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/ICSMGE2026-2312

Optimizing pile foundation design using machine learning surrogate model

Yoshimasa Shigeno, Tomohiro Tanikawa
Research and Develop Institute, Takenaka Corporation, Japan, shigeno.yoshimasa@takanaka.co.jp

ABSTRACT: This study investigates the optimizing of pile design using a machine learning-based surrogate model for a seismically isolated building with a piled-raft foundation and grid-form deep mixing walls (DMWs). The surrogate model was developed using a deep neural network (DNN) to predict the maximum pile forces during earthquakes. This replaces the time-consuming process of 3D soil-structure interaction (SSI) nonlinear dynamic analysis process. The DNN uses the physical properties of the piles and the input seismic motions as inputs, and the maximum cross-sectional force distribution as output values. The input seismic motions were indexed by characteristics that showed good loss convergence. The estimated pile cross-sectional force by the surrogate model was highly accurate, with an error of approximately less than 1.0%. Compared to finite element methods (FEM), the surrogate model can provide estimated values in a very short time, enabling the rapid evaluation of thousands of design alternatives. The minimum margin and the cost function were set, and the surrogate model was used to optimize the pile design. This approach successfully demonstrated its ability to quickly identify balanced, cost-effective pile designs that meet safety requirements by completing thousands of design evaluations in seconds.

KEYWORDS: Surrogate model, nonlinear dynamic analysis, piled raft, deep mixing walls.

1 INTRODUCTION

For a seismically isolated building that combines grid-form ground improvement with a piled-raft foundation, we created a surrogate model using machine learning to estimate the maximum distribution of pile sectional forces during earthquakes, aiming to rationalize pile design. Generally, pile design for seismic loads uses quasi-static methods. However, when a detailed analysis is required, such as consideration of three-dimensional behavior or improved soil body failure, nonlinear dynamic analysis with a three-dimensional soil-structure interaction (3D SSI) model is necessary. Nevertheless, since these analyses require significant computational time, examining numerous cases is difficult. Replacing the analysis component with a surrogate model enables rapid evaluation and facilitates the exploring more rational pile design solutions.

2 OUTLINE OF THE BUILDING AND ANALYSIS MODEL

2.1 *Building*

The subject of the examination is a 13-story base-isolated residential building constructed in downtown Tokyo in 2008. Figure 1 shows the ground profile and the vertical section of the building. The ground is very soft with silty clay up to GL-44 m, and liquefaction is feared to occur between GL-4.8 m and GL-7.9 m. Grid-type ground improvement is applied to prevent liquefaction under the raft. The foundation structure is a piled-raft foundation with this grid-type ground improvement. In this building, seismic waves are observed at the points shown in the figure, and seismic data with a maximum acceleration of 174.8 gal (EW direction) was obtained at point A (Figure 1) during the Great East Japan Earthquake in 2011.

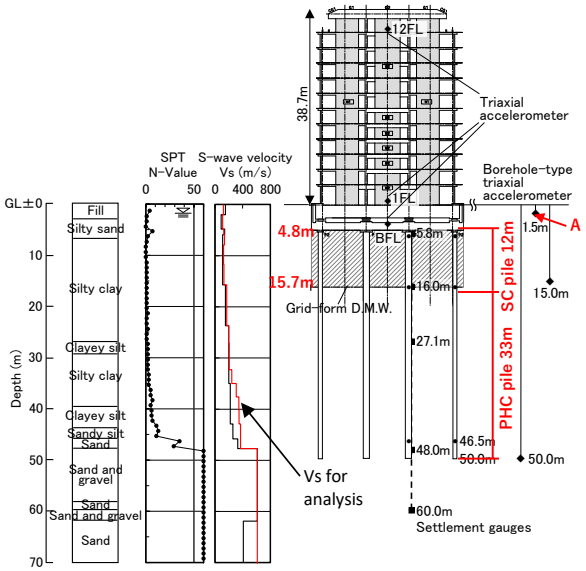

Figure 1. Ground profile and building structural elevation drawing.

2.2 *FEM model*

In the existing research, we have constructed a 3D SSI model that can reproduce the response of the observed seismic waves (Shigeno et al. 2017). Figure 2 shows the FEM model. Nonlinear models are used for the ground and the improved wall. The Ishihara-Yoshida model is used for the ground, and the elasto-plastic model developed by Namikawa and Mihara (2007), which considers tensile fracture and softening, is used for the improved wall. For details on the physical properties of the ground, refer to Shigeno et al. (2017), and the improved soil, refer to Namikawa and Mihara(2007) and Shigeno et al. (2024).

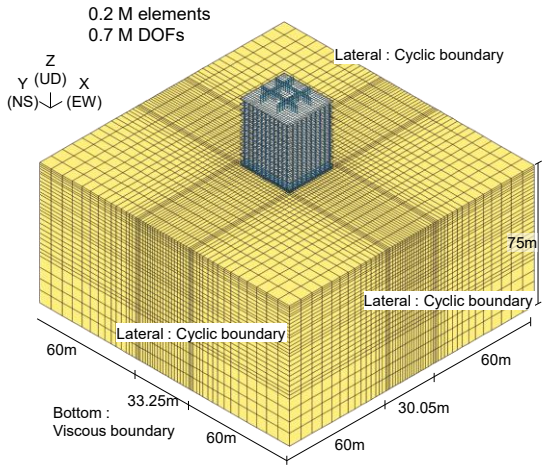

Figure 2. FEM model.

2.3 *Input seismic motion*

Artificial waves are used for strong seismic input motions called a “Level 2 earthquake”, which is officially notified in the Japanese building design code. This corresponds to the

"Maximum considered earthquake" in the U.S. code. The acceleration response spectrum is defined as shown in Figure 3. The waves were generated using the Kobe, Hachinohe, Kushiro, and random phase data. The Kobe phase data is based on the observations from the 1995 South Hyogo Prefecture earthquake in 1995. The Hachinohe data is from the 1968 Tokachi-oki earthquake. The Kushiro phase data is from the 1993 Kushiro-oki earthquake. The random phase data was generated using a random function. Figure 4 shows the input motion (2E) of the Kobe wave. The maximum acceleration is 355.7 gal. Seismic motions were input in the NS direction, and the distribution of the maximum values of pile cross-sectional forces were obtained.

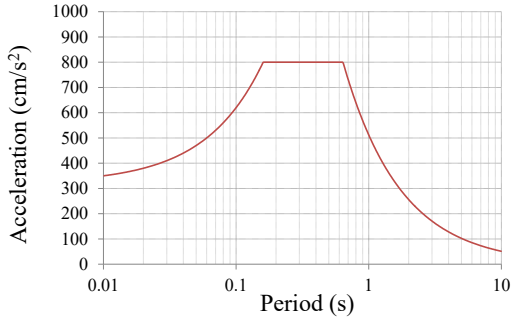

Figure 3. Acceleration response spectrum of Level 2 earthquake.

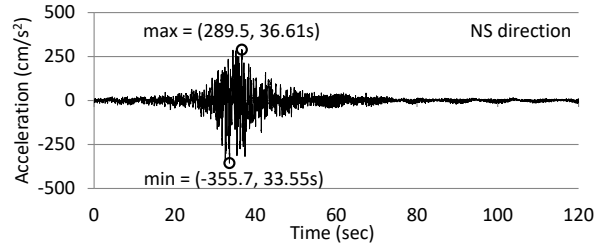

Figure 4. Input acceleration wave (Kobe wave) at a depth of 75 m (2E).

3 CONSTRUCTION OF SURROGATE MODEL

3.1 *Architecture of deep neural network*

A surrogate model for predicting the distribution of the maximum value of the pile section force using Deep Neural Network (DNN) was constructed by machine learning. The inputs to the surrogate model were the pile specifications and the indexed input seismic motions. The outputs were the distributions of the maximum pile section force values. Figure 5 shows the arrangement of the improved walls and piles. The piles are precast. Piles 1A, 1B and 3B represent the three groups of piles shown in the figure, which were classified based on the diameter and the pile arrangement. Pile groups 1A and 1B do not have strictly symmetrical spacing between the improved walls. However, it is reasonable to use piles 1A and 1B as representatives, because the greater spacing between the improved walls results in greater sectional forces.

Piles are divided into upper and lower piles as shown in Figure 1. The upper pile is a steel concrete (SC) pile. The lower pile is a prestressed high-strength concrete (PHC) pile. In this study, only the specifications of the upper piles were changed. For the upper pile, a reinforced prestressed concrete (PRC) pile was also considered. Since there are different ways of indexing the input seismic motions, we chose the method with the best machine learning error convergence process. The DNN had four intermediate layers and 100 neurons in each layer. The conceptual diagram of DNN is shown in Figure 6.

The input layer had 10 to 12 nodes with 9 physical properties of the piles (mass density, section area, and second moment of section) and 1 to 3 characteristic values of the input seismic motion. Each upper pile was divided into 7 elements in the FEM model. The bending moment was evaluated at two points for each element. Excluding duplicates, there were 8 bending moments. The shear forces were 7. There were 14 axial forces in both positive and negative directions. The total number of nodes in the output layer for three piles were 87.

The Mean Square Error (MSE) was used for the loss, which is the error between the predicted value and the correct value of the model. The MSE is defined by following equation.

$$MSE = \frac{1}{n}\sum(\hat{y} - y)^2 \quad (1)$$

where $\hat{y}$ is the predicted value, y is the actual value, and n is the number of points.

The ReLU function was used for the activation function of the neuron, which has many applications. The function is defined by the following equation.

$$f(x) = \max(0, x) \quad (2)$$

where x is the input value to each neuron.

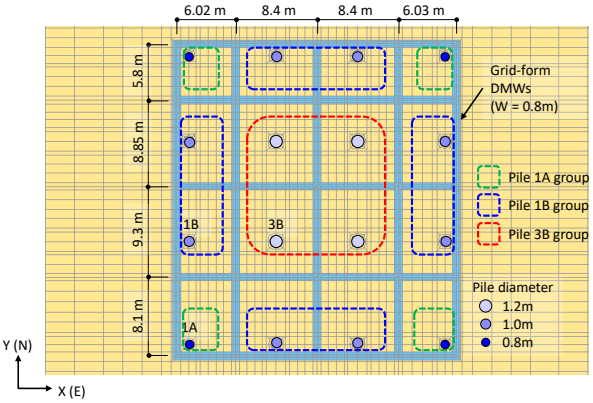

Figure 5. Foundation plan and pile grouping.

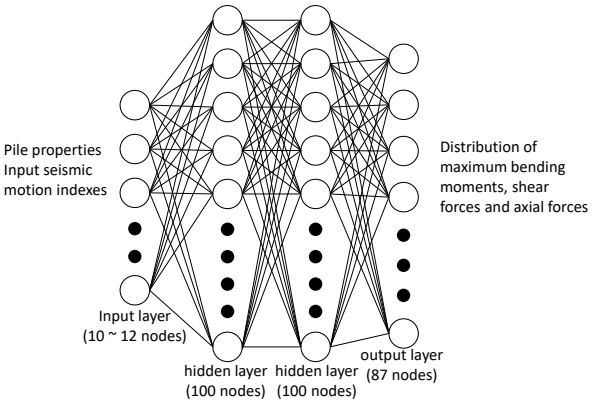

Figure 6. Deep neural network model.

3.2 *Tuning machine learning hyperparameters*

To prepare the dataset for learning, the 500 analyses were performed with all upper piles as SC piles. Then, the hyperparameters of the DNN were tuned assuming that the input seismic motion was characterized by the maximum relative displacement of the soil column model shown in Figure 8. The characterization of the input seismic motions is described in detail in the following section.

The hyperparameters were tuned at four points. These are dropout, batch size, standardization and optimization method.

Dropout is a method to prevent overlearning, which occurs when the DNN becomes overly fitted. Neurons in each layer are deactivated with a certain probability during training. However,

since no increase in loss was observed even when the dropout rate was set to zero, we decided not to use it.

The mini-batch method, in which partial data are randomly selected from the training data, is often used to improve the learning efficiency. The batch size is the number of data to be selected, which affects loss convergence. In this case, the number of training data was set to 400, and the number of test data was set to 100. For the batch size of training data, the loss convergence was good when 100 out of 400 data were selected.

Standardization is a method of improving loss convergence by normalizing the input and output data and weights between neurons. However, this method was not used because no improvement in convergence was observed during training.

The Adam, AdamW, and Adadelta methods were investigated as optimization methods to minimize the training loss. Details of these methods can be found in the PyTorch documentation (2025). The results are shown in Figure 7. The horizontal axis is the number of epochs. One epoch refers to the implementation of training on all the training data. Based on the degree of loss reduction, the Adam method was chosen.

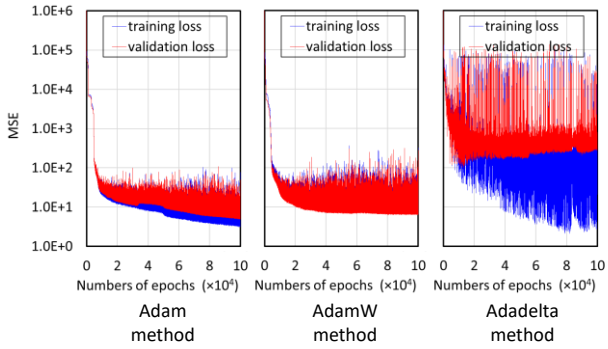

Figure 7. Loss convergence differences by optimization method.

3.3 *Indexing of input seismic motion*

Indexing the input seismic motion is examined next. The cross-sectional force of piles has a strong correlation with the ground response during earthquakes. For the ground response characteristics, the response of the soil column model was used due to its low analysis cost. Figure 8 shows the soil column model. The part of the improved wall part is was to a stabilized soil. The response at the top of the stabilized soil (point A in Figure 8) was used. The maximum acceleration, maximum relative displacement, and energy spectrum were chosen as indices. The energy spectra were averaged from 0.2 s to 1.0 s considering the first and second natural periods of the soil column model. Figure 9 shows the loss convergence of each index. Of these, the maximum relative displacement showed good loss convergence. This is due to the high correlation between the pile bending moment and maximum ground displacement. The case where the maximum shear strain of two elements under the stabilized soil shown in Figure 9 was added to the maximum relative displacement was also examined. This is because the shear force of the pile increases at the boundary between the soil and the improvement wall. However, the convergence was not significantly different from the case with the maximum relative displacement alone. Based on these analyses, the maximum relative displacement at the top of the soil column model was selected as the index of the input seismic motion.

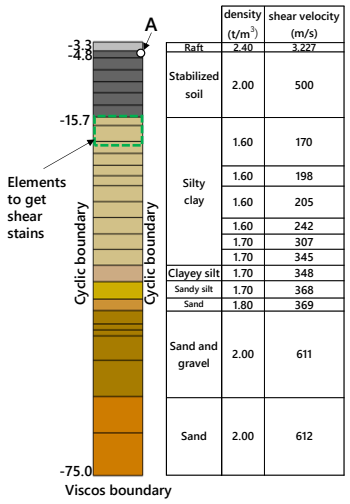

Figure 8. Soil column model for ground motion.

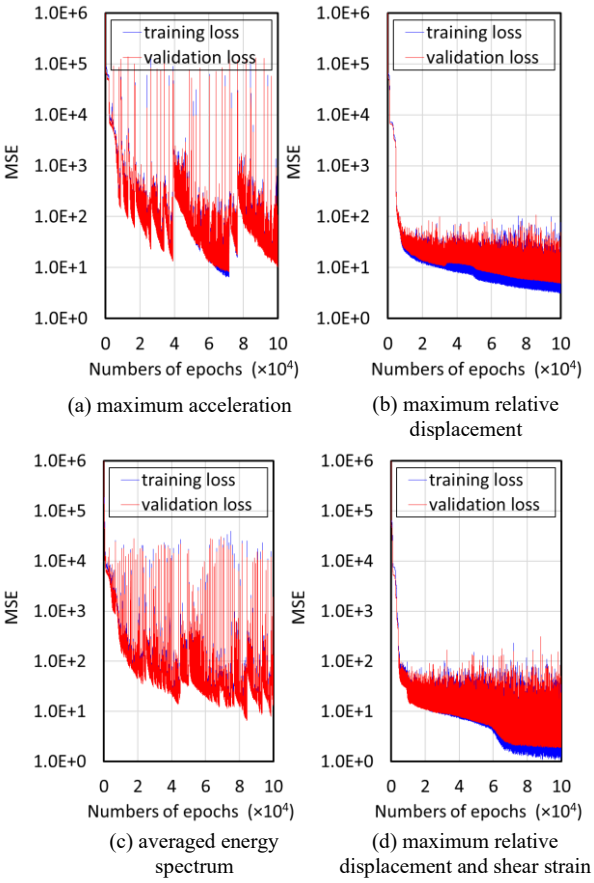

Figure 9. Loss convergence differences of by indexes of input seismic motions.

3.4 *Accuracy evaluation*

As shown in Figure 9, the loss convergence of the maximum relative displacement tends to decrease as the number of epochs increases from 100,000. Therefore, the number of epochs was increased to 200,000. Figure 10 shows the loss convergence of 200,000 epochs. The estimation error of each section force was examined from the results of the last 200 epochs of the testing phase. For the estimation error, the root mean squared percentage error (RMSPE) defined by the following equation was used.

$$RMSPE = \sqrt{\frac{1}{n}\sum\left(\frac{\hat{y}-y}{y}\right)^2} \quad (3)$$

where $\hat{y}$ is the predicted value, y is the actual value, and n is the number of data points.

The percentage error of the maximum bending moment was 0.27% on average with a 0.41% standard deviation. And 90.4% were less than 0.7%. The percentage error of the maximum shear stress was 0.58% on average with a 0.99% standard deviation. And 83.9% were less than 1.0%. The percentage error of the maximum axial force was 0.14% on average with a 0.13% standard deviation. And 96.4% were less than 0.4%. Figure 11 shows the distribution of the maximum bending moment in the case when the percentage error of the bending moment was the largest. This is sufficient accuracy for evaluating the foundation structure.

The execution time of the created surrogate model was 4.0×10^{-4} s (using 1 core). In contrast, the execution time of the FEM analysis was over one hour (using 192 cores). The CPU is Intel Xeon Gold 6442Y with 24 cores and a frequency of 2.6 GHz. The surrogate model calculation is by far the fastest, and the parameter study of 10,000 cases takes only 4.0 seconds.

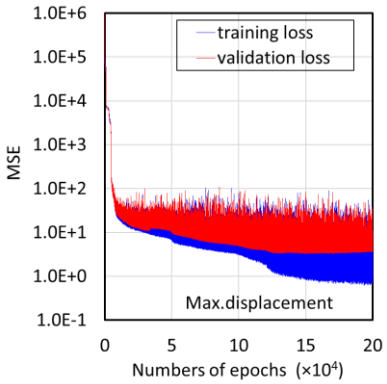

Figure 10. Loss convergence for 200,000 epochs with maximum relative displacement as an index of input seismic motions.

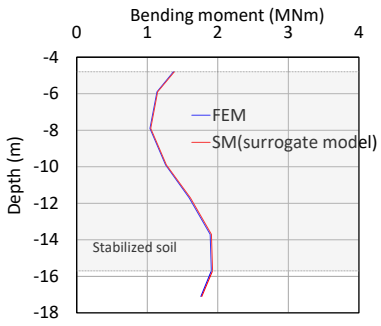

Figure 11. Maximum bending moment distribution with maximum estimation error.

4 PILE DESIGN OPTIMIZATION

4.1 *Parameter study using a surrogate model*

The design margin for the bending moment and shear force of piles with different specifications was estimated using the surrogate model. The design margin is defined as follows:

$$\text{Design margin} = \text{elastic limit}/\text{maximum force} \quad (4)$$

The bending moment margin was obtained using the bending moment (M)-axial force (N) interaction curve, as shown in Figure 12 and Equation°(4). The axial force was obtained by adding the estimated maximum positive and negative dynamic axial forces (dN_{min}, dN_{max}) to the long-term axial design force (N). The damage limit of the bending moment (M_{DL}) was obtained from the minimum value of the damage limit state (DLS) curve corresponding to the axial force. This was done by drawing a perpendicular line from the rectangle in Figure 12 to the DLS curve.

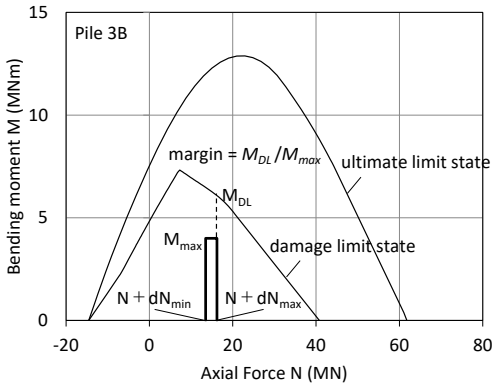

Figure 12. Bending moment-Axial force interaction curve for pile and calculation method of bending moment margin.

A preliminary parameter study was conducted on the dataset, including a case in which all the piles were PRC. The results showed that the Pile 3B group had a sufficient margin for replacing SC piles with PRC piles. Thus, 1,000 datasets were prepared with the Pile 3B group as PRC piles and the Pile 1A and Pile 1B groups as SC piles. A surrogate model was then developed. Figure 13 shows the result of machine learning.

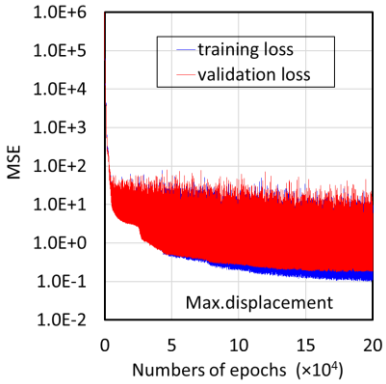

Figure 13. Loss convergence when the Pile 3B group consists of PRC piles and the Pile 1A, 1B group consists of SC piles.

The surrogate model was used to evaluate all cases in which the Pile 3B group was PRC piles and the Piles 1A and 1B groups were SC piles. The types of piles considered are listed in Table 1. The specifications for Piles 1A and 1B vary in steel pipe thickness by 1.0 mm within the range shown in the table. There are 20 total variations for each pile. The specifications for Pile 3B vary depending on the diameter of irregularly shaped steel bars, as shown in the table. There are 8 variations in total. Multiplying the number of pile specifications by the four input seismic motions yields 12,800 cases.

Table 1. Specifications of pile groups.

Pile group	Diameter (mm)	Pile type	Steel pipe thickness (mm)	Rebar diameter (mm)	Variations
1A	800	SC	6.0 ~ 25.0	–	20
1B	1,000	SC	6.0 ~ 25.0	–	20
3B	1,200	PRC	–	13,16,19 22,25,29 32,35	8

Figure 14 and Figure 15 show the evaluation results. The minimum and maximum design margins are plotted in different colors for each input wave. As can be seen, the minimum design margin exceeded 1.0 in most cases. This building can be said to have a high seismic resistance. Using SC piles for the Pile 1A and 1B groups certainly contributes to this resistance. The diameter of the piles and the improved soil strength also impact seismic resistance. However, since the diameter of the piles was determined by the long-term load, the strength of the improved soil has a more direct impact. Optimizing the design to include the improved soil strength remains a challenge for the future.

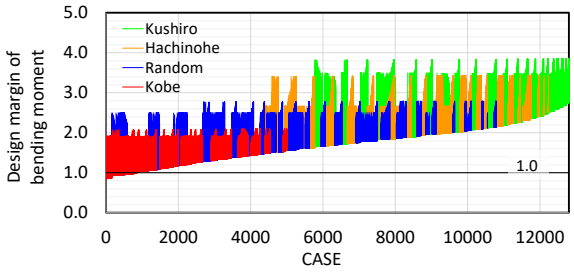

Figure 14. Minimum and maximum margins of bending moment estimated by surrogate model.

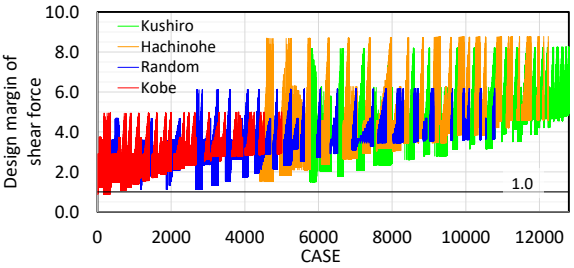

Figure 15. Minimum and maximum margins of shear force estimated by surrogate model.

In what cases did the margin ratio fall below 1.0? There were 768 cases out of 12,800. The results for the Kobe and Random waves were severe, and the margin was less than 1.0 only under these two waves. Figure 16 to Figure 18 show the relationship between the bending moment margin and the pile specifications for each pile in cases where the minimum margin was less than 1.0.

In Pile 1A, the bending moment margin was less than 1.0 when the steel plate thickness was 6.0 mm for all cases under the Kobe and Random waves. With a 7.0 mm steel plate, the bending moment margin was less than 1.0 for all cases under the Kobe wave. There were 480 cases in total. Figure 19 shows the estimated maximum bending moment distribution and M-N curves for the case with the minimum margin. This was the case where the steel pipe thickness of Pile 1A and Pile 1B were 6.0 mm, and Pile 3B had 13.0 mm rebars under Kobe wave. Pile 1A had a large bending moment at the lower end of the deep mixing walls (DMWs). Although Pile 1B had the largest bending moment, Pile 1A was more critical due to its smaller diameter and lower strength as shown by the M-N curve.

As shown in Figure 17, the margin was less than 1.0 in the cases where the steel pipe thickness of Pile 1B was 6.0 mm and 7.0 mm under the Kobe wave. There were 320 cases in total. Pile 1B also has the large bending moment at the bottom of the DMWs.

As shown in Figure 18, the margin of Pile 3B was large. The shape of the bending moment distribution differed from those of Piles 1A and 1B. The maximum occurred at GL-13.7 m, which is 2.0 m above the bottom of the DMWs. This is due to the pile layout. The pile diameter size and the shape of the bending moment distribution affect the large margin.

There are also cases in which the shear force margin was less than 1.0. In all cases, the steel pipe thickness of Pile 1B was 6.0 mm under the Kobe wave. These are included in the cases where the bending moment margin of Pile 1B was less than 1.0.

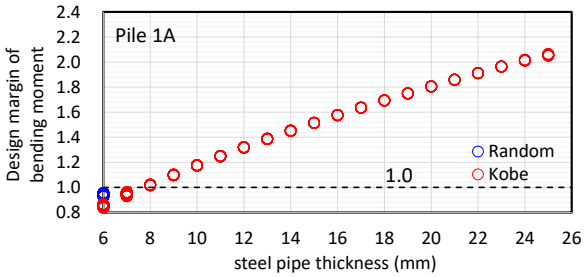

Figure 16. Bending moment margin for Pile 1A when the minimum margin ratio is less than 1.0.

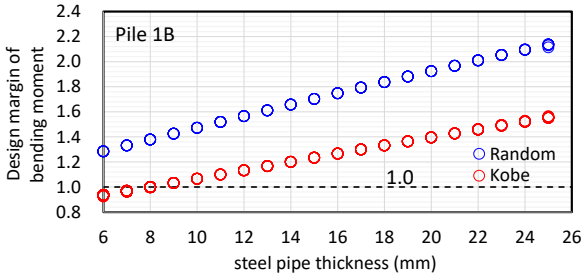

Figure 17. Bending moment margin for Pile 1B when the minimum margin ratio is less than 1.0.

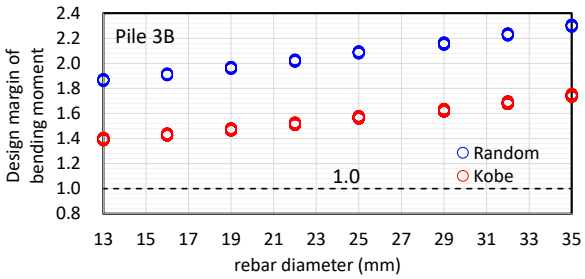

Figure 18. Bending moment margin for Pile 3B when the minimum margin ratio is less than 1.0.

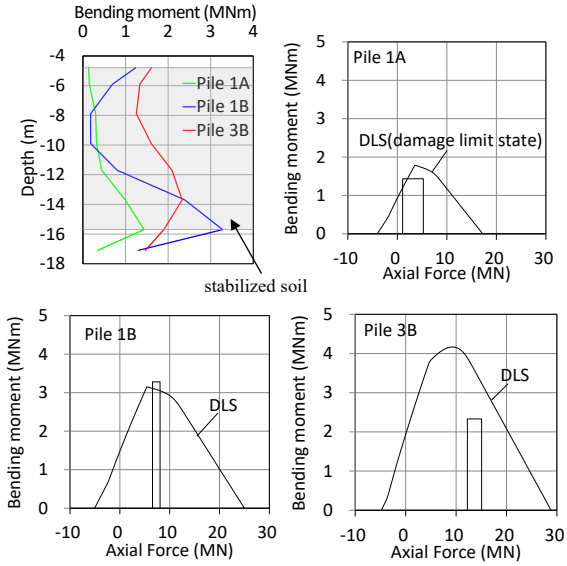

Figure 19. Bending moment distributions and M-N curves for the minimum margin case (Pile 1A, 1B: 6.0 mm steel plate, Pile 3B: 13.0 mm rebar, and Kobe wave) estimated using a surrogate model.

Table°2. Pile specifications for the eight lowest cost cases that satisfy the minimum margin

order	steel pipe thickness of 1A (mm)	steel pipe thickness of 1B (mm)	rebar diameter of 3B (mm)	total cross-sectional area of iron (m^2)	Min. margin of bending moment	Max. margin of bending moment	Min. margin of shear force	Max. margin of shear force
1	11.0	11.0	13.0	0.790	1.100	1.669	1.716	2.735
2	11.0	11.0	16.0	0.798	1.100	1.697	1.716	2.777
3	10.0	11.0	22.0	0.800	1.100	1.761	1.716	2.840
4	11.0	11.0	19.0	0.808	1.100	1.728	1.716	2.813
5	12.0	11.0	13.0	0.810	1.100	1.669	1.717	2.736
6	10.0	11.0	25.0	0.813	1.100	1.796	1.716	2.865
7	12.0	11.0	16.0	0.818	1.101	1.698	1.717	2.778
8	11.0	11.0	22.0	0.820	1.100	1.761	1.717	2.841

4.2 *Optimization*

To optimize the pile design, the minimum margin was set to 1.1. Designers can set the minimum margin to any value greater than 1.0. Optimization requires cost. The cost of a precast pile depends on the amount of steel used. For simplicity, we assumed that the cost of a unit mass of steel was the same. In this case, the steel cross-sectional area can be used as a cost function. Eight lower-cost cases that satisfy the minimum margin were selected and are shown in Table°2. The steel pipe thickness for Pile 1B was 11.0 mm for all the cases, reflecting the largest number of piles (8 piles) among the groups. The steel pipe thicknesses for the Pile 1A group ranged from 10.0 mm to 12.0 mm. The rebar diameters of the Pile 3B group varied. However, rebar diameters larger than 25.0 mm were excluded due to the cost.

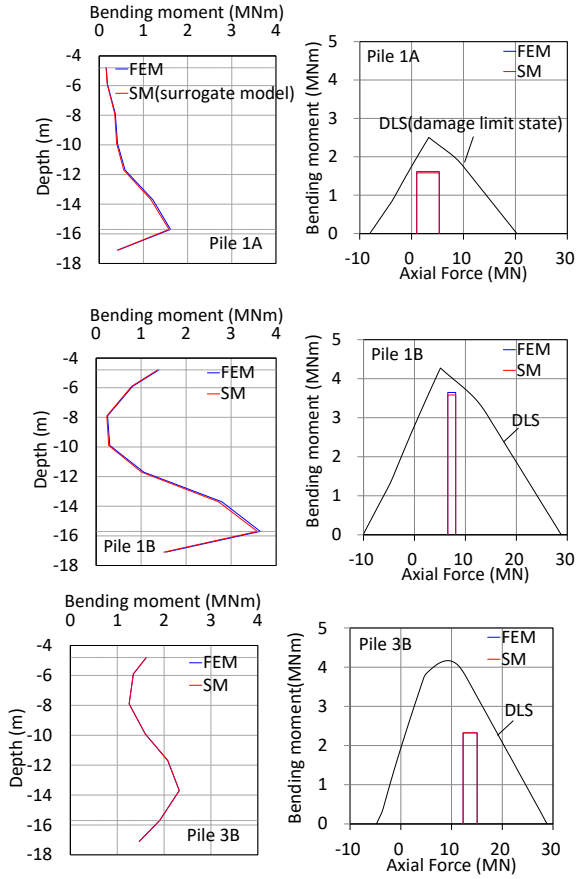

Figure 20. Comparison of evaluations using FEM and surrogate model for maximum bending moment distribution and M-N correlation diagrams in the minimum cost case.

Figure 20 shows a comparison of the maximum bending moment distribution in the FEM analysis and the surrogate model estimation of the minimum cost case. The difference between the two was very small, with a maximum of 0.04%. A comparison of the FEM and surrogate model in the M-N diagram is also shown. The difference was also very small.

For reference, Table 3 shows the design margin for the current design, and the steel cross-sectional area is 0.995 m^2. The present case has a higher cost and an unbalanced margin compared to the cases selected in Table°2.

Table 3. Design margin of current design case (Kobe wave).

Pile group	Diameter (mm)	Steel pipe thickness (mm)	Design margin of bending moment	Design margin of shear force
1A	800	9.0	1.15	1.90
1B	1,000	9.0	1.01	1.39
3B	1,200	12.0	2.05	5.21

5 CONCLUSIONS

A surrogate model was developed using machine learning to estimate the maximum cross-sectional force distribution of piles obtained from a nonlinear dynamic analysis. Large-scale earthquake ground motions were used in this analysis, which employed a 3D SSI model for a building with grid form DMWs and a piled raft foundation. The estimation error of the surrogate model was 0.27% on average with a standard deviation of 0.41% for the maximum bending moment. This level of accuracy is sufficient to identify trends in cross-sectional force changes when the pile specification is modified. Using this surrogate model, optimization of the pile design was attempted. A minimum margin was set, and the amount of steel material was used as a cost function to narrow down the design alternatives. The refinement results showed that a design satisfying the specified margin with a good margin balance and low cost could be extracted in a few seconds.

This case study shows that surrogate models can efficiently approximate the complex nonlinear behavior of geotechnical engineering problems. Surrogate models are also useful for replacing computationally expensive, detailed analyses and enabling many case studies and optimization studies.

6 REFERENCES

Namikawa, T. and Mihira, S. 2007. Elasto-plastic model for cement-treated sand. *Int. J. Numer. Anal. Mech. Geomech*, 31: 71-107.

PyTorch documentation, version 2.7, 2025, [Online] Available at: https://pytorch.org/docs/stable/index.html [Accessed 16th May 2025].

Shigeno, Y., Yamashita, K., Hamada, J., and Nakamura, N. 2017, Numerical evaluation of seismic performance of piled raft with grid-form DMWs under large earthquake loads. *Design and analysis of pile raft foundations -2017*, 109-127, Tamkang University Press, Taipei

Shigeno, Y. and Yamashita, K.2024. Seismic Performance of Piled Raft with Grid-form DMWs Considering Various Nonlinear Events in DMWs, *18th WCEE*, Milan, Italy.