

AI-based simulation surrogates for planning rainfall-induced landslide mitigation

Michele Baldassini^{1*,†}, Francesco Pistolesi^{1,†}, Ignacio Giomi^{2,‡}, Evelina Volpe^{2,‡} and Elisabetta Cattoni^{2,‡}

¹University of Pisa-Department of Information Engineering, Largo Lazzarino, 1, Pisa (PI), 56122, Italy

²eCampus University-Department of Theoretical and Applied Sciences, Via Isimbardi, 10, Novedrate (CO), 22060, Italy

Abstract

Landslides triggered by intense or prolonged rainfall are a growing threat in the context of climate change. Physically based models are commonly used to estimate slope stability under varying hydrogeological conditions. However, their high computational cost may limit their use in time-sensitive risk assessment and decision-making scenarios. We propose a novel application of artificial intelligence (AI) for planning in safety-critical domains. In particular, our approach accelerates the prediction of key slope stability indicators—the factor of safety, the depth of the sliding surface, and the final position of the water table—under various rainfall events. Using a large dataset of more than 16,000 simulations, we train regression models that quickly approximate the results of complex numerical analyses across various slope geometries, soil properties, and hydrological conditions. The proposed models can perform rapid risk assessment for rainfall-induced landslides, exploring various rainfall scenarios and slope responses in tight timeframes. This helps select mitigation strategies that prioritize safety while considering reliability, feasibility constraints, and the likelihood of rainfall events. Our approach is thus key for complex, real-world AI planning in hydrogeological risk management, in particular considering the context of climate change, which will increase the number of sites requiring effective protective measures and rapid intervention in years to come.

Keywords

Artificial intelligence, climate change, decision support system, disaster resilience planning, slope stability.

1. Introduction

Landslides triggered by intense or prolonged rainfall are becoming more frequent and severe due to climate change. These events threaten infrastructure, communities, and ecosystems in many regions. Recent studies confirm that rainfall is one of the most common triggering agents [1]. Precipitation significantly affects slope behavior by progressively saturating the upper soil layers. This reduces the soil resistance and leads to the formation of a failure zone [2]. In addition, water infiltration following intense rainfall events can raise the phreatic level and trigger instability. The slope equilibrium is affected by the opposite roles of the resisting force and the driving force. The ratio between the resisting and driving forces is defined as the factor of safety (FoS): the slope is considered unstable when the FoS is lower than or equal to 1.

Physically based models (PBMs) can simulate a slope behavior under various hydrological conditions, and are widely used to support risk assessment and planning. Hydro-mechanical models can evaluate the changes in the FoS , considering the effects of rainfall infiltration and changes in pore water pressure [3, 4]. PBMs are grounded in well-understood physical laws and are often used in both local- and regional-scale studies [5].

Recent developments in slope stability modeling have used various computational techniques to enhance predictive reliability under complex conditions, especially in unsaturated soils. In particular, a

ECAI Workshop on AI-based Planning for Complex Real-World Applications, October 25, 2025, Bologna, IT

*Corresponding author.

[†]These authors contributed equally.

[‡]These authors contributed equally.

✉ michele.baldassini@ing.unipi.it (M. Baldassini); francesco.pistolesi@unipi.it (F. Pistolesi); ignacio.giomi@uniecampus.it (I. Giomi); evelina.volpe@uniecampus.it (E. Volpe); elisabetta.cattoni@uniecampus.it (E. Cattoni)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

probabilistic calibration method for coupled hydro-mechanical modeling was proposed to demonstrate that the use of multiple sensors (e.g., displacement, pore pressure, water content) improves parameter estimation under rainfall infiltration [6]. A Bayesian framework helped calibrate physical landslide models by combining spatially variable parameters via random fields and approximate Bayesian computation [7]. Bayesian networks were also used to update the soil parameters and susceptibility to landslide through spatial and cross correlations based on observed slope performance [8].

However, PBMs are generally computationally expensive. They require time-consuming setup processes, including the preparation of configuration files, the setting of many parameters, and the tuning of boundary conditions. These tasks require expert knowledge and can delay the use of PBMs in operational contexts [9]. Computational time becomes a major limitation, in particular when using PBMs in time-critical applications, such as emergency planning, or when there is a need to explore many options to select the best mitigation strategy for a given slope and certain rainfall scenarios that are highly likely to occur in the slope area.

To address these limitations, recent studies have explored artificial intelligence (AI) models trained on numerical simulation outputs. For example, tree-based and neural models have been tested to estimate landslide-related indicators under specific site conditions [10, 11]. Some works have introduced models that simplify slope stability simulations in various geotechnical contexts [12]. Combining physical modeling and AI is a promising direction for geohazard prediction [13]. For example, various AI models were trained on synthetic data from limit equilibrium simulations to predict the factor of safety [14]. Limit analysis was integrated with neural networks to evaluate slope stability of inhomogeneous soils [15]. This approach was then extended to 3D slope stability by training neural models on dimensionless parameters derived from slope charts, enabling rapid 3D/2D factor of safety comparisons [16]. Convolutional neural networks were also used as a surrogate for the random field finite element method to capture spatial soil variability while reducing computational costs [17]. An extensive 2D parametric study (+4,000 simulations) combined hydraulic, mechanical, and geometric variables to generate a comprehensive dataset for early warning systems [18]. Finally, a multi-approach algorithm using a simplified physically based model (X-Slip) was proposed for spatial and temporal prediction of soil slips over large areas [19].

Although these approaches are interesting, there is a need for fast and reliable AI solutions that can approximate the output of PBMs under a broad range of conditions. In particular, there are few methods that can simultaneously predict multiple indicators.

This paper proposes an AI-based model to predict the *FoS*, the depth of the sliding surface, and the position of the water table. The *FoS* evaluates slope stability, whereas the other two indicators help suggest effective slope stabilization measures. To train the model, we generated a dataset of more than 16,000 simulations based on various combinations of slope geometries, soil types, and rainfall events. Each sample of the dataset was obtained by associating each combination with the *FoS*, depth of the sliding surface, and the position of the water table, returned by GeoStudio, a geotechnical software to perform slope stability and seepage analyses (using PBMs).

The trained regression models achieved high precision compared to the results of PBMs, with near-instant response time. Our method can thus perform rapid AI-based planning in multiple rainfall scenarios. In particular, it can help decision-makers select mitigation strategies that consider safety, cost, and time of implementation. This multi-criteria approach is key in a changing climate, where the number of at-risk slopes is expected to grow and where quick intervention will often be required. The experiments showed that the method predicts the three target indicators with an average precision close to 94.3%. Our method is a starting point for developing a modular AI-based planning tool. In particular, it could be updated to include new slope geometries, soil characteristics, and rainfall events that map to an area of interest. The idea is to use our method as a module of a decision support system, where users can select the properties of the slope and the rainfall event, and then obtain the corresponding hydrogeological risk based on the factor of safety. We conclude the paper by showing that experts can be given the possibility to select a set of AI-planned mitigation guidelines, including information regarding reliability, feasibility, implementation, and typical cost, based on the risk obtained and on the predicted depth of the sliding surface and water table position. This helps authorities work alongside

experts to choose the best safety plan for a given risk scenario, using a fast and easy-to-use tool.

This work delivers the results of part of the European Project "*SAFE-LAND – Mitigating the risk of flooding and landslides via artificial intelligence with a view to extreme climate events*," grant number 101140345, co-funded by the European Commission.

The structure of the paper is as follows: Section 2 introduces the problem and the dataset; Section 3 describes the model; Section 4 presents two application examples; Section 5 draws the conclusions.

2. Simulations and Dataset

2.1. Slope stability: basic concepts and parameterization

A *slope* (natural or man-made) is an inclined surface of soil or rock exposed to gravity. The gravitational force tends to move the soil downward: if it becomes higher than the resisting force, slope failure (i.e., a landslide) occurs. Slope failure is based on soil type, groundwater location, seepage, and slope geometry.

The soil type is characterized by *mechanical* and *hydraulic properties*. The shear strength under saturated conditions—i.e., when soil voids are completely occupied by water—depends on the mechanical properties, such as the effective cohesion (c'), the effective friction angle (ϕ'), and soil unit weight (γ). In unsaturated conditions—i.e., when the soil voids are partially filled by air and water—the shear strength also depends on hydraulic properties, including suction (negative pore pressure), which contributes to an increase in resistance.

The *geometry* of a slope is characterized by the slope angle (angle of inclination to the horizontal), the slope height and length, the depth of the soil layer above the bedrock (rock layer under the soil), and the depth of the bedrock (Fig.1).

The stability of a slope depends on the ratio of the shear strength of the soil (characterized by mechanical parameters) and the shear strength developed along a potential failure surface. Various limit equilibrium methods can evaluate the stability of a slope [20, 21, 22, 23]). These methods discretize the potential sliding mass into slices. The main differences among these approaches lie in the specific equations of statics they satisfy (moment equilibrium and/or force equilibrium), and in the interslice forces they consider. Formally, a landslide occurs when the shear stress applied along a failure surface overcomes the soil shear strength. Considered a potential failure surface, the *FoS* is defined as follows:

$$FoS = \tau / \tau_m \quad (1)$$

where τ is the shear strength of the soil and τ_m is the mobilize shear strength.

The reduction in shear strength—for example, due to rainfall infiltration, weathering, excess pore pressure—can act as triggering factors of slope movements.

2.2. Parametric analyses

A dataset was generated by considering diverse rainfall events and multiple slope geometries, each characterized by specific configurations of mechanical and hydraulic parameters, to evaluate the slope response. Within the dataset, each *slope-rainfall* pair was associated with three parameters obtained via simulations implemented using physically based models: i) the factor of safety (*FoS*); ii) the depth of the sliding surface z_s ; iii) the final position of the water table z_w^{final} . Our dataset was made up of 16,019 samples.

The Morgenstern-Price [22] method was selected as the limit equilibrium method. The saturated shear strength was described by the Mohr-Coulomb model, defined as:

$$\tau = c' + \sigma'_n \tan(\phi') \quad (2)$$

where σ_n is the effective normal stress at the shear plane, a function of the soil unit weight. We analyzed how the mechanical parameters in drained conditions influence the slope stability by varying the effective cohesion (c'), effective friction angle (ϕ'), and soil unit weight (γ) whose mean values (μ),

Table 1

Values of the mechanical properties of the soil

Mechanical properties	μ	Range	Δ
Effective cohesion, c' [kPa]	20	0-40	10
Effective Friction Angle, ϕ' [°]	25	5-45	10
Soil Unit Weight γ [kN/m ³]	18	12-24	3

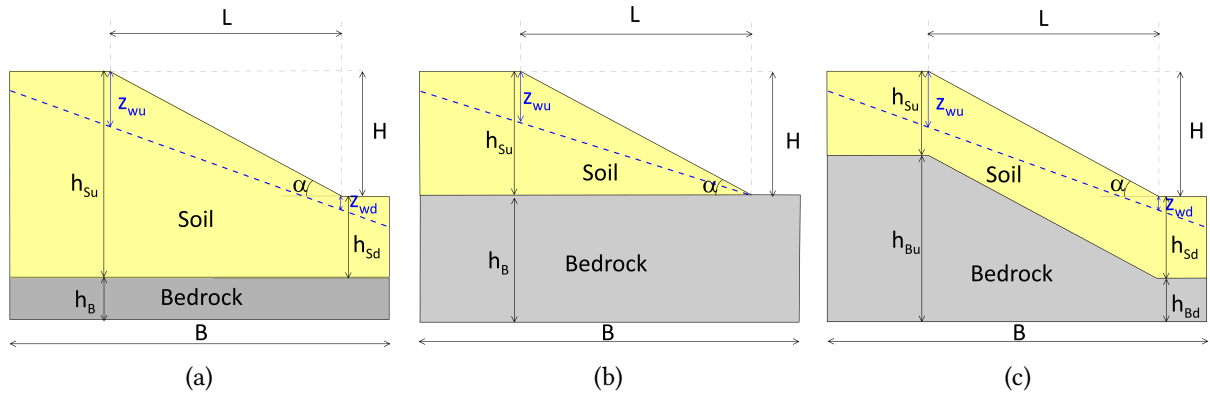
Table 2

Values of hydraulic soil properties

Soil type	K_{sat} [m/s]	van Genuchten (1980)		
		α [kPa ⁻¹]	n [-]	m [-]
1-High	1.00×10^{-4}	0.5	2	0.5
2-Medium	1.00×10^{-6}	0.08	1.7	0.412
3-Low	1.00×10^{-8}	0.02	1.5	0.333

ranges of variation, and step sizes (Δ) are in Table 1. These are representative values of the behavior of sandy, silty, and clayey soils [24, 25].

In *unsaturated soils*, the increase in resistance [26, 27] is crucial to evaluate the probability of landslide occurrence. For this reason, the increase in resistance due to suction was defined by the extended Mohr-Columb model proposed by Vanapalli et al. [28] by considering the change in the volumetric water content. The hydraulic parameters, namely the saturated hydraulic conductivity (K_s) and the van Genuchten parameters [29]—that describe the unsaturated behavior of the soil—are in Table 2. Three soil types were selected based on their value of K_s (*high*, *medium*, and *low*) corresponding to sandy, silty, and clayey soils, respectively. Fig. 1 shows different geometrical configurations of the slope that were used for generating the dataset used to train the AI models. The considered geometrical parameters were the slope angle (α), slope length (L), total length (B), slope height (H), total height- upstream (h_u), total height-downstream (h_d), soil depth-upstream (h_{Su}), soil depth-downstream (h_{Sd}), bedrock depth-upstream (h_{Bu}) and bedrock depth-downstream (h_{Bd}). Table 3 summarizes all the combinations. Three ratios between the soil strata and the bedrock were evaluated. The first two considered a horizontal bedrock with $h_{Su(90)} = 0.9 \cdot h_u$ and $h_{Sd(90)} = h_{Su(90)} - H$ (Fig. 1a), and $h_{Su(H)} = H$ and $h_{Sd(H)} = 0$ (Fig. 1b). The third case considered an inclined bedrock, where $h_{Su(25)} = 0.25 \cdot h_u$, and $h_{Sd(25)} = 0.25 \cdot h_d$ (Fig. 1c). These values are representative of typical failure mechanisms (*circular*, *toe*, and *base*).

**Figure 1:** Slope geometries.

2.3. Methodology

Slope stability was evaluated using GeoStudio, which enables time-dependent analysis by solving the water mass balance equation with the finite element method (FEM) in the SEEP/W module, followed by limit equilibrium analysis in the SLOPE/W module. Slope stability assessed through FEM-based models can capture complex failure mechanism. However, the Limit Equilibrium Method (LEM) is more computationally efficient for large-scale simulations and sensitivity studies [30, 31]. The SEEP/W module solved the mass balance equation by considering the water flow in time through saturated and unsaturated porous media. Different rainfall events and positions of the water table were considered to

Table 3

Values of geometrical parameters

α [°]	L [m]	B [m]	H [m]	h_u [m]	h_d [m]	h_{Su} [m]			h_{Sd} [m]		
						$h_{Su(90)}$	$h_{Su(25)}$	$h_{Su(H)}$	$h_{Sd(90)}$	$h_{Sd(25)}$	$h_{Sd(H)}$
20	20	100	7.3	25	17.7	22.5	6.3	7.3	15.2	4.4	0.0
	40	200	14.6	45	30.4	40.5	11.3	14.6	25.9	7.6	0.0
	80	400	29.1	90	60.9	81.0	22.5	29.1	51.9	15.2	0.0
30	20	100	11.5	35	23.5	31.5	8.8	11.5	20.0	5.9	0.0
	40	200	23.1	70	46.9	63.0	17.5	23.1	39.9	11.7	0.0
	80	400	46.2	140	93.8	126.0	35.0	46.2	79.8	23.5	0.0
40	20	100	16.8	50	33.2	45.0	12.5	16.8	28.2	8.3	0.0
	40	200	33.6	100	66.4	90.0	25.0	33.6	56.4	16.6	0.0
	80	400	67.1	200	132.9	180.0	50.0	67.1	112.9	33.2	0.0
50	20	100	23.8	70	46.2	63.0	17.5	23.8	39.2	11.5	0.0
	40	200	47.7	145	97.3	130.5	36.3	47.7	82.8	24.3	0.0
	80	400	95.3	290	194.7	261.0	72.5	95.3	165.7	48.7	0.0

define changes in the volumetric water content. Rainfall-induced instability was modeled by applying surface water fluxes ($\text{mm}^3/\text{h}/\text{mm}^2$) at the ground level, simulating low, medium, and high rainfall intensities (30-, 200-, and 500-year return periods). Table 4 presents the total 100-hour accumulated precipitation for each case, distributed using Chicago hyetographs with a central peak. Three groundwater table scenarios were considered: (1) high water table (z_{w1}^{init}), matching the ground surface; (2) low water table (z_{w2}^{init}), aligned with the bedrock surface; (3) intermediate water table (z_{w3}^{init}), located between the ground surface and the bedrock.

Table 4

Rainfall events

Intensity i	Return Period T_r [years]	Accumulated precipitation, h_v [mm]
i_1	30	315.90
i_2	200	480.50
i_3	500	586.00

The dataset considers 36 combinations of geometrical parameters (Table 3), 3 types of soil based on hydraulic characteristics ($K_{sat,1}$, $K_{sat,2}$, and $K_{sat,3}$), as reported in Table 2, 3 rainfall events (i_1, i_2 , and i_3) and a base scenario i_0 without precipitation (Table 4), 3 initial positions of the water table (z_{w1}^{init} , z_{w2}^{init} , and z_{w3}^{init}), and 13 combinations of mechanical parameter (Table 1). Table 5 summarizes the combinations of parameters for each row of Table 3. The outputs of each simulation were the factor of safety (FoS), the depth of the sliding surface (z_s), and the final position of the water table after the precipitation event (z_w^{final}). The FoS leads to the risk assessment (slope unstable if $FoS \leq 1$), whereas z_s and z_w^{final} help suggest stabilization measures.

3. Predicting Landslide with Artificial Intelligence

This section describes the training process of three regression models to predict the Factor of Safety (FoS), the depth of the sliding surface [m] (z_s), and the final position (i.e., depth from the ground surface) of the water table [m] (z_w^{final}). The maximum values of z_s and z_w^{final} were selected in order to define the mitigation measure. Then, the section discusses the importance of each feature to the predictions of each model.

Table 5

Parameter combination for the dataset

Geometry	h_{Su}	z_w^{init}	K_{sat}	Intensity, i	Simulation ID
α, L, B, H, h_u	$h_{Su}(i)$	z_{w1}^{init}	$K_{sat,1}$	i_0	1-13
			$K_{sat,2}$		
			$K_{sat,3}$		
α, L, B, H, h_u	$h_{Su}(i)$	z_{w2}^{init}	$K_{sat,1}$	i_0, i_1, i_2, i_3	14-65
			$K_{sat,2}$		65-117
			$K_{sat,3}$		118-169
α, L, B, H, h_u	$h_{Su}(i)$	z_{w3}^{init}	$K_{sat,1}$	i_0, i_1, i_2, i_3	170-221
			$K_{sat,2}$		222-273
			$K_{sat,3}$		274-325

3.1. Data Preparation

The dataset was partitioned into two mutually exclusive subsets: a training set comprising 70% of the data (11,213 samples), and a hold-out test set containing the remaining 30% (4,806 samples). The split was performed using a stratified random sampling to preserve the distribution of the target variable in both subsets. All regression models were evaluated on the hold-out test samples, which were not used during the training. This procedure ensured that model performance reflects the ability to generalize to unseen cases.

The parameters describing the slope rainfall interaction were preprocessed to extract the features used as inputs to the three regression models. Features were normalized using the z -score normalization to obtain zero mean and unit standard deviation. Normalization parameters (mean and standard deviation) were computed from the training data to prevent data leakage and ensure unbiased evaluation on the test set [32].

3.2. Feature Selection

To find the most relevant input parameters (*features*) for the regression models, we used the Sequential Feature Selection (SFS) algorithm [33]. SFS iteratively adds features that minimize a predefined performance criterion. Decision Trees (DTs) [34] were chosen as regression models for their ability to capture non-linear relationships and their interpretability, which is essential in engineering applications requiring transparent and explainable models. The DT is a supervised learning algorithm that recursively partitions the input space based on feature values. At each node, the algorithm selects the optimal feature and threshold to split the data, using a criterion as the Gini index. The splitting process continues recursively until a stopping criterion is satisfied, such as reaching a maximum tree depth. The resulting tree structure consists of internal nodes representing decision rules and leaf nodes that provide the final prediction. In regression tasks, the prediction at each leaf corresponds to the average value of the target variable in that subset.

The predictive contribution of each feature was evaluated using a 10-fold cross-validation scheme. Each fold comprised about 1,122 samples, with models trained on nine folds and validated on the remaining one. The mean R^2 computed across the test folds was used as a performance criterion. To ensure robustness, the SFS was repeated 30 times with different random seeds, and the set of selected features was obtained by averaging the results over all repetitions.

The most frequently selected features were as follows: z_{wu}^{init} , h_{Bd} , ϕ' , c' for predicting the FoS and the z_s , z_{wu}^{init} , h_{Bd} , k_{sat} , T_r for predicting the z_w^{final} . These features were then used to train the DTs.

3.3. Train and Test

Three DTs regressors were trained to predict the three target variables (i.e., z_s , z_s , z_w^{final}), all following the same hyperparameter optimization procedure. The hyperparameters considered were the maximum

Table 6Mean and standard deviation of R^2 scores over the 10-fold cross-validation.

	1	2	3	4	5	6	7	8	9	10
FoS	0.940 ± 0.015	0.935 ± 0.017	0.943 ± 0.012	0.930 ± 0.019	0.945 ± 0.016	0.936 ± 0.014	0.939 ± 0.013	0.931 ± 0.020	0.946 ± 0.017	0.937 ± 0.015
z_s	0.887 ± 0.026	0.894 ± 0.020	0.880 ± 0.023	0.889 ± 0.027	0.898 ± 0.021	0.882 ± 0.025	0.895 ± 0.018	0.884 ± 0.022	0.893 ± 0.020	0.890 ± 0.021
z_w^{final}	0.999 ± 0.004	1.000 ± 0.002	0.998 ± 0.005	1.000 ± 0.002	0.999 ± 0.004	1.000 ± 0.002	0.998 ± 0.004	0.999 ± 0.003	1.000 ± 0.002	1.000 ± 0.001
	11	12	13	14	15	16	17	18	19	20
FoS	0.938 ± 0.013	0.942 ± 0.016	0.934 ± 0.015	0.933 ± 0.016	0.940 ± 0.014	0.932 ± 0.018	0.944 ± 0.016	0.936 ± 0.013	0.937 ± 0.015	0.939 ± 0.014
z_s	0.886 ± 0.023	0.891 ± 0.022	0.883 ± 0.020	0.889 ± 0.023	0.887 ± 0.025	0.893 ± 0.017	0.885 ± 0.024	0.888 ± 0.020	0.890 ± 0.021	0.884 ± 0.020
z_w^{final}	0.999 ± 0.004	1.000 ± 0.002	0.999 ± 0.003	1.000 ± 0.002	1.000 ± 0.002	0.999 ± 0.003	1.000 ± 0.002	1.000 ± 0.001	0.999 ± 0.002	0.999 ± 0.002
	21	22	23	24	25	26	27	28	29	30
FoS	0.935 ± 0.016	0.943 ± 0.017	0.931 ± 0.018	0.938 ± 0.014	0.936 ± 0.016	0.934 ± 0.015	0.940 ± 0.015	0.937 ± 0.014	0.939 ± 0.015	0.935 ± 0.016
z_s	0.896 ± 0.018	0.882 ± 0.024	0.892 ± 0.021	0.889 ± 0.021	0.885 ± 0.023	0.887 ± 0.021	0.890 ± 0.019	0.883 ± 0.020	0.894 ± 0.023	0.888 ± 0.020
z_w^{final}	1.000 ± 0.003	0.998 ± 0.004	1.000 ± 0.002	1.000 ± 0.002	0.999 ± 0.004	0.998 ± 0.005	1.000 ± 0.001	1.000 ± 0.002	0.999 ± 0.004	1.000 ± 0.002

tree depth d , maximum number of splits s , minimum number of leaf nodes l , and number of features f . These hyperparameters were optimized via a grid search, evaluating multiple combinations to maximize model performance in terms of the R^2 score. The search space included the following ranges: $d \in \{5, \dots, 15\}$, $s \in \{3, \dots, 8\}$, $l \in \{3, \dots, 8\}$, and $f \in \{1, \dots, 6\}$. For each hyperparameter combination (d, s, l, f) , we performed 30 independent training sessions using the 10-fold cross-validation. Model performance was assessed using the mean and standard deviation of the R^2 score across folds.

The configurations that achieved the highest mean R^2 across all runs were: $d = 16$, $s = 9$, $l = 7$, and $f = 4$ for predicting the FoS, $d = 11$, $s = 7$, $l = 5$, and $f = 3$ for predicting z_s , $d = 8$, $s = 5$, $l = 3$, and $f = 2$ for predicting z_w^{final} .

Table 6 summarizes the mean and standard deviation of the R^2 obtained from each run of 10-fold cross-validation. The performance is shown separately for the three DTs predicting the FoS, the z_s , and the z_w^{final} . The results confirmed the stability and robustness of the models, with FoS and z_s achieving high R^2 values (averaging around 0.94 and 0.89, respectively), and z_w^{final} consistently reaching highly accurate scores across all runs.

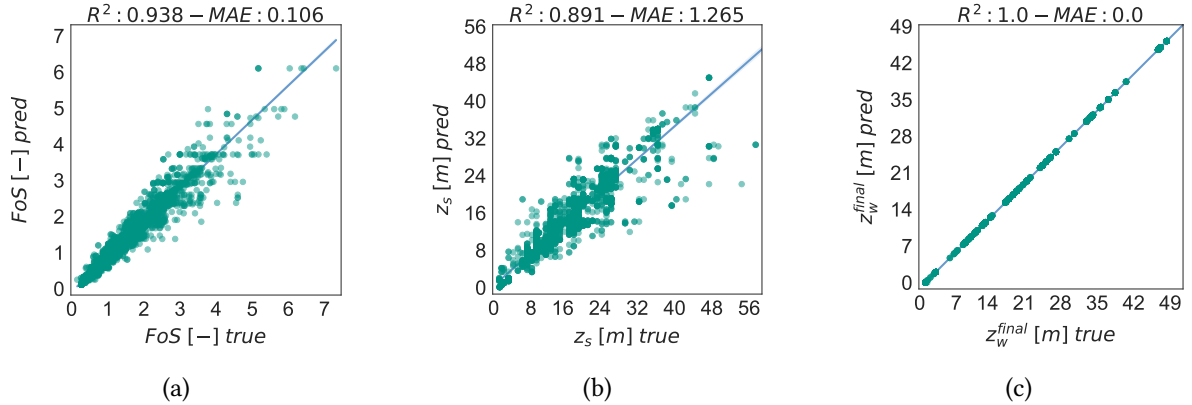


Figure 2: Regression plot of the Factor of Safety (FoS) (a), the depth of sliding surface [m] (z_s) (b), and the maximum final piezometric surface depth [m] (z_w^{final}) (c).

3.4. Model Evaluation

After finding the best architecture for each DT model, we trained each regressors using the entire training set. Then, we tested the resulting three regressors using the samples in the hold-out test set.

To visually examine the error, Fig. 2 shows the regression plots regarding the FoS (a), z_s (b), and z_w^{final} (c). Each sample is represented by a green dot whose coordinates represent the values of FoS , z_s , and z_w^{final} (abscissa)—obtained by GeoStudio—and the corresponding predicted values (ordinate). Each plot in Fig. 2 also shows the regression line (in blue) determined using the least squares method.

The model achieved high accuracy for all outputs to predict (targets), with R^2 of 0.938, 0.891, and 1.0 for FoS, z_s , and z_w^{final} , respectively. The FoS model shows a point distribution close to the regression

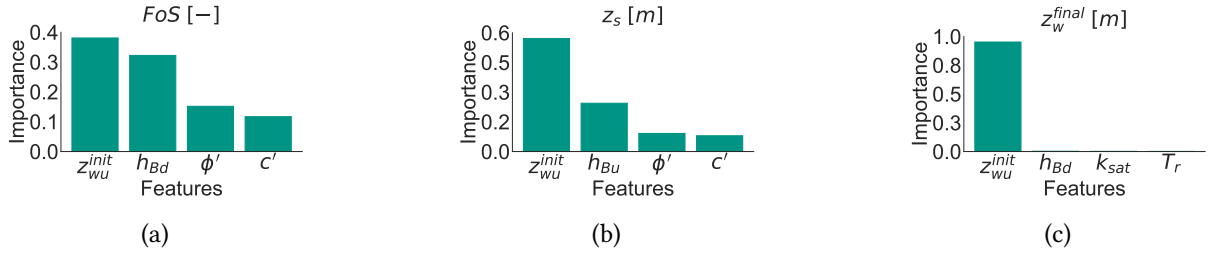


Figure 3: Feature importance plot of the Factor of Safety (FoS) (a), the depth of sliding surface [m] (z_s) (b), and the maximum final piezometric surface depth [m] (z_w^{final}) (c).

line, with minimal dispersion and a mean absolute error (MAE) of 0.106. Predictions for z_s also shows a similar trend, but with slightly higher variance, particularly for larger depth values. The z_w^{final} regressor is the most precise, with a high fidelity to the values obtained by GeoStudio.

Fig. 3 shows three histograms that represent the feature importance for each regression model. For all targets, the initial elevation of the water table (z_{wu}^{init}) was the most influential predictor, highlighting both its key role in slope stability and the importance of considering the unsaturated behavior.

In the prediction of FoS (Fig. 3a), both z_{wu}^{init} and the bedrock depth (h_{Bd}) contributed significantly, whereas effective friction angle (ϕ') and cohesion (c') played a secondary role, in line with geotechnical expectations. Regarding the depth of the sliding surface z_s , Fig. 3b shows that z_{wu}^{init} dominates again, followed by the bedrock depth-upstream (h_{Bu}), whereas mechanical parameters had a minor influence.

Finally, regarding the prediction of z_w^{final} , Fig. 3c highlights that z_{wu}^{init} was the most important only. This confirms that the final position of the phreatic level is highly influenced by its initial position (z_{wu}^{init}), and marginally by rainfall.

These results are in line with the consolidated strong dependence of the outputs on the hydro-mechanical parameters.

4. Application examples

Our system can be used to assess the hazard conditions of a slope and suggest appropriate mitigation measures. This enables AI planning for decision-making in slope stabilization.

For unstable slopes ($FoS \leq 1$), the system helps select the most suitable stabilization intervention mainly based on the predicted values of z_s and z_w^{final} . This selection can be made using predefined *effectiveness matrices* each suggesting a possible stabilization technique: piles, diaphragm walls, soil nailing, and strand anchors (see Tables 7 to 10).

For each technique, the corresponding matrix quantifies the degree of mitigation by using an *effectiveness score* based on the interaction between the depths of the sliding surface and the piezometric level (water table). Effectiveness scores are discretized as follows:

- 1 = highly effective (green)
- 0.5 = quite effective (orange)
- 0.25 = moderately effective (yellow)
- 0 = ineffective (white).

When multiple stabilization measures achieve the same score, the selection requires inspecting the *applicability matrix* (see Tab. 11). This matrix considers additional practical aspects (reliability, feasibility, ease of implementation, and indicative cost) by assigning each intervention an applicability score (S):

- $S > 3$ (green): highly recommended measure
- $2 < S \leq 3$ (orange): suggested measure
- $1 < S \leq 2$ (yellow): less suitable measure
- $S < 1$ (white): not recommended measure.

To demonstrate how the system integrates predictive modeling and AI planning, we discuss two examples involving two representative unstable slopes. As a first example, we consider an unstable slope whose indicators returned by the AI models were $FoS = 0.84$, depth of the sliding surface $z_s = 13.5$ m, and depth of the phreatic level (z_w^{final}) of 13.5 m (low depth of the water level). With reference to the effectiveness matrices of Tab. 7 to 10, interventions DW and SA are associated with an effectiveness index equal to 1: they are both highly effective. Based on the applicability matrix (see Tab. 11), DW is characterized by a higher applicability score than SA (3 vs 2): diaphragm walls are thus preferred. As a second example, consider an unstable slope characterized by $FoS = 0.95$, with $z_s = 5$ m and $z_w^{final} = 23$ m. The effectiveness matrices of Tab. 7 to 10 identify piles as the most effective measure, thereby allowing us to propose this stabilization measure without referring to the applicability matrix.

These two application examples show that our system can easily suggest stabilization planning measures based on quantitative risk indicators obtained by the AI models. The integration with stabilization effectiveness and applicability matrices helps select the most appropriate mitigation strategies, balancing technical performance, feasibility, implementation effort, and economic impact.

Potential applications of the system include AI planning in the early design phases to quickly compare multiple alternative stabilization measures, considering many hydrogeological scenarios. The system could also be used in monitoring frameworks, where collections of updated input data could be used to dynamically reassess slope stability and recommend stabilization measures as conditions evolve.

Table 7
Effectiveness matrix of PILES (P)

PILES (P)			Depth of the piezometric level		
			High	Low	Absent
			0.5	1	1
Depth of sliding surface	Superficial (<1.0 m)	0	0	0	0
	Shallow (1 to 3 m)	0.5	0.25	0.5	0.5
	Medium (3 to 8 m)	1	0.5	1	1
	Deep (8 to 15 m)	0.5	0.25	0.5	0.5
	Very deep (>15 m)	0	0	0	0

Table 8
Effectiveness matrix of DIAPHRAGM WALLS (DW)

DIAPHRAGM WALLS (DW)			Depth of the piezometric level		
			High	Low	Absent
			0.5	1	1
Depth of sliding surface	Superficial (<1.0 m)	0	0	0	0
	Shallow (1 to 3 m)	0	0	0	0
	Medium (3 to 8 m)	0.5	0.25	0.5	0.5
	Deep (8 to 15 m)	1	0.5	1	1
	Very deep (>15 m)	0.5	0.25	0.5	0.5

Table 9
Effectiveness matrix of SOIL NAILING (SN)

SOIL NAILING (SN)			Depth of the piezometric level		
			High	Low	Absent
			0.5	1	1
Depth of sliding surface	Superficial (<1.0 m)	1	0	0.5	1
	Shallow (1 to 3 m)	1	0	0.5	1
	Medium (3 to 8 m)	0.5	0	0.25	0.5
	Deep (8 to 15 m)	0	0	0	0
	Very deep (>15 m)	0	0	0	0

Table 10
Effectiveness matrix of STRAND ANCHORS (SA)

STRAND ANCHORS (SA)			Depth of the piezometric level		
			High	Low	Absent
			0.5	1	1
Depth of sliding surface	Superficial (<1.0 m)	0	0	0	0
	Shallow (1 to 3 m)	0	0	0	0
	Medium (3 to 8 m)	0.5	0.25	0.5	0.5
	Deep (8 to 15 m)	1	0.5	1	1
	Very deep (>15 m)	0.5	0.25	0.5	0.5

Table 11
Example of applicability matrix for stabilization measures

	P	DW	SN	SA
Reliability	1	1	0.5	0.5
Feasibility	1	1	0.5	0.5
Implementation	0.5	0.5	0.5	0.5
Typical cost	0.5	0.5	0.5	0.5
Total score (S)	3	3	2	2

5. Conclusions

This paper has presented an AI-based model for multi-criteria AI planning. Our system accelerates the assessment of rainfall-induced landslide risk and helps select effective measures based on reliability,

feasibility, ease of implementation, and cost. Trained using a dataset +16,000 physically based simulations, our model quickly predicts key indicators of slope response to rainfall—the factor of safety, the depth of the sliding surface, and the position of the water table. The ability to rapidly explore a wide range of rainfall scenarios and slope responses can help evaluate multiple mitigation strategies, thereby supporting real-world planning and decision-making in hydrogeological risk management.

Our results show that AI can complement physically based simulations, bridging the gap between high-fidelity modeling and operational usability. Future work will focus on extending the generalization capabilities to more complex scenarios, including uncertainty quantification, and exploring collaboration with optimization and planning algorithms for automated decision support.

6. Acknowledgments

This work is part of the European project “SAFE-LAND – Mitigating the risk of flooding and landslides via artificial intelligence with a view to extreme climate events”, grant n. 101140345, co-funded by the European Commission. The authors acknowledge the Directorate-General for European Civil Protection and Humanitarian Aid Operations (ECHO), ECHO.B – Disaster Preparedness and Prevention.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] A. C. Mondini, F. Guzzetti, M. Melillo, A. Pievatolo, Short to long term space-time prediction of rain-induced landslides under uncertainty, *Science of The Total Environment* 984 (2025) 179453.
- [2] P. Sitarenios, F. Casini, A. Askarinejad, S. Springman, Hydro-mechanical analysis of a surficial landslide triggered by artificial rainfall: the ruedlingen field experiment, *Géotechnique* 71 (2021) 96–109.
- [3] Y. Li, X. Hu, H. Zhang, H. Zheng, N. Li, Displacement prediction and failure mechanism analysis of rainfall-induced colluvial landslides, *Journal of Hydrology* 660 (2025) 133361.
- [4] M. P. A. Gatto, S. Misiano, L. Montrasio, Space-time prediction of rainfall-induced shallow landslides through artificial neural networks in comparison with the slip model, *Engineering Geology* 344 (2025) 107822.
- [5] S.-H. Jiang, H.-H. Jie, J. Xie, J. Huang, C.-B. Zhou, Probabilistic back-analysis of rainfall-induced landslides for slope reliability prediction with multi-source information, *Journal of Rock Mechanics and Geotechnical Engineering* 16 (2024) 3575–3594.
- [6] L. Zhang, F. Wu, Y. Zheng, L. Chen, J. Zhang, X. Li, Probabilistic calibration of a coupled hydro-mechanical slope stability model with integration of multiple observations, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 12 (2018) 169–182.
- [7] I. Depina, E. A. Oguz, V. Thakur, Novel bayesian framework for calibration of spatially distributed physical-based landslide prediction models, *Computers and Geotechnics* 125 (2020) 103660.
- [8] H. Wang, T. Xiao, X. Li, L. Zhang, L. Zhang, A novel physically-based model for updating landslide susceptibility, *Engineering Geology* 251 (2019) 71–80.
- [9] A. Troncone, L. Pugliese, G. Lamanna, E. Conte, Prediction of rainfall-induced landslide movements in the presence of stabilizing piles, *Engineering Geology* 288 (2021) 106143.
- [10] H. Wen, F. Yan, J. Huang, Y. Li, Interpretable machine learning models and decision-making mechanisms for landslide hazard assessment under different rainfall conditions, *Expert Systems with Applications* 270 (2025) 126582.
- [11] H.-Z. Cui, B. Tong, T. Wang, J. Dou, J. Ji, A hybrid data-driven approach for rainfall-induced landslide susceptibility mapping: Physically-based probabilistic model with convolutional neural network, *Journal of Rock Mechanics and Geotechnical Engineering* (2024).

- [12] H. Cui, V. Medina, M. Hürlimann, J. Ji, Fast physically-based probabilistic modelling of rainfall-induced shallow landslide susceptibility at the regional scale considering geotechnical uncertainties and different hydrological conditions, *Computers and Geotechnics* 172 (2024) 106400.
- [13] W. Gong, S. Zhang, C. H. Juang, H. Tang, S. P. Pudasaini, Displacement prediction of landslides at slope-scale: Review of physics-based and data-driven approaches, *Earth-Science Reviews* 258 (2024) 104948.
- [14] D. Tien Bui, H. Moayedi, M. Gör, A. Jaafari, L. K. Foong, Predicting slope stability failure through machine learning paradigms, *ISPRS International Journal of Geo-Information* 8 (2019) 395.
- [15] Z. Qian, A.-J. Li, W. Chen, A. Lyamin, J. Jiang, An artificial neural network approach to inhomogeneous soil slope stability predictions based on limit analysis methods, *Soils and foundations* 59 (2019) 556–569.
- [16] J. Meng, H. Mattsson, J. Laue, Three-dimensional slope stability predictions using artificial neural networks, *International Journal for Numerical and Analytical Methods in Geomechanics* 45 (2021) 1988–2000.
- [17] Z.-Z. Wang, S. H. Goh, Novel approach to efficient slope reliability analysis in spatially variable soils, *Engineering Geology* 281 (2021) 105989.
- [18] S. A. Espinosa F, M. H. El Naggar, Parametric study of rainfall-induced instability in fine-grained sandy soil, *Geotechnics* 4 (2024) 1159–1174.
- [19] M. P. A. Gatto, L. Montrasio, X-slip: A slip-based multi-approach algorithm to predict the spatial-temporal triggering of rainfall-induced shallow landslides over large areas, *Computers and Geotechnics* 154 (2023) 105175.
- [20] W. Fellenius, Calculation of the stability of earth dams, in: *Proc. of the second congress on large dams*, volume 4, 1936, pp. 445–463.
- [21] A. W. Bishop, The use of the slip circle in the stability analysis of slopes, *Geotechnique* 5 (1955) 7–17.
- [22] N. R. U. Morgenstern, V. E. Price, The analysis of the stability of general slip surfaces, *Geotechnique* 15 (1965) 79–93.
- [23] E. Spencer, A method of analysis of the stability of embankments assuming parallel inter-slice forces, *Geotechnique* 17 (1967) 11–26.
- [24] J. K. Mitchell, K. Soga, *Fundamentals of soil behavior*, volume 3, John Wiley & Sons, 2005.
- [25] M. Budhu, *Soil mechanics fundamentals*, John Wiley & Sons, 2015.
- [26] E. E. Alonso, A. Gens, A. Josa, A constitutive model for partially saturated soils, *Géotechnique* 40 (1990) 405–430.
- [27] D. G. Fredlund, H. Rahardjo, *Soil mechanics for unsaturated soils*, John Wiley & Sons, 1993.
- [28] S. K. Vanapalli, D. G. Fredlund, D. E. Pufahl, A. W. Clifton, Model for the prediction of shear strength with respect to soil suction, *Canadian Geotechnical Journal* 33 (1996) 379–392.
- [29] M. T. V. Genuchten, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, *Soil Science Society of America Journal* 44 (1980) 892–898.
- [30] D. Q. Li, T. Xiao, Z. J. Cao, K. K. Phoon, C. B. Zhou, Efficient and consistent reliability analysis of soil slope stability using both limit equilibrium analysis and finite element analysis, *Applied Mathematical Modelling* 40 (2016) 5216–5229.
- [31] F. Vahedifard, D. Leshchinsky, K. Mortezaei, N. Lu, Effective stress-based limit-equilibrium analysis for homogeneous unsaturated slopes, *International Journal of Geomechanics* 16 (2016) D4016003.
- [32] P. Harrington, *Machine Learning in Action*, Manning Publications Co., 2012.
- [33] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.* 3 (2003) 1157–1182.
- [34] J. R. Quinlan, Induction of decision trees, *Machine Learning* 1 (1986) 81–106.