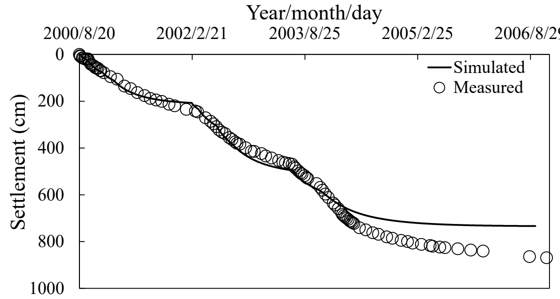

Figure 13. Time history of settlement due to land reclamation.

5 CONCLUSIONS

The main findings of this study are as follows:

1. By combining an Artificial Neural Network and Monte Carlo cross-validation, the continuous *e*-log *p*' relationship and the relationship between coefficient of permeability and void ratio can be estimated at any arbitrary location.
2. A consolidation settlement simulation utilizing the estimated *e*-log *p*' relationship and relationship between coefficient of permeability and void ratio can accurately reproduce the measured results from the field.
3. The proposed framework provides a reliable and powerful tool for geotechnical design and analysis, offering a significant advancement in characterizing heterogeneous ground conditions.

REFERENCES

Ching, J. & Yoshida, I.: Data-Drive Site Characterization for Benchmark Examples: Sparse Bayesian Learning versus Gaussian Process Regression, *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems*, https://doi.org/10.1061/AJRUA6.RUENG-983.

Furudoi, T. 2010. The Second Phase Construction of Kansai International Airport Considering the Large and Long-Term Settlement of the Clay Deposits", *Soils & Foundations*, 50(6), 805-816.

Hasegawa, N., Matsui, T., Tanaka, Y., Takahashi, Y., and Nambu, M. 2006. In-situ consolidation behavior of Pleistocene clay below seabed at Kobe Airport, *Journal of the Japan Society of Civil Engineers C,* 62(4), pp. 780-792, (in Japanese).

Hasegawa, N., Matsui, T., Tanaka, Y., Takahashi, Y., and Nambu, M. 2007. Consolidation Properties of Holocene Layers below Seabed at Kobe Airport, *Journal of the Japan Society of Civil Engineers C,* 63(4), pp. 923-935, (in Japanese).

Xu, Q. S. and Liang, Y. Z., 2001. Monte Carlo cross validation, *Chemometrics and Intelligent Laboratory Systems*, 56, 1-11.

Zhang, P., Yin, Z. Y. & Sheil, B. 2023. Interpretable data-driven constitutive modelling of soils with sparse data, *Computers and Geotechnics*,160. https://doi.org/10.1016/j.compgeo.2023.105511.