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/ICSMGE2026-793

Improving Reliability of Slope Stability Predictions through Physics-Informed Machine Learning

Ekansh Agarwal
Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada

Te Pei
Department of Civil Engineering, Stony Brook University, Stony Brook, NY 11794, USA

Ning Luo
Bechtel Corporation, Reston, VA 20190, USA, nluo@bechtel.com

ABSTRACT: Physics-informed machine learning (PIML) offers a promising pathway for geotechnical analysis by blending domain knowledge with data-driven models. This study proposes a PIML framework for predicting the factor of safety (FoS) of homogeneous slopes, using high-fidelity finite element method (FEM) simulations as ground truth and Bishop's simplified method as a physics-based regularizer. A 3D FEM model was developed in Abaqus to generate 2,000 slope stability cases, with soil and geometric parameters sampled from realistic statistical distributions. The PIML network incorporated Bishop-derived FoS values into its loss function, balancing fidelity to FEM data with adherence to geotechnical theory. Performance was evaluated across multiple dataset sizes, training fractions, and physics-weighting parameters. Results show that physics-informed models achieved notable gains in predictive accuracy for small datasets, with R^2 improvements up to 0.05-0.06 over purely data-driven models and exhibited smoother convergence and reduced variance. The framework demonstrates strong potential as a practical, generalizable, and interpretable tool for slope stability analysis, particularly under data-scarce conditions.

KEYWORDS: Physics-informed machine learning, finite-element analysis, slope stability, Bishop, neural network.

1 INTRODUCTION

Geotechnical structures like shallow foundations, slopes, and retaining walls are essential for the stability and safety of civil infrastructure. Their design depends on both the structural dimensions and the properties of the supporting soil. Achieving accuracy in design, which is usually measured using the factor of safety, is crucial. However, there's always a need to balance accuracy with efficiency. High-precision design often requires detailed lab testing, minimizing human errors, and using advanced numerical techniques like finite element analysis to capture complex soil behavior, irregular shapes, and layered profiles. But these methods are computationally expensive and require significant expertise, making them challenging to implement for routine design. To make the design process more practical and easier to use, engineers often rely on traditional methods, also called physics-based models. These methods are based on fundamental mechanical principles, but they often involve simplifying assumptions that can reduce their reliability in real-world scenarios. For example, slope stability is commonly analyzed using limit equilibrium methods (LEMs). However, LEMs too often rely on assumed failure surfaces and simplified force models, which may not reflect actual slope behavior under conditions like pore pressure changes, soil anisotropy, or progressive failure (Griffiths and Lane, 1999).

In recent years, machine learning (ML) models have become more common in geotechnical engineering as an alternative to traditional methods. These data-driven models can capture complex, non-linear relationships between soil and structural features and can make accurate predictions efficiently. Several studies have applied ML techniques to estimate bearing capacity (Liu et al., 2024), analyze slope stability (Agarwal et al., 2021; Lei et al., 2024), tunnels (Verma et al., 2022), and predict wall pressures (Nguyen et al., 2024). However, the main drawback of these models is that they lack any built-in understanding of physical principles. They rely entirely on the data provided during training and validation. Hence, they often struggle with overfitting and may perform poorly when applied to new or limited data, something that is quite common in geotechnical work due to the high cost and variability of soil testing (Phoon et al., 2019). This highlights the need for a method that is based on physics, can handle complex behavior, and is efficient enough for routine use. Physics-informed machine learning (PIML) offers a promising solution. These models combine physical laws with ML algorithms, helping the models learn patterns that are both data-driven and consistent with geotechnical knowledge. This leads to better generalization and accuracy, even when data is limited (Cuomo et al., 2019; Karniadakis et al., 2021; Pei et al., 2023; Pei and Qiu, 2024; Pei and Luo, 2025; Agarwal et al., 2025).

In this study, we propose a PIML framework that can be used for the slope stability analysis of simple homogeneous slopes and can be extended for heterogeneous complex slopes. The prime idea is to present and show the effect of including physics in machine learning algorithms for slope stability. To simplify the algorithm, results previously calculated using Bishop's method of slope stability are used as a physics-based reference for comparison with the purely data-driven model. This helps ensure that the model's predictions remain physically realistic, even when it is applied to sparse or unfamiliar data. To train, validate, and test the PIML models, we developed and validated the detailed three-dimensional (3D) finite element model (FEM) for soil slope. The results from PIML are compared with the pure data-driven ML model. PIML framework consistently provides more accurate, robust, and physically consistent predictions in the case of data-scarce training which demonstrates the potential of PIML not only as a powerful predictive tool but also as a practical design approach.

The following sections describe the development of the physics-based model, the finite element simulation, and the dataset used in this study. We then outline the methodology of the proposed framework, followed by the results and discussion.

2 DATASET PREPARATION AND FINITE ELEMENT MODELING

2.1 *Overview of methodology*

To create synthetic datasets for the slope, input parameters were sampled probabilistically based on statistically reasonable distributions (as shown in Table 1) to ensure the datasets were both realistic and diverse for training, validation, and testing. Cohesion (c) and friction angle (ϕ) were assumed to be negatively correlated with a coefficient of -0.5, to generate more realistic samples and avoid convergence issues in the finite element analysis. A total of 2,000 samples were generated and the output variable, factor of safety (FoS), was computed using the validated 3D FE model developed in Abaqus 2023, which serve as the ground truth. Using Abaqus for generating the dataset is advantageous as it provides detailed and accurate simulations of complex soil-structure interactions, essential for robust NN training. To maintain consistency and efficiency, soil behavior in the FE model was defined using the simple elastoplastic Mohr-Coulomb failure criterion, and the influence of the groundwater table was neglected. Python scripting was used to automate the model, allowing systematic variation of geometry and soil parameters, application of boundary conditions, meshing, and extraction of results. Specific modeling details are provided in the following subsections.

2.2 *Generation of dataset*

Five input parameters were varied: unit weight (γ), c, ϕ, slope height (H), and slope inclination angle with horizontal (β). The soil properties were generated using a lognormal distribution. The slope height was sampled uniformly, while the slope angle was sampled from a Beta distribution. The Beta distribution was chosen because it allows flexible control over the shape of the distribution within a bounded interval, making it suitable for modeling slope angles that typically fall within a specific range in real-world scenarios. A summary of the generated input and output features for the soil slope is provided in Table 1. The selected ranges of soil strength and geometric parameters were chosen to represent typical conditions of small to medium natural slopes composed of cohesive-frictional soils, commonly reported in literature (e.g., Chen et al., 2014; Griffiths and Lane, 1999). The mean values and statistical variations reflect realistic field ranges observed in silty and clayey soils with moderate stability. Similarly, slope heights (4-6 m) and angles (25°-40°) were adopted to ensure representative geometries for conventional slope-stability analyses, while maintaining numerical convergence and consistency with previous benchmark studies.

Table 1. Summary statistics of the generated dataset

Soil slope						
	γ (kN/m^3)	c (kPa)	ϕ (°)	H (m)	β (°)	FoS: FEM
Mean	18.00	14.89	20.05	5.00	32.46	2.05
Std.	0.91	3.72	2.02	0.57	3.42	0.38
Min.	15.16	5.75	14.26	4.00	25.26	1.19
Max.	21.12	29.63	28.31	6.00	39.84	3.69

Soil geometry was adopted from Chen et al. (2014), which consists of a simple homogeneous slope model (Figure 1a). The researchers used the strength reduction method (SRM) to determine the FoS and compared it with results from LEM. We replicated this setup in Abaqus using Python scripting, implementing a bisection algorithm to automate the SRM process. The shear strength parameters were gradually reduced until failure occurred, indicated by non-convergence of the numerical solution. FoS is defined as the critical strength reduction factor (SRF), the highest SRF for which the model still converges, marking the threshold of slope stability just before failure. A Python script was developed to read the Abaqus status file and automatically extract this last converged SRF value, which is taken as FoS for the slope. The simulation was performed in two steps: a geostatic step to establish initial stress conditions, followed by a static step that used these stresses as predefined fields to conduct the strength reduction analysis. We validated our model against results from Chen et al. (2014), and the factor of safety matched closely (Figure 1b), giving us confidence in using this model for large-scale data generation. Given the computational demands of SRM, high-performance computing (HPC) resources were utilized. Jobs were submitted via Python scripts and executed on a computing cluster using the Abaqus solver, followed by postprocessing.

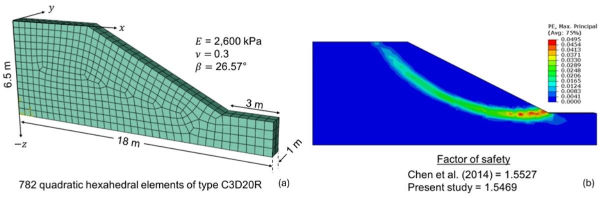

Figure 1. (a) Slope mesh developed in Abaqus; (b) Contour plot for equivalent plastic strain at critical SRF using SRM-based FE analysis.

2.3 *Bishop's model*

To incorporate geotechnical domain knowledge as a physics-based regularizer in our PIML framework, we implemented an optimized version of Bishop's simplified method of slices. The code performs a two-stage filtered search to identify the critical slip surface with the minimum FoS for slope stability analysis. Given soil properties and slope geometry, the algorithm first conducts a coarse grid search over potential circle centers and radii, followed by a refined local search around the best candidate. For each configuration, the method calculates FoS using an iterative scheme that accounts for normal forces and mobilized shear resistance. We used this implementation to compute FoS across a dataset of 2000 parametric slope cases, treating the results as a physics-derived regularization parameter. These FoS values, rooted in classical soil mechanics, are used to guide and constrain the learning of our PIML model, with high-fidelity FEM simulations serving as ground truth for supervised training.

3 METHODOLOGY

Stability analysis of soil slopes can be framed as a supervised regression problem. The process begins with sample generation and dataset partitioning, where preprocessing ensures clean and reliable data, and the dataset is split into training and validation sets for unbiased model evaluation. This is followed by model development, where machine learning models are constructed and their hyperparameters optimized to effectively capture the complex relationships between input features and structural stability. Finally, model performance is assessed using appropriate evaluation metrics to gauge accuracy, robustness, and generalizability across varying conditions. In this study, we developed a PIML model to predict the stability of soil slope. Its performance was compared against a baseline fully connected neural network NN (no physics) to evaluate the improvements in predictive accuracy achieved by embedding physical knowledge into the learning process.

3.1 *Physics-informed machine learning model*

The proposed PIML framework integrates numerical simulation data with domain knowledge derived from limit equilibrium analysis to predict the FoS of soil slopes. A fully connected feed-forward NN was designed with two hidden layers, each consisting of 128 neurons. Configurations with different neurons per layer starting from 64 and going up to 256 were tested to evaluate the effect of network depth and complexity; however, 128 neurons provided a good balance between accuracy and efficiency. The hyperbolic tangent (tanh) activation function was primarily adopted due to its smooth and bounded nonlinear mapping, which is well-suited for approximating continuous geomechanical responses. To assess sensitivity, alternative activations such as sigmoid and swish were also evaluated; however, tanh suited our model well. The distinguishing feature of the PIML framework is the incorporation of a physics-guided loss function, enabling the model to balance fidelity to the high-fidelity FEM-generated data with adherence to physical trends embedded in Bishop's simplified method. In this study, the slope stability problem was formulated as a supervised regression task, where the target variable was the FoS obtained from FEM simulations. To enhance physical interpretability and improve the generalization of NN predictions, a PIML framework was adopted. The total training loss was defined as the weighted sum of a data-driven term and a physics-informed regularization term:

$$L_{total} = L_{data} + \beta L_{physics} \tag{1}$$

where, $L_{data} = \frac{1}{N}\sum\left(y_{FEM} - y_{predicted}\right)^2$ is the mean squared error (MSE) between the predicted FoS and FEM outputs, and $L_{physics} = \frac{1}{N}\sum\left(y_{Bishop} - y_{predicted}\right)^2$ is the MSE between the predicted FoS and values obtained from Bishop's method. The hyperparameter β governs the contribution of the physics term relative to the data term.

We chose not to hard-code Bishop's solver directly into the PIML training because its two-stage search and iterative calculations are not differentiable and would slow down the training significantly. Running it for every batch would also hurt HPC efficiency. Instead, we calculated the FoS values beforehand and used them as a soft regularizer, which keeps the physics insight while making the training faster and more stable. This provided flexibility, allowing the model to learn beyond the limitations of Bishop's assumptions. The framework ensured that the model did not simply replicate Bishop's method but instead balanced between FEM accuracy and Bishop's physical interpretability. Also, this hybrid approach offered a practical advantage in geotechnical engineering: Bishop's results are widely recognized and interpretable by practitioners, thus enabling the model's predictions to be reliable. The novelty of this work lies in the first systematic integration of Bishop's limit equilibrium outputs into a neural network for slope stability analysis. Entirely data-driven applications of machine learning in geotechnical engineering often lack physical plausibility and struggle in data-scarce conditions, which are common due to the high cost and variability of soil testing. By introducing a physics-informed regularization term, the proposed model enhances robustness and reduces the risk of overfitting, ensuring that predictions remain consistent with established geotechnical understanding. A schematic diagram of the employed PIML framework is shown in Figure 2.

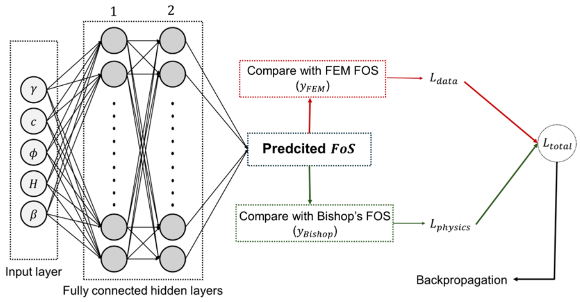

Figure 2. Schematic diagram of the PIML model used in this study.

3.2 *Validation*

The FEM dataset, comprising 2,000 samples, was standardized using z-score normalization for inputs and independently scaled outputs to ensure numerical stability. The PIML model was trained using the Adam optimizer (learning rate = 0.001) for 600 epochs, with β = 0.5 weighting the Bishop loss term. Performance was evaluated using two strategies: (i) a 70-30 random split and (ii) 10-fold cross-validation (CV), both assessing the coefficient of determination (R^2) and root mean squared error (RMSE) on test data.

Figure 3 shows that both strategies achieved rapid convergence, with R^2 exceeding 0.95 after ~50 epochs and RMSE stabilizing below 0.08. The mean R^2 across all epochs was 0.96 (train) and 0.97 (test) for the random split, and 0.96 for both training and testing for 10-fold CV, indicating strong generalization. The close alignment between the random split and CV curves demonstrates that the model's predictive accuracy is robust to data partitioning strategy.

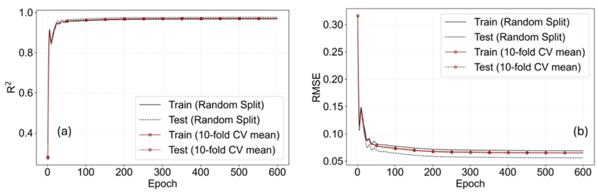

Figure 3. PIML Convergence under random split and 10-fold cross-validation.

4 RESULTS AND MODEL EVALUATION

4.1 *Effects of physics weighting*

The performance of the proposed physics-informed model was evaluated across different values of the physics weighting parameter β. Heatmap plots (Figure 4) illustrate the variation in validation R^2 across combinations of training fractions and β values for dataset sizes of 100, 200, 500, and the full dataset. A consistent trend emerged: in data-scarce settings (100-200 samples), moderate physics weighting (β between 0.1 and 0.5) produced notable gains in R^2 compared to purely data-driven models (β=0). For example, with only 100 data points and a 50% training split, the R^2 score increased from 0.84 (β=0) to 0.92 (β=0.5), indicating a strong benefit from the inclusion of physics guidance. This improvement can be attributed to the regularizing effect of embedding physics, in our case, Bishop's slope stability FoS.

For larger datasets (500 samples and full size), the differences across β values were less pronounced, with all models achieving R^2 values above 0.96. This suggests that as the volume of FEM data increases, the network becomes sufficiently capable of capturing complex nonlinear relationships on its own. However, even in these cases, physics-informed models avoided erratic early-epoch fluctuations and converged more smoothly, reflecting enhanced training stability.

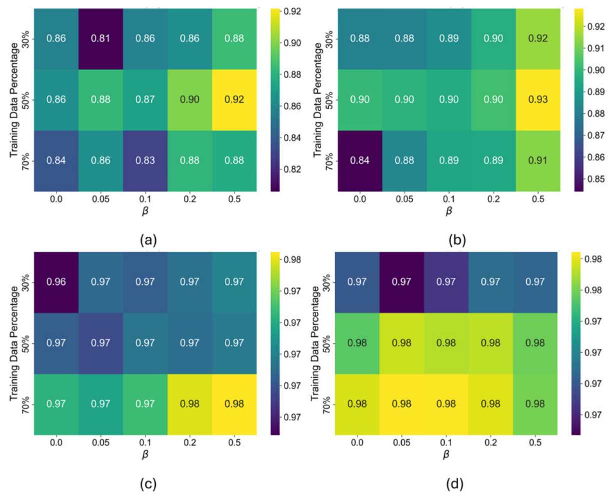

Figure 4. Effect of physics weighting (β) on validation R^2.

This pattern is further illustrated in Figure 5, where convergence curves for 100-sample (50% training set) and full (70% training set) datasets show that, in small-data scenarios, moderate β values lead to higher and more stable R^2, while for the full dataset, all β values converge quickly to near-perfect accuracy with minimal differences.

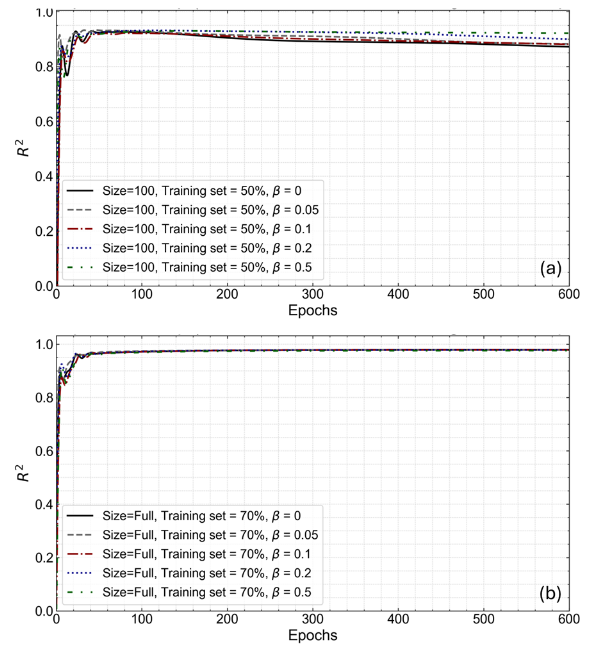

Figure 5. Convergence of R^2 over epochs for different β in (a) 100-sample dataset and (b) full dataset.

4.2 *Sensitivity to training set size*

We also examined the effect of varying the training data fraction (0.3, 0.5, 0.7) for each dataset size to assess the sensitivity of model performance (Figure 4). Surprisingly, for dataset sizes of 100 and 200, the 50% training split often resulted in better validation R^2 scores compared to the 70% split. For example, at 200 samples and $\beta = 0.5$, the model with a 50% training fraction achieved $R^2 = 0.93$, outperforming the $R^2 = 0.91$ observed with 70% training data. This counterintuitive outcome can be explained by the dynamics of small datasets. A higher training fraction leaves a smaller validation set, which may not be representative of the data distribution and may cause unreliable validation metrics due to noise or sampling bias. Additionally, in low-data regimes, increasing the training size may lead to overfitting, especially with high-capacity neural networks. On the other hand, for larger datasets (500 samples and full set), R^2 values remained high (> 0.96) across all training fractions, reflecting improved robustness and generalization when sufficient data is available. These observations underscore the importance of carefully selecting training-validation splits in low-resource scenarios, particularly when using physics-informed learning.

Figure 6 shows the convergence of R^2 and validation RMSE for representative PIML models across dataset sizes. As expected, larger datasets achieve higher R^2 and lower RMSE, reflecting the benefit of more training samples. However, even with as few as 100-200 samples, the PIML framework attains strong predictive performance (R^2>0.9) and stable convergence, underscoring its effectiveness in data-scarce scenarios. This highlights the framework's practical value for geotechnical problems where large, high-quality datasets are often unavailable.

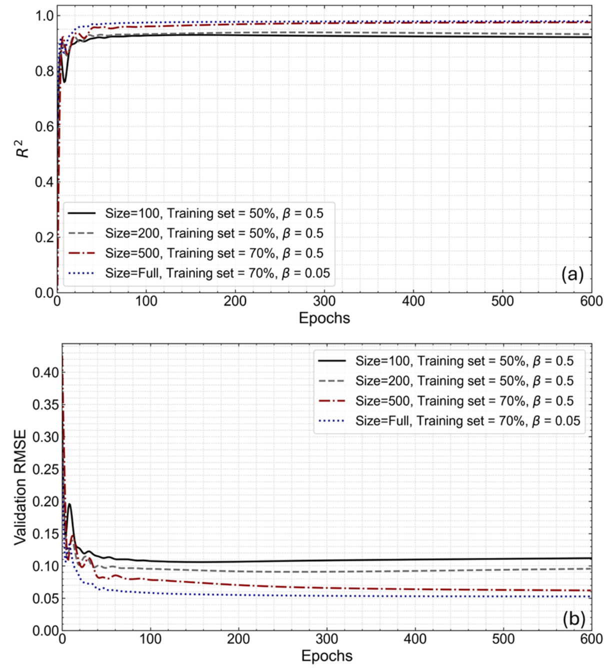

Figure 6. Convergence of (a) R^2 and (b) validation RMSE, for PIML models across different dataset sizes.

4.3 *Best performing models*

For each dataset size, the best-performing model was identified in terms of final validation R^2. Predicted versus actual FoS scatter plots (Figure 7) were plotted to demonstrate the prediction accuracy of the models. Across all dataset sizes, points clustered closely around the 1:1 line, with the closest alignment observed for the full dataset ($R^2 \approx 0.98$). In smaller datasets, purely data-driven models displayed slightly greater scatter, while physics-informed models reduced this spread, ensuring closer adherence to FEM-derived ground truth. Across all evaluations, the highest R^2 values were achieved using larger datasets (500 samples or full), moderate physics weighting (β = 0.1-0.2), and balanced training fractions.

The best performing model achieved $R^2 = 0.98$ using the full dataset with 70% training and $\beta = 0.1$. In low-data settings, the physics-informed approach delivered strong results. With only 100 samples and $\beta = 0.5$, the model reached $R^2 = 0.92$, remarkably close to the performance of models trained on 5 times more data. This demonstrates the practical feasibility and sample efficiency of the proposed PIML approach for slope stability prediction.

These results confirm that embedding physics not only improves predictive performance but also enhances generalization to unseen inputs, even when explicit physics supervision is absent during inference. Overall, the best-performing models emerged from a synergy of sufficient data, balanced training split, and a moderate physics regularization term.

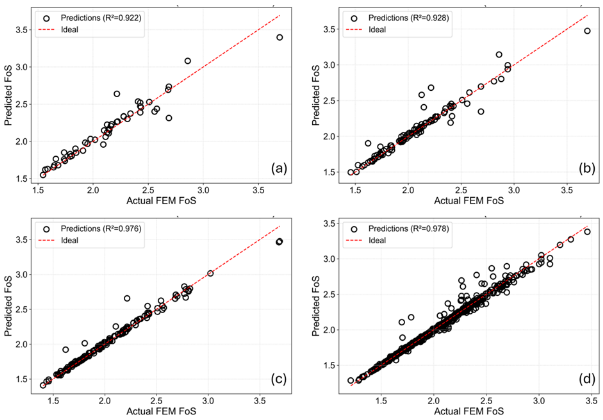

Figure 7. Predicted versus actual FoS for the best-performing PIML models across dataset sizes.

4.4 *Discussion*

Overall, the results confirm that embedding Bishop's FoS as a physics-based regularizer offers tangible benefits for slope stability prediction under data-scarce conditions. The improvements, while modest in absolute terms (0.03-0.06 increase in R^2 for 100-200 sample sets), are significant from a geotechnical practice perspective, where even small gains in reliability can reduce risk in safety-critical designs. The approach is novel in that it leverages Bishop's outputs as a lightweight physics constraint, rather than embedding the entire limit equilibrium formulation, striking a balance between efficiency and physical consistency. Importantly, the PIML framework did not over-constrain the model to Bishop's assumptions. Instead, it enabled the neural network to interpolate between high-fidelity FEM ground truth and well-established geotechnical theory. With larger datasets, the benefits were smaller but still evident in smoother convergence and reduced variance across runs. Thus, the proposed approach demonstrates the value of physics-informed learning as a practical, generalizable, and trustworthy tool for slope stability analysis.

5 CONCLUSIONS

This study presented a PIML framework for slope stability analysis, integrating the outputs of Bishop's simplified method into a neural network trained on high-fidelity FEM data. The approach demonstrated that even a relatively lightweight physics constraint can substantially improve predictive reliability when training data are scarce, a common scenario in geotechnical engineering due to the high cost and variability of soil testing. By embedding Bishop's FoS as a soft regularizer in the loss function, the model achieved a better balance between accuracy and physical consistency than purely data-driven approaches. In the smallest dataset tested (100 samples), the PIML model achieved performance close to models trained on several times more data, demonstrating strong sample efficiency. For larger datasets, the benefits were seen in smoother convergence and reduced variance, confirming that physics-informed training contributes to stability and robustness even when abundant data are available.

Beyond statistical improvements, this work shows that incorporating domain knowledge into machine learning workflows can bridge the gap between black-box predictions and the interpretability required for engineering design. The methodology is adaptable, computationally efficient, and compatible with standard geotechnical analysis tools, making it a practical option for broader application in slope stability evaluation and other soil-structure interaction problems. This combination of accuracy, interpretability, and practicality supports the inclusion of PIML approaches in future design workflows and paves the way for their application to more complex and variable field conditions.

However, certain limitations must be acknowledged. The present framework focuses on simple homogeneous slopes modeled with the Mohr-Coulomb constitutive law, limiting its applicability to more complex scenarios such as layered soils, anisotropy, progressive failure, or pore pressure effects. The physics-informed term is based on Bishop's FoS, which, despite its wide acceptance, assumes predefined circular slip surfaces, potentially reducing accuracy for highly irregular geometries or stress conditions. Future work could improve physical applicability by directly incorporating governing partial differential equations (PDEs) for soil equilibrium and stress-strain compatibility, enabling constraints beyond Bishop's assumptions. Embedding advanced constitutive models, such as critical state frameworks, would capture strain-hardening/softening, dilatancy, and other complex behaviors. Extending the framework to heterogeneous and anisotropic slopes, integrating coupled hydro-mechanical effects, and adding uncertainty quantification would further enhance robustness and reliability for practical geotechnical design.

6 ACKNOWLEDGEMENTS

The study is supported by the new faculty start-up fund, Texas A&M University-Corpus Christi (TAMU-CC). EA and NL gratefully acknowledge TAMU-CC for providing the necessary resources and support. TP's effort is partially supported by the new faculty start-up fund at Stony Brook University and by the U.S. National Science Foundation under Award No. 2425802.

7 DATA AVAILABILITY STATEMENT

Some data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request. These data include the source code for the present study, FEM Python scripts, and the results used to generate all figures and tables.

8 REFERENCES

ABAQUS. 2023. Analysis User's Manual. Dassault Systèmes Simulia, Inc., United States.

Agarwal, E., Pei, T., and Luo, N. 2025. Towards Robust and Adaptive Bearing Capacity Prediction for Dry Shallow Foundations Using Physics-Informed Machine Learning. In *Geo-Extreme 2025:Remote Sensing, Instrumentation, Big Data, and Decision Making*. https://doi.org/10.1061/9780784486498.011.

Agarwal, E., Verma, A.K., Pain, A. and Sarkar, S., 2021, December. Probabilistic Analysis of a Geosynthetic Reinforced Soil Retaining Wall Under Seismic Conditions Using Genetic Programming. In *Indian Geotechnical Conference* (pp. 217-226). Singapore: Springer Nature Singapore.

Chen, X., Wu, Y., Yu, Y., Liu, J., Xu, X.F., and Ren, J. 2014. A two-grid search scheme for large-scale 3-D finite element analyses of slope stability. *Computers and Geotechnics* 62, 203-215.

Cuomo, S., Di Cola, V.S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. 2022. Scientific machine learning through physics–

informed neural networks: Where we are and what's next. *Journal of Scientific Computing* 92(3), 88.

Griffiths, D.V., and Lane, P.A. 1999. Slope stability analysis by finite elements. *Géotechnique* 49(3), 387-403.

Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. 2021. Physics-informed machine learning. *Nature Reviews Physics* 3(6), 422-440.

Lei, D., Zhang, Y., Lu, Z., Lin, H., Fang, B., and Jiang, Z. 2024. Slope stability prediction using principal component analysis and hybrid machine learning approaches. *Applied Sciences* 14(15), 6526.

Liu, Y., and Liang, Y. 2024. Integrated machine learning for modeling bearing capacity of shallow foundations. *Scientific Reports* 14(1), 8319.

Nguyen, T., Shiau, J., and Ly, D.K. 2024. Enhanced earth pressure determination with negative wall-soil friction using soft computing. *Computers and Geotechnics* 167, 106086.

Phoon, K.K., Ching, J., and Wang, Y. 2019. Managing risk in geotechnical engineering from data to digitalization. *Proc. 7th International Symposium on Geotechnical Safety and Risk* (ISGSR 2019), Singapore, Research Publishing.

Phoon, K.K., and Kulhawy, F.H. 1999. Evaluation of geotechnical property variability. *Canadian geotechnical journal*, *36*(4), 625-639.

Pei, T., Qiu, T., and Shen, C. 2023. Applying knowledge-guided machine learning to slope stability prediction. *Journal of Geotechnical and Geoenvironmental Engineering* 149(10).

Pei, T., and Qiu, T. 2024. Machine learning with monotonic constraint for geotechnical engineering applications: an example of slope stability prediction. *Acta Geotechnica* 19(6), 3863-3882.

Pei, T., and Luo, N. 2025. Prediction of ultimate bearing capacity of shallow foundations using monotonically constrained neural networks. *Proc. Geotechnical Frontiers 2025 Conference*, Louisvielle, 11-20.

Verma, A.K., Pain, A., Agarwal, E. and Pradhan, D., 2022. Reliability assessment of tunnels using machine learning algorithms. *Indian Geotechnical Journal*, *52*(4), pp.780-798.