Proceedings of the XVIII ECSMGE 2024
GEOTECHNICAL ENGINEERING CHALLENGES
TO MEET CURRENT AND EMERGING NEEDS OF SOCIETY

ISBN 978-1-032-54816-6
DOI 10.1201/9781003431749-280

ECSMGE 24
XVIII EUROPEAN CONFERENCE ON SOIL MECHANICS AND GEOTECHNICAL ENGINEERING

Probability-based hybrid landslide assessment system integrating landslide susceptibility and numerical modeling

Système d'évaluation hybride des glissements de terrain basé sur la probabilité intégrant la susceptibilité aux glissements de terrain et la modélisation numérique

Y.-H. Chen*, K.-W. Liao
Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, Taiwan
**r12622015@ntu.edu.tw*

ABSTRACT: Landslides have emerged as a critical disaster type due to the increasing frequency of extreme rainfall events attributed to climate change. In recent years, research on landslide analysis and slope stability has grown. The most common research methods in landslide studies fall into two categories: numerical modeling and machine learning models. Both approaches hold potential for contributing to landslide disaster management, although research combining these methods remains limited. Therefore, the primary objective of this study is to harness the strengths of both approaches, which involves the establishment of landslide susceptibility maps and the subsequent development of probability-based numerical model, providing valuable insights for disaster prevention on high-risk areas. The results from this preliminary phase indicate that the categorization boundaries in the machine learning-based landslide susceptibility maps are primarily influenced by the proportion of historical landslide areas. However, given that this factor does not represent the main characteristics of the target objects, it should be excluded in future research. The forthcoming stages will emphasize the examination of slope and dip slope area proportions while incorporating additional landslide-related geomorphic and hydrological factors and enhancing data resolution to improve predictive model accuracy. In the context of probability-based numerical model, the research highlights the sensitivity of cohesion and the internal friction angle to the probability of slope system failure. Real-world slope and soil parameter experiments are planned for the next phase to establish a more realistic probability-based numerical model.

RÉSUMÉ: Les glissements de terrain ont émergé comme un type de catastrophe critique en raison de l'augmentation de la fréquence des événements pluvieux extrêmes attribués au changement climatique. Ces dernières années, la recherche sur l'analyse des glissements de terrain et la stabilité des pentes a augmenté. Les méthodes de recherche les plus courantes dans les études sur les glissements de terrain se répartissent en deux catégories : la modélisation numérique et les modèles d'apprentissage automatique. Les deux approches ont le potentiel de contribuer à la gestion des catastrophes liées aux glissements de terrain, bien que la recherche combinant ces méthodes reste limitée. Par conséquent, l'objectif principal de cette étude est de tirer parti des forces des deux approches, ce qui implique l'établissement de cartes de susceptibilité aux glissements de terrain et le développement ultérieur de modèle numérique basés sur la probabilité, fournissant des informations précieuses pour la prévention des catastrophes dans les zones à haut risque. Les résultats de cette phase préliminaire indiquent que les limites de catégorisation dans les cartes de susceptibilité aux glissements de terrain basées sur l'apprentissage automatique sont principalement influencées par la proportion de zones de glissements de terrain historiques. Cependant, étant donné que ce facteur ne représente pas les principales caractéristiques des objets cibles, il devrait être exclu des recherches futures. Les prochaines étapes mettront l'accent sur l'examen des proportions de pentes et de zones de pentes inclinées tout en incorporant des facteurs géomorphologiques et hydrologiques supplémentaires liés aux glissements de terrain et en améliorant la résolution des données pour améliorer la précision du modèle prédictif. Dans le contexte des modèle numérique basés sur la probabilité, la recherche met en évidence la sensibilité de la cohésion et de l'angle de frottement interne à la probabilité de défaillance du système de pente. Des expériences sur le terrain réel des paramètres de pente et de sol sont prévues pour la prochaine phase afin d'établir un modèle numérique basé sur la probabilité plus réaliste.

Keywords: Landslide management; susceptibility maps; reliability analysis; sensitivity analysis

1 INTRODUCTION

Landslides pose a significant threat to human life and property, making them one of the most critical disaster types. Moreover, with the increasing frequency of extreme rainfall events due to climate change, landslides triggered by heavy precipitation demand more attention and management (Crozier, 2010). Over the years, there has been a multitude of research in slope stability analysis and landslide disaster mitigation. The most common research methods fall into two categories: (1) numerical modeling and (2) machine learning models. In this study, the term "numerical models" refers to methods based on a deterministic approach. These methods rely on complex mathematical computations that require computer calculations. Examples include Limit Equilibrium Methods (LEMs), Finite Element Method (FEM), and the Material Point Method (MPM). Numerical models can simulate the movement of slopes and study the mechanisms of failure, but typically limited to relatively small areas. On the other hand, machine learning models use geospatial and hydrological data to generate risk assessment maps on a broader scale, employing a data-driven approach. This method offers the advantage of identifying areas with potential high risk (Merghadi et al., 2020).

Both machine learning and numerical models have their merits and can significantly contribute to landslide disaster management. However, there is limited research that combines both methods for analysis. Hence, the ultimate aim of this study is to leverage the strengths of both approaches. Initially, a machine learning model will be used to establish landslide susceptibility maps. Subsequently, probability-based numerical model using a probability of failure approach will be constructed for areas with a potential high risk, offering valuable insights for disaster prevention.

The research consists of three phases. The first phase involves creating preliminary machine learning and probability-based numerical model for simpler cases. Factors affecting landslide susceptibility will be analysed, and adjustments made based on their impact. The future direction for model complexity will also be discussed.

In the second phase, the most suitable combination of factors identified in the first phase will be incorporated into more complex model designs, considering additional real-world conditions in the study area. The representative model will be employed to identify hazard zones and conduct probability-based numerical model analysis. Finally, the third phase involves model validation and optimization and an attempt to generalize this hybrid model for landslide disaster applications in other regions.

This draft paper primarily focuses on the outcomes of the first phase and presents specific plans for future research work based on these results.

2 METHODS

2.1 Study area

The Jhuoshuei River is Taiwan's longest river, approximately 187 km in length, and it flows through the central region of Taiwan, spanning four counties and cities. Its upper and middle reaches are characterized by fractured bedrock and heavy rainfall in mountainous areas, resulting in high sediment content in the river. The lower reaches are densely populated and heavily involved in agricultural activities. Taking into account slope vulnerability and the considerations for resident safety, this area was chosen for the first phase of the study.

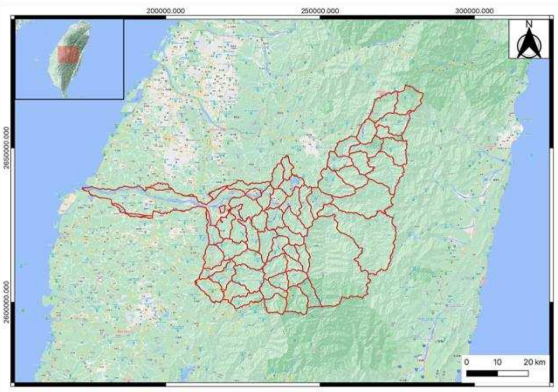

Figure 1. Study watersheds and their slope units.

2.2 Establishment of Machine Learning Models

The study area was divided into 53 sub-watersheds using watershed analysis methods (Figure 1). Three major extreme rainfall typhoons in recent years, namely, Typhoons Saola, Soulik, and Meranti, were selected as the main events for landslide analysis. Geospatial and hydrological factors were chosen as parameters for building the machine learning models and included slope, dip slope area proportion, historical landslide area proportion, and rainfall from individual typhoon events.

Using the landslide records corresponding to the selected typhoon events as known landslide outcomes, Support Vector Machine (SVM) was employed to create landslide prediction machine learning models. The dataset was divided into a training set (3/4 of the data) and a testing set (the remaining 1/4). Two types of classification boundaries, linear and non-linear (RBF), were utilized to compare the classification performance of the two SVM models.

2.3 Identification of key factors in machine learning models

As a classification model, visual representation of SVM classification boundaries can be created by pairwise combinations of factors. This allows for a comparison of the determinative impact of different factors on landslide classification outcomes. For instance, if the classification boundary appears to be vertical, it signifies that the classification results are entirely dependent on the horizontal-axis factor. Conversely, if the boundary is horizontal, it implies that the classification results are solely dependent on the vertical-axis factor.

By comparing the pairwise combinations, it is possible to rank and assess the significance of each factor in determining landslide prediction outcomes within this model. ranking is instrumental in determining the level of emphasis to be placed on each factor.

2.4 Establishment of probability-based numerical model

This research focuses on landslides with large displacement, and traditional slope stability analysis methods like the Limit Equilibrium Method (LEM) and Finite Element Method (FEM) have limitations in simulating large displacements and cannot describe the motion processes adequately (Kaur and Sharma, 2016; Soga et al., 2016). In contrast, the Material Point Method (MPM) can provide better simulations of the motion processes and offer in-depth insights into the physical mechanisms of slope failure (Wyser et al., 2020).

In this study, Anura3D software was chosen for the MPM analysis of slope stability. To simplify the problem in the initial stages of the research, a basic geometric scenario was selected for analysis (Anura3D MPM Research Community, 2022a). The geometric is illustrated in Figure 2.

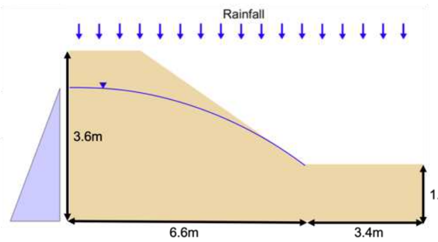

Figure 2. Geometric of selected problem (Anura3D MPM Research Community, 2022a).

To quantitatively assess the slope's safety level, this study introduces reliability analysis. This involves using AK-PSO-HHs (Thedy and Liao, 2023) to calculate failure probabilities under various conditions based on the MPM results. Note that the accuracy of this method has been validated, for details, please refer to Thedy and Liao, 2023.

In reliability analysis, it is essential to provide a clear definition of „failure“. As the shear stress-based safety factor cannot fully represent the occurrence of large landslides, this study chooses to define failure in terms of „displacement“, which is a more direct indicator of landslide occurrence. In practical management, displacement or displacement rates are often used to establish warning thresholds.

2.5 Identification of key factors in probability-based numerical model

This study analyses seven factors, encompassing cohesion (C), internal friction angle (ϕ), wet density (ρ), Young's modulus (E), along with hydrological parameters, including infiltration rate (I_R), groundwater level (H), and the correlation coefficient between cohesion and internal friction angle (r). These parameters are selected within reasonable ranges, their respective probability distribution types and statistical parameters are detailed in Table 1.

A sensitivity analysis of the system's failure probability concerning various factors is conducted following established methodologies (Malkawi et al., 2000). By adjusting the coefficients of variation (COV) for these seven factors within their reasonable ranges, the impact of each factor's variability on the system's failure probability is explored. This comparative analysis identifies the sensitivity of each parameter, and parameters demonstrating sensitivity within the system are considered as key factors of interest.

Table 1. Parameter information.

Parameter	Probability distribution	Mean	Range
C (kPa)	Lognormal	2.5	1.1~5.3
ϕ (°)	Lognormal	27	19.2~41.2
E (kPa)	Lognormal	30000	8450~11540
ρ (kg/m^3)	Lognormal	2650	2175~3250
I_R (m/s)	Uniform	0.0002	0.0001~ 0.0004
H (m)	Beta	2.2	1.5~2.5
r (-)	–	0	-0.9~0.9

3 RESULTS

3.1 SVM model performance

Performance assessment of the models includes confusion matrices for both linear and non-linear classification boundaries (Table 2, Table 3). Three primary validation indices, True Positive Rate (TPR), True Negative Rate (TNR), and Accuracy, are utilized to gauge model effectiveness. The findings consistently favour the linear model over the non-linear one across all metrics. Subsequently, emphasis will be placed on the superior linear model in the ensuing discussion.

Table 2. Linear SVM Model performance.

	Actual landslides	Actual no landslides
Predicted Landslides	11	6
Predicted no landslides	2	21
TPR	84.6%	
TNR	77.8%	
Accuracy	80%	

Table 3. Nonlinear SVM Model performance.

	Actual landslides	Actual no landslides
Predicted Landslides	9	8
Predicted no landslides	8	15
TPR	52.9%	
TNR	65.2%	
Accuracy	60%	

3.2 Factor analysis in the SVM model

To establish a landslide prediction model, the factors used are combined pairwise. A linear classification boundary model is employed to visualize the relationships between the classification boundary and various factors, aiming to compare the significance of each factor in predicting „landslide occurrence“, as shown in Figure 3.

In the combination of dip slope area proportion and slope, the model categorizes all data points as „no landslide“ (Figure 3a), resulting in an overall accuracy of only 57.5%. Conversely, when combining historical landslide area proportion with dip slope area proportion (Figure 3b), the overall accuracy reaches 75%.

However, the classification boundary is solely dependent on historical landslide area proportion, as one factor. A similar situation is observed in the combination of historical landslide area proportion with slope (Figure 3c).

Finally, by combining rainfall of each event and historical landslide area proportion, the overall predictive accuracy reaches 77.5% (Figure 3d), approaching the predictive accuracy of the model with four integrated factors (80%). Nevertheless, the classification boundary remains primarily influenced by historical landslide rates, with only slight adjustments based on rainfall.

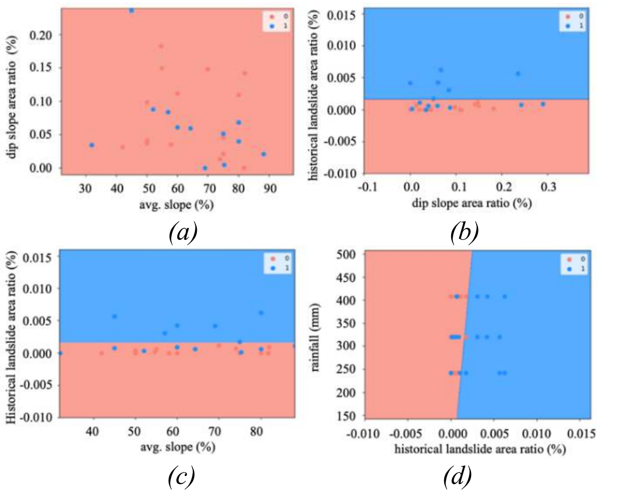

Figure 3. Pairwise comparison SVM classification boundaries. (a) Combination of dip slope area proportion and slope. (b) Historical landslide area proportion and dip slope area proportion. (c) Historical landslide area proportion and slope. (d) Rainfall of each event and historical landslide area proportion.

3.3 Factor sensitivity in the probabilistic numerical model

In this section, the approach outlined by Malkawi et al. (2000) is employed to create sensitivity assessment plots of the variation in failure probability concerning each parameter.

The results of this simulation exhibit a linear relationship in all cases, except for the negative correlation between friction angle and cohesion. The sensitivity of each parameter can be assessed by examining the slope of the regression line. In this research case, the most influential factors in terms of their impact on the system's failure probability are COV($\emptyset$) with a slope of 0.3077 and COV(C) with a slope of 0.2492. Following these are the positive correlation (r_+) between cohesion and friction angle with a slope of 0.028, COV(H) with a slope of 0.0175, and COV(I_R) with the smallest slope of 0.0023. COV(ρ) and COV(E) exhibit minimal sensitivity in this context.

4 DISCUSSION

4.1 Factors of interest in the SVM model

The impact of slope and dip slope area proportion on classification boundaries, as observed in pairwise evaluations, seems to be minimal (Figure 3).

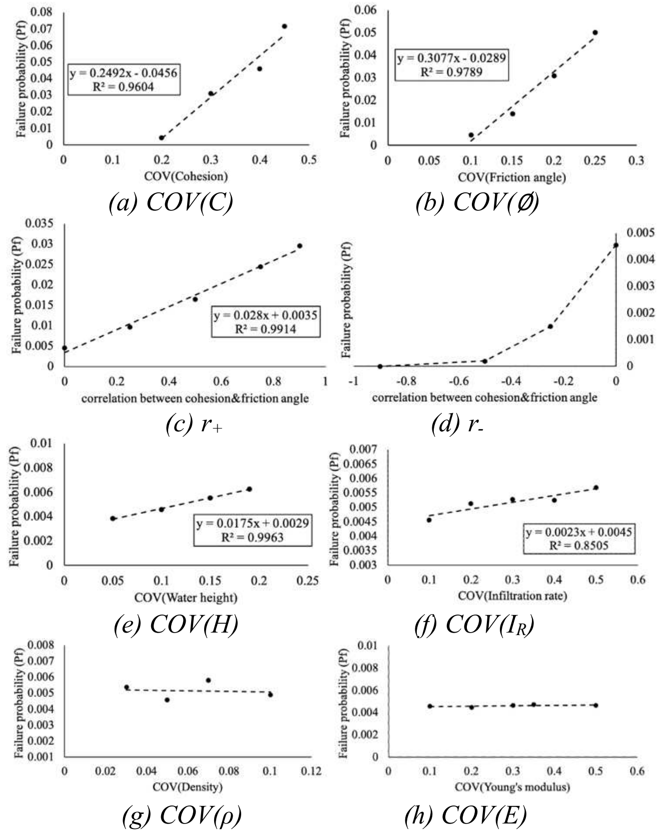

Figure 4. Sensitivity analysis of each parameter.

However, it is essential not to draw conclusions regarding the insignificance of these two geomorphic factors. The historical landslide area proportion, which takes precedence, represents the proportion of landslide-prone areas within sub-watersheds. Disasters tend to recur in regions where critical conditions exist, making the „historical landslide area proportion" an integrative measure that accounts for all contributing landslide factors. Thus, incorporating it along with natural background conditions, such as slope and dip slope area proportion, as predictive factors in the model, can dilute the effects of individual natural conditions and potentially lead to misleading conclusions about their importance.

In the final phase of this study, the primary objective is to identify high-potential landslide areas that have been previously overlooked. These areas typically lack significant historical landslide records. While the historical landslide area proportion predominantly influences the classification in this research phase, considering its impact on other factors, along with its misalignment with the target characteristics, suggests its exclusion from the list of factors of interest in subsequent research phases. Moreover, the predictive performance highlights the need to incorporate additional landslide-related geomorphic and hydrological factors, such as rock type, rock fragmentation, and groundwater levels, as factors of interest. This comprehensive approach will provide a more complete understanding of the influence of various spatial factors on landslide occurrence.

4.2 SVM model refinement directions

Based on the current phase's results, recommendations for enhancing the machine learning model in the next phase are: (1) Continue investigating rainfall of each event, and focusing on diluted factors like slope and dip slope area proportion while diversifying relevant factors for improved predictions; (2) Increasing the spatial resolution of factors to better represent sub-watershed characteristics; (3) Considering a smaller study area, focusing on a few key sub-watersheds due to enhanced data resolution.

4.3 Factors of interest in the probability-based numerical model

Based on the results of parameter sensitivity analysis (Figure 4), cohesion (*C*), the internal friction angle (Ø), and their correlation were identified as the most sensitive factors influencing system failure probability. Therefore, they have been included as attention factors.

While groundwater level *(H)* and infiltration rate (I_R) showed some sensitivity in the model, their impact on slope stability, particularly the influence of „water" in the context, is not as significant as expected.

Hence, their exact implications require further investigation, and they will be retained as attention factors for subsequent stages of research.

On the other hand, soil moisture density (ρ) and Young's modulus (*E*), which were found to be insensitive in this stage, may be excluded from the attention factors in sensitivity analysis during the subsequent phases.

4.4 Further discussion on the groundwater level in the numerical model

This study examines the sensitivity of groundwater levels in slope safety evaluation, essentially altering the soil submerged range. The changes in groundwater level arises from alterations in excess pore water pressure, leading to a reduction in effective stress and, consequently, triggering slope instability. The material properties adopted in this study are single point two-phase. Specifically, pore water and the acceleration of soil is set as unknowns to solve the governing equations. Ultimately, this calculation yields the physical behavior of saturated soil (Anura3D MPM Research Community, 2022b). The effectiveness of analyzing the impact of groundwater level changes on landslides using the MPM has been validated through the study by Troncone et al., 2020.

Our founding here is that the sensitivity of groundwater level is relatively smaller than soil strength parameters. However, this doesn't imply its insignificance to landslides. Since the Mohr-Coulomb criterion directly relies on soil strength parameters, it is reasonable for these two parameters to exhibit the highest sensitivity. Besides these two parameters, groundwater level is the most significant factor affecting slope stability, where soil infiltration naturally also becomes an important factor influencing slope stability. Thus, in a situation where a slope undergoes rainfall infiltration without soil strength parameter changes, groundwater level demonstrates the highest sensitivity to slope stability.

4.5 Probability-based numerical model refinement directions

The initial phase of this research focused on simplifying the analysis of case study, primarily applicable to artificial slopes with a height of 3.6 meters or less. Caution is needed when extending this model to slopes exceeding this range. Hence, future phases should focus on making the numerical model's geometry more complex, resembling real-world scenarios or directly conducting motion analyses for specific slope conditions. Regarding parameter values, using actual soil experiment data from the analysis area is essential to realistically reflect the impact and sensitivity of various considerations on slope failure mechanisms.

5 CONCLUSIONS AND FUTURE WORKS

This study is the initial phase of the three-tiered objectives, concentrating on simpler scenarios to establish initial machine learning and probabilistic-numerical model. Based on these findings, the second phase should consider the following adjustments:

1. Enrich the factors in the landslide prediction machine learning model. In addition to the factors explored in this phase, such as rainfall of each event, slope and dip slope area proportion, incorporate more landslide-related geophysical and hydrological factors to enhance predictive performance.
2. Narrow down the research scope of the landslide prediction machine learning model for improved spatial data resolution.
3. Refine the probability-based numerical model by real-world slope geometry and soil parameters. In sensitivity analyses, emphasize discussions about the variation in soil strength parameters (C, ϕ) and the sensitivity of hydrological factors (H, I_R) within the system.

REFERENCES

Anura3D MPM Research Community. (2022a). *Anura3D MPM Software Tutorial Manual*, [online] Available at: https://github.com/Anura3D/Anura3D_OpenSource/tree/main/manuals, accessed: 02/02/2024.

Anura3D MPM Research Community. (2022b). *Anura3D MPM Software, Scientific Manual*, [online] Available at: https://github.com/Anura3D/Anura3D_OpenSource/blob/main/manuals/ScientificManual_2022.pdf, accessed: 02/02/2024.

Crozier, M.J. (2010) Deciphering the effect of climate change on landslide activity: A review, *Geomorphology*, 124(3–4), pp.260-267, https://doi.org/10.1016/j.geomorph.2010.04.009.

Kaur, A. and Sharma, R.K. (2016) Slope stability analysis techniques: A review. *International Journal of Engineering Applied Sciences and Technology*, 1(4), pp. 52-57, Available at: https://www.ijeast.com/papers/52-57,Teasma104,IJEAST.pdf , accessed: 02/02/2024.

Malkawi, A.I.H., Hassan, W.F and Abdulla, F.A. (2000) Uncertainty and reliability analysis applied to slope stability. *Structural Safety*, 22(2), pp. 161-187, https://doi.org/10.1016/S0167-4730(00)00006-0.

Merghadi, A., Yunus, A. P., Dou, J., Whiteley. J., ThaiPham, B., Bui, D. T., Avtar, R. and Abderrahmane, B. (2020) Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. *Earth-Science Reviews*, 207, 103225, https://doi.org/10.1016/j.earscirev.2020.103225

Soga, K., Alonso, E., Yerro, A., Kumar, K. and Bandara, S. (2016) Trends in large-deformation analysis of landslide mass movements with particular emphasis on the material point method. *Géotechnique*, 66(3), pp. 248-273, https://doi.org/10.1680/jgeot.15.LM.005

Thedy, J. and Liao, K. W. (2023) Adaptive Kriging adopting PSO with Hollow-Hypersphere space in structural reliability assessment. *Probabilistic Engineering Mechanics*, 74, 103513, https://doi.org/10.1016/j.probengmech.2023.103513

Troncone, A., Pugliese, L., and Conte, E. (2020) Run-Out Simulation of a Landslide Triggered by an increase in the Groundwater Level Using the Material Point Method. *Water*, 12(10), 2817, https://doi.org/10.3390/w12102817

Wyser, E., Alkhimenkov, Y., Jaboyedoff, M. and Podladchikov, Y.Y. (2020) A fast and efficient MATLAB-based MPM solver: fMPMM-solver v1.1. *Geoscientific Model Development*, 13(12), pp. 6265-6284, https://doi.org/10.5194/gmd-13-6265-2020