Computers and Geotechnics 195 (2026) 108104

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo

Research paper

Accelerating stochastic simulation of post-failure landslide runout using a random graph neural network-based simulator

Yongjin Choi [a,b], Seungjun Lee [c,d,*]

[a] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, 30332, GA, USA
[b] KAIST InnoCORE PRISM-AI Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea
[c] Division of Physical Metrology, Korea Research Institute of Standards and Science (KRISS), Daejeon, 34113, Republic of Korea
[d] Department of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, 44919, Republic of Korea

ARTICLE INFO

Keywords:
Landslide runout
Graph neural network
Material point method
Random field
Probabilistic hazard assessment

ABSTRACT

Stochastic simulation of post-failure landslide behavior under spatially variable soil properties is essential for reliable hazard assessment, yet it remains challenging due to the inherent uncertainty of geomaterials and the high computational cost of high-fidelity numerical methods like the Material Point Method (MPM). This study proposes a Random Graph Network Simulator (RGNS) to accelerate probabilistic simulations of landslide post-failure processes in spatially heterogeneous soils. The RGNS leverages graph neural networks to emulate MPM-based granular flow behavior by learning local interaction laws among material points within a latent graph representation, enabling efficient, physics-consistent, and generalizable simulation of landslide runout dynamics. The RGNS is trained on a limited set of MPM simulations incorporating Gaussian random fields and is validated against multiple slope geometries and heterogeneity parameters outside the training configuration. The results demonstrate that RGNS accurately reproduces the probabilistic distributions of key post-failure metrics, including runout distance, sliding volume, and influence distance, with coefficients of determination mostly exceeding 0.93. With its computational efficiency, RGNS enables large-scale Monte Carlo (MC) simulations; specifically, 10,000 MC realizations of landslide runout are completed in approximately 3–4 days of dynamic simulation, compared to over 400 days using a conventional MPM solver. The application of RGNS to exceedance probability-based landslide hazard zoning reveals that deterministic analyses based on mean soil properties may significantly underestimate hazard extents in low-probability, high-consequence scenarios. These results demonstrate that RGNS provides a practical means for efficiently quantifying post-failure uncertainty under spatially variable soil conditions, offering a promising pathway for accelerating probabilistic landslide hazard assessment and risk-informed decision-making.

1. Introduction

Slope failures and their associated post-failure runout events pose significant hazards to civil infrastructure. Accurately assessing and managing these risks is challenging due to inherent uncertainties arising from complex terrain conditions, heterogeneous material properties, and the highly dynamic nature of slope movements. Among these uncertainties, the spatial variability of soil properties is a primary factor influencing both slope failure initiation and runout behavior (Wang et al., 2016; Griffiths et al., 2009; Liu et al., 2023).

Stochastic numerical modeling has been widely adopted to quantify the uncertainties in slope stability under spatially varying soil properties. In particular, researchers have developed random limit equilibrium methods (RLEM) (Cao et al., 2017; Cho, 2010; Jiang et al., 2015, 2017; Li et al., 2016a) and random finite element methods (RFEM) (Griffiths et al., 2009; Griffiths and Fenton, 2004; Xiao et al., 2016; Huang et al., 2010; Li et al., 2016a). These methods extend traditional LEM and FEM by integrating stochastic frameworks, such as Monte Carlo (MC) simulations, to evaluate the probabilistic nature of slope failures stemming from spatially varying soil properties represented by random field models.

Although RLEM and RFEM effectively capture uncertainties in slope failure initiation, they face limitations when simulating post-failure behaviors such as runout and progressive failure. These post-failure processes are crucial for assessing the consequences of slope failures, yet traditional methods struggle to model the large deformations involved in post-failure (Soga et al., 2016). LEM, for instance, is based on the force and moment equilibrium on a predefined failure surface, and does not account for kinematic and stress–strain evolution (Liu

* Corresponding author.
E-mail address: seungjun@kriss.re.kr (S. Lee).

https://doi.org/10.1016/j.compgeo.2026.108104
Received 13 January 2026; Received in revised form 10 March 2026; Accepted 26 March 2026
Available online 30 March 2026
0266-352X/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

et al., 2019). While FEM considers stress–strain behavior and failure surface formation, it suffers from mesh distortion when handling large deformations (Sordo et al., 2024). Consequently, neither method is suitable for accurately modeling landslide post-failure dynamics.

To overcome this limitation, the Material Point Method (MPM) has been increasingly used for large deformation analyses in slope failures (Ceccato et al., 2024; Troncone et al., 2023a,b). MPM employs a hybrid Eulerian–Lagrangian framework, allowing it to track large deformations without mesh distortion and making it well-suited for simulating post-failure dynamics such as runout. However, MPM remains computationally expensive due to the repetitive mapping between background meshes and material points throughout simulations. This computational burden becomes even more pronounced when MPM is applied in stochastic analyses that require numerous simulations to account for the random variability of model parameters.

Wang et al. (2016) proposed the random MPM (RMPM) approach to address post-failure uncertainties accounting for spatial variability of soil properties in progressive slope failures. Liu et al. (2019) extended RMPM to investigate uncertainties related to slope stability and failure modes. To mitigate the high computational cost associated with RMPM, they incorporated the RLEM as a preliminary filter to identify failure occurrences before applying RMPM for large deformation analysis. Y. Jiang et al. (2024) proposed the Stochastic Material Point Method (CSMPM), which improves computational efficiency by employing multi-resolution grids within MPM, facilitating more practical probabilistic assessments of slope failures. While previous studies have demonstrated the effectiveness of MPM-based approaches for modeling post-failure landslide behavior (Troncone et al., 2025), they all point out the MPM's significant computational burden. This remains the major bottleneck for the large-scale stochastic process.

To address the computational bottleneck associated with MPM-based uncertainty quantification, researchers have explored statistical and machine learning (ML)-based surrogate models as efficient alternatives. These models establish a nonlinear functional mapping between key influence factors—such as geometry, boundary conditions, and material properties—and associated landslide risks, including runout behavior (Zeng et al., 2021; Ju et al., 2022; Li et al., 2016b). By approximating the complex relationships governing slope failures into low-dimensional mapping, surrogate models significantly reduce computational costs while maintaining reasonable accuracy.

Despite their efficiency, this low-dimensional mapping oversimplifies the granular flow process, resulting in the loss of the full post-failure dynamics. Additionally, these models lack generalizability beyond their training datasets because they do not inherently incorporate the governing physical laws of slope failures. As a result, while ML-based models can predict travel distances and probable paths of granular flows based on terrain and material properties within the training distribution (Durante and Rathje, 2021; Ju et al., 2022; Zeng et al., 2021), they cannot reconstruct the full granular flow process during the landslide. Similarly, response surface-based surrogate models (Zhou et al., 2021; Cho, 2009; Li et al., 2015) facilitate computationally efficient slope reliability analysis, but they are not designed to simulate the full failure process, and the adaptability to complex geometries is limited. These shortcomings highlight the need for surrogate models capable of capturing granular flow behavior more comprehensively.

Recent advancements in learned physics simulators (Battaglia et al., 2016, 2018; Sanchez-Gonzalez et al., 2020) present a promising solution to these challenges. Among these, graph neural network (GNN)-based simulators (GNS) have emerged as an effective surrogate model for MPM in granular flow simulations. Unlike traditional surrogates that rely on static input–output mappings, GNS learns to model the full temporal evolution of granular flows by explicitly modeling local material interaction laws that govern the system dynamics within a graph-based framework. By leveraging these learned physical interactions, GNS can accurately simulate complex flow dynamics and generalize to previously unseen configurations, including novel geometries and boundary conditions. This generalizability stems from inductive biases (Battaglia et al., 2018) inherent in the GNS architecture—specifically, translational and permutation invariance and the multi-step message-passing mechanism—which allow the model to learn local, relational dynamics that are independent of the coordinate system. By iteratively updating node states based on neighboring interactions, GNS captures learned conservation principles governing the system's dynamics. Furthermore, GNS achieves substantial computational efficiency by operating on latent representations of local interactions rather than explicitly solving global governing equations.

Recent studies have demonstrated the effectiveness of GNS for granular flow problems. Choi and Kumar (2024a) and Zhao et al. (2025) developed GNS for granular flow simulations and showed that it can accelerate simulation by orders of magnitude compared to traditional numerical solvers. Choi and Kumar (2024b) introduced a differentiable GNS with automatic differentiation, which offers an efficient framework for solving inverse and design problems in granular flows. Due to its high efficiency and generalizability, GNS presents itself as a promising surrogate model for granular flow simulators like MPM, which can accelerate probabilistic slope post-failure analysis. However, the extent to which GNS can reliably model granular flows with spatially heterogeneous material properties—especially in the context of stochastic processes—remains an open question.

This study proposes a random GNS (RGNS) approach, illustrated in Fig. 1, to accelerate the uncertainty quantification of post-failure slope behavior under spatially varying material properties. In line with RMPM and RFEM frameworks, the stochasticity of RGNS arises from

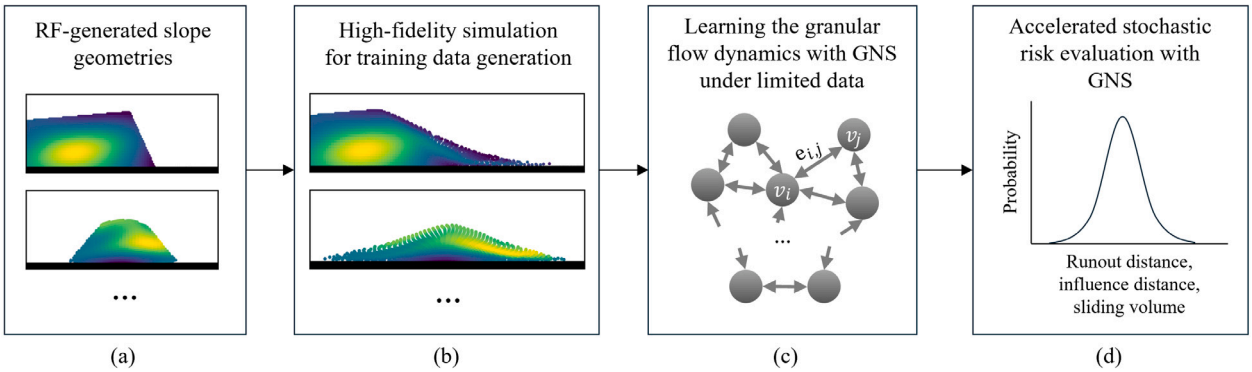

Fig. 1. Schematic overview of the Random Graph Neural Simulator (RGNS) framework for accelerating stochastic slope post-failure risk analysis. (a) Slope geometries with spatially varying soil properties are generated using random field (RF) theory to construct training datasets. (b) The material point method (MPM) is used to simulate slope runout processes for the generated realizations. (c) Resulting granular flow trajectories are used to train the graph network-based simulator (GNS), which learns the underlying flow dynamics and enables physics-consistent simulation of slope failure. (d) Monte Carlo (MC) simulations are then performed using the trained GNS to efficiently characterize the probabilistic post-failure response under spatially variable soil conditions.

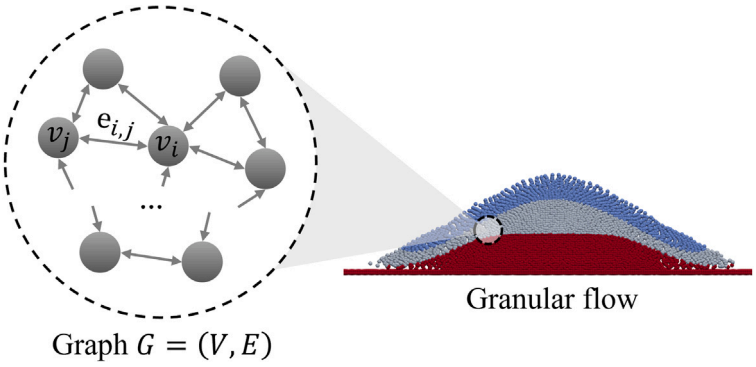

Fig. 2. Graph representation of granular flow. Graph $G = (V, E)$ consists of a set of vertices $V = \{v_1, v_2, \dots, v_i\}$, and a set of edges connecting a pair of vertices $E = \{e_{1,1}, e_{1,2}, \dots, e_{i,j}\}$.

the representation of spatial heterogeneity, while the GNS provides a deterministic simulation backbone that emulates the underlying flow dynamics. We first train the GNS on a limited set of slope runout failure trajectory data generated by high-fidelity simulations (i.e., MPM), which include spatially varying soil properties generated using random field theory. Then, we perform stochastic evaluations on previously unseen slope configurations, including unseen geometries and different random field parameters, to evaluate the statistical response of the key post-failure characteristics, such as runout distance, sliding volume, and influence distance. Additionally, we demonstrate a practical application of RGNS to accelerate the evaluation of the post-failure risk uncertainty, which is impractical with the conventional MPM due to computational overhead.

The manuscript is organized as follows: Sections 2.1 and 2.2 provide a brief review of GNS. Section 2.3 presents the stochastic method used in this study for representing slopes with spatially varying soil properties. Section 3 describes the training of RGNS. Sections 4 and 5 show our results and application to risk assessment.

2. Methodology

2.1. Graph representation of granular system

GNS represents a physical state of granular flow as a graph, $G = (V, E)$ (see Fig. 2). Graphs consist of a set of vertices $V = \{v_1, v_2, \dots, v_i\}$ representing the soil particles or discretized regions, and a set of edges connecting a pair of vertices $E = \{e_{1,1}, e_{1,2}, \dots, e_{i,j}\}$ representing the interaction between the particles or the discretized regions. In MPM, the model that we use to generate training data, vertices correspond to material points, and their interactions are encoded through the edges.

2.2. Graph neural network simulator

The Graph Network Simulator functions as an efficient and generalizable surrogate model that operates on graphs to emulate high-fidelity numerical simulators of granular media. GNS (see Fig. 3) takes the simulation domain's current state X_t, which contains particle positions, velocities, boundary information, and material characteristics, and predicts the subsequent state X_{t+1} (Eq. (1)).

$$X_{t+1} = GNS(X_t) \tag{1}$$

The simulation of granular flow using GNS follows an autoregressive approach, where predictions are iteratively fed back as input to forecast subsequent timesteps until reaching the desired timestep k (i.e., $X_0 \rightarrow$ **GNS** $\rightarrow X_1 \rightarrow$ **GNS** $\rightarrow \dots \rightarrow X_{k-1} \rightarrow$ **GNS** $\rightarrow X_k$). This sequential prediction process is referred to as a rollout. Subsequent paragraphs provide a brief overview of GNS's structure shown in Fig. 3. For more comprehensive details, refer to Sanchez-Gonzalez et al. (2020) and Choi and Kumar (2024a).

To perform the computation in Eq. (1), GNS uses two components: dynamics approximator $\mathcal{D}_\Theta$ (Eq. (2)) and update function $\mathcal{U}$ (Eq. (3)).

$$Y_t = \mathcal{D}_\Theta(X_t) \tag{2}$$

$$X_{t+1} = \mathcal{U}(X_t, Y_t) \tag{3}$$

The dynamics approximator $\mathcal{D}_\Theta$ is a learned function parameterized by Θ. It takes the current state X_t and predicts the dynamics Y_t of the granular flow using encoder–processor–decoder structure. The encoder ($G = \text{Encoder}(X_t)$) transforms the current state X_t into a latent graph $G = (V, E)$, which represents the particle interactions. The encoder uses multi-layer perceptrons (MLPs) with two hidden layers, each containing 128 units.

The processor ($G' = \text{Processor}(G)$) performs n rounds of message passing, which propagates information of graph vertices along the edges and returns the updated graph $G' = (V', E')$. This models energy or momentum transfer between particles. We adopt the message-passing formulation introduced in Battaglia et al. (2016). We perform $n = 10$ message-passing steps, which enables information to propagate beyond immediate neighbors and capture collective interaction effects. This choice follows the systematic sensitivity analysis of Sanchez-Gonzalez et al. (2020), where increasing the number of message-passing steps was shown to improve rollout accuracy by incorporating longer-range and more complex particle interactions, while incurring an increase in computational cost. Based on this analysis, Sanchez-Gonzalez et al. (2020) adopted $n = 10$ as a practical balance between interaction range and efficiency.

The decoder ($Y_t = \text{Decoder}(G')$) extracts the dynamics Y_t from the updated graph. The update function $\mathcal{U}$ ($X_{t+1} = \mathcal{U}(X_t, Y_t)$) then applies the predicted dynamics Y_t to transition the current state X_t to the next state X_{t+1}. In this setup, $\mathcal{U}$ operates similarly to explicit Euler integration in numerical solvers, where Y_t corresponds to the second derivative of the state, representing acceleration. The decoder uses the same MLP structure as the encoder.

These neural networks from the encoder–processor–decoder compose the learnable parameter set Θ of $\mathcal{D}_\Theta$. Θ is trained by minimizing the mean squared error (MSE) between the ground truth particle accelerations A_t and the predicted dynamics Y_t at the given particle state X_t (Eq. (4)).

$$\mathcal{L}(\Theta) = \frac{1}{M} \sum_{i=1}^{M} \left\| A_t^{(i)} - Y_t^{(i)} \right\|^2 \tag{4}$$

where M is the number of particles, $A_t^{(i)}$ and $Y_t^{(i)}$ denote the ground-truth and predicted accelerations of particle i, respectively.

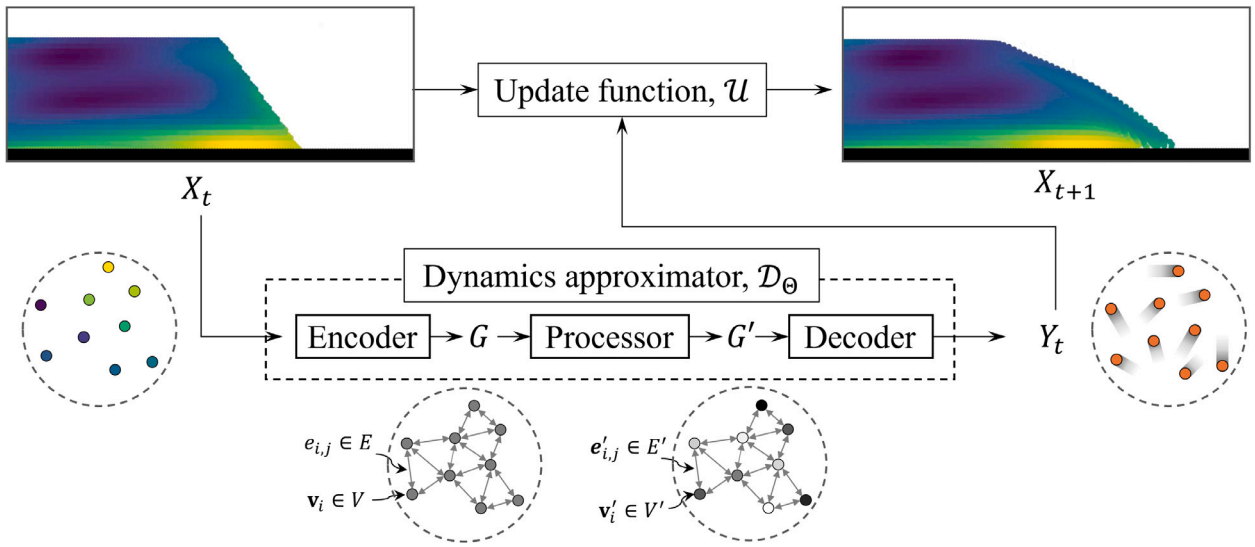

Fig. 3. Graph network simulator and its structure.

2.3. Random field generation

The spatial variability of soil properties is modeled as Gaussian random fields (GRFs) defined on a two-dimensional discretized domain. The field is represented by the vector:

$$\mathbf{H} = \{H(x_i, y_j)\}_{i=1,\ldots,n_x;\, j=1,\ldots,n_y} \tag{5}$$

where $H(x_i, y_j)$ denotes the field value at the grid point with coordinates (x_i, y_j), and n_x and n_y are the number of discretization points in the x- and y-directions, respectively. The total number of points is $N = n_x n_y$, so that $\mathbf{H} \in \mathbb{R}^{N\times 1}$. A GRF is characterized by its mean μ, variance σ^2, and an autocorrelation function (ACF) that specifies the spatial correlation between any two locations.

To model the correlation structure, we adopt the single exponential (SNX) autocorrelation function. We adopt the SNX autocorrelation function because it is a widely established baseline in geotechnical random-field analyses, particularly for modeling spatial variability in natural slopes when high-resolution, site-specific spatial data are unavailable (Jiang et al., 2022; Li et al., 2015). We note that alternative correlation models exist (e.g., the squared exponential (SQX) or second-order Markov (SMK) functions), and they produce smoother random fields and may influence quantitative predictions. Li et al. (2015) report that the choice of ACF does not have a critical influence on probability-of-failure estimates. However, its impact on probabilistic post-failure responses remains unclear and is an open research topic.

The correlation between two points (x_i, y_i) and (x_j, y_j) in SNX is given by Eq. (6).

$$\rho\big(H(x_i, y_i), H(x_j, y_j)\big) = \exp\left[-2\left(\frac{\Delta x}{\ell_x} + \frac{\Delta y}{\ell_y}\right)\right] \tag{6}$$

where $\Delta x = |x_i - x_j|$ and $\Delta y = |y_i - y_j|$ are the spatial separations in the horizontal and vertical directions, respectively. The parameters ℓ_x and ℓ_y denote the correlation lengths in the x- and y-directions. Larger values of ℓ_x or ℓ_y imply smoother fields with longer-range spatial dependence, while smaller values yield rougher realizations with short-range correlation. Therefore, ℓ_x and ℓ_y control the level of soil heterogeneity for the respective direction.

From this ACF, the covariance matrix $\mathbf{C}$ is defined as

$$C_{ij} = \sigma^2\, \rho\big(H(x_i, y_i), H(x_j, y_j)\big) \tag{7}$$

To generate realizations of the random field, we employ the Cholesky matrix decomposition (CMD) (Jiang et al., 2014; Liu et al., 2017):

$$\mathbf{C} = \mathbf{L}\mathbf{L}^\top. \tag{8}$$

where $\mathbf{L} \in \mathbb{R}^{N\times N}$ is a lower triangular matrix and $\mathbf{L}^\top$ is its transpose. We then draw a standard normal vector $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$, with $\mathbf{I}_N$ denoting the $N \times N$ identity matrix, and transform it as:

$$\mathbf{H} = \mu\mathbf{1} + \mathbf{L}\mathbf{z}, \tag{9}$$

where $\mathbf{1} \in \mathbb{R}^{N\times 1}$ is a vector of ones. $\mathbf{H}$ represents the anisotropic Gaussian random field over the domain with the prescribed μ, σ^2, and spatial correlation prescribed by ACF.

Finally, the vector $\mathbf{H}$ is reshaped into a two-dimensional array of size $n_x \times n_y$ and mapped onto the material points at their corresponding spatial locations, thereby assigning spatially varying soil properties within the numerical model for slopes.

Although CMD is widely used for random field generation due to its conceptual simplicity and ease of implementation (Li et al., 2015; Jiang et al., 2022), the computational cost scales with the number of discretization points. For very fine discretizations, and large or 3D domains, alternative approaches such as spectral representation methods (S.-H. Jiang et al., 2024) or Karhunen–Loève expansions (Cho, 2010) can provide improved scalability and reduced computational overhead. However, in our 2D setting, the random field generation accounts for only a negligible fraction of the total simulation time compared to the post failure runout computation (discussed later in Section 4.3). Therefore, CMD does not constitute a computational bottleneck in our workflow.

3. Data and training setup

We generate a dataset of slope runout trajectories using the CB-Geo Material Point Method (MPM) code (Kumar et al., 2019) to train the GNS. We adopt an explicit MPM time-integration scheme with a time step of 5×10^{-5} s and sample the simulation states every 0.025 s to construct the training data. Table 1 summarizes the MPM analysis configuration.

Fig. 4 shows two base pre-failure slope geometries that represent typical slopes and embankments in the training data. To increase geometric diversity, we randomly perturb the vertex coordinates of these base geometries within the prescribed minimum and maximum ranges in the x- and y-directions (indicated by the square brackets in the figure) using uniform distributions. These perturbations generate a wide range of slope configurations and corresponding post-failure runout behaviors, which inform model training.

The spatially varying material fields (friction angle, ϕ) are generated using the random field approach described in Section 2.3, with

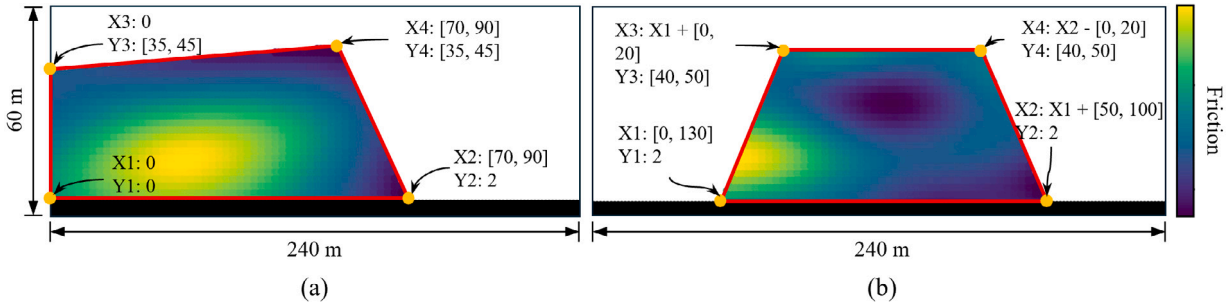

Fig. 4. Configurations of pre-failure slope geometries for generating the training datasets. The figure is not in scale. The square brackets represent the min and max variation ranges of the vertices' x and y locations.

Table 1
MPM configuration and statistical parameters of friction angle random field for training data.

Property		Fig. 4(a)	Fig. 4(b)
MPM	Domain size (m)	240 × 60	
	Cell size (m)	2 × 2	
	# material points per cell	4	
	dt (s)	5e−5	
Random field (ϕ)	μ (°)	20–35[a]	20–35[a]
	σ (°)	5	5
	ℓ_x (m)	40	30
	ℓ_y (m)	25	15

[a] The values represent uniform sampling ranges.

parameters summarized in Table 1. We later assess the generalization of GNS on random fields drawn from different parameter sets. The MPM analysis configuration is also provided in Table 1. In total, 1000 slope geometries and their corresponding runout trajectories are generated for training. Additional 50 slopes are assigned for validation data.

The GNS training requires the selection of the following hyperparameters: connectivity radius, and number of message passing steps, which have been identified as the most influential parameters to the performance (Sanchez-Gonzalez et al., 2020). The connectivity radius specifies the maximum distance within which two material points are connected by a graph edge from a vertex. Larger connectivity radius captures more local interactions, but entails increased computational cost. We set the radius to 2.0 times the MPM cell size provided in Table 1 based on the analysis from Zhao et al. (2025). The number of message passing steps defines how many rounds of message passing occur between connected graph vertices. Increasing this value improves the model's capacity to capture long-range interactions but also raises computational cost. We use 10 steps, which balance the learning capacity and computation cost (Choi and Kumar, 2024a; Sanchez-Gonzalez et al., 2020).

For model training, we employ the Adaptive Moment Estimation (ADAM) optimizer (Kingma, 2014) with an initial learning rate of $lr_{init} = 1 \times 10^{-4}$, scaled by the number of GPUs (n_{gpu}). In our setup, $n_{gpu} = 2$ with a batch size of 2 per GPU. The learning rate follows an exponential decay schedule $lr(i) = lr_{\text{init}} \cdot 0.1^{\,i/5\times10^6} \cdot n_{\text{gpu}}$, where i is the training step, which is also adopted by the previous studies (Choi and Kumar, 2024a; Zhao et al., 2025).

Fig. 5 shows the learning curves. The training loss is evaluated at every training step, and the validation loss is evaluated every 100,000 steps as the average over 100 randomly sampled examples from the validation set. Both the training and validation losses decrease steadily during the initial phase and then gradually plateau, while maintaining consistent downward trends, indicating stable optimization without signs of overfitting. Because no meaningful improvement is observed beyond 2.5 million steps, training is manually terminated at that point.

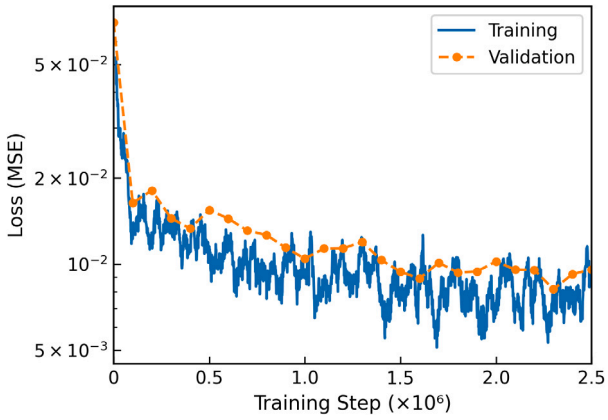

Fig. 5. Learning history of GNS.

4. Stochastic performance evaluation

4.1. Evaluation metrics and definitions

To evaluate the stochastic behavior of slope runout and its associated deformation process, three representative post-failure metrics were adopted to characterize the extent and magnitude of the failed mass movement: runout distance (R), sliding volume (V), and influence distance (I) (Fig. 6). Together, these metrics define the primary quantities of interest used to evaluate the stochastic characteristics of post-failure deformation (Ma et al., 2022).

Specifically, the runout distance was defined as the horizontal travel distance of the flow front, measured from its initial position to the location corresponding to the 99th-percentile of the displaced mass. This definition effectively filters out disconnected or isolated material points and yields a representative measure of overall mobility, consistent with those in Vicari et al. (2025) and Li et al. (2021). While the absolute runout distance may vary with the chosen percentile (e.g., 95% and 99%), a brief sensitivity check suggested that the predictive performance and overall trends are not substantially affected. We therefore use the 99th percentile to better capture the runout extent relevant to hazard assessment. The sliding volume quantifies the total mobilized material exceeding a threshold displacement, while the influence distance measures the horizontally retreated extent of the mobilized zone from the slope crest. A displacement threshold of 1 m was adopted following the previous studies (Liu et al., 2019; Ma et al., 2022). These definitions delineate the mobilized and stable zones, enabling a quantitative comparison of probabilistic characteristics between the MPM and RGNS simulations.

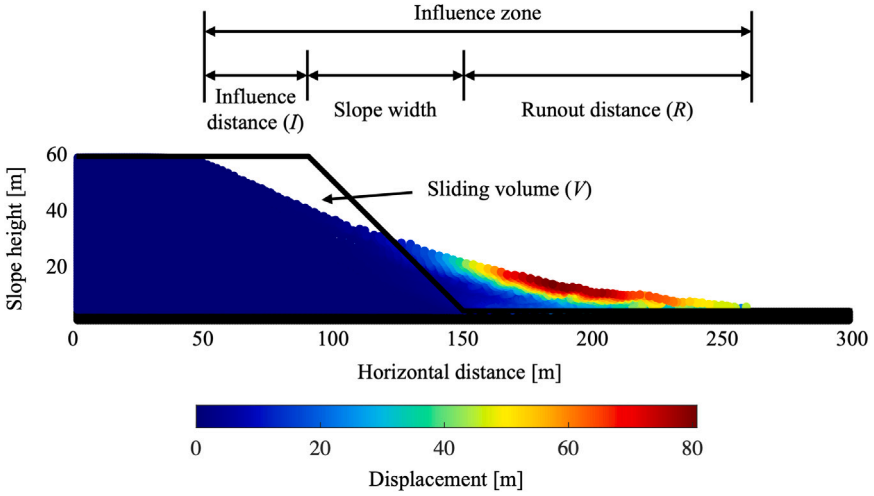

Fig. 6. Schematic illustration of post-failure deformation for landslide runout analysis.

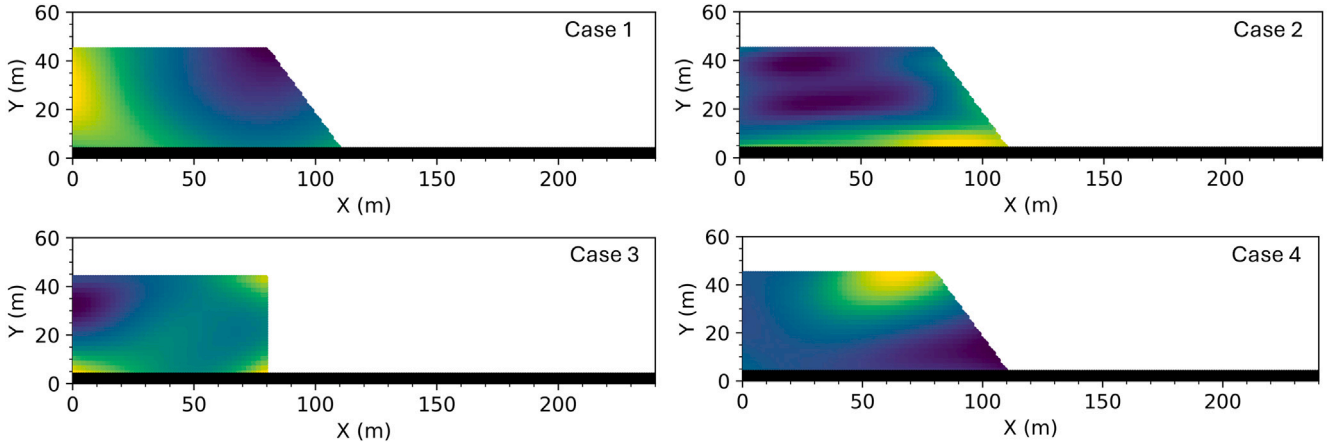

Fig. 7. Test case geometries and an example random field realizations. This video provides simulation examples from RGNS and MPM for our test cases: https://youtu.be/ZgxtTIuGenY.

4.2. Stochastic test cases for generalization analysis

The trained RGNS was evaluated on three stochastic test cases (Fig. 7 and Table 2) designed to examine its generalization capability beyond the training distribution in terms of geometry and random field parameters. All test cases were generated using the random-field-based approach described in Section 2.3, but they differ in both geometry and random-field characteristics:

- **Case 1:** An elongated slope with a larger vertical correlation length (ℓ_y = 30 m), producing smoother vertical variability than in the training data.
- **Case 2:** An elongated slope with a smaller ℓ_y = 10 m, yielding stronger vertical heterogeneity than the training data.
- **Case 3:** A 90° slope representing a granular-column-type failure, designed to test the model's generalization to a different geometry.
- **Case 4:** An elongated slope with the same geometry and spatial correlation structure as Case 1, but with a higher mean friction angle ($\mu = 35°$) and lower standard deviation ($\sigma = 3°$).

For each case, we generated 250 realizations of the slopes with spatially varying friction-angle fields, and the corresponding runout behaviors were simulated using the MPM and trained RGNS. This video provides simulation examples from MPM and RGNS for our test cases: https://youtu.be/ZgxtTIuGenY.

Table 2
Configurations of test cases with varying geometries and random field parameters.

Test case	Geometry	Random field parameters			
		μ (°)	σ (°)	ℓ_x (m)	ℓ_y (m)
1	Longer slope	30	5	50	30
2	Longer slope	30	5	40	10
3	90° slope	30	5	50	30
4	Longer slope	35	3	50	30

Previous studies model spatial variability using different strength parameters, including undrained shear strength, c, and ϕ (Jiang et al., 2022). This study focuses on ϕ. This choice is to be consistent with the current development status of GNS-based granular flow modeling, which has primarily been established for frictional materials (Choi and Kumar, 2024a; Zhao et al., 2025).

4.3. Comparison between RGNS and MPM simulations

The comparison between MPM and RGNS predictions was conducted in terms of the statistical distributions of the runout distance R, sliding volume V, and influence distance I. Figs. 8 to 11 and Table 3 collectively demonstrate that the RGNS captures the stochastic distributions of the three post-failure metrics across all test cases.

To quantitatively assess the agreement between the RGNS and MPM simulations, two statistical indicators were employed: the coefficient of determination (R^2) and the normalized root mean square error (NRMSE), defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}, \quad \text{NRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}}{y_{\max} - y_{\min}} \tag{10}$$

where y_i and $\hat{y}_i$ denote the MPM and RGNS results for the ith realization, respectively; $\bar{y}$ is the mean of the MPM results, and $y_{\max}$ and $y_{\min}$ are the maximum and minimum values of the MPM dataset. Higher R^2 and lower NRMSE values indicate better consistency between the surrogate and the high-fidelity reference model.

Generally, RGNS successfully reproduces both the mean and variance of the MPM-derived results, with a slight systematic underestimation observed in some instances. Quantitatively, the R^2 values for all three metrics exceed 0.89 in every test case, confirming a strong correlation between the surrogate and the high-fidelity simulations. In particular, the runout distance shows $R^2 > 0.96$ with an NRMSE below 0.141, indicating that the RGNS accurately captures the flow front of the failed mass. The sliding volume exhibits the best agreement, with R^2 up to 0.986 and NRMSE less than approximately 0.04, whereas the influence distance shows slightly larger variability ($R^2 = 0.89$–0.96).

The probability distributions of the key post-failure metrics predicted by RGNS are in close agreement with those obtained from the MPM simulations, as illustrated in Figs. 8 to 11. The histogram-based distributions of RGNS predictions generally follow the MPM reference results, demonstrating consistent probabilistic trends across all test cases. A consistent leftward shift is observed in the runout distance distribution, reflecting the small underestimation noted above. To quantify this bias, we analyzed the paired runout differences across realizations and computed bootstrap 95% confidence intervals for the mean signed error (Efron and Tibshirani, 1993). The intervals exclude zero for all test cases, indicating a systematic underestimation of runout distance, with relative bias values on the order of 13–15%. This discrepancy is attributable to the surrogate nature of RGNS, which approximates high-fidelity dynamics through learned representations and can therefore introduce systematic biases. Such errors can be more pronounced for metrics that depend on long-horizon predictions, such as runout distance, because the autoregressive formulation of the GNS (Eq. (1)) leads to error accumulation over extended timesteps (Lucchetti et al., 2025).

Mechanistically, particles that travel longer distances experience longer rollout trajectories and may therefore accumulate slightly larger absolute prediction errors. Consistent with this, additional analysis in this study showed a clear positive association between runout distance and the final-state particle-position MSE across all four test cases, with Spearman correlation coefficients ranging from 0.72 to 0.88. Nevertheless, the stochastic simulations considered in this study remain in good agreement with the MPM reference results for the key runout metrics (see Table 3), suggesting that the impact of cumulative error remains limited within the range of landslide runout distances examined here. Additionally, the prediction maintains overall physical consistency as discussed with Fig. 12. A more systematic investigation would be required to rigorously quantify how cumulative errors scale with increasing runout distance. As a practical consideration, a simple post-hoc scaling could reduce the mean bias; however, we do not introduce a correction factor here. Its appropriate form is likely to depend on the simulation configurations (e.g., geometry, material properties) and runout regime, and establishing a general calibration would require validation under a wider range of simulation configurations, which is beyond the scope of this study.

Note that, while previous studies have demonstrated the generalizability of GNS to unseen geometries (Sanchez-Gonzalez et al., 2020; Choi et al., 2025; Zhao et al., 2025), the present results further suggest that GNS shows potential to maintain consistent probabilistic performance under material heterogeneity characterized by random-field parameter settings not included in the training distribution. This capability can be attributed to the fact that GNS learns local interaction laws between soil bodies that govern granular flow dynamics, rather than relying on direct mappings tied to specific realizations. The observed generalization, however, is assessed within the range of heterogeneity conditions considered in this study.

GNS is a highly efficient granular flow simulator. Previous studies have reported computational speedups of approximately 100× to 1000× relative to MPM, depending on the hardware acceleration used for the MPM solver (e.g., CPU- or GPU-parallelized implementations) (Zhao et al., 2025; Choi and Kumar, 2024a). In this study, we directly compared GNS with the MPM solver used herein (CB-Geo MPM; Kumar et al., 2019) using test case 1, which comprises 5855 material points. The CB-Geo MPM simulations were executed using a CPU-parallelized implementation on Intel Cascade Lake processors with 56 cores (128 GB RAM), while the GNS simulations were run on a single NVIDIA A100 GPU (40 GB memory). RGNS completed the simulation in approximately 31 s, whereas MPM required 3675 s, corresponding to a speedup of about 120×. This efficiency could be further optimized under different GNS model configurations, such as the number of message-passing steps or the connectivity radius (Choi and Kumar, 2024a).

Because the CB-Geo MPM implementation used in this study does not support GPU implementation, a direct same GPU hardware comparison is not feasible. Nevertheless, prior studies provide useful benchmarks for contextualizing the expected performance difference. Zhao et al. (2021) compares GNS with their GPU-accelerated MPM and reports performance gains of approximately 50–100× in the same device, depending on the model configuration. Choi and Kumar (2024a) reports that GPU-accelerated MPM implementations (Arduino et al., 2021) achieve speedups of roughly 5–10× relative to the CPU parallelized CB-Geo MPM. This information suggests that a meaningful performance advantage of GNS over a GPU-accelerated MPM would still be expected.

To further examine the physical consistency of the RGNS predictions beyond statistical agreement, we compare the energy evolutions during the slope runout failure between RGNS and MPM simulations. Fig. 12 shows the normalized energy evolutions of MPM and GNS over time for the best, median, and worst case predictions for test case 4, selected based on particle-position MSE. The same plots for the other test cases are included in Appendix A. The results show that RGNS well reproduces the general energy evolution trends observed in the MPM simulations: the initial potential energy gradually decreases as the failed mass descends, while kinetic energy first increases due to acceleration and subsequently dissipates as the motion stabilizes. Correspondingly, the dissipated energy increases monotonically, reflecting energy loss through frictional dissipation during granular flow. For the best and median prediction cases, RGNS closely follows the MPM energy trajectories. Although the worst case prediction shows a larger deviation at the final timestep, the overall trends remain reasonably consistent. These results indicate that RGNS preserves the overall physical energy evolution trends governing flow dynamics, providing additional evidence of the physical consistency of the RGNS predictions.

5. Application: Accelerated landslide hazard zoning based on exceedance probability

As a practical application, RGNS is applied to quantify the exceedance probability of post-failure displacements under spatially variable soil conditions represented by random fields. Hazard exceedance is evaluated in terms of runout and influence distances, which indicate whether post-failure deformation reaches predefined critical locations. The overall procedure consists of three components: (1) generation of random field samples representing heterogeneous soil properties; (2) RGNS-based simulation of post-failure granular flow for each realization; and (3) statistical evaluation of the resulting indicators to compute exceedance probabilities.

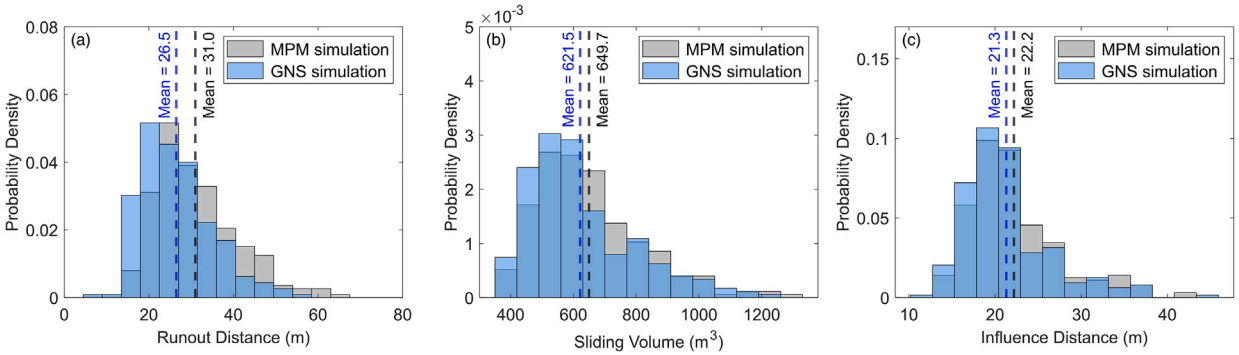

Fig. 8. Histogram comparison between MPM and GNS simulations for test case 1.

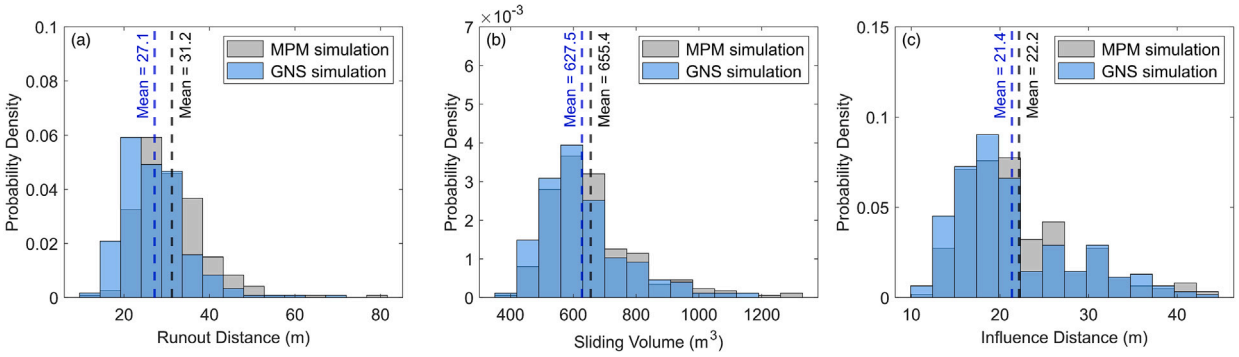

Fig. 9. Histogram comparison between MPM and GNS simulations for test case 2.

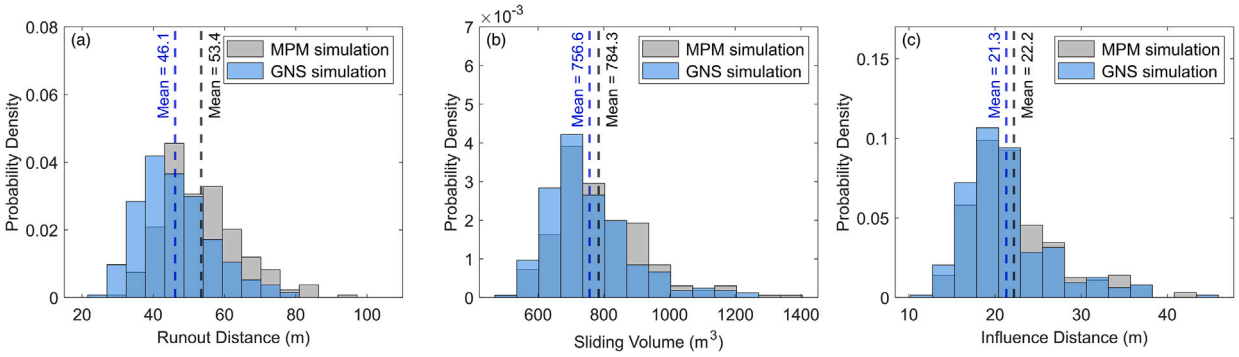

Fig. 10. Histogram comparison between MPM and GNS simulations for test case 3.

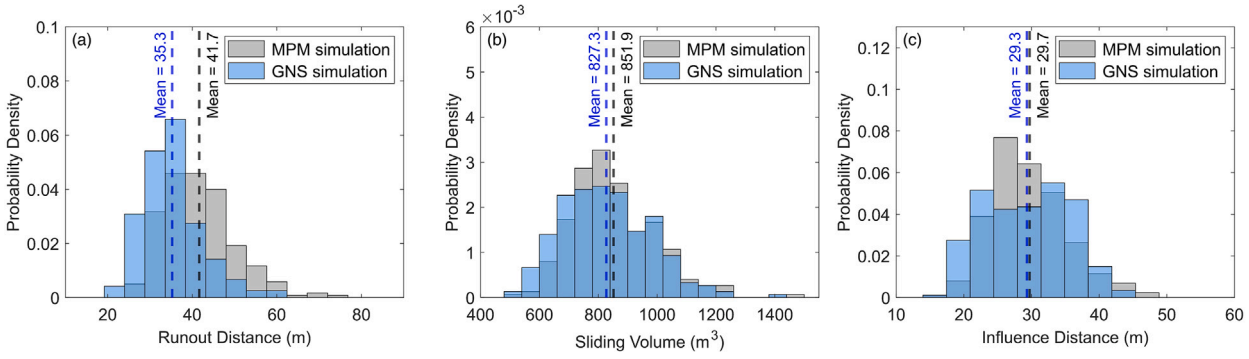

Fig. 11. Histogram comparison between MPM and GNS simulations for test case 4.

Table 3
Comparison of GNS predictions with MPM reference results across four stochastic test cases.

Metric	Test case 1		Test case 2		Test case 3		Test case 4	
	R^2	NRMSE	R^2	NRMSE	R^2	NRMSE	R^2	NRMSE
Runout distance (R)	0.983	0.080	0.963	0.067	0.964	0.117	0.974	0.141
Sliding volume (V)	0.986	0.037	0.967	0.043	0.983	0.039	0.971	0.038
Influence distance (I)	0.962	0.048	0.890	0.066	0.935	0.056	0.905	0.062

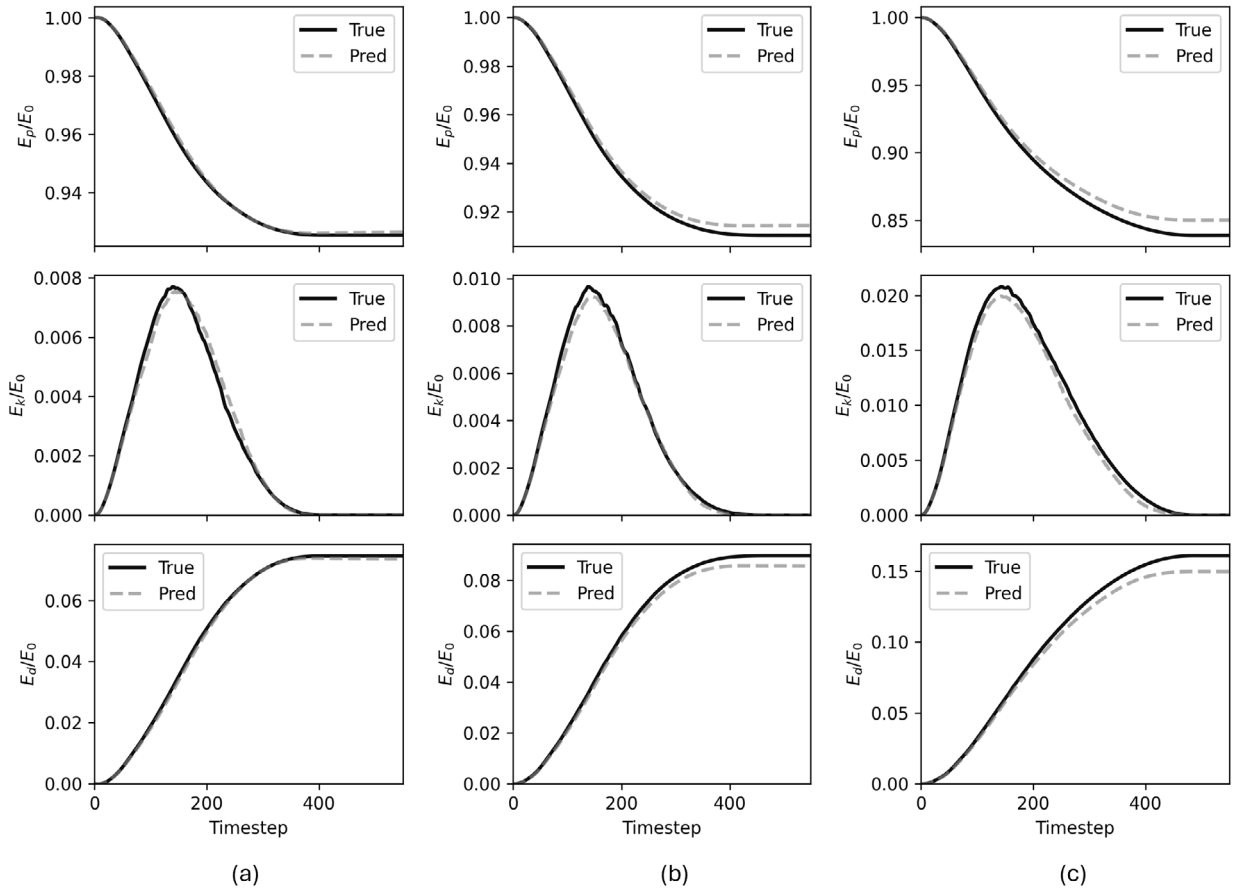

Fig. 12. Normalized energy evolution of MPM (True) and GNS (Pred) over time for (a) best, (b) median, and (c) worst case predictions for test case 4. Each row corresponds to potential energy E_p, kinetic energy E_k, and dissipation energy E_d, respectively, normalized by the initial potential energy E_0. The energy components are defined as: $E_p = \sum_{i=1}^{M} m_i g h_i$, $E_k = \frac{1}{2}\sum_{i=1}^{M} m_i v_i^2$, and $E_d = E_0 - (E_p + E_k)$, where M is the total number of material points, and m_i, h_i, and v_i represent the mass, height, and velocity of material point i, respectively.

To evaluate whether post-failure displacements exceed critical distances, the hazard assessment adopts a limit-state formulation in which each random realization is compared against predefined threshold values. The thresholds I_s and R_s represent the allowable influence distance and the allowable runout distance, respectively. These typically correspond to the locations of critical structures or zones of interest, such as roads, railways, bridges, or residential areas. For each realization, denoted by $\mathbf{x}$, the RGNS simulation provides the influence distance $I(\mathbf{x})$ and runout distance $R(\mathbf{x})$, which are evaluated using the following limit-state functions:

$$G_I(\mathbf{x}) = I_s - I(\mathbf{x}) \tag{11}$$

$$G_R(\mathbf{x}) = R_s - R(\mathbf{x}) \tag{12}$$

When $G_I(\mathbf{x}) < 0$ or $G_R(\mathbf{x}) < 0$, the predicted displacement exceeds the respective threshold, indicating that the landslide would reach the specified location. The exceedance probabilities are then expressed as

$$P(I > I_s) = P\left[G_I(\mathbf{x}) \le 0\right] = \int_{G_I(\mathbf{x})\le 0} f_{\mathbf{x}}(\mathbf{x})\, d\mathbf{x} \tag{13}$$

$$P(R > R_s) = P\left[G_R(\mathbf{x}) \le 0\right] = \int_{G_R(\mathbf{x})\le 0} g_{\mathbf{x}}(\mathbf{x})\, d\mathbf{x} \tag{14}$$

where $f_{\mathbf{x}}(\mathbf{x})$ and $g_{\mathbf{x}}(\mathbf{x})$ denote the joint probability density functions of the random field variables. In this study, these probabilities are estimated via MCS as the proportion of realizations exceeding the limit state, and are subsequently used to quantify the hazard level at the locations of interest.

The accelerated probabilistic assessment is applied to test case 1, characterized by horizontal and vertical correlation lengths of ℓ_x = 50 m and ℓ_y = 30 m, respectively. A total of 10,000 random field realizations are generated, and RGNS is employed to simulate the corresponding post-failure behavior for each realization within the MC framework. In general, the sample size required for reliable probabilistic evaluation depends on problem-specific factors such as slope geometry, domain size, the choice of random variables, and the degree of spatial heterogeneity. Previous studies empirically observed that several hundred to a few thousand simulations are required to achieve statistical stabilization (Wang et al., 2016; Liu et al., 2019, 2021; Ma et al., 2022; Liu et al., 2023). To determine an adequate sample

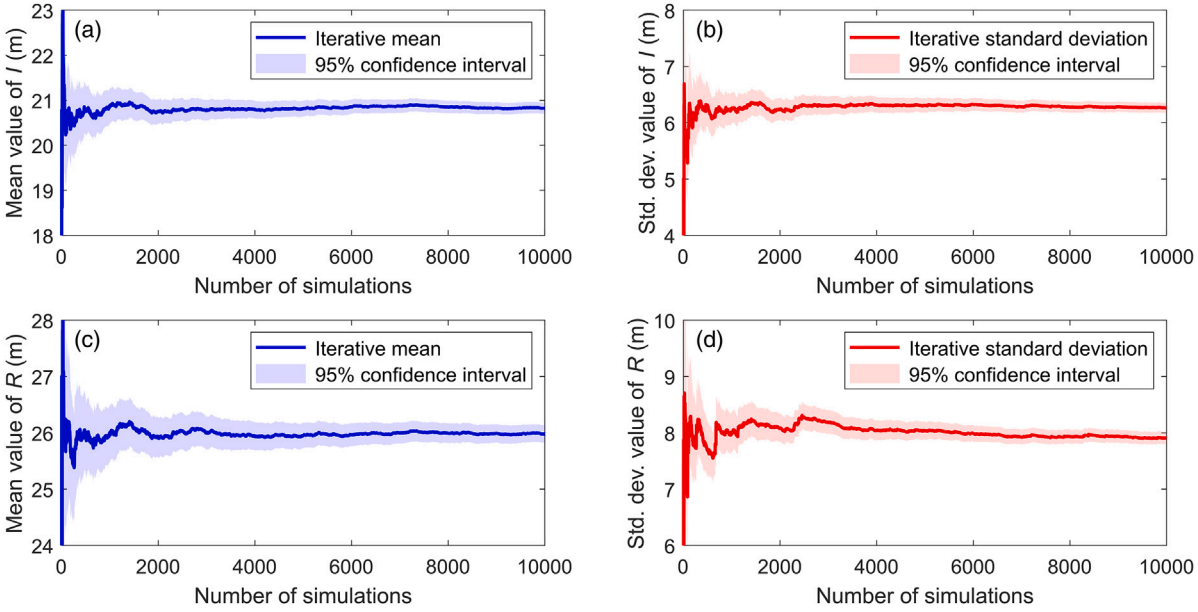

Fig. 13. Convergence monitoring of MC simulations for I and R.

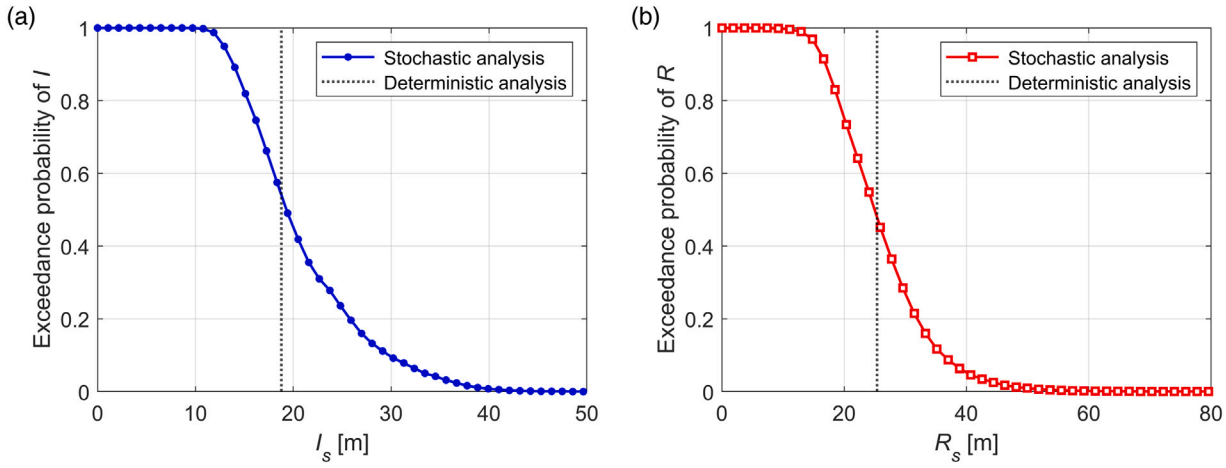

Fig. 14. Exceedance probability curves for (a) influence distance and (b) runout distance.

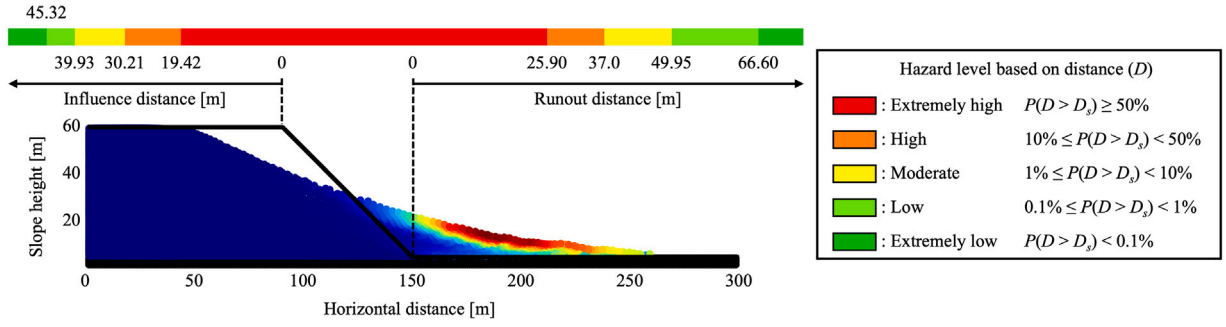

Fig. 15. Probabilistic landslide hazard zoning based on exceedance probability of post-failure distance.

size for our problem, we monitor the evolution of the running mean and standard deviation of I and R as in previous studies. Fig. 13 illustrates the convergence behavior with the 95% confidence intervals. The fluctuations in both indicators gradually diminish as the number of simulations increases. Following Ma et al. (2022), we assessed convergence using an absolute tolerance of 10^{-3} on successive estimates of the mean and variance; the two statistics were monitored independently, and convergence was declared only when both satisfied the tolerance. Convergence is observed at approximately 6000 simulations for I and 7500 for R. Although convergence is achieved prior to reaching the full sample size, all $10,000$ simulations are utilized in the probabilistic evaluation to ensure stable statistics and maintain consistency across subsequent analyses. While the convergence analysis indicates that substantially more simulations are required than those reported in the

previous studies (Wang et al., 2016; Liu et al., 2019, 2021; Ma et al., 2022; Liu et al., 2023), the RGNS framework facilitates this large-scale stochastic evaluation.

Fig. 14 presents the exceedance probability curves derived from RGNS-based MC simulations. For the influence distance I measured from the slope crest, the median exceedance threshold (i.e., 50% exceedance probability) occurs at approximately 19.42 m, which is close to but slightly higher than the deterministic prediction based on the mean soil properties (i.e., 18.78 m). However, lower exceedance probabilities correspond to substantially greater distances; for instance, thresholds of approximately 30.21 m and 39.93 m are associated with 10% and 1% exceedance probabilities, respectively. This indicates that retreating deformation may extend significantly beyond deterministic estimates under critical realizations of spatial soil variability. A similar trend is observed for the runout distance, where the median exceedance threshold of approximately 25.90 m remains close to the deterministic result (i.e., 25.37 m). In contrast, runout distances of approximately 37.0 m and 49.95 m correspond to 10% and 1% exceedance probabilities, respectively. In both metrics, the exceedance probability decreases rapidly as the threshold distance increases and eventually approaches zero. These quantitative differences between deterministic and probabilistic predictions demonstrate that relying solely on mean-field analyses may underestimate the spatial extent of potential landslide impacts, particularly in the low-probability tail of the response distribution.

Fig. 15 illustrates the probabilistic landslide hazard zoning map derived from the exceedance probability of the post-failure distance (D). Here, D serves as a generic metric representing either the influence distance or the runout distance, depending on the specific area of concern. For any given spatial point, the threshold distance D_s is defined as the horizontal distance from the source to that location. The hazard level is then quantified by the exceedance probability $P(D > D_s)$, representing the likelihood that post-failure deformation will reach or exceed the specified point. The classification of these hazard levels follows the exceedance-probability-based framework proposed by Ma et al. (2022), which adopts the event likelihood descriptors introduced by Lacasse and Nadim (2011). Accordingly, five hazard levels—extremely high, high, moderate, low, and extremely low—are defined based on specific probability ranges. This systematic approach ensures a consistent and interpretable mapping between quantitative exceedance probabilities and qualitative hazard descriptors.

This zoning framework translates the exceedance probability curves shown in Fig. 14 into spatially intuitive hazard information. When coupled with the high computational efficiency of RGNS, it enables probabilistic hazard zoning to be performed at a practically feasible scale. Specifically, while conducting 10,000 stochastic simulations for robust uncertainty quantification would require over 400 days using high-fidelity MPM alone, the RGNS completes the same analysis within a few days (i.e., approximately 3–4 days). This acceleration allows for the stable estimation of exceedance probabilities, particularly in the low-probability tails of the response distribution, which is critical for risk-informed decision-making in spatially heterogeneous soils.

We note that the reported computational times refer to the dynamic runout simulation stage only and do not include preprocessing steps such as random field generation. For reference, generating 10,000 random field realizations requires approximately 73 min, which is negligible compared to the total simulation time. Therefore, the reported acceleration primarily reflects the difference in solving the post-failure dynamics between MPM and RGNS.

6. Conclusion

This study proposes a Random Graph Neural Simulator (RGNS) for accelerating stochastic simulations of post-failure landslide behavior under spatially variable soil properties. By learning granular flow dynamics from high-fidelity Material Point Method (MPM) simulations that incorporate random fields, RGNS provides an efficient surrogate framework for large-scale uncertainty quantification of landslide runout processes.

We evaluate the stochastic performance of RGNS across multiple test cases, including not only slope geometries, but also material heterogeneity characterized by random-field parameter settings unseen during the training. RGNS accurately reproduces the probabilistic distributions of key post-failure metrics, including runout distance, sliding volume, and influence distance. Across the evaluated test cases, coefficients of determination generally exceed 0.93, with values above 0.96 for runout distance, demonstrating strong agreement with MPM reference simulations.

Owing to its high computational efficiency, RGNS enables large-scale probabilistic analyses that are impractical using conventional MPM. In this study, 10,000 stochastic realizations were simulated to obtain stable estimates of exceedance probabilities. While such analysis would require more than 400 days using MPM alone, RGNS completed the same task within approximately 3 to 4 days. This efficiency facilitates the application of exceedance-probability-based hazard zoning frameworks, particularly for quantifying low-probability, high-consequence scenarios that are commonly underestimated by deterministic analyses.

Despite its advantages, the proposed RGNS framework has several limitations. The approach entails upfront costs, as it requires high-fidelity MPM training data and associated one to two days of training time (Choi and Kumar, 2024a). Its simulation capability is confined to the material models and their parameter ranges of the training dataset. Although RGNS demonstrates good generalization across unseen geometries and spatial statistical characteristics of material fields, it does not generalize across fundamentally different constitutive behaviors. For example, the current RGNS is trained to represent Mohr–Coulomb material behavior only. In addition, predictions in the extreme tail of the response distribution should be treated with caution, as rare-event estimates may be more sensitive to accumulated errors during autoregressive rollout. The present study is also limited to dry materials and does not explicitly account for hydro-mechanical coupling. Addressing these aspects would require substantial extensions to both the training data and model network formulation development. Future research can address these limitations by incorporating additional physical processes and broader material representations through expanded training datasets and improved modeling architectures.

CRediT authorship contribution statement

Yongjin Choi: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Seungjun Lee:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the InnoCORE program of the Ministry of Science and ICT (N10250154).

Appendix A. Energy evolutions for test cases 1 to 3

The energy evolutions for test cases 1 to 3 are shown in Figs. A.1–A.3.

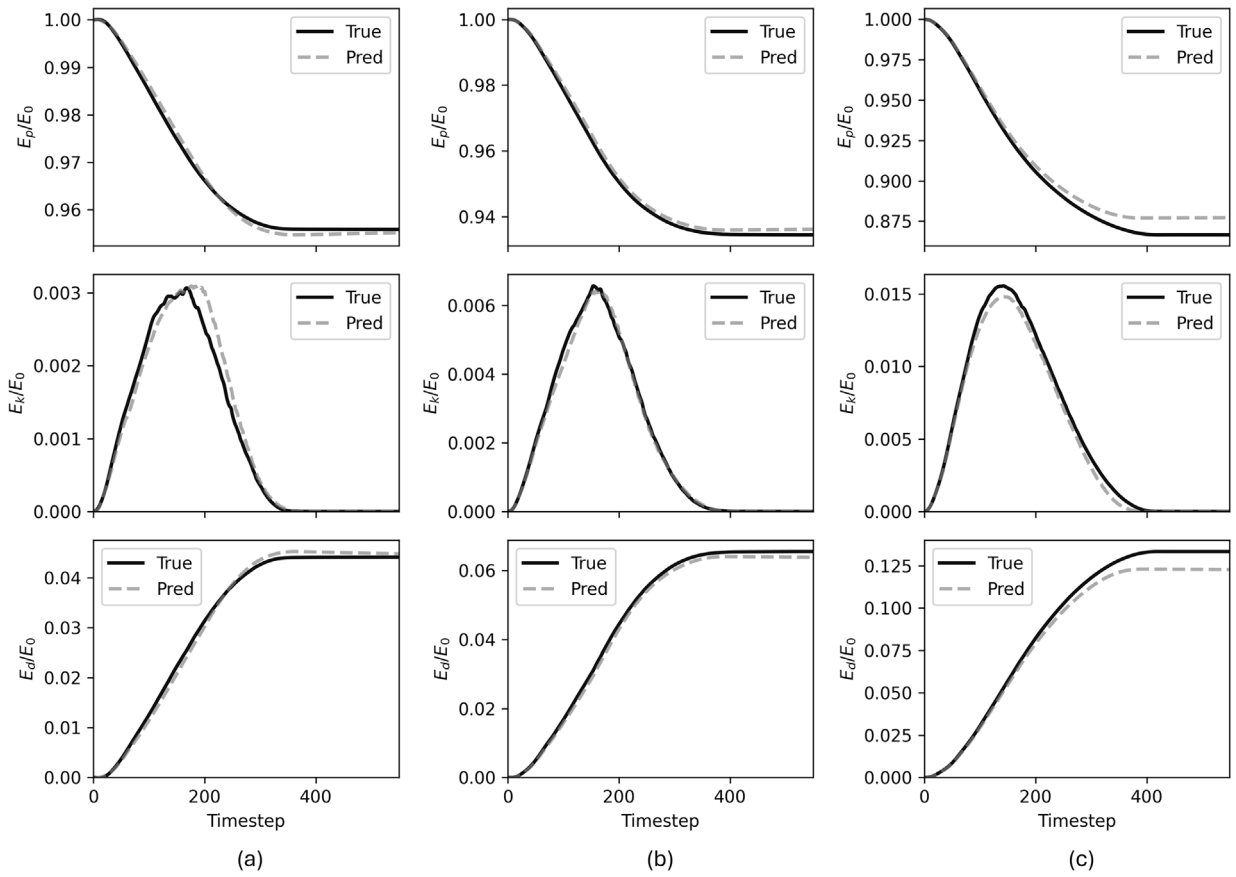

Fig. A.1. Normalized energy evolution over time of MPM (True) and GNS (Pred) for (a) best, (b) median, and (c) worst case predictions for test case 1. Each row corresponds to potential energy E_p, kinetic energy E_k, and dissipation energy E_d, respectively, normalized by the initial potential energy E_0.

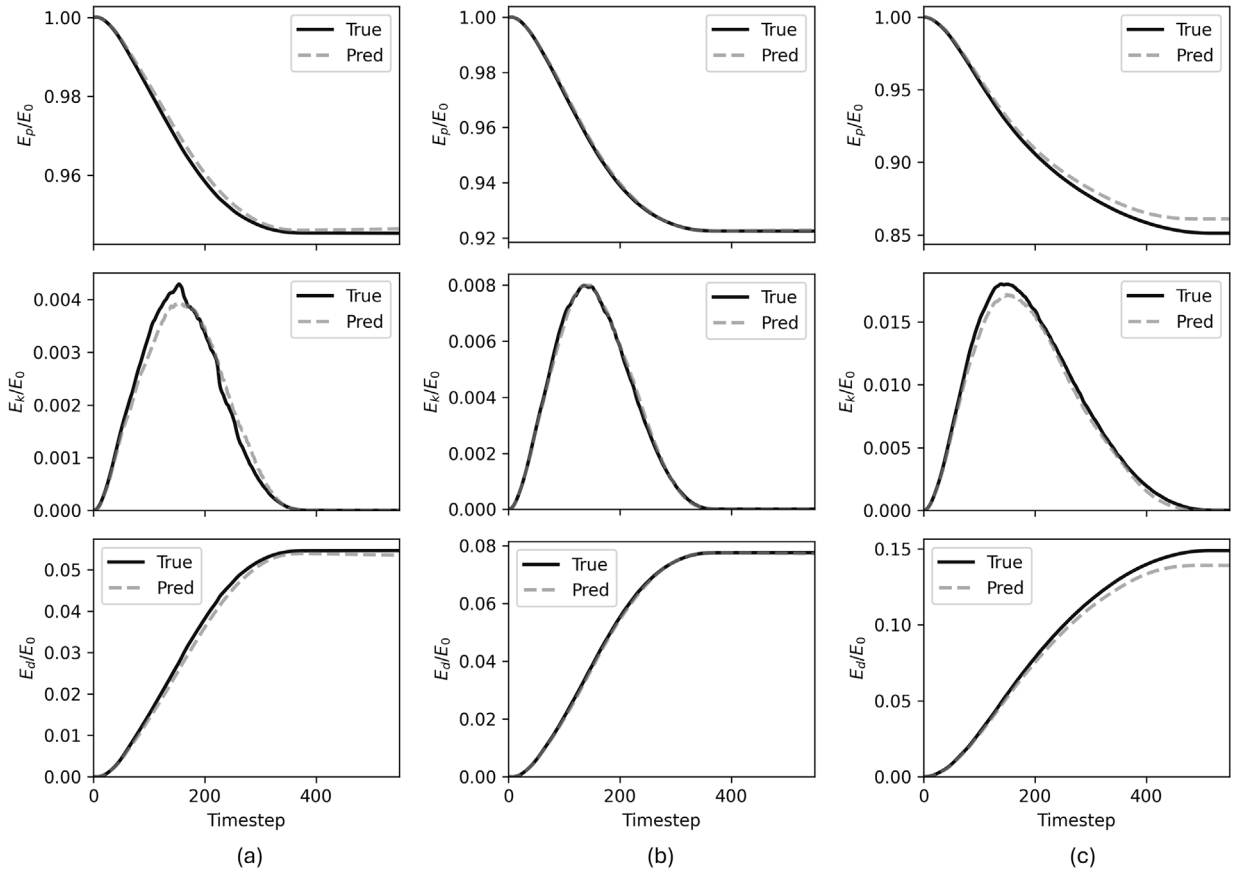

Fig. A.2. Normalized energy evolution over time of MPM (True) and GNS (Pred) for (a) best, (b) median, and (c) worst case predictions for test case 2. Each row corresponds to potential energy E_p, kinetic energy E_k, and dissipation energy E_d, respectively, normalized by the initial potential energy E_0.

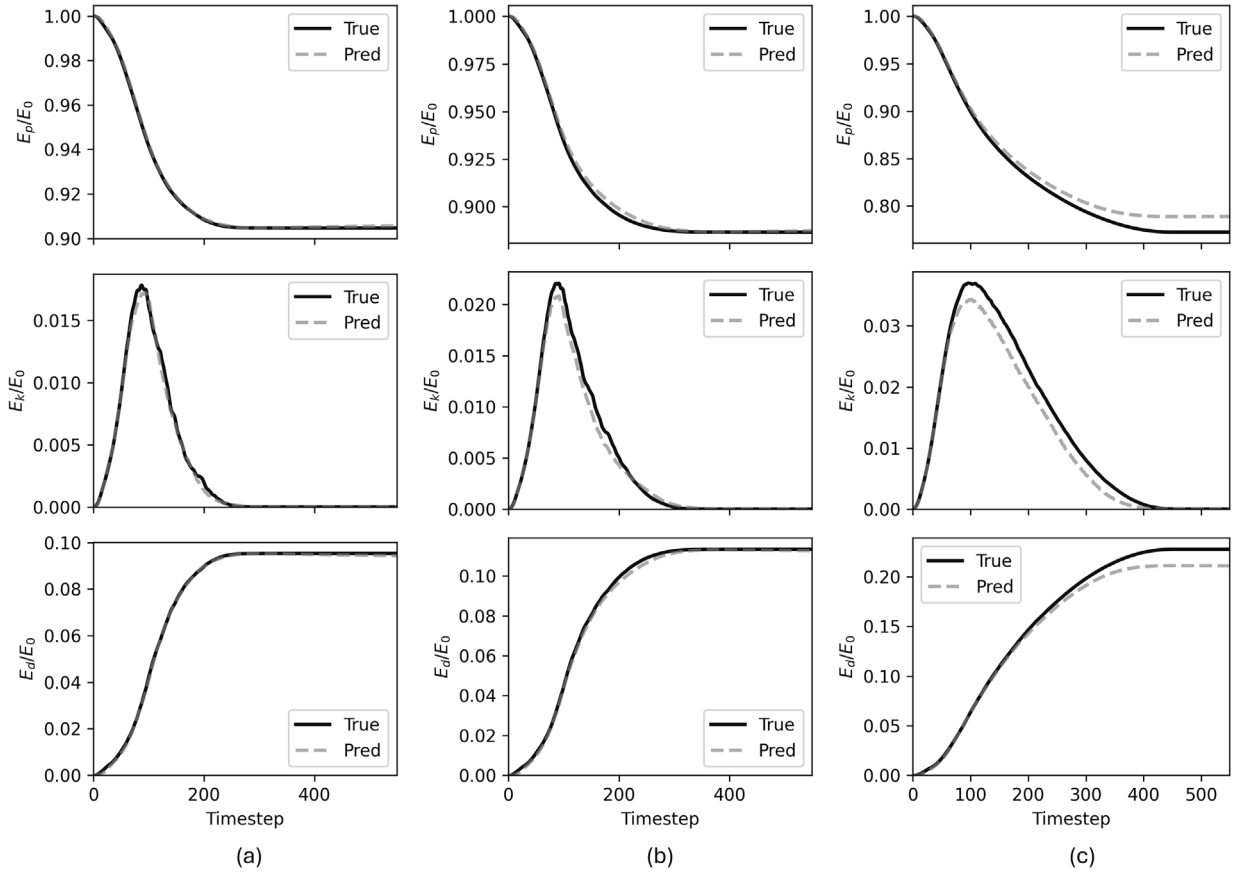

Fig. A.3. Normalized energy evolution over time of MPM (True) and GNS (Pred) for (a) best, (b) median, and (c) worst case predictions for test case 3. Each row corresponds to potential energy E_p, kinetic energy E_k, and dissipation energy E_d, respectively, normalized by the initial potential energy E_0.

Data availability

Data will be made available on request.

References

Arduino, P., Bonus, J., Motley, M., Eberhard, M., 2021. Tsunami-driven debris effects on structures using a multi-gpu mpm tool. Mecánica Comput. 38, 3.

Battaglia, P.W., Hamrick, J.B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., 2018. Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261.

Battaglia, P., Pascanu, R., Lai, M., Jimenez Rezende, D., et al., 2016. Interaction networks for learning about objects, relations and physics. Adv. Neural Inf. Process. Syst. 29.

Cao, Z., Wang, Y., Li, D., Cao, Z., Wang, Y., Li, D., 2017. Practical reliability analysis of slope stability by advanced Monte Carlo simulations in a spreadsheet. Probabilistic Approaches Geotech. Site Charact. Slope Stab. Anal. 147–167.

Ceccato, F., Yerro, A., Di Carluccio, G., 2024. Simulating landslides with the material point method: Best practices, potentialities, and challenges. Eng. Geol. 338, http://dx.doi.org/10.1016/j.enggeo.2024.107614.

Cho, S.E., 2009. Probabilistic stability analyses of slopes using the ANN-based response surface. Comput. Geotech. 36 (5), 787–797.

Cho, S.E., 2010. Probabilistic assessment of slope stability that considers the spatial variability of soil properties. J. Geotech. Geoenvironmental Eng. 136 (7), 975–984.

Choi, Y., Kumar, K., 2024a. Graph neural network-based surrogate model for granular flows. Comput. Geotech. 166, 106015.

Choi, Y., Kumar, K., 2024b. Inverse analysis of granular flows using differentiable graph neural network simulator. Comput. Geotech. 171, 106374.

Choi, Y., Macedo, J., Liu, C., 2025. Differentiable graph neural network simulator for forward and inverse modeling of multi-layered slope system with multiple material properties. arXiv preprint arXiv:2504.15938.

Durante, M.G., Rathje, E.M., 2021. An exploration of the use of machine learning to predict lateral spreading. Earthq. Spectra 37 (4), 2288–2314.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman and Hall/CRC, http://dx.doi.org/10.1201/9780429246593.

Griffiths, D., Fenton, G.A., 2004. Probabilistic slope stability analysis by finite elements. J. Geotech. Geoenvironmental Eng. 130 (5), 507–518.

Griffiths, D., Huang, J., Fenton, G.A., 2009. Influence of spatial variability on slope reliability using 2-D random fields. J. Geotech. Geoenvironmental Eng. 135 (10), 1367–1378.

Huang, J., Griffiths, D., Fenton, G.A., 2010. System reliability of slopes by RFEM. Soils Found. 50 (3), 343–353.

Jiang, Y., Byrne, E., Glassey, J., Chen, X., 2024. Integrating graph neural network-based surrogate modeling with inverse design for granular flows. Ind. Eng. Chem. Res..

Jiang, S.-H., Huang, J., Griffiths, D., Deng, Z.-P., 2022. Advances in reliability and risk analyses of slopes in spatially variable soils: A state-of-the-art review. Comput. Geotech. 141, 104498.

Jiang, S.-H., Huang, J., Yao, C., Yang, J., 2017. Quantitative risk assessment of slope failure in 2-D spatially variable soils by limit equilibrium method. Appl. Math. Model. 47, 710–725.

Jiang, S.-H., Li, D.-Q., Cao, Z.-J., Zhou, C.-B., Phoon, K.-K., 2015. Efficient system reliability analysis of slope stability in spatially variable soils using Monte Carlo simulation. J. Geotech. Geoenvironmental Eng. 141 (2), 04014096.

Jiang, S.-H., Li, J.-P., Ma, G.-T., Rezania, M., 2024. Probabilistic assessment of 3D slope failures in spatially variable soils by cooperative stochastic material point method. Comput. Geotech. 172, 106413.

Jiang, S.-H., Li, D.-Q., Zhang, L.-M., Zhou, C.-B., 2014. Slope reliability analysis considering spatially variable shear strength parameters using a non-intrusive stochastic finite element method. Eng. Geol. 168, 120–128.

Ju, L.-Y., Xiao, T., He, J., Wang, H.-J., Zhang, L.-M., 2022. Predicting landslide runout paths using terrain matching-targeted machine learning. Eng. Geol. 311, 106902.

Kingma, D.P., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kumar, K., Salmond, J., Kularathna, S., Wilkes, C., Tjung, E., Biscontin, G., Soga, K., 2019. Scalable and modular material point method for large-scale simulations. arXiv preprint arXiv:1909.13380.

Lacasse, S., Nadim, F., 2011. Learning to live with geohazards: From research to practice. In: Geo-Risk 2011: Risk Assessment and Management. American Society of Civil Engineers, pp. 64–116. http://dx.doi.org/10.1061/41183(418)4.

Li, D.-Q., Jiang, S.-H., Cao, Z.-J., Zhou, W., Zhou, C.-B., Zhang, L.-M., 2015. A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties. Eng. Geol. 187, 60–72.

Li, X., Sovilla, B., Jiang, C., Gaume, J., 2021. Three-dimensional and real-scale modeling of flow regimes in dense snow avalanches. Landslides 18 (10), 3393–3406.

Li, D.-Q., Xiao, T., Cao, Z.-J., Phoon, K.-K., Zhou, C.-B., 2016a. Efficient and consistent reliability analysis of soil slope stability using both limit equilibrium analysis and finite element analysis. Appl. Math. Model. 40 (9–10), 5216–5229.

Li, D.-Q., Zheng, D., Cao, Z.-J., Tang, X.-S., Phoon, K.-K., 2016b. Response surface methods for slope reliability analysis: review and comparison. Eng. Geol. 203, 3–14.

Liu, L.-L., Cheng, Y.-M., Jiang, S.-H., Zhang, S.-H., Wang, X.-M., Wu, Z.-H., 2017. Effects of spatial autocorrelation structure of permeability on seepage through an embankment on a soil foundation. Comput. Geotech. 87, 62–75.

Liu, L.-L., Liang, C.-Q., Huang, L., Wang, B., 2023. Parametric analysis for the large deformation characteristics of unstable slopes with linearly increasing soil strength by the random material point method. Comput. Geotech. 162, 105661.

Liu, K., Wang, Y., Huang, M., Yuan, W.-H., 2021. Postfailure analysis of slopes by random generalized interpolation material point method. Int. J. Geomech. 21 (3), 04021015.

Liu, X., Wang, Y., Li, D.-Q., 2019. Investigation of slope failure mode evolution during large deformation in spatially variable soils by random limit equilibrium and material point methods. Comput. Geotech. 111, 301–312.

Lucchetti, A., Cadini, F., Giglio, M., Lomazzi, L., 2025. Graph network-based structural simulator: Graph neural networks for structural dynamics. arXiv preprint arXiv: 2510.25683.

Ma, G., Rezania, M., Mousavi Nezhad, M., 2022. Stochastic assessment of landslide influence zone by material point method and generalized geotechnical random field theory. Int. J. Geomech. 22 (4), 04022002.

Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., Battaglia, P., 2020. Learning to simulate complex physics with graph networks. In: International Conference on Machine Learning. PMLR, pp. 8459–8468.

Soga, K., Alonso, E., Yerro, A., Kumar, K., Bandara, S., 2016. Trends in large-deformation analysis of landslide mass movements with particular emphasis on the material point method. Géotechnique 66 (3), 248–273.

Sordo, B., Rathje, E., Kumar, K., 2024. Sequential hybrid finite element and material point method to simulate slope failures. Comput. Geotech. 173, 106525.

Troncone, A., Pugliese, L., Parise, A., Conte, E., 2023a. Analysis of a landslide in sensitive clays using the material point method. Geotech. Res. 10 (2), 67–77.

Troncone, A., Pugliese, L., Parise, A., Conte, E., 2023b. A practical approach for predicting landslide retrogression and run-out distances in sensitive clays. Eng. Geol. 326, 107313.

Troncone, A., Pugliese, L., Parise, A., Mazzuca, P., Conte, E., 2025. Post-failure stage analysis of flow-type landslides using different numerical techniques. Comput. Geotech. 182, 107152.

Vicari, H., Tran, Q.-A., Juel, M.M., Gaume, J., 2025. The role of dilatancy and permeability of erodible wet bed sediments in affecting erosion and runout of a granular flow: Two-phase MPM–CFD simulations. Comput. Geotech. 185, 107307.

Wang, B., Hicks, M., Vardon, P., 2016. Slope failure analysis using the random material point method. Géotech. Lett. 6 (2), 113–118.

Xiao, T., Li, D.-Q., Cao, Z.-J., Au, S.-K., Phoon, K.-K., 2016. Three-dimensional slope reliability and risk assessment using auxiliary random finite element method. Comput. Geotech. 79, 146–158.

Zeng, P., Sun, X.P., Xu, Q., Li, T.B., Zhang, T.L., 2021. 3D probabilistic landslide run-out hazard evaluation for quantitative risk assessment purposes. Eng. Geol. 293, http://dx.doi.org/10.1016/j.enggeo.2021.106303.

Zhao, S., Chen, H., Zhao, J., 2025. A physical-information-flow-constrained temporal graph neural network-based simulator for granular materials. Comput. Methods Appl. Mech. Engrg. 433, 117536.

Zhao, W., Xia, X., Su, X., Liang, Q., Liu, X., Ju, N., 2021. Movement process analysis of the high-speed long-runout Shuicheng landslide over 3-D complex terrain using a depth-averaged numerical model. Landslides 18 (9), 3213–3226.

Zhou, Z., Li, D.-Q., Xiao, T., Cao, Z.-J., Du, W., 2021. Response surface guided adaptive slope reliability analysis in spatially varying soils. Comput. Geotech. 132, 103966.