Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore. ©2025 The Author(s), CC-BY 4.0 licence.
doi: 10.3850/GRF-25280825_isgsr-071-P150-cd

A SURROGATE MODEL FOR UNCERTAINTY QUANTIFICATION FOR THE REINFORCED SOIL FOOTING PROBLEM

Reza Jamshidi Chenari

Senior Geotechnical Engineer, TREK Geotechnical Inc., Winnipeg, Manitoba, Canada. E-mail: jamshidi_reza@yahoo.com

Richard J. Bathurst

Professor, Royal Military College of Canada, Ontario, Canada. E-mail: bathurst-r@rmc.ca

Surrogate modeling, also known as metamodeling, has several applications in the analysis of reinforced soil structures. Reinforced soil design often involves dealing with uncertainties related to material properties and project conditions. Surrogate models can facilitate probabilistic analysis by allowing the rapid evaluation of different scenarios. This helps in assessing the reliability and safety of a candidate design under uncertain conditions. In recent times, there has been an increased emphasis on understanding the influence of soil spatial variation on soil-structure interactions in geotechnical engineering problems. Traditional probabilistic theories often fall short in accurately representing spatial variability due to computational demand. As a result, there is growing interest in stochastic numerical modeling by generating a large number of realizations of soil parameters and incorporating them into numerical analyses to predict stability outcomes. However, this method is computationally demanding, especially using Monte Carlo simulations to predict uncertainty and probability of failure. To confidently estimate the probability of failure, or reliability index, in reliability-based analysis and design, a sufficient number of realizations is necessary. However, this is often impractical due to computational demand. A solution to reduce the number of required Monte Carlo simulations to achieve a confident estimate of reliability index is to use artificial intelligence, particularly artificial neural networks (ANNs). This study adopts an efficient ANN algorithm, specifically the Radial Basis Function (RBF), to predict the ultimate bearing capacity of a shallow footing placed on a geogrid-reinforced granular fill overlying a very soft clay deposit using a large synthetic database of footing load-settlement results previously reported by the writers. Assuming the foundation soil is spatially variable, the goal was to forecast the associated ultimate load outcomes across numerous realizations. The study underscores the potential of surrogate modeling to enhance the efficiency and accuracy of stochastic soil-structure interaction modeling.

Keywords: artificial neural network (ANN), surrogate modeling, radial basis function (RBF), soil variability, reliability-based design, reinforced soil footing.

1. Introduction

The Monte Carlo Simulation (MCS) method, a fully probabilistic analysis method, is widely used to address uncertainties arising from the spatial variability of parameters in geotechnical analysis and design. This technique involves generating many realizations to assess structure behavior under varying conditions. However, a significant challenge with MCS is the substantial computational effort required to perform a large number of simulations, which is necessary to achieve accurate predictions of the failure probability or reliability index for the structure. Jamshidi Chenari *et al.* (2024) have highlighted the importance of the number of realizations when conducting Monte Carlo simulations for stability analysis in geomechanics. Jamshidi Chenari and Bathurst (2023a) also highlighted this issue when modeling a shallow foundation on reinforced soils. In a very recent study, Jamshidi Chenari *et al.* (2024) showed that the minimum number of Monte Carlo simulations required depends on the consequence level or target probability of failure.An extensive dataset developed by Jamshidi Chenari and Bathurst (2023a) was used as the basis for ANN analyses, with a single case employed in the 2024 study to validate the ANN algorithm.This paper, however, uses a larger fraction of this dataset in an efficient ANN algorithm to perform feature selection analysis as well as validate the robustness of the ANN tool in rapidly generating highly accurate and reliable predictions of output parameters.

2. The MCS Method

The consequence of failure and the target probability of failure are linked in geotechnical engineering, forming the foundation for effective risk management and design strategies. The Canadian Foundation Engineering Manual (CFEM 2023) and the Canadian Highway Bridge Design Code (CHBDC, CSA 2019) provide guidelines for target probabilities of failure based on the level of consequence. Serviceability Limit States (SLS) are generally associated with higher allowable failure probabilities compared to Ultimate Limit States (ULS) because SLS failures typically result in inconvenience rather than catastrophic outcomes (CFEM 2023). Depending on the chosen target probability of failure, conducting a substantial number of stochastic stability analyses may be necessary to achieve a reliable estimation of failure probability. Table 1 presents the minimum required number of realizations corresponding to various target probabilities of failure (P_t), assuminga target coefficient of variation on estimated failure probability $\mathrm{COV}_{\bar{P}} = 0.1$.

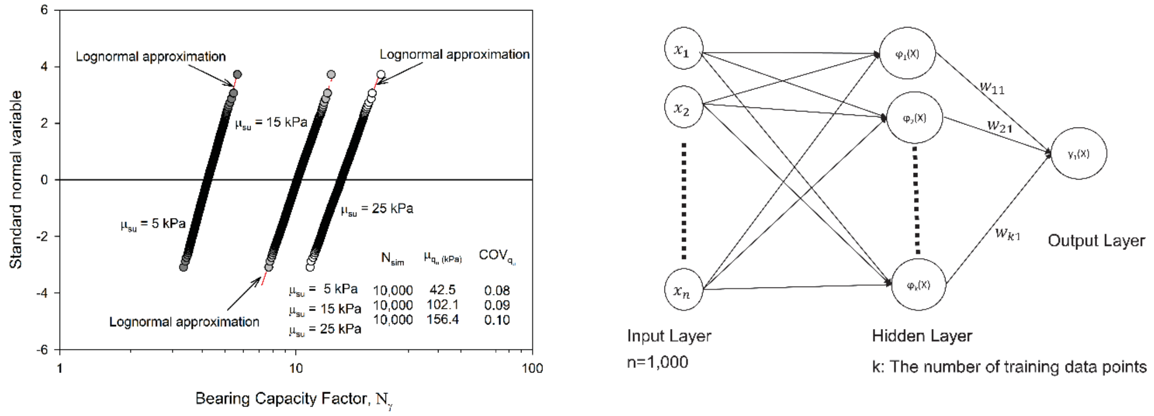
Table 1. The minimum number of realizations (N_{min}) required for different target probabilities of failure and$COV_{\bar{P}} =$ 0.1.

	ULS		SLS	
Consequence of failure	P_t	N_{min}	P_t	N_{min}
Low	1/1,000	100,000	1/100	10,000
Typical	1/5,000	500,000	1/500	50,000
High	1/10,000	1,000,000	1/1,000	100,000

The key takeaway from the data presented in Table 1 is that the minimum required number of realizations often exceeds initial expectations. Researchers typically use 1,000 or fewer realizations to determine the first and second moments—namely, the mean and COV of parameters such as bearing capacity and factor of safety (Griffiths *et al.* 2002; Ranjbar Pouya *et al.* 2014). While this approach is adequate for estimating the mean and COV of the parameters under consideration, a much larger number of realizations is required for precise estimation of the probability of failure (P_f). However, to maintain manageable total runtimes, researchers may opt for a lower number of realizations (NoR) initially, followed by integration into an efficient ANN algorithm. The following sectionsexplore one such approach in detail.

3. Data Structure and Surrogate Modeling

For probabilistic analysis, where numerous numerical simulations are required to capture the probabilistic behavior of geo-structures, data generation serves as the backbone of surrogate modeling. Surrogate modeling is an engineering approach used to predict the outcomes of a process without modeling the real process itself. In this study, data from a large number of finite difference method (FDM) numerical analysesusing the program FLAC 2D, originally published by Jamshidi Chenari and Bathurst (2023a), are used to develop a machine learning-based algorithm capable of handling the extensive requirements of Monte Carlo simulations. The rectangular domain under study consists of 1,000 zones populated with shear strength parameter values across various realizations. They employed five different mean undrained cohesion values (s_u), ranging from 5 kPa to 25 kPa, to represent very soft to soft clay deposits. Three specific cases were selected for surrogate modeling purposes in this paper. These include mean undrained shear strength values of 5, 15, and 25 kPa, together with COV of s_u of 50% and an isotropic correlation length of 1 m, representing significant spatial variability compared to random homogeneous cases. Figure 1(a) depicts the three datasets used in the surrogate modeling for this study as well as the architecture of the adopted ANN algorithm.



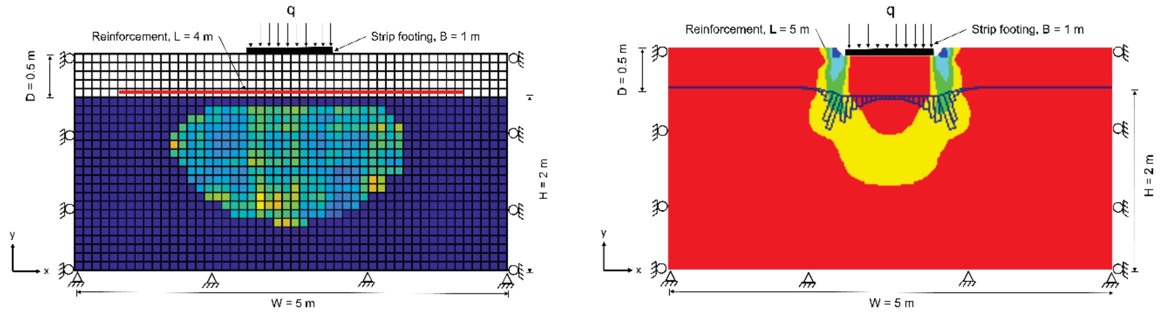
(a) (b)

Fig. 1. Data structure used in surrogate modeling: a) probabilistic numerical analysis results; b) architecture of the adopted ANN algorithm.

A Radial-Basis Functions (RBF) algorithm (e.g., Jamshidi Chenari *et al.* 2024), was employed to represent the ANN algorithm in the surrogate modelingwith a Gaussian activation function. Figure 1(b) illustrates the architecture of the RBF algorithm. Data preparation in this study involved vectorizing each 20×50 input matrixinto a 1000×1-column vector. Each vector comprises 1,000 undrained shear strength (s_u) values for a realization, with the output vector representing the ultimate bearing capacity (q_u), which is used for reliability index and probability of failure calculations. The first 1,000 runs were used for the training and testing phases of the proposed RBF algorithm. The remaining 9,000 data points from the three cases were reserved for the forward prediction stageto predict data for the remaining unexposed points. K-Fold Cross-Validation (CV) partitioning was employed to evaluate and enhance the performance and generality of the surrogate model.

4. Feature Selection

Feature selection is a vital step in machine learning and data analysis to identify the most relevant features that contribute to predicting the output variable(s) (the ultimate bearing capacity in this study). Correlation analysis is a straightforward and efficient method often used for this purpose, especially when the relationships between features and output(s) are expected to be linear. The goal is to identify the zones of highest importance in terms of their contribution to the ultimate bearing capacity of the overlying shallow footing.Using a correlation analysis, the cross-correlation coefficient was calculated for each property vector and the output vector in a zone, representing the ultimate bearing capacity of the overlying shallow footing for all 10,000 realizations. The correlation coefficients were sorted, and the 300 zones of highest importance are plotted in Figure 2(a). On the other hand, Figure 2(b) presents a plastic shear strain contour plot from Jamshidi Chenari and Bathurst (2023b), illustrating plastic zones beneath a footing seated on a reinforced granular fill over a soft clay deposit. By focusing only on the zones within the underlying soft clay deposit, it is observed that the pattern of the most important zones, identified through this simple correlation analysis, fairly matches the high shear strain zones in the underlying soft clay deposit.

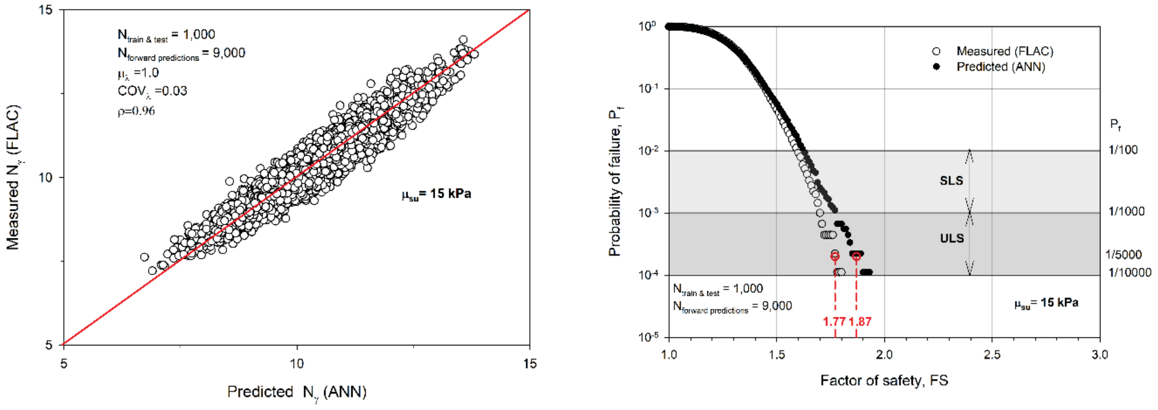


(a) (b)

Fig. 2.Results of correlation analysis for μ_{su}=15 kPa, ô$_x$=ô$_y$=1.0 m, CoV$_{su}$=50%, compared with the numerical analysis results: a) machine learning-based correlation analysis; b)shear strain increment contour plot.

5. Results and Discussion

Three different cases with very soft to soft clay deposits, corresponding to mean undrained shear strength values of 5, 15, and 25 kPa, were examined to evaluate the performance of the proposed ANN algorithm in predicting ultimate bearing capacity values for different realizations. For the first two cases, 10,000 realizations were available for RBF algorithm calculations. However, for the soft to firm clay deposit (μ_{su}=25kPa), the first 2,709 realizations were unintentionally overwritten with new realizations corresponding to a different correlation length. Consequently, only the remaining 7,291 data points were usedin this study.Figure 3 illustrates the measured versus predicted bearing capacity coefficients for the case of μ_{su}=15 kPa, along with the corresponding probability of failure plots. The resistance model bias (λ), defined as the ratio of measured (original) to predicted ultimate bearing capacity values, has a mean bias of 1.0, with a very small coefficient of variation (COV$_\lambda$) of approximately 3%. This highlights the strong predictive ability of the designed ANN algorithm. The Pearson correlation coefficient between the predicted and measured datasets is 0.96, confirming a strong correlation between the two. Figure 3(b) depicts the variation of the probability of failure of the shallow footing versus the design factor of safety. Different target $P_f$values are used to delineate grey and dark grey areas representing different consequence levels associated with the serviceability and ultimate limit states, respectively,in Table 1.However, unlike the results for very soft clay deposits (presented in Jamshidi Chenari *et al.* 2024), some deviations between the original and predicted results are observed. While these differences are minor, they mainly occur in areas with fewer data points, particularly in the lower tail of the ultimate bearing capacity distribution. In certain cases, as discussed by Jamshidi Chenari and Bathurst (2023a), the original data diverges from the main trend due to the progressively limited data available at high FS values. This divergence persists in the predicted trend but remains within a manageable range. For a typical failure consequence, the recommended design factor of safety is 1.8 based on the original trend and 1.9 based on the predicted trend. This indicates that the predicted trend errs on the side of caution, favoring safer designs.



(a) (b)

Fig. 3. Predicted versus original plots for μ_{su}=15 kPa using the RBF algorithm trained with NoR=1,000: a) bearing capacity coefficient (N_γ); b) probability of failure (P_f).

6. Conclusion

This study used an efficient ANN model (the RBF algorithm) to predict the ultimate bearing capacity of a shallow footing supported by a geogrid-reinforced granular fill overlying a very soft to soft clay deposit. The spatial variability of the foundation soil was accounted for to support reliability-based design of the footing. Traditional probabilistic numerical stability analyses often require a very largenumber of realizations, which can be computationally prohibitive, particularly for scenarios with severe failure consequences. To address this challenge, the study implemented a streamlined and effective ANN algorithm, combined with a limited number of numerical simulations, to accurately predict outcomes for additional realizations. This approach assumes that the unseen realizations belong to the same statistical distribution as the training dataset. The key findings from this investigation include:

1. The number of realizations required can vary widely based on the consequences of failure, making probabilistic numerical approaches inefficient when an extremely large number of simulations is necessary.
2. The RBF algorithm proved highly efficient when adequately trained with a dataset of sufficient size.
3. The RBF algorithm integrates effectively into reliability index computations, providing more precise probability of failure predictions.
4. By conducting a selected number of stochastic simulations using commercial software (e.g., FLAC 2D), it is possible to train ANN models, such as the RBF algorithm, to make accurate large-scale predictions, significantly enhancing the feasibility of reliability analyses.

References

Bathurst, R. J., and Jamshidi Chenari, R. (2023). Bearing capacity of strip footings seated on granular layers over spatially variable undrained soft clay. *In Geo-Risk 2023*, pp. 357–366.

Bathurst, R. J., and Jamshidi Chenari, R. (2024). Bearing capacity of strip footings seated on unreinforced and reinforced granular layers over random homogeneous and isotropic spatially variable undrained soft clay. *In Proceedings of the 5th Pan-American Conference on Geosynthetics (GeoAmericas 2024)*.

Canadian Standards Association (CSA) (2019). Canadian Highway Bridge Design Code. CAN/CSA-S6-19, Mississauga, Ontario, Canada.

Canadian Geotechnical Society (2023). Canadian Foundation Engineering Manual (5th ed.).

Jamshidi Chenari, R., and Bathurst, R. J. (2023a). Bearing capacity of strip footings seated on unreinforced and geosynthetic-reinforced granular layers over spatially variable soft clay deposits. *Journal of Geotechnical and Geoenvironmental Engineering*, 149(6), 04023034.

Jamshidi Chenari, R., and Bathurst, R. J. (2023b). Influence of geosynthetic stiffness on bearing capacity of strip footings seated on thin reinforced granular layers over undrained soft clay. *Geotextiles and Geomembranes*, 51(1), 43–55.

Jamshidi Chenari, R., Banirostam, T., and Toloui, A. (2024). Application of artificial intelligence in stochastic soil modeling: An example in reinforced soil modeling. *In Proceedings of GeoMontreal 2024*.

Griffiths, D. V., Fenton, G. A., and Manoharan, N. (2002). Bearing capacity of rough rigid strip footing on cohesive soil: Probabilistic study. *Journal of Geotechnical and Geoenvironmental Engineering*, 128(9), 743–755.

Ranjbar Pouya, K., Zhalehjoo, N., and Jamshidi Chenari, R. (2014). Influence of random heterogeneity of cross-correlated strength parameters on bearing capacity of shallow foundations. *Indian Geotechnical Journal*, 44, 427–435.