Proc. of the 8th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Jinsong Huang, D.V. Griffiths, Shui-Hua Jiang, Anna Giacomini and Richard Kelly
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5182-7_08-015-cd

ISGSR 2022

Prediction of Slope Failure through Integrating Statistical Design of Experiments and Artificial Neural Networks

Yuderka Trinidad González[1], Vernon R. Schaefer[2], and Derrick K Rollins[3]

[1]Department of Architecture and Civil Engineering, University of Bath, Bath, UK.
E-mail: yatg20@bath.ac.uk
[2]Department of Construction and Environmental Engineering, Iowa State University, Ames, USA, vern@iastate.edu
[3]Department of Chemical and Biological Engineering, Department of Statistics, Iowa State University, Ames, USA, drollins@iastate.edu

Abstract: This study proposes the integration of analytical geotechnical methods and statistical tools to develop a prediction model for the factor of safety (*FS*) of homogenous soil slopes as a function of soil and geometry properties. The proposed model is developed by combining statistical design of experiment (DOE), artificial neural networks (ANN), and limit equilibrium (LE) analysis to generate suitable combinations of input factors and for data analysis. The model adequacy as a prediction tool for preliminary design is evaluated by measuring accuracy using a case histories dataset. The performance results indicate that the predicted values of *FS* have a high correlation with the computer-simulated values (analytical values), indicating that the developed model compares to the use of performing analysis in LE software without the need for special packages. However, all the compared tools fall into the low accuracy zone if the threshold between stability and failure is set to 1. For achieving high accuracy (area under the curve >0.85) from the proposed classifier, a safety margin of 20% should be used. In other words, for a 1.20 *FS* threshold between stability and failure, the proposed model classifies with high accuracy for the analyzed cases.

Keywords: Algorithms; Data processing; Factor of safety; Slope stability; Artificial neural networks; Statistical design of experiment (DOE).

1 Introduction

In many areas, slope instability is a major threat. In the geotechnical field, slope stability is a key analysis conducted to verify the safety of cuts, embankments, and/or natural slopes against failure. In this regard, estimating the factor of safety (*FS*) of a slope is one of the main tasks when performing slope stability analyses. There are several analytical methods available to perform a slope stability analysis, with limit equilibrium (LE) and numerical methods (generally, finite element (FE) and finite difference (FD)) being the most common. Several authors such as Memon 2018; Trinidad González 2017; Cheng et al. 2007; Griffiths and Lane 1999; Duncan 1999; Christian 1999; and have presented the basics of the analytical techniques and some comparisons between methods. However, the use of special software is required to perform such analyses. For this reason, authors such as Jin-kui and Wei-wei 2018; Kostić et al. (2016); Liu et al. 2014; Das et al. 2011; Samui and Kothari 2011; Zhao 2008 have implemented statistical tools, support vector machines, and genetic algorithms to create mathematical expressions for the *FS*. These models aimed to develop reliable estimation tools with a simpler application than traditional analytical methods. Kostić et al. (2016) generated synthetic information by applying response surface techniques as a data generator and a solution using the LE approach (Spencer method). This study proposes the integration of analytical geotechnical methods and statistical tools to develop a prediction model for *FS* of homogenous soil slopes as a function of soil and geometry properties. The proposed model is developed by combining statistical design of experiment (DOE), artificial neural networks (ANN), and LE analysis to generate suitable combinations of input factors and for data analysis. The aim of the proposed model is a simplification of performing slope stability analysis without the need for special software while providing the reliability of the prediction tool when compared to the field condition. Hence the main objectives of this paper are: (1) to develop a prediction model for *FS* of homogeneous slopes with an augmented input space to overcome the constraints of the existing models; and (2) to provide the correlation between the model and the field condition, which was not previously done for models developed with synthetic data.

2 Applied Methods

2.1 Pre-processing and data generation

DOE and LE analysis are blended to generate a suitable combination of input factors for later data analysis. A multifactor (i.e., input), full factorial design is used as a synthetic data generator (combining slope geometry and soil properties). Six factors are selected for the analysis following findings from previous research regarding the factors controlling the mechanism of failure (Sah et al. (1994), Sakellariou and Ferentinou (2005), Yang et al.

(2004), Ahangar-Asr et al. (2010), Samui and Kothari (2011) and Manouchehrian et al. (2014), Kostić et al. (2016), Trinidad González et al. (2020)). The multifactor, full factorial design generated 4,032 combinations of slope and soil properties in a complex, highly dimensional input space. Each factor and its levels (e.g., L1 is Level 1) are given in Table 1.

Table 1. Soil and geometry properties for generating the multilevel full factorial design

Factor	L1	L2	L3	L4	L5	L6	L7
Slope height (H, m)	6	15.5	25	34.5	44	53.5	63
Slope inclination (β, °)	10	30	50	70			
Effective Cohesion (c', kPa)	5	20	35	50			
Effective friction angle (φ', °)	10	22	34	46			
Pore pressure coefficient (r_u)	0	0.3	0.6				
Soil unit weight (γ, kN/m^3)	12	19	26				

The results of the 4,032 ($7x4^3x3^2$) input combinations are then used to perform LE analyses to determine the *FS* for each condition. *FS* from the LE analysis corresponds to the "observed" or "modeled" response, while the *FS* from the generated ANN model corresponds to the "fitted" response. For the stability analysis, LE analyses are conducted. Many different types of software have been developed based on limit equilibrium methods such as proprietary software and specifically programmed loops that enable an analysis of a given situation using different procedures. In this study, Slide 2018 (Rocscience Inc. 2018) and the Spencer method are used. The Spencer method satisfies all the requirements for static equilibrium (Duncan et al. 2014). Non-circular mode of failure with auto refine search is used as a search method. The boundary conditions are set sufficiently far not to influence the *FS*. A sketch of the section of a slope with the input properties for the LE analyses is shown in Figure 1.

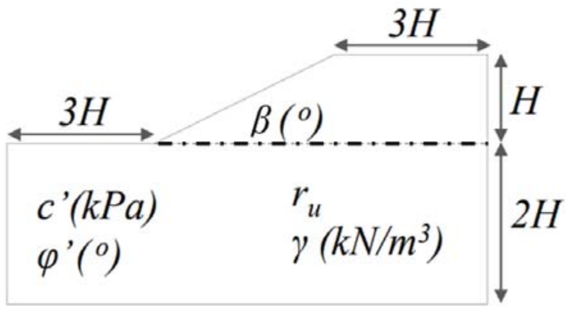

Figure 1. Graphical representation of inputs for LE analyses

2.2 Fitting approach and performance measurements

After pre-processing, the ANN prediction model is generated. ANN are non-linear regression or classification model structures consisting of an input layer (the properties or variables), the hidden layers (or hidden units that relates the input and output from a linear combination of the inputs), and the output layer (the fitted response or responses of the system). The ANN used in this study is a fully connected, multiple-layer perceptron with an input layer of six neurons, a hidden layer of eight neurons, and an output layer of one neuron. The model architecture is presented in Figure 2. For the data classification, 1/3 of the dataset was assigned to the validation set.

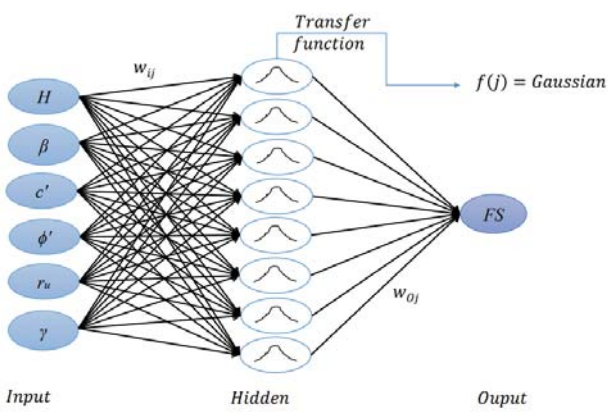

Figure 2. Schematic illustration of ANN structure.

The mathematical representation of the model is given by Equation 1 as

$$\hat{y} = \beta_0 + \sum_{j=1}^{n}\left[w_{Oj} . f_j\left(b_{Aj} + \sum_{i=1}^{n} w_{ij}x_i\right)\right] \tag{1}$$

where β_0 is the bias of the output layer, w_{Oj} is the connection weight between neuron j of the hidden layer (j = *1 to 8*) and the output layer; b_{Aj} is the bias at neuron j of the hidden layer; w_{ij} is the connection weights between the input variable i (*for i = 1 to 6*) and neuron j of the hidden layer; x_i is input i, f_j is the activation function at the hidden layer. The activation function used in this study is the identity transformation (Gaussian) that for the variable j (neurons (j = *1 to 8*) is defined as

$$f(j) = e^{(b_{Aj} + \sum_{i=1}^{n} w_{ij}x_i)} \tag{2}$$

During the training stage, the weights are adjusted using wOj by learning from the training data set. The sum of the adjusted weights is then added to the bias of the output layer β_0. A quasi-Newton method, BFGS (Broyden–Fletcher–Goldfarb–Shanno algorithm), iterates for optimization of the penalty parameter in the training stage. Simultaneously, the BFGS algorithm monitors the likelihood function of the validation set (an independent set that is used to check model performance following Trinidad González et al. 2021; Ghasemi et al. 2019; Cheng and Titterington 1994). A cross-validation technique is used to test the effectiveness of the model and avoid overfitting (Ghasemi et al. 2019). The holdout procedure is applied to determine which data are randomly assigned to either training or validation sets. The performance of the final model is measured based on the statistics summarized in Table 2.

Table 2. Statistics for performance evaluation of the generated ANN model

Statistics	Defined as
The "average difference" as the estimate of the model bias	$AD = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})$
The "average absolute difference" as the averaged absolute distance between fitted and simulated values	$AAD = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}\rvert$
The correlation between the simulated and fitted values	$r_{fit} = \frac{n\sum_{i=1}^{n} y_i\hat{y} - (\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n}\hat{y})}{\sqrt{n\sum_{i=1}^{n} {y_i}^2 - (\sum_{i=1}^{n} y_i)^2}\sqrt{n\sum_{i=1}^{n}\hat{y}^2 - (\sum_{i=1}^{n}\hat{y})^2}}$
Generalized coefficient of determination	$r^2_{Generalized} = 1 - e^{\frac{2}{n}(L\hat{\beta} - L_{Null})}$

where n is the number of input vectors, y_i is the i^{th} simulated response, $\hat{y}_i$ is the i^{th} fitted response, and $L\hat{\beta}$ is the negative loglikelihood of the set using the model parameters on the training data.

3 Results and Discussions

3.1 Model performance

The general components of the model equation (the estimates) are given as follows:

$$\hat{y} = 1.71 + 7.19x\, f_1 + 40.78\, x\, f_2 \pm 3.41\ x\, f_3 + 4.075\, x\, f_4 + 1.25x\, f_5 \pm 14.72\, x\, f_6 + 10.73x\, f_7 + 4.63\, x\, f_8 \tag{3}$$

The summary of the results from the performance evaluation is presented in Table 3.

Table 3. Summary of Statistics for performance evaluation

Statistics	Training	Validation
AD	0.007	0.01
AAD	0.05	0.06
r_{fit}	0.998	0.998
$*R^2$	0.997	-

$* R^2 = r^2_{Generalized}$

The results indicate that the predicted values of *FS* have a high correlation with the computer-simulated values as well as low *AD* and *AAD*. Hence, the prediction tool modeled the response well. The observed vs. predicted plots for both training and validation sets are presented in Figure 3.

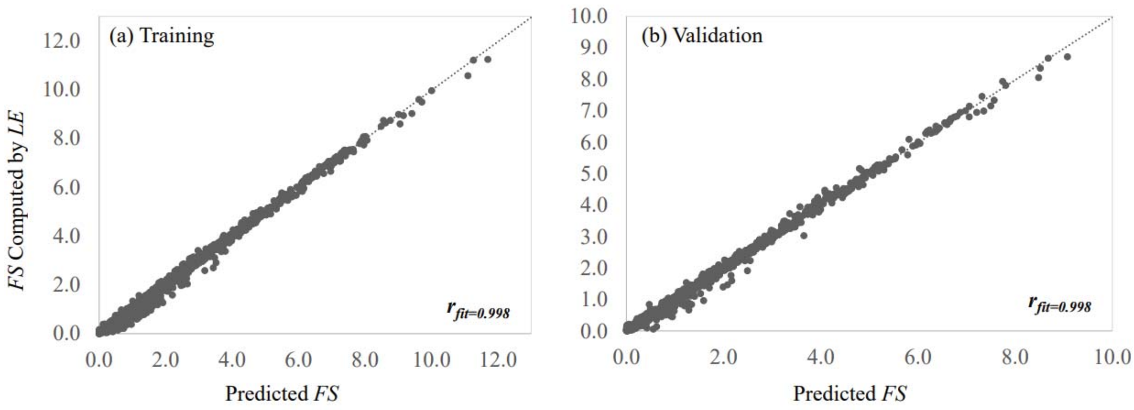

Figure 3. Observed vs. predicted plot for the best ANN model

An input analysis is conducted to verify the standardized effect of the input properties on the response. Results are presented in Figure 4, indicating that for the input space evaluated, the inputs with larger effects on the response are the inclination angle and the friction angle. The results are in agreement with the ones presented by Sakellariou and Ferentinou (2005) when performing a parametric study in an ANN model generated using a genetic algorithm, and Jin-kui and Wei-wei (2018) when conducting a sensitivity analysis applying the basics of orthogonal design. Results presented in Figure 4 also indicate that interactions among the inputs also affect the response. Main effects come from changes in the response due to changes in the inputs independently, while interaction effects come from changes in the response due to the pairing of multiple factors.

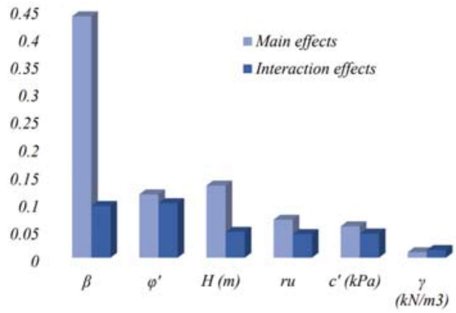

Figure 4. Assessment of variable importance (input's effect over the response)

3.2 Application and Comparisons with Previous Models

The comparison follows (1) quantification of prediction capabilities in a wider input space when compared to analytical solutions and (2) quantification of accuracy when compared to the field condition. The comparisons are conducted with independent datasets not used for the model generation. The models presented by Kostić et al. (2016), Manouchehrian et al. (2014), Ahangar-Asr et al. (2010), Yang et al. (2004) are compared with the proposed model (TG Model). All but Kostić et al. (2016) models used the same data set of slopes presented by Sah et al. (1994). Kostić et al. (2016) model was created with a synthetic dataset like in the present study. The constraints to avoid extrapolation are specified in Table 4.

Table 4. Range of inputs from the compared prediction models updated from Kostić et al. (2016)

Input	TG Model	Kostić et al. (2016)	Manouchehrian et al. (2014)	Ahangar-Asr et al. (2010)	Yang et al. (2004)
β	10-70	25-70	16-53	16-53	16-53
φ'	10-46	10-50	0-45	0-45	0-45
H (m)	6-63	6-10	3.6-214	3.66-214	3.66-214
r_u	0-0.6	0-0.5	0-0.5	0-0.5	0.11-0.5
c' (kPa)	5-50	0-50	0-50	0-150.05	0-150.05
γ (kN/m3)	12-26	16-20	12-28.44	12-28.44	12-28.44

Twenty random combinations of slopes are selected and modeled with LE analyses to determine the *FS* using the Spencer method. The summary of the performance from the comparison is presented in Table 5. The results indicate that for the models created from the field database, the r_{fit} value is considerably low compared to the models from the synthetic database (Kostić et al. (2016) and TG model). This is expected because Kostić et al. (2016) and the TG model are developed applying analytical solutions. The results are used to demonstrate that the proposed model prediction capabilities compare to performing LE analysis without the need for special packages. However, an agreement with analytical solutions does not necessarily guarantee agreement with the field condition. Hence, the purpose of the second comparison is to evaluate how the mathematical tools perform when predicting the real condition of stable and unstable slopes that is the end goal of prediction.

Table 5. Summary of Statistics for performance evaluation, testing sets

Statistics	TG Model	Model 2016	Model 2014	Model 2010	Model 2004
AD	0.62	0.63	0.84	0.34	0.28
AAD	0.34	0.37	0.64	0.29	0.23
r_{fit}	0.99	0.99	0.64	0.48	0.73

To determine the model accuracy compared to the field condition, 51 cases are selected from the ones presented by Sah et al. (1994); Manouchehrian et al. (2014). The cases are coded to zero when the slope is failed and one when the slope is stable. TP is true positive, failed slopes classified as failed. TN is true negative, stable slopes classified as stable. FP is false positive, slopes classified as failed that are stable. FN is a false negative, slopes classified as stable that are, in fact, failed. From the results, a receiver operating characteristic (ROC) curve and the area under the curve (AUC) are determined for each prediction tool. The ROC curve is a graphical plot that illustrates the diagnostic ability of a binary classifier system when the discrimination threshold is varied. The ROC is created by plotting TP_{rate} or Sensitivity versus FP_{rate} or *1 - Specificity*. AUC measures true positive rate and false positive rate trade-off, testing the quality of the value generated by a classifier (prediction tool) then comparing the value to a threshold. Different thresholds for boundary stability and failure are set until an AUC>0.85 is achieved. The closer a curve is to the point (0, 1), the more accurate a predictor is. According to D'Agostino et al. (2018), as a rule of thumb, AUC above 0.85 means high classification accuracy, one between 0.75 and 0.85 moderate accuracies, and one less than 0.75 low accuracies. The studied *FS* thresholds are 1.0, 1.1, 1.2. The AUC of the ROC shows the ability of the test to distinguish between classes (in this study, failed and stable slopes). The results are summarized in Figure 5.

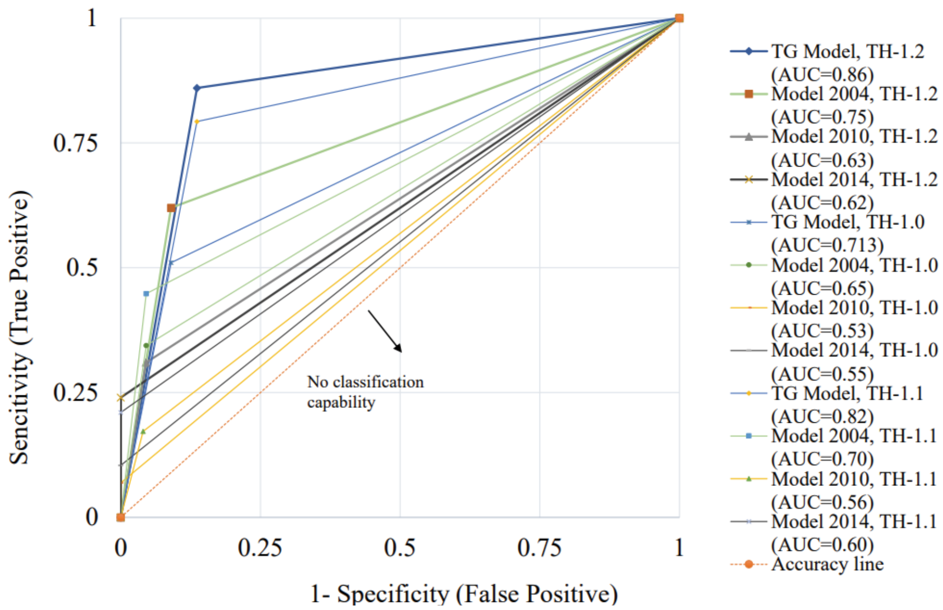

Figure 5. Receiver operating characteristic (ROC) curve and the area under the curve (AUC) to assess the ability of the prediction tools to classify slopes as failed or stable

The results from Figure 5 indicate that, in general, for all studied prediction models, the proposed model (TG model) achieved a higher AUC. When the threshold between failure and stability is set to *FS* of 1, the larger AUC, corresponding to the TG model, is around 0.71. This translates to a low accuracy indicating that high agreement between analytical methods and the prediction models does not guarantee a high agreement with the field condition. The threshold between failure and stability is varied to verify the minimum *FS* needed to achieve high classification accuracy from the compared prediction tools. AUC of 0.86 is achieved by applying a 20% margin of safety to the use of the proposed model (TG model). These results are useful to set confidence intervals to the results of predictions using models for preliminary design.

4 Conclusions

A prediction model for the factor of safety (*FS*) of homogeneous soil slopes is developed by combining statistical design of experiment (DOE), artificial neural networks (ANN), and limit equilibrium (LE) analysis to generate suitable combinations of input factors and for data analysis. Six factors are selected for developing the proposed model (TG Model). The performance results indicate that the predicted values of *FS* have a high correlation with the computer-simulated values (analytical values). Hence, the prediction tool modeled the response well. These results indicate that the developed model compares to the use of performing analysis in LE software without the need for special packages. The proposed model has high accuracy in predicting *FS* when compared to the field condition when a safety margin of 20% to denote the boundary between stability and failure is used. The applicability of the TG model can be broadened by adding levels to the input factors (i.e., lower, and higher heights, lower and higher friction, and cohesion) or input factors for layered slopes. It should be noted that the proposed approach should be used as a preliminary prediction tool and for simplified approximations within the range of parameters for which it was developed.

Acknowledgments

This work was supported by the Civil Construction and Environmental Engineering department at Iowa State University.

References

Ahangar-Asr, A., Faramarzi, A., and Javadi, A. A. (2010). "A new approach for prediction of the stability of soil and rock slopes." *Engineering Computations: Int J for Computer-Aided Engineering*, 27(7), 878–893.

Cheng, B., and Titterington, D. (1994). "Neural networks: a review from a statistical perspective." *Statistical Science*, 2–30.

Cheng, Y. M., Lansivaara, T., and Wei, W. B. (2007). "Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods." *Computers and Geotechnics*, 34(3), 137–150.

Christian, J. T. (1999). "What Has the Finite Element Method Done for (or to) Geotechnical Engineering?" *Civil engineering practice, Journal of the Boston Society of Civil Engineers*, 14(2), 73–74

Das, S. K., Biswal, R. K., Sivakugan, N., and Das, B. (2011). "Classification of slopes and prediction of factor of safety using differential evolution neural networks." *Environmental Earth Sciences*, 64(1), 201–210.

Duncan, J. M. (1999). "Applying the finite element to practical use in geotechnical engineering." *Civil engineering practice, Journal of the Boston Society of Civil Engineers*, 14(2), 75–80.

Duncan, J. M., Wright, S. G., and Brandon, T. L. (2014). Soil Strength and Slope Stability. *John Wiley & Sons*.

Ghasemi, P., Aslani, M., Rollins, D. K., and Williams, R. C. (2019). "Principal component analysis-based predictive modeling and optimization of permanent deformation in asphalt pavement: elimination of correlated inputs and extrapolation in modeling." *Structural and Multidisciplinary Optimization*, 59(4), 1335–1353.

Griffiths, D. V., and Lane, P. a. (1999). "Slope stability analysis by finite elements." *Geotechnique*, 49(3), 387–403.

Jin-kui, R., and Wei-wei, Z. (2018). "Sensitivity Analysis of Influencing Factors of Building Slope Stability Based on Orthogonal Design and Finite Element Calculation." *E3S Web of Conferences*, 53, 03076.EDP Sciences.

Kostić, S., Vasović, N., and Sunarić, D. (2016). "Slope stability analysis based on experimental design." *International Journal of Geomechanics*, 16(5), 04016009.

Liu, Z., Shao, J., Xu, W., Chen, H., and Zhang, Y. (2014). "An extreme learning machine approach for slope stability evaluation and prediction." *Natural Hazards*, 73(2), 787–804.

Manouchehrian, A., Gholamnejad, J., and Sharifzadeh, M. (2014). "Development of a model for analysis of slope stability for circular mode failure using genetic algorithm." *Environmental Earth Sciences*, 71(3), 1267–1277.

Memon, Y. (2018). "A comparison between limit equilibrium and finite element methods for slope stability analysis." *Missouri University of Science & Technology, Rolla, Missouri*. 1-18. DOI: 10.13140/RG.2.2.16932.53124.

Rocscience Inc. (2016). "Slide 7.0 2D limit equilibrium slope stability analysis program."

Sah, N., Sheorey, P. R., and Upadhyaya, L. (1994). "Maximum likelihood estimation of slope stability." *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 31(I), 47-53. Pergamon.

Sakellariou, M. G., and Ferentinou, M. D. (2005). "A study of slope stability prediction using neural network." *Geotechnical and Geological Engineering*, 23, 419–445.

Samui, P., and Kothari, D. P. (2011). "Utilization of a least square support vector machine (LSSVM) for slope stability analysis." *Scientia Iranica*, 18(1), 53–58.

Trinidad González, Y. (2017). "Comparison of design/analysis methods for pile reinforced slopes." *Graduate Thesis 15631. Iowa State University*.

Trinidad González, Y., Schaefer, V. R., and Rollins, D. K. (2020). "Statistical assessment of factor of safety for pile-reinforced slopes." *Geotechnical and Geoenvironmental Engineering*, 146(9), 1–11.

Trinidad González, Y., Schaefer, V. R., and Rollins, D. K. (2021). "Statistical Insights Regarding Fully Softened Shear Strength." *Geotechnical and Geological Engineering*, 39(2), 1155–1171.

Yang, C., Tham Feng, X. ., Wang, Y. J., and Lee, P. K. K. (2004). "Two-stepped evolutionary algorithm and its application to stability analysis of slopes."." *Journal of Computing in Civil Engineering*, 18(2), 145–153.

Zhao, H. B. (2008). "Slope reliability analysis using a support vector machine." *Computers and Geotechnics*, 35(3), 459–467.