Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Effect of Geotechnical Properties on Rainfall-Induced Landslide Probability: Infinite Slope Analysis and a Mechanics-Verified Hybrid Deep Learning Surrogate

Muhammed Yeasin Arafat

yeasinarafat1804177@gmail.com

Bangladesh University of Engineering and Technology

Research Article

Keywords: Rainfall-induced landslide, Infinite slope, Reliability analysis, Hybrid deep learning, Perched water table, Green–Ampt infiltration

Posted Date: November 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11429117/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Effect of Geotechnical Properties on Rainfall-Induced Landslide Probability: Infinite Slope Analysis and a Mechanics-Verified Hybrid Deep Learning Surrogate

Article type: Original Paper

Muhammed Yeasin Arafat

Department of Civil Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh.

yeasinarafat1804177@gmail.com

Abstract

Slopes are still judged on a single factor of safety (FoS), although two slopes with the same FoS can owe their strength to very different mixes of cohesion and friction. This study examines how much that matters in rain. Antecedent rain and a 24-h storm were routed into the slope with a slope-corrected Green–Ampt model; where the water perched above a less permeable layer, its pore pressure was applied slice by slice in a full-search Bishop analysis and on a planar slide along the layer. A physics-guided hybrid deep-learning surrogate (Hybrid-6), fitted to 21,000 Bishop solutions, made 3.3×10^7 stability evaluations practical after its sensitivities and failure probabilities were checked against direct Bishop runs. For nine slopes of identical dry FoS (1.50), a 300-mm storm after 300 mm of antecedent rain increased the failure probability of friction-rich slopes 4.3–4.6 times more than that of cohesion-rich slopes (bootstrap 95 % intervals within 4.0–5.0), and every rain-induced failure was translational. Applying the same water as a uniform pore-pressure ratio on the critical circle blurred this difference and reversed it at two of three slope angles. Within the conditions analysed, the planar FoS of the saturated layer described the storm-induced increase in failure probability ($R^2 = 0.90$, 54 cases), though less well for a withheld layer depth ($Q^2 = 0.57$). For the Rangamati (Bangladesh) landslides of 13 June 2017, the model limited the vulnerable slip depth to 1.4–2.1 m, and its predicted failure window contained the reported time of failure.

Keywords Rainfall-induced landslide; Infinite slope; Reliability analysis; Hybrid deep learning; Perched water table; Green–Ampt infiltration.

1 Introduction

Rainfall is the most frequent trigger of landslides on natural and engineered slopes. Water that infiltrates raises the pore pressure on a potential slip surface, lowers the effective normal stress and so removes part of the frictional shear strength (Iverson 2000). How strongly a given storm destabilises a slope depends on its geotechnical properties, namely the effective cohesion c', friction angle φ', unit weight γ, slope geometry and the soil's capacity to take in water, and on the uncertainty in each of them (Christian et al. 1994; Phoon and Kulhawy 1999). In practice, a slope is still accepted or rejected on one deterministic factor of safety (FoS). Yet two slopes with the same FoS can hold it through very different mixes of cohesion and friction, and pore pressure acts only on the frictional part. This paper asks how geotechnical properties and their uncertainty control the probability of rainfall-induced failure, and how much of that probability a single dry FoS hides.

1.1 Previous work

Rain fall infiltration and slope stability. Physically based models couple infiltration with the infinite-slope equation, from steady-state hydrology (Montgomery and Dietrich 1994) to transient pressure diffusion (Iverson 2000; Baum et al. 2008) and Green–Ampt wetting fronts (Green and Ampt 1911), which Chen and Young (2006) extended to sloping surfaces. Studies in this journal have established the roles of antecedent rainfall (Rahardjo et al. 2001), rainfall intensity, pattern and hydraulic conductivity (Huat et al. 2006; Xue and Gavin 2008; Lin and Zhong 2019; Paul et al. 2025) and soil saturation (Montrasio et al. 2018), and have refined the Green–Ampt model for progressive failure (Chen et al. 2021) and for non-uniform initial moisture (Zhu et al. 2026). Critical rainfall curves (Kawajiri et al. 2022) and empirical intensity–duration thresholds (Caine 1980; Guzzetti et al. 2008) summarise these processes for warning. Most of these studies are deterministic. They show when a given slope fails, but not how the failure probability of slopes with the same FoS depends on what carries their strength.

Probabilistic slope stability. Reliability analysis replaces the FoS with a probability of failure P_f that reflects property uncertainty (Christian et al. 1994; Duncan 2000; Baecher and Christian 2003), and slopes with similar FoS can differ widely in P_f (El-Ramly et al. 2002). Spatial averaging and random fields change both P_f and the failure mechanism (Vanmarcke 1977; Griffiths and Fenton 2004; Griffiths et al. 2009). Under rainfall, spatially variable permeability changes the timing and probability of failure (Santoso et al. 2011; Cho 2014). Recent work in this journal has used radially degraded random fields for rainfall-induced slope reliability (Wang et al. 2024) and Bayesian Markov chain Monte Carlo analysis of slope stability (Fattahi and Zandy Ilghani 2020). Because each analysis needs thousands to millions of stability evaluations, surrogates such as response surfaces (Li et al. 2015; Santos et al. 2025) and Gaussian-process regression (Samui et al. 2019) are widely used.

Machine learning for slope stability. Neural networks and tree ensembles trained on case-history or generated databases now predict the FoS or the stability class with high reported accuracy (Sakellariou and Ferentinou 2005; Zhou et al. 2019; Nouri et al. 2021; Marrapu et al. 2021). Recent reviews in this journal survey machine-learning and deep-learning applications to slope stability (Arif et al. 2025; Li et al. 2025) and coupled and hybrid frameworks for rainfall-induced slope failure (Viswanath et al. 2026). Phoon and Zhang (2023) pointed to recurring weaknesses of such models: data of unclear provenance, validation by accuracy alone, and little use for risk. Physics-guided learning, which constrains a model's response to known mechanics, is emerging as a remedy (Karniadakis et al. 2021; Rathor and Khandelwal 2026). It has rarely been checked, however, against the sensitivities and failure probabilities of the limit-equilibrium method it replaces.

1.2 Research gap and contributions

Three gaps follow. (i) Rainfall studies rarely compare slopes of the same FoS but different strength composition, so it is not known whether, and by how much, the cohesion–friction mix

changes rainfall risk. (ii) Probabilistic studies often convert infiltration into a uniform pore-pressure ratio along the critical dry circle, although the water collects in a band above a less permeable layer; the mechanism that fails in rain may then not be the one that governs the dry design. (iii) Machine-learning surrogates are seldom verified in the quantities a reliability analysis uses: sensitivities, critical states and failure probabilities.

The paper makes four contributions:

- A mechanics-verified hybrid deep-learning surrogate (Hybrid-6) of the simplified Bishop method. Six experts share a differentiable infinite-slope layer and a mechanics sign loss, and a second learned term adds an explicit perched-water pore-pressure field. The surrogate is used only where its sensitivities, critical pore-pressure ratios and Monte Carlo failure probabilities agree with direct Bishop solutions on the same samples, and it makes 3.3×10^7 evaluations affordable.
- A controlled comparison of slopes with equal dry FoS but cohesion-rich, balanced and friction-rich strength, under 5-day antecedent rain and a 24-h storm, with the pore pressure applied where it acts and both circular and translational mechanisms checked. The comparison quantifies how much the strength mix changes rainfall risk, shows that a uniform pore-pressure ratio can reverse the result, and ranks which site data are worth collecting.
- A saturated-layer screening relation for the conditions analysed. The planar FoS at the depth of the less permeable layer, with the perched water table at the ground surface, describes the storm-induced increase in P_f across 54 cases and ranks slopes that share the same dry system FoS. It is a descriptive, interpolative relation, not a general predictive model.
- A consistency check against a field event, the Rangamati (Bangladesh) landslides of 13 June 2017, where the framework excludes deep failure, bounds the vulnerable slip depth and produces a failure window that can be compared with the reported time of failure.

The study is limited to rainfall loading. Figure 1 sketches the problem and the sequence of the analysis.

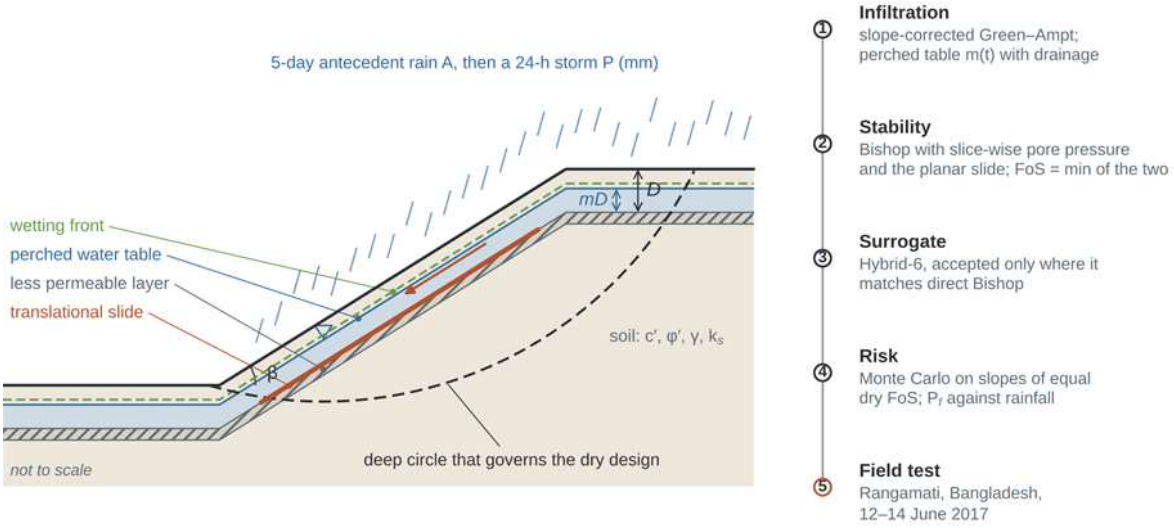


Fig. 1 Conceptual model of the study. Antecedent rain and a 24-h storm infiltrate the slope (slope-corrected Green–Ampt), the water perches above a less permeable layer at depth D and raises a water table of height mD, and stability is checked both on a deep circle (simplified Bishop with the slice-wise pore pressure) and on a translational slide along the layer. The numbered column gives the sequence of the analysis.

2 Materials and methods

2.1 Data

2.1.1 Slope stability dataset

The study uses a dataset of 120,000 slopes. They span height H 5–50 m, angle β 15–65°, c′ 5–120 kPa, φ′ 15–45°, γ 16–24 kN/m³ and pore-pressure ratio r_u 0.02–0.6. Each record also gives daily rainfall, groundwater and site descriptors, and a static FoS (0.25–5.0). The data are split 70 / 7.5 / 7.5 / 15 % (training, validation, calibration, test), stratified on FoS class. The 7,718 rows on the 0.25 floor or 5.0 ceiling are treated as censored, which leaves 112,282 uncensored rows.

The dataset's pore-pressure ratio is defined on the slope-normal overburden, $u = r_u \gamma z \cos^2\beta$. All models use Bishop's vertical ratio $r_u,v = u/(\gamma h) = r_u \cos^2\beta$. Rainfall, groundwater and site

descriptors add nothing once r_u is known (R^2 of ln FoS 0.9935 without them, 0.9934 with them), so daily rainfall is used only for a secondary, empirical rainfall → r_u link.

2.1.2 Static Bishop benchmark

We generated 30,000 static slopes by Latin hypercube sampling (H 5–50 m, β 15–60°, c′ 5–120 kPa, φ′ 15–45°, γ 16–24 kN/m³, r_u,v 0–0.55) and solved them with a vectorised simplified-Bishop solver, which was verified against 2,000 Morgenstern–Price solutions (Morgenstern and Price 1965; r = 0.995). The benchmark is split like the slope dataset: 21,000 slopes for training, 2,250 for validation, 2,250 for conformal calibration and 4,500 for testing. An independent factorial of 600 Bishop slopes × r_u,v = 0, 0.1, …, 0.4 (3,000 solutions) serves as a second test set.

2.1.3 Real case histories

Published compilations of static case histories (Sah et al. 1994; Zhou et al. 2019) give 363 cases after de-duplication (177 failed, 359 near-duplicate groups). They are used only as an external discrimination test.

2.2 Mechanics

2.2.1 Infinite slope with pore pressure

For a planar slip surface at vertical depth z parallel to the ground, with pore pressure expressed by Bishop's ratio,

$$F_{inf} = \frac{c'}{\gamma z \sin\beta\cos\beta} + \left(1 - \frac{r_{u,v}}{\cos^2\beta}\right)\frac{\tan\phi'}{\tan\beta} \qquad (1)$$

$$\frac{\partial F_{inf}}{\partial r_{u,v}} = -\frac{\tan\phi'}{\sin\beta\cos\beta}, \qquad \frac{\partial \ln F_{inf}}{\partial r_{u,v}} = -\frac{\tan\phi'}{F_{inf}\sin\beta\cos\beta} \qquad (2)$$

The absolute sensitivity does not contain c′; the relative one does, through F_{inf}. This gives three hypotheses. **H1:** cohesion barely changes the FoS lost per unit pore pressure. **H2:** that loss is

set by tanφ′/(sinβ cosβ). **H3:** cohesion nevertheless reduces the relative loss. Setting $F_{inf} = 1$ gives the critical ratio

$$r_{u,v}^{cr} = \cos^2\beta \left[1 - \left(1 - \frac{N}{\sin\beta\cos\beta}\right)\frac{\tan\beta}{\tan\phi'}\right], \qquad N = \frac{c'}{\gamma z} \qquad (3)$$

With slope-parallel seepage and a water table at height m z above the slip plane, $r_{u,v} = m\,\gamma_w \cos^2\beta/\gamma$, and soil below the table has saturated unit weight γ_{sat} (Skempton and DeLory 1957).

2.2.2 Rainfall infiltration with antecedent rain

Infiltration follows the slope-corrected Green–Ampt model (Green and Ampt 1911; Chen and Young 2006), which has been widely combined with the infinite slope for rainfall-induced instability (Chen et al. 2021; Zhu et al. 2026). Rain normal to the surface is I cos β. After cumulative infiltration F, measured normal to the slope, the infiltration capacity is

$$f = k_s\left(\cos\beta + \frac{\psi_f\,\Delta\theta}{F}\right) \qquad (4)$$

and the infiltration rate is min(I cos β, f). It is integrated explicitly (Δt = 15 min) through a time-varying rain record, which avoids the constant-intensity assumption of the analytical solution. For constant rain the stepped and analytical solutions agree to 0.001 mm. Parameters k_s, ψ_f and θ_e are those of Rawls et al. (1983), and the moisture deficit is $\Delta\theta = 0.5\,\theta_e$. Above a Green–Ampt wetting front the pressure head is close to zero. Matric suction is ignored, which is conservative (Fredlund and Rahardjo 1993; Lu and Godt 2008). The unit weight above the perched table is kept at γ, so the weight of the water held in the wetted zone is neglected; the wetted zone carries no positive pore pressure.

Positive pore pressure develops when infiltrated water perches on a less permeable layer at vertical depth D. Water in excess of the unsaturated pore volume above an antecedent table m_0, $\Delta\theta\, D \cos\beta\,(1 - m_0)$, raises the table to height mD. Perched water drains slope-parallel by Darcy flow, $q = k_s \sin\beta\, h_w$ per unit width, where h_w is the perched thickness above m_0. Spread over the slope length $L = H/\sin\beta$, this gives the storage balance

$$\Delta\theta\, D\cos\beta\, \frac{dm}{dt} = i_{excess}(t) - \frac{k_s \sin\beta\,(m - m_0)\, D\cos\beta}{L}, \qquad 0 \le m \le 1 \qquad (5)$$

The rainfall scenario is an antecedent depth A spread uniformly over 120 h, followed by a 24-h event of depth P. A and P are given in millimetres, as in rainfall thresholds (Caine 1980; Guzzetti et al. 2008). The highest m during the event is used.

2.2.3 Where the pore pressure acts: two mechanisms

A common shortcut converts the perched table into one pore-pressure ratio $r_u = m\,\gamma_w \cos^2\beta/\gamma$ and applies it uniformly along a deep circle. This is inconsistent: the water sits in a band of thickness mD above the layer at depth D, while a deep circle passes mostly below it. We therefore apply the pore pressure where it is and check two mechanisms.

Circular slips with an explicit pore-pressure field. In the simplified Bishop method, the pore pressure at each slice base at vertical depth d is

$$u = \begin{cases} \gamma_w\,[d - D(1-m)]\cos^2\beta_{loc} & D(1-m) < d \le D \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

Soil within the saturated band weighs γ_{sat}. The search covers the deep toe and base circles of the reference solver and, in addition, shallow circles that enter and leave anywhere on the crest, face or toe, so the critical surface is re-found for every wet state.

Translational slide on the perched layer. A circle cannot follow a planar layer: on a 40-m slope with the layer at 1.5 m and m = 1, the best circle gives FoS 0.92 against 0.77 for a plane. The planar mechanism is therefore checked with the infinite slope at depth D. This is exact for a long slide on the layer and conservative for shorter ones, because end resistance is ignored. The system FoS is the lower of the two, and failure means system FoS < 1.

2.2.4 Simplified Bishop as reference

The simplified Bishop method (Bishop 1955) with a circular-surface search is the mechanics reference. It supplies the benchmark FoS and the reference value in every check: sensitivities

by central finite differences, critical r_u,v by bisection, elasticities, and Monte Carlo failure probabilities on the same random samples as the surrogate. These checks are numerical verification of the surrogate against the method it replaces; they are not physical validation of the method itself.

2.3 Hybrid-6 surrogate

2.3.1 Architecture

Every network takes six standardised inputs (H, β, γ, c′, φ′, r_u, v). A differentiable layer computes ln F_{inf} at z = H and at z = λH, ln N and ln(tanφ′/tanβ), and appends them as four features, so that autodiff sensitivities d ln F/dx pass through the physics. The prior depth ratio λ = 0.18 minimises the variance of ln(F_{Bishop}/F_{inf}) on the Bishop training rows. Six experts with complementary architectures are used: E1 LSTM-GRU-DNN-XGBoost, E2 CNN-XGBoost, E3 an infinite-slope residual network, E4 physics-guided CNN-DNN-attention, E5 CNN-BiLSTM-SE-attention and E6 PLE-ResNet-attention (Gorishniy et al. 2022; Table S1). Baselines are four tree ensembles including XGBoost (Chen and Guestrin 2016), a plain DNN and the calibrated infinite slope.

2.3.2 Training and selection

The deep networks are pretrained on the uncensored training rows of the slope dataset (78,561 rows) and fine-tuned on the 21,000 Bishop training rows with early stopping (E3 is trained on Bishop only). The loss adds a mechanics sign penalty, evaluated by automatic differentiation:

$$L = \mathrm{MSE} + \sum_x \overline{\mathrm{ReLU}\left(-s_x \, \partial \ln \hat{F} / \partial x\right)^2}, \qquad s_x = +1 \text{ for } c', \phi'; \; -1 \text{ for } H, \beta, \gamma, r_{u,v} \qquad (7)$$

Four ways of combining the experts were scored on cross-fitted validation predictions: non-negative ridge stacking (Wolpert 1992; Breiman 1996), attention fusion with and without a physics penalty, and a plain average. Stacking had the lowest validation error (RMSE 0.0102

in ln F, against 0.0109–0.0137) and was fixed before the test set was opened. Its weights are E1 0.62, E4 0.28, E3 0.09, E6 0.01, E2 0 and E5 0.

2.3.3 Circular mechanism under a perched table

Hybrid-6 predicts the dry Bishop FoS. The effect of the perched pore-pressure field on the circular mechanism is added by a second learned term, $\ln F_{pw} = \ln F_{\text{Hybrid-6}}(\text{dry}) + r(x, D, m)$. The term r is fitted by gradient boosting, constrained to decrease with m, with infinite-slope features at the layer depth. It is trained on a separate benchmark of 24,000 Latin-hypercube slopes (H 5–50 m, β 20–50°, c′ 0.3–60 kPa, φ′ 15–45°, γ 16–22 kN/m³, D 1–4 m, m 0–1) solved with the explicit pore-pressure field and the extended search, split 70 / 15 / 15 %. On its test split it reproduces Bishop with an RMSE in ln F of 0.0109 (mean error 0.64 %) and classifies FoS < 1 correctly for 99.6 % of slopes. The uniform-r_u shortcut, evaluated on the same slopes, has a mean error of 12 %, underestimates FoS by 10 % at the median and misclassifies 9 % of slopes.

2.3.4 Checks

Split-conformal intervals on ln F (Shafer and Vovk 2008; Lei et al. 2018), global and per FoS class × slope band, are calibrated on 2,250 calibration slopes and evaluated on the 4,500 test slopes. Physical consistency is measured by wrong-direction responses under five single-property perturbations of 3,000 slopes and by agreement of d ln F/dx with direct Bishop on 300 test slopes.

2.3.5 Why a surrogate: compute cost

The reference solver evaluates 37 slopes per second per core (28 with the pore-pressure field, the extended search and the planar check). The reliability analyses need 3.3×10^7 FoS evaluations (Table 1), about 320 core-hours of direct computation. On the same two cores the surrogates need 41 min. A compiled or vectorised Bishop code would narrow this gap but would

not close it. For the dry Bishop FoS, the single best expert, E1, is as accurate as Hybrid-6 within the test-set confidence intervals (Section 3.2) and 31 % faster, so it is a simpler alternative.

2.4 Analyses

2.4.1 Geotechnical controls under pore pressure

H1–H3. For each of the 600 factorial slopes, dF/dr_u,v and $d \ln F/dr_u$,v are least-squares slopes over r_u,v = 0–0.4, computed in the same way for direct Bishop, Hybrid-6 and the infinite slope. They are related to the properties by Spearman correlation and standardised regression.

Dry versus wet. FoS at r_u,v = 0 and 0.3 is compared by property quintile (4,000 test slopes), together with elasticities $d \ln F/d \ln x$ (150 slopes, Hybrid-6 and direct Bishop) and total Sobol indices of ln F (Jansen 1999; Saltelli et al. 2010).

With the explicit pore-pressure field. H1–H3 are repeated by direct computation on 600 random slopes (H 6–40 m, β 20–50°, c′ 2–60 kPa, φ′ 18–42°, γ 16–22 kN/m³), with the system FoS of Section 2.2.3 at m = 0 and 1 on a layer at 2.7 m.

Critical rainfall. For three soils of equal dry infinite-slope FoS on a 35° slope, the water-table ratio at failure m_{cr} follows from Eq. 1, and the rainfall needed to reach it is found by stepping Eq. 4 and Eq. 5 through constant-intensity storms of 24–168 h for four textures and two antecedent state .

2.4.2 Reliability analysis

Slopes. Nine slopes share H = 15 m and γ = 19 kN/m³ and combine three strength mixes (cohesion-rich φ′ = 22°, balanced 30°, friction-rich 38°) with β = 30°, 35° and 40°. For each, c′ is solved so that the dry system FoS (full-search Bishop and planar slide on the layer at D = 2.7 m) is exactly 1.50. The deep circle governs the dry FoS of the cohesion-rich slopes and the planar slide that of the others. Because the no-rain failure probabilities still differ, the storm-induced increase ΔP_f is reported alongside P_f.

Uncertainty. c' is lognormal (COV 0.30). φ' is normal (COV 0.10), correlated with c' at $\rho = -0.5$. γ is normal (COV 0.05). k_s is lognormal (COV 1.0) and $\Delta\theta$ normal (COV 0.20), both around sandy-loam values; loam is a sensitivity case. These COVs lie within the ranges of Phoon and Kulhawy (1999). The antecedent table m_0 is drawn from the dry-weather slopes of the dataset ($P < 20$ mm). The main analysis places the perched layer at $D = 2.7$ m, the middle of the range tested below; for these slopes it coincides with the depth λH at which Bishop and the infinite slope agree best.

Fragility. For each slope, $A \in \{0, 150, 300\}$ mm and $P = 0$–400 mm, 20,000 samples give $P_f = P(\text{system FoS} < 1)$ from the surrogates, with common random numbers across rainfall levels. The Monte Carlo coefficient of variation of P_f is 1.1–3.7 % at the main design points. The friction-rich to cohesion-rich ratio of ΔP_f is given with a 95 % interval from 2,000 bootstrap resamples of the 20,000 samples; this interval reflects Monte Carlo sampling error only, not uncertainty in the input distributions or the model. Direct Monte Carlo (full-search Bishop with the pore-pressure field and the planar check) on the same samples ($n = 1{,}500$) checks 15 points. The same samples are also evaluated with the uniform-r_u shortcut for comparison.

Layer depth. The layer depth is not known in general. The analysis is therefore repeated with $D = 1.5$ and 4.0 m, re-solving c' for each depth so that the dry system FoS stays 1.50 ($A = 150$ and 300 mm; $P = 0$, 200 and 300 mm). Where the planar slide governs every failure, the system-level direct checks above test only the closed-form planar mechanism. The circular surrogate is therefore also checked on its own, against direct circular solutions on the same samples, for all nine slopes.

Importance and value of information. At a design state ($A = 300$ mm, $P = 200$ mm), P_f is recomputed with each variable fixed at its mean, with its COV halved (a proxy for better site investigation), and with the c'–φ' correlation set to 0 and -0.3 instead of -0.5.

Spatial averaging. Single random variables represent a slip surface with perfectly correlated strength. To bound the effect of spatial variability, the COVs of c' and φ' are multiplied by Vanmarcke's (1977) variance-reduction factor Γ for a Markov correlation over a slip-surface length $L \approx 1.2H/\sin\beta$ (31 m at $\beta = 35°$). Scales of fluctuation $\theta = 40$, 20 and 10 m give $\Gamma = 0.79$, 0.67 and 0.52. This is applied to the $\beta = 35°$ slopes with $A = 300$ mm.

FoS versus P_f. For 300 random slopes (H 5–40 m, β 20–50°, γ 16–22 kN/m³, c' 5–60 kPa, φ' 18–40°, r_u,v = 0.2), the mean-value FoS is compared with P_f from 2,000 samples at 0.5, 1 and 1.5 times the COVs above.

2.4.3 Saturated-layer screening relation

The fragility results are summarised by a closed-form index that needs only standard site data. For a planar slide on the less permeable layer at depth D, the infinite-slope FoS with slope-parallel seepage and the perched water table risen to the ground surface (m = 1) is

$$F_{sat} = \frac{c'}{\gamma_{sat} D \sin\beta \cos\beta} + \left(1 - \frac{\gamma_w}{\gamma_{sat}}\right)\frac{\tan\phi'}{\tan\beta} \qquad (8)$$

with the soil in the saturated band at γ_{sat}. F_{sat} is the lowest FoS the storm can produce on that plane. For the nine equal-FoS slopes at each of three layer depths (1.5, 2.7 and 4.0 m) and two antecedent depths (150 and 300 mm), the storm-induced increase ΔP_f at P = 300 mm from the Monte Carlo analysis is regressed on F_{sat}, D and an indicator of the antecedent depth:

$$\ln \Delta P_f = a_0 + a_1 F_{sat} + a_2 D + a_3\, \delta_{A=300} \qquad (9)$$

ΔP_f is floored at 10^{-3} (six cases). Because the 54 cases come from a designed set rather than a random sample, the fit is assessed in several ways: leave-one-out cross-validation; leave-one-group-out cross-validation by layer depth, slope angle, strength mix and slope; refits with floors of 10^{-4}–5×10^{-3} and without the 15 cases with $\Delta P_f < 0.005$; influence diagnostics; and comparison with the same regression on the dry planar FoS at the layer.

2.4.4 Field case: Rangamati, June 2017

The field case uses daily totals at Rangamati (24 h to 06:00) reported by the Bangladesh Meteorological Department: 165 mm on 12 June, 343 mm on 13 June and 180 mm on 14 June. Rain is spread uniformly within each day because hourly records are not available. Deep failure is checked with Bishop and Hybrid-6 and shallow translational failure with the infinite slope (depths 0.5–4 m). Infiltration, perched water and drainage follow the model of Section 2.2.2 (sandy-loam parameters, slope length 58 m), stepped through the record with $\Delta t = 3$ min. A Monte Carlo run ($n = 20{,}000$) samples c', φ', γ, k_s and $\Delta\theta$ as above, the slip depth uniformly on 1–3 m and m_0 uniformly on 0–0.3 to reflect the rain of the preceding days. $P_f(t)$ is conditioned on the slope standing before the storm. The field data that enter the analysis are the slope geometry, the laboratory strength and unit weight, the daily rainfall and the reported time of failure; slip depths, failure times and probabilities are model outputs.

3 Results

3.1 The dataset FoS is an infinite slope; Bishop agrees with it only at shallow depth

The dataset FoS is the infinite-slope FoS with $z = H$ (Fig. 2a). On the 16,848 uncensored test slopes the median ratio is 1.003 and the Spearman correlation 0.997, and the two agree on FoS < 1 for 97.5 % of slopes (all 112,282 uncensored rows: 1.002 and 0.996). Bishop behaves differently (Fig. 2b). Its best agreement with the infinite slope is at $z \approx 0.3H$, and at $z = H$ the infinite slope underestimates the Bishop FoS by a factor of 1.83.

3.2 Hybrid-6 is accurate, physically consistent and calibrated

On the 4,500 Bishop test slopes, Hybrid-6 predicts ln FoS with an RMSE of 0.0092 (95 % CI 0.0071–0.0114), a mean absolute FoS error of 0.37 % and the correct FoS class for 99.8 % of

slopes (Fig. 3a). No failing slope is predicted as stable. The best tree baseline (CatBoost) has an RMSE 1.7 times higher and the calibrated infinite slope 19 times higher. The RMSE on the independent Bishop factorial is 0.0038. The single expert E1 is not distinguishable from Hybrid-6 in accuracy: its RMSE is 0.0095 (CI 0.0073–0.0121) and its mean absolute error 0.35 %, slightly lower than that of Hybrid-6. Two further training seeds of Hybrid-6 give RMSEs of 0.0084–0.0097.

Hybrid-6 differs from E1 mainly in physical consistency. Under five single-property perturbations of 3,000 slopes, it moves the wrong way in 0.14 % of cases, against 0.33 % for E1 and 0.39–0.67 % for the tree baselines (Fig. 3b). Two other experts (E4 and E6) have lower rates still (0.11 % and 0.07 %) but are less accurate. All rates are small in absolute terms: 0.14 % is about 21 of 15,000 perturbed responses. The sensitivities of Hybrid-6 agree in sign with direct Bishop for 99.7–100 % of slopes, and the elasticities follow Bishop closely for all six inputs (Spearman ρ 0.97–1.00; Fig. 4a).

The 90 % conformal interval is ±0.68 % of FoS, with an overall coverage of 0.898. Coverage is uneven between groups: 0.79 for failing slopes and 0.85 for slopes steeper than 45°. Class-conditional (Mondrian) calibration brings every group to 0.88–0.90 (Fig. 4b).

On the 363 real case histories, discrimination is modest. Every learner and the calibrated infinite slope reach an AUC of 0.67–0.68 (95 % CIs about 0.62–0.73), and the uncalibrated infinite slope at $z = H$ reaches 0.63 (Fig. 3c). Because all models perform alike, the limit seems to lie in the six-variable description of a real slope, not in the learner.

3.3 Pore pressure: friction and slope angle set the loss, cohesion buffers it (H1–H3)

On the 600 factorial slopes, direct Bishop gives a median $\partial F/\partial r_{u,v}$ of −1.65 (Fig. 5).

- **H1.** Across cohesion quartiles the absolute loss changes only from −1.46 to −1.77 per unit r_u,v while c′ rises from 5 to 120 kPa (Fig. 5a). In a standardised regression c′ has a coefficient of −0.12, against −0.57 for φ′ and +0.39 for β.
- **H2.** The group tanφ′/(sinβ cosβ) explains 92.7 % of the variance of Bishop's $\partial F/\partial r_u,v$ (Fig. 5b).
- **H3.** The relative loss $-\partial \ln F/\partial r_u,v$ falls from 1.22 in the lowest cohesion quartile to 0.68 in the highest (Fig. 5c).

Hybrid-6 reproduces all three within 0.03. The infinite slope gets H2 by construction but misses Bishop's small cohesion and height effects, which come from the deeper critical circles of cohesive and high slopes.

With the explicit pore-pressure field. When the water is placed in a perched band (direct computation, 600 slopes), H3 becomes stronger and H1 holds only within a mechanism. The FoS lost on saturation falls from 0.43 in the lowest cohesion quartile to 0.06 in the highest, and the relative loss from 41 % to 3 % (Spearman with c′ −0.89). Cohesion therefore protects in a second way: it moves the critical surface below the wet band, where the water does not reach it. Where the planar slide governs, the absolute loss barely depends on c′ (Spearman −0.14), as H1 predicts. The group tanφ′/(sinβ cosβ) now explains less of the loss ($R^2 = 0.41$), because the depth of the critical surface relative to the band also matters.

3.4 Rain changes which properties control stability

Dry, the median elasticity of FoS ranks β (0.76) > φ′ (0.70) > H ≈ c′ ≈ γ (0.42) (direct Bishop; Hybrid-6 within 0.01). At $r_u,v = 0.3$, the elasticity of φ′ falls to 0.54, level with the others (Fig. 6a), and its total Sobol index halves from 0.15 to 0.075 while that of c′ rises from 0.31 to 0.37 (Fig. 6b). The FoS lost from $r_u,v = 0$ to 0.3 rises with φ′ (0.27 to 0.67 across quintiles) and barely with c′, while the relative loss falls from 31 % to 18 % from the lowest to the highest c′ quintile (Fig. 7).

The same contrast appears in the rainfall needed to fail a shallow slip (Fig. 8). For three soils of equal dry infinite-slope FoS (1.4) on a 35° slope, the cohesion-rich soil fails when the perched table reaches m = 0.97, the balanced soil at 0.80 and the friction-rich soil at 0.67 (Fig. 8a). On a 1-m slip in sandy loam this means 346 mm in 72 h for the friction-rich soil against 405 mm for the cohesion-rich soil (Fig. 8b). Texture matters as much as strength. From a dry start, loam and clay loam cannot take in enough water in 72 h to fail any of the three soils at dry FoS 1.4 (loam can fail them at dry FoS 1.2), and a wetter start ($m_0 = 0.2$) cuts the critical rainfall by 82–97 mm. Critical intensity–duration curves are lowest for friction-rich soils and wet starts (Fig. 8c).

3.5 Rainfall fragility: equal FoS, unequal risk

All nine slopes have a dry system FoS of 1.50 and similar no-rain failure probabilities (P_f 0.004–0.013, except 0.023 for the friction-rich 30° slope). Their response to rain depends first on antecedent conditions (Fig. 9a–c). From a dry start, even a 300-mm, 24-h storm leaves P_f below 0.035, because the water cannot fill the pore space above a 2.7-m layer in one day.

Mechanism. With the layer at 2.7 m, the planar slide on the perched layer has a lower FoS than the best circle in every failed sample, so all rain-triggered failures are translational. The circular mechanism with the explicit pore-pressure field falls below FoS 1 in at most 8.7 % of samples (friction-rich 30° slope, A = 300, P = 300 mm) and in at most 0.7 % for the other slopes, against system probabilities of 0.04–0.29. The deep circle governs the dry design of the cohesion-rich slopes but not their failure in rain.

Strength mix. The curves separate by strength mix (Table 2). After 300 mm of antecedent rain, a 300-mm storm gives the friction-rich slopes P_f = 0.12–0.29, against 0.04–0.06 for the cohesion-rich slopes: 3.4–4.6 times higher at every slope angle. Measured as the increase ΔP_f caused by the storm (relative to P = 0 after the same antecedent rain), which removes the differences in starting risk, the ratio is 4.6 at $\beta = 30°$ (95 % interval 4.3–5.0), 4.3 at 35° (4.0–

4.7) and 4.4 at 40° (4.0–4.8). With 150 mm of antecedent rain the ratio is 2.3–4.5. On loam, infiltration is slower and P_f stays below 0.10, but the order of the mixes is unchanged.

The shortcut misleads. Applying the same water as a uniform r_u over the whole circle (Fig. 9d; last column of Table 2) changes the picture. It overstates the risk of cohesion-rich slopes, whose deep circles it loads with water that is not there, and understates that of friction-rich slopes, whose failure plane it misses. With the shortcut, the friction-rich to cohesion-rich ratio at A = 300, P = 300 mm is 0.5–2.1, and below 1 at β = 35° and 40°.

Direct checks. At all 15 check points the surrogate and direct Monte Carlo give the same P_f (for example 0.117 for the friction-rich 35° slope at A = 300, P = 200 mm, 95 % interval 0.100–0.133), with 100 % verdict agreement and an RMSE in ln F of at most 0.0012. This agreement is partly by construction: on these samples every failure is planar, and the planar mechanism is closed form in both calculations. The uniform-r_u shortcut on the same samples falls outside the 95 % interval at 12 of the 15 points. The circular surrogate, the part that is actually approximated, was therefore tested on its own against direct circular solutions with the pore-pressure field on the same samples. Its circular P_f lies within the 95 % interval at all nine slopes, with an RMSE in ln F of 0.002–0.016.

Layer depth. The size of the rain effect depends strongly on where the water perches, but the ranking of the strength mixes does not. Measured as ΔP_f at A = 300, P = 300 mm, friction-rich slopes are more vulnerable than cohesion-rich ones at every depth: by a factor of more than 100 with the layer at 1.5 m, 4.3–4.6 at 2.7 m and 1.4–2.8 at 4 m. The factor at 1.5 m is imprecise, because the cohesion-rich increases there are 0.0002–0.0008 and close to the Monte Carlo resolution. The largest increase falls from 0.46 to 0.24 to 0.05 as the layer deepens, because 300 mm of antecedent rain fills less of a deeper unsaturated zone. At 4 m the total P_f can be higher for cohesion-rich slopes (0.045–0.060 against 0.027–0.034 at β = 35–40°), because their no-

rain risk on the deep planar mechanism is higher. Total P_f therefore ranks slopes only together with their starting risk, whereas ΔP_f isolates what the rain does.

3.6 A saturated-layer check describes the storm-induced risk

The dry system FoS cannot separate the nine slopes, because it is 1.50 for all of them by construction. The dry planar FoS at the layer separates them partly: it differs between slopes at 1.5 and 2.7 m, but at 4 m it is also 1.50 for all nine, while ΔP_f ranges from 0.003 to 0.053 (0.015–0.053 after 300 mm of antecedent rain; Fig. 10a). Within each layer depth, F_{sat} ranks ΔP_f almost perfectly (Spearman $\rho = -1.00$ at 1.5 and 2.7 m and −0.98 at 4 m), whereas the dry planar FoS gives $\rho = -1.00$, −0.84 and no ranking at all (Fig. 10b).

Across all 54 cases the fitted relation is

$$\ln \Delta P_f = 4.55 - 5.13\, F_{sat} - 1.03\, D + 0.94\, \delta_{A=300} \qquad (10)$$

with $R^2 = 0.90$ in $\ln \Delta P_f$ and a root-mean-square error of 0.53 (0.59 in leave-one-out prediction; Fig. 10c). The 95 % confidence intervals of the coefficients are −5.64 to −4.62 for F_{sat}, −1.20 to −0.86 for D and 0.64 to 1.25 for the antecedent indicator. Read literally, each 0.1 decrease in F_{sat} multiplies ΔP_f by about 1.7, each additional metre of layer depth divides it by about 2.8, and 300 mm instead of 150 mm of antecedent rain multiplies it by about 2.6.

How far the relation can be trusted depends on the kind of prediction. Leave-one-out Q^2 is 0.88, and leaving out a whole slope angle, slope or strength mix gives Q^2 = 0.87, 0.85 and 0.77. Leaving out a whole layer depth, however, gives only Q^2 = 0.57: the relation interpolates between the conditions it was fitted to, but does not extrapolate reliably to a layer depth outside them. The fit is sensitive to how very small increases are treated: with floors from 10^{-4} to 5×10^{-3}, R^2 ranges from 0.94 to 0.85 and the F_{sat} coefficient from −6.1 to −3.8, and without the 15 cases with $\Delta P_f < 0.005$ the values are 0.89 and −4.9. The multiplier per 0.1 of F_{sat} is therefore uncertain, roughly 1.5–1.9. Four of the 54 cases have Cook's distance above 4/n (largest 0.65), and the residuals show no clear departure from normality (Shapiro–Wilk p = 0.06).

F_{sat} and the dry planar FoS are strongly correlated across the 54 cases (r = 0.97). Used in place of F_{sat}, the dry planar FoS gives $R^2 = 0.88$, and with both terms $R^2 = 0.91$. The advantage of F_{sat} over the 54 cases taken together is therefore small. Its clear advantage is within a layer depth, where it ranks slopes that the dry planar FoS ranks poorly or not at all. The relation is a descriptive summary of this study's Monte Carlo results for a 15-m slope in sandy loam with a 300-mm, 24-h storm; it is not a substitute for a probabilistic analysis and has not been tested on other slopes or field data.

3.7 Which uncertainty drives the risk, and what is worth measuring

At the design state (A = 300 mm, P = 200 mm; base P_f 0.030–0.21), the answer depends on what carries the slope (Fig. 11a, b).

- **Cohesion-rich slopes.** Knowing c′ removes 100 % of P_f, and halving its COV removes 91–100 %.
- **Friction-rich slopes.** Strength data help little. Knowing φ′ lowers P_f by at most 17 % and raises it on two of the three slopes; knowing c′ removes 0–44 %. Knowing the moisture deficit Δθ or the antecedent table m_0 removes 27–38 % each, and halving the COV of Δθ removes 20–29 %.
- **Balanced slopes** lie between: c′ data remove 28–99 % and hydraulic data 25–37 %.
- **Unit weight** matters little (≤ 4 %).

Because c′ and φ′ are negatively correlated, fixing φ′ raises P_f of cohesion-rich slopes by 35–43 %. Fixing k_s at its mean raises P_f by 21–43 %, because its mean exceeds its median. Single-variable importance measures that ignore correlation or skew can therefore mislead. The correlation itself matters: with independent c′ and φ′, P_f rises by 5–54 % (most for cohesion-rich slopes), and with $\rho = -0.3$ by 3–23 %. The assumed $\rho = -0.5$ is therefore not conservative and should be site-specific.

Spatial averaging lowers P_f unevenly (Fig. 11c). At P = 300 mm with θ = 20 m, P_f falls by 76 % for the cohesion-rich 35° slope (0.047 to 0.011) and by only 9 % for the friction-rich slope (0.170 to 0.155). The ratio between them grows from 3.6 to 14, and to 61 at θ = 10 m. Cohesion-rich slopes fail in the strength tail, which averaging removes. Friction-rich slopes fail because

484 water reaches the layer, which averaging of strength does not change. These figures come from
485 a variance-reduction bound, not from a random-field analysis.

486 3.8 Mean factor of safety and failure probability

487 Across 300 random slopes at r_u,v = 0.2, a mean-value FoS of 1.25–1.35 corresponds to a median
488 P_f of 0.0065 and a 90th percentile of 0.037 with typical COVs. With COVs 1.5 times larger
489 these become 0.064 and 0.12 (Fig. 12). To keep $P_f \leq 0.01$ for 90 % of slopes, the required mean-
490 value FoS is 1.05 for low, 1.20 for typical and 1.45 for high COVs. A single target such as FoS
491 = 1.3 is therefore over-conservative where properties are well known and unsafe where they are
492 not, as noted by Duncan (2000) and El-Ramly et al. (2002).

493 3.9 Field case: Rangamati, Bangladesh, 12–14 June 2017

494 On the morning of 13 June 2017, landslides across Rangamati District killed about 120 people
495 after heavy rain over the preceding days (Abedin et al. 2020). The studied road-cut slope in silty
496 sand (SM) has H = 46 m, $\beta = 53°$, $c' = 11$ kPa, $\varphi' = 26.3°$ and $\gamma = 17.5$ kN/m³ see also Islam et
497 al. (2021). Figure 13 shows the rainfall record and the slope.
498 **Deep failure is excluded.** A deep circular failure has a Bishop FoS of 0.62 even when dry
499 (Hybrid-6: 0.615) and would need $c' \approx 47$ kPa to stand. Since the slope stood, these laboratory
500 properties describe a weathered surface mantle, and failure had to be shallow.
501 **There is a critical depth window.** The dry infinite-slope FoS is 1.68 at 1 m, 1.24 at 1.5 m and
502 1.03 at 2 m, and falls below 1 beyond 2.09 m. Slips of 1 m cannot fail even when saturated (FoS
503 1.36). F_{sat} (Eq. 8) falls below 1 at 1.44 m, so in the model only depths between 1.44 m and 2.09
504 m are vulnerable to rain (Fig. 13b).
505 **Timing.** With mean properties and a dry start, a 2-m slip fails at 03:36 on 13 June and a 1.75-
506 m slip at 10:39 (Fig. 14b). In the Monte Carlo run, 47 % of samples stand before the storm.
507 Conditional on standing, P_f is 0.003 after the 165-mm day and 0.19 by 06:00 on 13 June (Fig.

508 14c). The median failure time is 02:54 on 13 June, and the interquartile range, from 19:22 on
509 12 June to 14:15 on 13 June, contains the reported morning failures. Drainage changes these
510 results by less than 1 % over three days. The outcome is controlled mostly by k_s ($|\Delta P_f| = 0.12$ if
511 known exactly), then by slip depth (0.03) and c′ (0.02) (Fig. 14d). The event exceeded the
512 regional 72-h threshold of about 210 mm (Khatun et al. 2023) early on 12 June.

513 # 4 Discussion

514 ## 4.1 What controls rainfall-induced risk

515 Slopes with the same dry FoS carry different rainfall risk, and the difference follows predictably
516 from their strength mix. Pore pressure acts only on the frictional term, so on a given surface the
517 FoS lost per unit pore pressure is set by $\tan\varphi'/(\sin\beta \cos\beta)$ and is almost independent of cohesion
518 (H1, H2), while cohesion adds FoS that rain cannot remove (H3). With a realistic pore-pressure
519 field, cohesion also protects in a second way, by pushing the critical surface below the band
520 where water collects. Both effects carry through to the risk. In the slopes analysed, the storm
521 raised the failure probability of friction-rich slopes more than that of cohesion-rich slopes of
522 the same dry FoS at every layer depth tested, by 4.3–4.6 times with the layer at 2.7 m. That
523 interval reflects Monte Carlo error only; the size of the ratio also depends on the assumed input
524 distributions, the layer depth and the soil texture.
525 Hydrology decides when and where this contrast matters. A single storm on a dry slope cannot
526 bring enough water to a layer a few metres deep, so antecedent rain is the precondition for fail
527 ure, as Rahardjo et al. (2001, 2007) found and as regional thresholds reflect (Khatun et
528 al. 2023). Where the water collects decides the mechanism: perched water loads a shallow
529 plane, not a deep circle, so rain-triggered failures are translational even where a deep circle
530 governs the dry design. This agrees with the shallow failure planes found in deterministic

infiltration analyses in this journal (Montrasio et al. 2018; Zhu et al. 2026), and adds that the strength mix decides how probable such failure is. Applying the water as a uniform r_u over the critical dry circle misses the mechanism and can reverse the ranking of slopes.

4.2 A second number next to the factor of safety

The results suggest a simple addition to practice. Next to the dry FoS, the designer can report the saturated planar FoS F_{sat} at the depth of the shallowest less permeable layer (Eq. 8). It needs no hydrological model, only c', φ', γ_{sat}, β and D, all of which a stability check already requires. In this study F_{sat} ranked the storm-induced risk of slopes within each layer depth, including at 4 m, where the dry planar FoS could not separate them. For the Rangamati slope, the depths at which F_{sat} and the dry FoS fall below 1 bound the vulnerable slip depths identified by the time-dependent analysis. A slope with dry FoS ≥ 1 but $F_{sat} < 1$ on a layer that antecedent rain can reach deserves attention as rainfall-sensitive. Two cautions apply. Across all 54 cases the dry planar FoS explains almost as much as F_{sat} (R^2 0.88 against 0.90), so the gain lies in within-depth ranking rather than in overall fit. And the numerical coefficients of Eq. 10 are specific to the conditions analysed and do not transfer reliably even to an unseen layer depth ($Q^2 = 0.57$).

4.3 What is worth measuring

Site investigation should follow the failure mode. For cohesion-rich slopes, c' data remove almost all of the risk. For friction-rich slopes, strength data help little, and monitoring or measuring the moisture state (antecedent water table, moisture deficit) is worth more. Because of the c'–φ' correlation, better knowledge of the secondary property alone can even raise the estimated risk, and the correlation coefficient itself changes P_f by up to half. Deterministic targets should also scale with uncertainty (Section 3.8), which adds a mechanism to earlier evidence that FoS and P_f can diverge (Duncan 2000; El-Ramly et al. 2002).

4.4 Role of the infinite slope and of the surrogate

The infinite slope captures the qualitative physics, and for slides on a perched layer it is the governing mechanism. As a substitute for circular failures, however, its φ′ sensitivity is half that of Bishop and it misclassifies 14 % of critical-r_u cells (Online Resource 1, Section S15 and Fig. S1). Hybrid-6 with the perched-water term closes this gap at about 1/235 of the cost of direct computation. In the risk analysis its role is to evaluate the circular mechanism for every sample, which is what shows that the deep circle does not govern; it was checked on its own for that role. Earlier machine-learning studies of slope stability in this journal reported accuracy on held-out data (Sakellariou and Ferentinou 2005; Nouri et al. 2021; Marrapu et al. 2021) or used surrogates for reliability without checking their sensitivities (Samui et al. 2019; Santos et al. 2025). The value of Hybrid-6 lies less in test accuracy, where E1 alone is as good, than in the verified agreement of its sensitivities, critical states and failure probabilities with Bishop and in its lower rate of wrong-direction responses, in line with the call of Phoon and Zhang (2023) to judge geotechnical machine learning by fitness for purpose. For the dry Bishop FoS, E1 is an equally valid and simpler choice. Neither model discriminates real case histories better than the infinite slope (AUC about 0.67–0.68), which is a reminder that agreement with a reference method is not agreement with nature.

4.5 Limitations

- Mechanics: Slopes are assumed homogeneous except for the less permeable layer (1.5–4 m), whose depth requires site-specific data. Three-dimensional effects, matric suction, vegetation, and water weight above the water table are neglected.
- Pore pressure: Sections 3.3–3.4 use uniform r_u for parametric analysis; rainfall simulations use a separate hydrological model.
- Uncertainty: Spatial variability is approximated using variance reduction. Reported confidence intervals capture Monte Carlo error only.

- Verification and validation: The surrogate are verified against only Bishop's method. Case-history discrimination remains modest (AUC = 0.67–0.68).
- Screening relation: The relation is based on 54 designed cases and extrapolates poorly across layer depths ($Q^2 = 0.57$) recalibration is required for other conditions.

5 Conclusions

- A hybrid deep-learning surrogate can stand in for limit-equilibrium analysis in a reliability study if it is judged on the quantities that study uses. Hybrid-6 reproduced the Bishop FoS with a mean error of 0.37 % (0.64 % with the perched-water term), and its sensitivities, critical pore-pressure ratios and circular-mechanism failure probabilities matched direct computation within sampling error, at about 1/235 of the cost. In accuracy alone, the single expert E1 was as good. The FoS of the 120,000-slope dataset proved to be the infinite-slope FoS with $z = H$, whereas Bishop agrees with the infinite slope only at shallow depth ($z \approx 0.3H$).
- On a given surface, the FoS lost per unit pore pressure is governed by $\tan\varphi'/(\sin\beta \cos\beta)$ and hardly by cohesion; cohesion instead cuts the relative loss by about 44 %. With a realistic pore-pressure field it also pushes the critical surface below the wet band.
- Antecedent rain is the precondition for failure on a less permeable layer a few metres deep. Once the pore pressure is placed where it acts, every rain-induced failure in the cases analysed was a shallow translational slide, and the storm raised the failure probability of friction-rich slopes more than that of cohesion-rich slopes of the same dry FoS at every layer depth tested (4.3–4.6 times with the layer at 2.7 m). A uniform pore-pressure ratio over a deep circle hides this contrast and can reverse it.
- Within each layer depth, the saturated planar FoS ranks the storm-induced risk where the dry FoS cannot. With the layer depth and the antecedent state, it describes 90 % of the variance of $\ln \Delta P_f$ over the 54 cases, but it extrapolates poorly to an unseen layer depth and should be treated as a screening aid for the conditions analysed.
- What is worth measuring depends on what carries the slope: c' data remove almost all of the risk of cohesion-rich slopes, whereas the hydraulic state matters more than strength for friction-rich slopes. The c'–φ' correlation changes P_f by up to half, and a single target FoS does not give uniform safety.

- For the Rangamati landslides of 13 June 2017, the framework rules out deep failure, bounds the vulnerable slip depth to 1.44–2.09 m and gives a failure window that contains the reported time of failure.

References

Abedin J, Rabby YW, Hasan I, Akter H (2020) An investigation of the characteristics, causes, and consequences of June 13, 2017, landslides in Rangamati District Bangladesh. Geoenviron Disasters 7:23. https://doi.org/10.1186/s40677-020-00161-z

Arif A, Zhang C, Sajib MH, Uddin MN, Habibullah M, Feng R, Feng M, Rahman MS, Zhang Y (2025) Rock slope stability prediction: a review of machine learning techniques. Geotech Geol Eng 43:124. https://doi.org/10.1007/s10706-025-03091-5

Baecher GB, Christian JT (2003) Reliability and statistics in geotechnical engineering. Wiley, Chichester

Baum RL, Savage WZ, Godt JW (2008) TRIGRS—a Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0. U.S. Geological Survey Open-File Report 2008-1159

Bishop AW (1955) The use of the slip circle in the stability analysis of slopes. Géotechnique 5(1):7–17.

Breiman L (1996) Stacked regressions. Mach Learn 24(1):49–64.

Caine N (1980) The rainfall intensity–duration control of shallow landslides and debris flows. Geogr Ann A 62(1–2):23–27

Chen B, Ren Q, Wang F, Zhang R, Liu B, Xie W (2021) Progressive failure analysis of slope water damage based on improved Green-Ampt infiltration model. Geotech Geol Eng 39:5109–5118. https://doi.org/10.1007/s10706-021-01816-w

Chen L, Young MH (2006) Green-Ampt infiltration model for sloping surfaces. Water Resour Res 42:W07420. https://doi.org/10.1029/2005WR004468

635 Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of the
636 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM,
637 New York, pp 785–794.
638 Cho SE (2014) Probabilistic stability analysis of rainfall-induced landslides considering spatial
639 variability of permeability. Eng Geol 171:11–20.
640 Christian JT, Ladd CC, Baecher GB (1994) Reliability applied to slope stability analysis. J
641 Geotech Eng 120(12):2180–2207.
642 Duncan JM (2000) Factors of safety and reliability in geotechnical engineering. J Geotech
643 Geoenviron Eng 126(4):307–316.
644 El-Ramly H, Morgenstern NR, Cruden DM (2002) Probabilistic slope stability analysis for
645 practice. Can Geotech J 39(3):665–683.
646 Fattahi H, Zandy Ilghani N (2020) Slope stability analysis using Bayesian Markov chain Monte
647 Carlo method. Geotech Geol Eng 38:2609–2618. https://doi.org/10.1007/s10706-019-01172-w
648 Fredlund DG, Rahardjo H (1993) Soil mechanics for unsaturated soils. Wiley, New York
649 Gorishniy Y, Rubachev I, Babenko A (2022) On embeddings for numerical features in tabular
650 deep learning. Adv Neural Inf Process Syst 35
651 Green WH, Ampt GA (1911) Studies on soil physics. Part I. The flow of air and water through
652 soils. J Agric Sci 4(1):1–24.
653 Griffiths DV, Fenton GA (2004) Probabilistic slope stability analysis by finite elements. J
654 Geotech Geoenviron Eng 130(5):507–518.
655 Griffiths DV, Huang J, Fenton GA (2009) Influence of spatial variability on slope reliability
656 using 2-D random fields. J Geotech Geoenviron Eng 135(10):1367–1378.
657 Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity–duration control of
658 shallow landslides and debris flows: an update. Landslides 5(1):3–17.

Huat BBK, Ali FH, Low TH (2006) Water infiltration characteristics of unsaturated soil slope and its effect on suction and stability. Geotech Geol Eng 24:1293–1306. https://doi.org/10.1007/s10706-005-1881-8

Islam MS, Begum A, Hasan MM (2021) Slope stability analysis of the Rangamati District using geotechnical and geochemical parameters. Nat Hazards 108(2):1659–1686. https://doi.org/10.1007/s11069-021-04750-5

Iverson RM (2000) Landslide triggering by rain infiltration. Water Resour Res 36(7):1897–1910.

Jansen MJW (1999) Analysis of variance designs for model output. Comput Phys Commun 117(1–2):35–43.

Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. Nat Rev Phys 3(6):422–440.

Kawajiri S, Kawaguchi T, Nakamura D, Yamashita S (2022) Application of critical rainfall curve to rainfall and snowmelt induced embankment slope collapse. Geotech Geol Eng 40:3897–3912. https://doi.org/10.1007/s10706-022-02097-7

Khatun M, Hossain ATMS, Sayem HM (2023) Climate variability & establishment of rainfall threshold line for landslide hazards in Rangamati, Bangladesh. Open J Geol 13(9):959–979. https://doi.org/10.4236/ojg.2023.139041

Lei J, G'Sell M, Rinaldo A, Tibshirani RJ, Wasserman L (2018) Distribution-free predictive inference for regression. J Am Stat Assoc 113(523):1094–1111.

Li D, Wang Z, Guo H, Zhang Y, Cheng X, Yu Q (2025) Deep learning in slope stability analysis: evolution, challenges, and future directions. Geotech Geol Eng 43:460. https://doi.org/10.1007/s10706-025-03424-4

Li DQ, Jiang SH, Cao ZJ, Zhou W, Zhou CB, Zhang LM (2015) A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. Eng Geol 187:60–72.

Lin H, Zhong W (2019) Influence of rainfall intensity and its pattern on the stability of unsaturated soil slope. Geotech Geol Eng 37:615–623. https://doi.org/10.1007/s10706-018-0631-7

Lu N, Godt J (2008) Infinite slope stability under steady unsaturated seepage conditions. Water Resour Res 44:W11404.

Marrapu BM, Kukunuri A, Jakka RS (2021) Improvement in prediction of slope stability & relative importance factors using ANN. Geotech Geol Eng 39:5879–5894. https://doi.org/10.1007/s10706-021-01872-2

Montgomery DR, Dietrich WE (1994) A physically based model for the topographic control on shallow landsliding. Water Resour Res 30(4):1153–1171.

Montrasio L, Valentino R, Meisina C (2018) Soil saturation and stability analysis of a test site slope using the shallow landslide instability prediction (SLIP) model. Geotech Geol Eng 36:2331–2342. https://doi.org/10.1007/s10706-018-0465-3

Morgenstern NR, Price VE (1965) The analysis of the stability of general slip surfaces. Géotechnique 15(1):79–93.

Nouri M, Sihag P, Salmasi F, Abraham J (2021) Prediction of homogeneous earthen slope safety factors using the forest and tree based modelling. Geotech Geol Eng 39:2849–2862. https://doi.org/10.1007/s10706-020-01659-x

Paul A, Islam MS, Jotisankasa A, Progga ST (2025) Influence of rainfall infiltration and hydraulic conductivity on hillslope stability. Geotech Geol Eng 43:300. https://doi.org/10.1007/s10706-025-03269-x

706 Phoon KK, Kulhawy FH (1999) Characterization of geotechnical variability. Can Geotech J
707 36(4):612–624.
708 Phoon KK, Zhang W (2023) Future of machine learning in geotechnics. Georisk 17(1):7–22.
709 https://doi.org/10.1080/17499518.2022.2087884
710 Rahardjo H, Li XW, Toll DG, Leong EC (2001) The effect of antecedent rainfall on slope
711 stability. Geotech Geol Eng 19:371–399. https://doi.org/10.1023/A:1013129725263
712 Rahardjo H, Ong TH, Rezaur RB, Leong EC (2007) Factors controlling instability of
713 homogeneous soil slopes under rainfall. J Geotech Geoenviron Eng 133(12):1532–1543.
714 Rathor APS, Khandelwal D (2026) Physics-guided machine learning for liquefaction-induced
715 settlement prediction. Geotech Geol Eng 44:259. https://doi.org/10.1007/s10706-026-03775-6
716 Rawls WJ, Brakensiek DL, Miller N (1983) Green-Ampt infiltration parameters from soils data.
717 J Hydraul Eng 109(1):62–70.
718 Sah NK, Sheorey PR, Upadhyaya LN (1994) Maximum likelihood estimation of slope stability.
719 Int J Rock Mech Min Sci Geomech Abstr 31(1):47–53.
720 Sakellariou MG, Ferentinou MD (2005) A study of slope stability prediction using neural
721 networks. Geotech Geol Eng 23:419–445. https://doi.org/10.1007/s10706-004-8680-5
722 Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based
723 sensitivity analysis of model output. Design and estimator for the total sensitivity index.
724 Comput Phys Commun 181(2):259–270.
725 Samui P, Kumar R, Yadav U, Kumari S, Bui DT (2019) Reliability analysis of slope safety
726 factor by using GPR and GP. Geotech Geol Eng 37:2245–2254.
727 https://doi.org/10.1007/s10706-018-0697-2
728 Santos L, Resende C, Martins K, Quevedo R, Lopez M (2025) Comparative analysis of slope
729 stability prediction for earth dams using response surface method, statistical models, and ANN.
730 Geotech Geol Eng 43:168. https://doi.org/10.1007/s10706-025-03138-7

731 Santoso AM, Phoon KK, Quek ST (2011) Effects of soil spatial variability on rainfall-induced
732 landslides. Comput Struct 89(11–12):893–900.
733 Shafer G, Vovk V (2008) A tutorial on conformal prediction. J Mach Learn Res 9:371–421
734 Skempton AW, DeLory FA (1957) Stability of natural slopes in London Clay. In: Proceedings
735 of the 4th international conference on soil mechanics and foundation engineering, vol 2.
736 Butterworths, London, pp 378–381
737 Vanmarcke EH (1977) Probabilistic modeling of soil profiles. J Geotech Eng Div
738 103(11):1227–1246.
739 Viswanath A, Ramesh S, Mangalathu S (2026) Rainfall-induced slope stability modeling: a
740 systematic review of coupled and hybrid frameworks. Geotech Geol Eng 44:281.
741 https://doi.org/10.1007/s10706-026-03803-5
742 Wang C, Li L, Kou H, Cheng Y, Meng D (2024) Rainfall induced slope reliability analysis
743 using radial degraded random fields. Geotech Geol Eng 42:6523–6548.
744 https://doi.org/10.1007/s10706-024-02905-2
745 Wolpert DH (1992) Stacked generalization. Neural Netw 5(2):241–259.
746 Xue J, Gavin K (2008) Effect of rainfall intensity on infiltration into partly saturated slopes.
747 Geotech Geol Eng 26:199–209. https://doi.org/10.1007/s10706-007-9157-0
748 Zhou J, Li E, Yang S, Wang M, Shi X, Yao S, Mitri HS (2019) Slope stability prediction for
749 circular mode failure using gradient boosting machine approach based on an updated database
750 of case histories. Saf Sci 118:505–518.
751 Zhu H, Wu H, Shu J, Zhang Z, Yang Q, Chen X, Zhu P (2026) Improved Green–Ampt model
752 for evaluating rainfall infiltration and slope stability. Geotech Geol Eng 44:33.
753 https://doi.org/10.1007/s10706-025-03540-1

Statements and Declarations

Funding No Funding

Competing interest's N/A

Author contributions Muhammed Yeasin Arafat contributes all.

Data availability

Data and all other thing are available in https://doi.org/10.5281/zenodo.23470375

AI Uses Ai is used for language enhancement.

Table captions

Table 1 Cost of the analyses on two CPU cores (throughputs in Online Resource 1, Table S10)

Table 2 P_f after a 24-h storm of P mm following 300 mm of antecedent rain (sandy loam, n = 20,000). Dry system FoS = 1.50 for all slopes. Last column: the same samples evaluated with a uniform r_u over the circle

Figure captions

Fig. 1 Conceptual model of the study. Antecedent rain and a 24-h storm infiltrate the slope (slope-corrected Green–Ampt), the water perches above a less permeable layer at depth D and raises a water table of height mD, and stability is checked both on a deep circle (simplified Bishop with the slice-wise pore pressure) and on a translational slide along the layer. The numbered column gives the sequence of the analysis. Not to scale

Fig. 2 The slope dataset and the infinite slope. **a** Dataset FoS against the infinite-slope FoS with z = H for the 16,848 uncensored test slopes (log axes; dashed line 1:1). **b** Median ratio of reference FoS to infinite-slope FoS against the slip-depth ratio z/H assumed in the infinite slope,

775 for the dataset FoS and for simplified Bishop (factorial set, 3,000 solutions); the vertical line
776 marks the prior depth z = 0.18H used inside Hybrid-6
777 **Fig. 3** Accuracy, physical consistency and real-case discrimination of the surrogates. **a** Test
778 RMSE of ln FoS on 4,500 Bishop slopes with 95 % bootstrap intervals. **b** Share of wrong-
779 direction responses under five single-property perturbations of 3,000 slopes; zero values are
780 drawn at 0.03 % for visibility. **c** ROC curves on 363 published case histories (177 failed) for
781 Hybrid-6, E1, XGBoost and the calibrated infinite slope; the diagonal is a random classifier
782 **Fig. 4** Agreement of Hybrid-6 with direct Bishop. **a** Elasticity $\partial \ln F / \partial \ln x$ of Hybrid-6 against
783 central finite differences of direct Bishop for six inputs on 300 test slopes (dashed line 1:1; ρ is
784 the Spearman correlation of the derivatives). **b** Observed coverage of the 90 % conformal
785 interval by FoS class and slope-angle band, for global and class-conditional (Mondrian)
786 calibration; n is the number of test slopes in each group
787 **Fig. 5** Sensitivity of FoS to the pore-pressure ratio on the 600 factorial slopes. **a** FoS lost per
788 unit $r_{u,v}$ against cohesion. **b** The same loss against the group $\tan \varphi' / (\sin \beta \cos \beta)$ for Bishop
789 (squares) and Hybrid-6 (dots); R^2 refers to Bishop. **c** Relative loss $-\partial \ln F / \partial r_{u,v}$ against cohesion.
790 In **a** and **c**, small symbols are individual Bishop slopes and lines are medians of equal-count
791 cohesion classes for Bishop, Hybrid-6 and the infinite slope at $z = \lambda H$
792 **Fig. 6** Which properties control the FoS, dry and wet. **a** Median absolute elasticity $|\partial \ln F / \partial \ln x|$
793 from direct Bishop on 150 slopes, dry (open bars) and at $r_{u,v} = 0.3$ (filled bars); vertical ticks
794 mark Hybrid-6. **b** Total Sobol indices of ln F (Hybrid-6), dry and at $r_{u,v} = 0.3$
795 **Fig. 7** FoS lost between $r_{u,v} = 0$ and 0.3 by quintile of each property (Hybrid-6, 4,000 test
796 slopes). **a** Absolute loss. **b** Relative loss
797 **Fig. 8** Critical conditions for a 1-m translational slip on a 35° slope, for three soils of equal dry
798 FoS. **a** Water-table ratio at failure, for dry FoS 1.2 and 1.4. **b** Critical 72-h rainfall for four
799 textures (dry FoS 1.4); open symbols: dry start ($m_0 = 0$); filled symbols: wetter start ($m_0 = 0.2$);

crosses: no failure within 72 h. **c** Critical intensity against duration for silt loam (log axes); solid lines: $m_0 = 0$; dashed lines: $m_0 = 0.2$

Fig. 9 Rainfall fragility of the equal-FoS slopes (dry system FoS 1.50, layer at 2.7 m, sandy loam, n = 20,000). **a**–**c** Probability of failure against 24-h storm rainfall after 0, 150 and 300 mm of antecedent rain; lines and markers: $\beta = 35°$; shaded bands: range over $\beta = 30–40°$; black symbols with bars: direct Monte Carlo checks (n = 1,500) with 95 % intervals (filled: $\beta = 35°$). **d** $\beta = 35°$, A = 300 mm: pore pressure applied where it acts (solid) and the same water as a uniform r_u over the critical circle (dotted)

Fig. 10 Storm-induced increase in the probability of failure (A = 150 or 300 mm, P = 300 mm) for the nine equal-FoS slopes with the layer at 1.5, 2.7 and 4.0 m (54 cases; log axes; values below 10^{-3} plotted at 10^{-3}). **a** Against the dry planar FoS at the layer. **b** Against the saturated planar FoS at the layer, F_{sat}; lines: Eq. 10 for A = 300 mm. **c** ΔP_f from Eq. 10 against the Monte Carlo value (dashed line 1:1), with the fit and cross-validation statistics; Q^2 values are out-of-sample

Fig. 11 Value of information and spatial averaging at the design state (A = 300 mm, P = 200 mm unless stated). **a** Reduction of P_f if each variable were known exactly. **b** Reduction of P_f if the COV of each variable were halved. Bars: mean over $\beta = 30°$, 35° and 40°; whiskers: range; negative values come from the $c'–\varphi'$ correlation and the skew of k_s. **c** P_f against 24-h storm rainfall at $\beta = 35°$, A = 300 mm, without averaging and with Vanmarcke variance reduction for scales of fluctuation $\theta = 20$ and 10 m

Fig. 12 Mean-value FoS against Monte Carlo probability of failure for 300 random slopes ($r_{u,v}$ = 0.2, n = 2,000 per slope) at three levels of property uncertainty. The shaded band marks FoS 1.25–1.35; dashed lines mark $P_f = 0.1$ and 0.01; slopes with no failure in 2,000 samples are plotted at the bottom of the axis

Fig. 13 The Rangamati case. **a** Daily rainfall at Rangamati, 11–14 June 2017 (24 h to 06:00), with the reported time of the fatal slides (Abedin et al. 2020); the regional 72-h threshold is from Khatun et al. (2023). **b** Section of the studied road-cut slope with the model-derived bounds of the vulnerable translational slip depth: 1.44 m, where $F_{sat} = 1$, and 2.09 m, where the dry FoS = 1 (depths drawn four times larger)

Fig. 14 Model response of the Rangamati slope through the June 2017 storm. **a** Cumulative rainfall. **b** FoS of translational slips at four depths for mean properties, a dry start and slope-parallel drainage. **c** Monte Carlo probability of failure conditional on the slope standing before the rain (n = 20,000); dashed line: median failure time; light band: interquartile range of the failure time. **d** Absolute change in P_f at 06:00 on 13 June if each variable were known exactly. The sand-coloured band in **a**–**c** marks the reported fatal slides

Tables

Table 1 Cost of the analyses on two CPU cores (throughputs in Online Resource 1, Table S10)

Analysis	FoS evaluations	Direct (core-h)	Surrogate (min)
Fragility, antecedent + event (Section 3.5)	18.4 M	184	23.5
Spatial averaging (Section 3.7)	4.1 M	41	5.2
Importance, value of information, correlation (Section 3.7)	2.5 M	25	3.2
Layer-depth sensitivity (Section 3.5)	3.2 M	32	4.2
Empirical-link bound (Online Resource 1)	3.1 M	23	2.9
FoS versus P_f (Section 3.8)	1.8 M	14	1.7
Critical r_u grids, elasticities, Sobol	0.09 M	0.7	0.1
Total	**33.1 M**	**319**	**41**

Table 2 P_f after a 24-h storm of P mm following 300 mm of antecedent rain (sandy loam, n = 20,000). Dry system FoS = 1.50 for all slopes. Last column: the same samples evaluated with a uniform r_u over the circle

β	Strength mix	c′ (kPa) / φ′ (°)	P = 0 (antecedent only)	P = 200 mm	P = 300 mm	ΔP_f (300 mm)	Uniform r_u (300 mm)
30°	cohesion-rich	19.6 / 22	0.010	0.044	0.062	0.052	0.090
30°	balanced	11.1 / 30	0.012	0.093	0.138	0.126	0.089
30°	friction-rich	3.3 / 38	0.045	0.212	0.286	0.240	0.187
35°	cohesion-rich	25.3 / 22	0.011	0.036	0.047	0.036	0.059
35°	balanced	16.3 / 30	0.015	0.079	0.110	0.094	0.054
35°	friction-rich	9.3 / 38	0.015	0.117	0.170	0.155	0.056
40°	cohesion-rich	30.6 / 22	0.011	0.030	0.036	0.025	0.047
40°	balanced	20.5 / 30	0.021	0.080	0.105	0.084	0.046
40°	friction-rich	14.4 / 38	0.014	0.086	0.123	0.109	0.025

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 07ESM3OnlineResource3Data.xlsx