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.
doi: 10.3850/GRF-25280825_isgsr-063-P012-cd

ISGSR2025

3D-CNN-BASED SURROGATE MODELING AND DATA AUGMENTATION FOR 3D SLOPE RELIABILITY IN SPATIALLY VARIABLE SOILS

Chongzhi WU

School of Civil Engineering, Chongqing University, Chongqing 400045, P. R. China. E-mail: iwuchongzhi@cqu.edu.cn

Ze Zhou WANG

Marie Skłodowska-Curie Fellow, Department of Engineering, University of Cambridge, Cambridge CB3 0FA, United Kingdom. E-mail: wangzz@u.nus.edu

Siang Huat GOH

Department of Civil & Environmental Engineering, National University of Singapore, Singapore 117576. E-mail: gohsianghuat@nus.edu.sg

Wengang ZHANG*

School of Civil Engineering, Chongqing University, Chongqing 400045, P. R. China. E-mail: zhangwg@cqu.edu.cn

While the significance of spatially variable soil properties in slope stability assessment has been well understood, the implementation of three-dimensional (3D) probabilistic slope reliability assessment is still bottlenecked by its excessive computational time. This paper presents a novel surrogate model based on Convolutional Neural Networks (3D CNN) to replace the computationally demanding 3D random field finite element method (RFEM) for Monte-Carlo simulations. To enhance model performance with limited training data, a data augmentation technique has been developed based on the shear strength reduction (SSR) method. The methodology's effectiveness is illustrated using a 3D slope case with undrained shear strength as the random variable.

Keywords: 3D slope reliability assessment; Spatial variability; 3D random finite element method; Surrogate model; 3D Convolutional Neural Network; Data augmentation.

1. Introduction

The spatial variability of soil properties is a major source of geotechnical uncertainties which significantly affects the reliability of geotechnical systems(Phoon & Kulhawy, 1999). Soil properties exhibit three-dimensional (3D) variability and preserve anisotropy in both vertical and horizontal directions(Griffiths et al., 2009). While various methodologies have been developed to assess slope reliability and risk, many rely on two-dimensional (2D) models assuming plane strain conditions(Wang et al., 2011). Such 2D simplification could lead to over- or under-estimation of slope segment reliability and may contradict actual observed failure surfaces in slope engineering(Xiao et al., 2016). Therefore, 3D slope reliability analysis is essential for more accurate assessments.

Monte Carlo Simulation (MCS), particularly through the random finite element method (RFEM), has become the most widely used approach for slope reliability analysis. RFEM couples finite element analysis with MCS to incorporate spatial variability of soil properties and can automatically locate critical slip surfaces. However, this method faces computational limitations due to the large sample size required, especially in 3D analyses where the computational burden becomes orders of magnitude greater than 2D analyses(Jiang et al., 2023).

Various surrogate models have been developed to address the computational burden in 2D analyses. However, traditional methods struggle with high-dimensional random fields in 3D problems. Convolutional Neural Network (CNN) emerges as a promising alternative due to its ability to process complex spatial information(Wang et al., 2021), though its effectiveness is limited by the availability of sufficient training data.

The challenges become more pronounced in 3D slope reliability analysis due to increased computational complexity. The substantial computational demands limit the number of training samples, while 3D random fields require more data to capture their complexity effectively. Therefore, there is a critical need to develop both an accurate computational method for 3D spatial information and a methodology to enhance the practicality of 3D reliability surrogate models.

This study proposes a practical methodology for 3D slope reliability assessment by combining 3D RFEM, 3D CNN, and a data augmentation strategy. This represents the first application of 3D CNN as a surrogate model for slope reliability assessment, complemented by a data augmentation strategy based on the shear strength reduction method to enhance model training efficiency.

2. Methodology framework

The proposed methodology for 3D slope reliability analysis consists of four phases: (i) 3D random field finite element modelling for generating training data, (ii) SSR-based data augmentation, (iii) training of the 3D-CNN-based surrogate model, and (iv) probabilistic slope reliability assessment.

First, an initial setof random field samples is generated based on soil properties statistics (mean, standard deviation, and autocorrelation distance). These 3D random fields are mapped into the finite element model, and stability analysis is conducted using the SSR method within PLAXIS 3D to obtain corresponding FS values.

Subsequently, these samples are augmented using the SSR-based strategy, as shown in Fig. 1, which involves scaling the original random fields to generate additional training samples while preserving their spatial correlation characteristics. The original and augmented datasets are then split into training and validation sets.

A 3D CNN architecture modified from VGG network is adapted to process the 3D random fields and trained to establish relationships between FS values and input random fields. Finally, the trained model predicts FS values for new sets of 3D random fields generated via MCS for probability assessment.

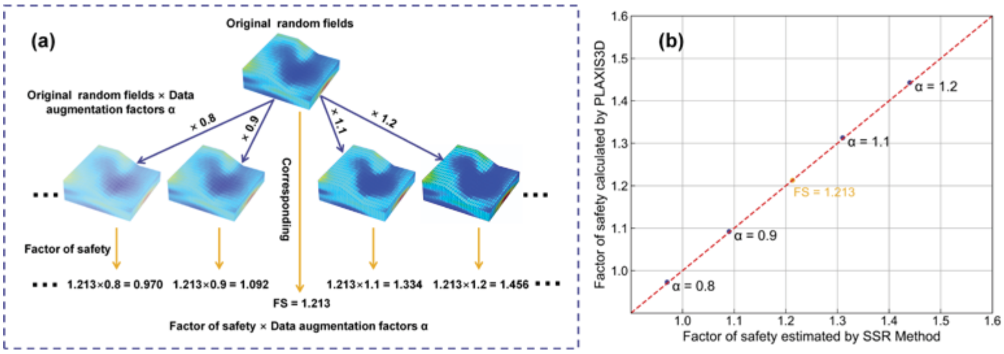

Fig. 1.(a) An illustration of SSR-based data augmentation implementation. (b) Validation of the proposed data augmentation technique.

3. Illustrative example

A 3D slope model was developed using Plaxis3D 23.01, with geometry and mesh discretization shown in Fig. 2. The soil was modeled using an elastic-perfectly-plastic model with Mohr-Coulomb yield criterion for undrained analysis. Only the undrained shear strength C_u was considered spatially variable,with a mean of 23 kPa, coefficient of variation of 0.3, and scales of fluctuation of 20 m horizontally and 4 m vertically. The model was discretized into 6825 rectangular prisms, resulting in 84316 elements and 119991 nodes. Random fields were generated using GSTools with Gaussian autocorrelation function and transformed to lognormal distribution. A total of 2000 sets of 3D random fields were generated and analyzed using the shear strength reduction method.In this study, the computation time for each original 3D RFEM model was approximately 1.5 hours. Further details can be found in(Wu et al., 2024), which provides a comprehensive introduction to this methodology.

3.1.*Architecture of proposed 3D CNN*

The 3D CNN input tensor are [n, 1, 20, 30, 15], representing n realizations of 3D random fields, with dimensions aligned with the slope geometry shown in Fig. 3. The architecture, adapted from VGG, consists of feature extraction layers (three convolutional blocks with ReLU activation and 3D average pooling) and regression layers (two fully connected layers). The model uses MAPE as loss function and Adam optimizer, implemented using PyTorch.

3.2.*Data augmentation strategy*

The study uses 500 samples for training and 1500 for testing, with varying training sample sizes (75, 125, 250, 400, and 500) to investigate model performance. The SSR-based data augmentation strategy expands these datasets using different scaling factors applied to random field variables. With a step size of 0.1, the scaling ranges (0.9-1.1, 0.8-1.2, 0.7-1.3, and 0.6-1.4) generate 3, 5, 7, and 9 different scaling factors respectively, effectively increasing the training data by the corresponding multiples. The data augmentation strategy effectively alters the distribution patterns of the training datasets and significantly increases the proportion of training samples with FS<1.0, addressing the critical data imbalance issue in reliability assessment. Note that this proportion represents the distribution of augmented training data, while the true failure probability (4.85%) is evaluated on the independent test set.

4. Results

4.1.*Performance of 3D CNN with different training sizes*

The model's performance improves with increasing training samples, reflected by higher R² values and lower RMSE (Fig. 4a). With 500 training samples, the model achieves R² of 0.95. However, the failure probability predictions consistently underestimate the true value of 4.85%, particularly with smaller training sets (Fig. 4b).

4.2.*Effect of data augmentation*

The data augmentation strategy shows varying impacts on model performance. For the case with 75 training samples, R² increases significantly from 0.812 to 0.866 after augmentation by a factor of 3, though further augmentation yields limited improvement (Fig. 5a). More importantly, the predicted failure probabilities converge toward the true value (4.85%) as the augmentation factor increases, with the relative error reducing from ~60% to ~5% for the 75-sample case (Fig. 5b).

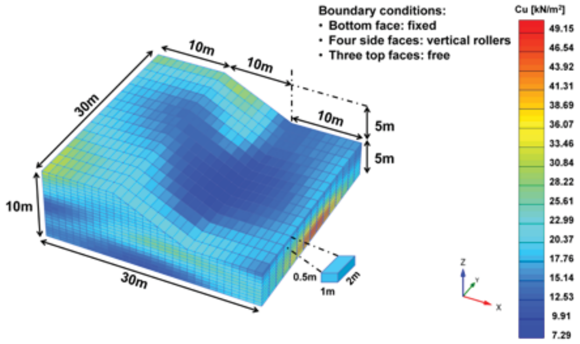

Fig. 2.Typical distribution of undrained strength in the 3D slope model.

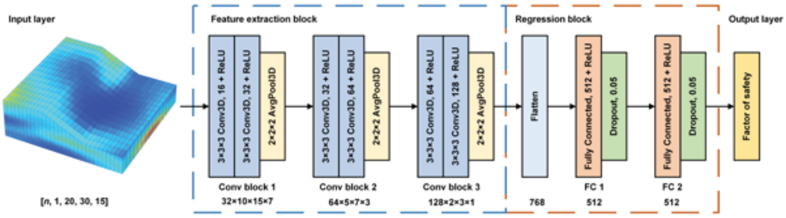

Fig. 3.The architecture of the proposed 3D CNN model.

4.3.*Comparison with other models*

Comparative analysis with 2D CNN and XGBoost reveals that 3D CNN consistently outperforms both alternatives (Fig. 6). The performance ranking is 3D CNN > 2D CNN > XGBoost. While all models initially underestimate failure probability, data augmentation improves predictions, with 3D CNN showing superior performance particularly with limited initial training data.

4.4.*Robustness validation*

Ten repeated experiments with 75 training samples confirm the effectiveness of data augmentation. The results demonstrate improved R² values and more consistent predictions after augmentation (Fig. 7a). Failure probability predictions consistently converge toward the true value with increased augmentation factors, while showing reduced uncertainty (Fig. 7b).

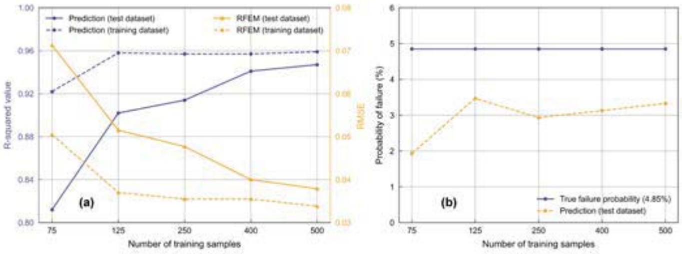

Fig. 4.Impact of five training sample sizes on R^2, RMSE and failure probability prediction.

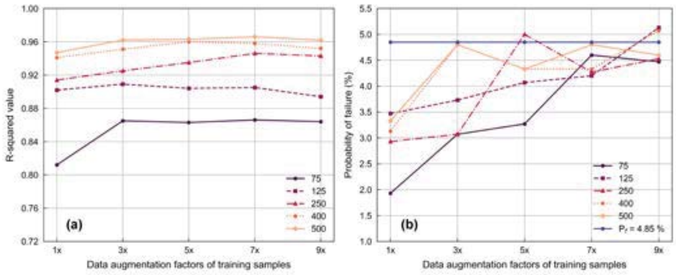

Fig. 5.The R^2 and failure probability of 3D CNN models trained on five training sample sizes with varying data augmentation factors.

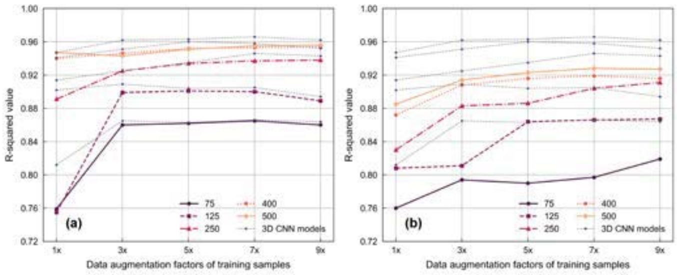

Fig. 6. Effects of data augmentation on the R^2 of (a) 2D CNN and (b) XGBoost trained using five training sample sizes with varying augmentation factor.

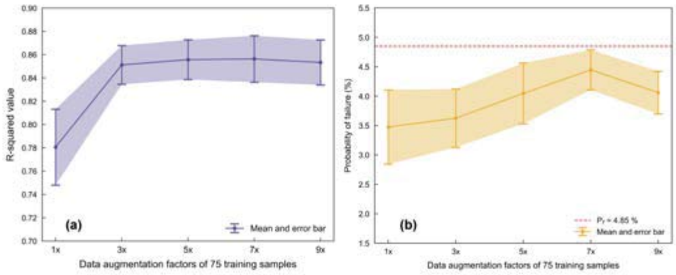

Fig. 7. Impact of data augmentation factors on the R^2 and failure probability predictions based on ten repeated experiments.

5. Conclusions

This study proposes a practical methodology for efficient 3D slope reliability assessment in spatially variable soils, combining 3D CNN as a surrogate model with a SSR-based data augmentation strategy. The key findings are:

(1) 3D CNN effectively captures spatial information within 3D random fields and accurately predicts slope stability factors.
(2) Both 2D and 3D CNNs outperform XGBoost in learning spatial distributions, with 3D CNN showing superior performance due to its ability to directly process 3D spatial data.
(3) The SSR-based data augmentation strategy enhances model accuracy, particularly with limited training data, and effectively addresses data imbalance issues in failure probability prediction.
(4) The combined methodology achieves high accuracy using only 5% of the full MCS sample set, demonstrating its potential for efficient and robust 3D reliability analysis.

References

Griffiths, D. V., J. Huang, and G. A. Fenton (2009). On the reliability of earth slopes in three dimensions. Proceedings of the Royal Society *A: Mathematical, Physical and Engineering Sciences* 465, 3145–3164.

Jiang, S.-H., G.-Y. Zhu, Z. Z. Wang, Z.-T. Huang, and J. Huang (2023). Data augmentation for CNN-based probabilistic slope stability analysis in spatially variable soils. *Computers and Geotechnics* 160, 105501.

Phoon, K.-K. and F. H. Kulhawy (1999). Evaluation of geotechnical property variability. *Canadian Geotechnical Journal* 36, 625–639.

Wang, Y., Z. Cao, and S.-K. Au (2011). Practical reliability analysis of slope stability by advanced Monte Carlo simulations in a spreadsheet. *Canadian Geotechnical Journal* 48, 162–172.

Wang, Z. Z., C. Xiao, S. H. Goh, and M.-X. Deng (2021). Metamodel-based reliability analysis in spatially variable soils using convolutional neural networks. *Journal of Geotechnical and Geoenvironmental Engineering* 147, 04021003.

Wu, C., Z. Z. Wang, S. H. Goh, and W. Zhang (2024). Comparing 2D and 3D slope stability in spatially variable soils using random finite-element method. *Computers and Geotechnics* 170, 106324.

Xiao, T., D.-Q. Li, Z.-J. Cao, S.-K. Au, and K.-K. Phoon (2016). Three-dimensional slope reliability and risk assessment using auxiliary random finite element method. *Computers and Geotechnics* 79, 146–158.