



# Hybrid probabilistic modeling for reliability assessment of Mosul and Dohuk earth dams under geotechnical uncertainty

Munadhil Ali Hasan

Department of Civil Engineering, Faculty of Engineering, University of Qom, Qom, Iran.

\*Corresponding author E-mail: [munadhilalihasan@gmail.com](mailto:munadhilalihasan@gmail.com)

## Abstract

The conventional deterministic slope-stability analyses of earth dams cannot quantify failure probability arising from natural variability and measurement uncertainty in geotechnical parameters. In this study, a probabilistic framework is devised that combines Monte Carlo simulation (MCS) using limit-equilibrium analysis (GeoStudio SEEP/W-SLOPE/W) and machine-learning surrogate models to assess the reliability indices and failure probabilities for Mosul Dam and Dohuk Dam northern Iraq under their normal functioning rapid drawdown seismic loading and the maximum flood occurrence. Due to the absence of recent field investigations or laboratory tests, we compiled geotechnical parameters from published reports and technical sources. We derived statistical moments from reported ranges using the method of moments. Further, we justified probability distributions from geotechnical reliability literature and available goodness-of-fit checks. Ten thousand MCS realizations were generated for every dam and loading condition. The predictive accuracy for the XGBoost resulted in the highest score for the study with a computed  $R^2 = 0.984$ ,  $MAE = 0.018$  and  $RMSE = 0.024$ . Furthermore, the reliability update was fully achieved by enabling the model in only 6.5 s. The seismic reliability of Mosul Dam has been found to be critically low with a value of  $\beta = 1.95$  and  $P_f = 2.56 \times 10^{-2}$ . This low value is mainly controlled by variability in foundation hydraulic conductivity. Under all scenarios, the Dohuk Dam was still satisfactory but rapid drawdown  $\beta$  reduced from 4.85 to 3.05. The framework enables screening-level, risk-informed monitoring and prioritization of operations, but it calls for site-specific validation prior to design-level decision-making.

**Keywords:** Earth dams; Geotechnical uncertainty; Machine learning surrogate models; Monte Carlo simulation; Probabilistic reliability.

## 1. Introduction

Earth dams are among the most prevalent water-retaining structures in the world. They are well suited to the materials found locally, are economical to construct and can adapt to a variety of topographic and geological formation conditions. Earth and rockfill dams in Iraq are important national infrastructure for irrigation, municipal water supplies, hydropower generation, and flood control. The interdependencies of the embankment fill, the foundation soils, the seepage regime and the external load influence the integrity of these dams. Moreover, these characteristics vary in space and time. The catastrophic downstream flooding, fatalities, and economic disruption caused by the failure of a major earth dam necessitate an assessment of stability for dam safety management [1], [6].

Conventional deterministic methods for assessing dam slope-stability quantify safety using a single factor of safety (FS) that is derived from fixed, representative soil parameters. The current standard design codes are almost exclusively based on deterministic methodologies. They do not provide any measure of probability of failure nor can they adequately consider the natural variability and measurement uncertainty of geotechnical parameters like cohesion,  $c'$ , friction angle,  $\phi'$ , and unit weight  $\gamma$  and hydraulic conductivity. Probabilistic reliability analysis overcomes this limitation by modeling soil parameters as random variables represented by probability distributions, allowing for the calculation of the reliability index ( $\beta$ ) and the failure probability ( $P_f$ ) that directly facilitate risk-based decision-making [2], [8], [13].

The Mosul Dam, located on the Tigris River about 50 km north of the city of Mosul, is a 113 m-high zoned earth-fill dam impounding Iraq's largest reservoir (11.1 billion  $m^3$ ). The alternating layers of gypsum, anhydrite, marl and limestone are assigned to Fat'ha (Lower Fars) formation. These soluble minerals continuously dissolve due to seepage flow which has



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. © 2025 The Authors

<https://www.muthuni-ojs.org/index.php/mjet/index>

formed cavities and sinkholes under the dam after impoundment, which happened in 1986. Thus, grouting has been continued at various stages. The process of dissolution causes the foundation to possess a high degree of spatial variability in terms of permeability and strengths, making ordinary deterministic analysis particularly inappropriate for this structure [7], [15].

The Dohuk Dam is a homogeneous earth-fill dam, 60 meters height, on the Dohuk River in the Kurdistan Region of Iraq. It is the main water supply for the city of Dohuk. After over forty years of operation, sedimentation has reduced effective storage capacity, while rapid drawdown events during the dam's operational regime create adverse pore-pressure conditions on the upstream slope. Due to declining storage capacity and transient seepage during drawdown, a reliability challenge arises that requires coupled saturated-unsaturated flow analysis and probabilistic assessment of slope-stability under time-varying boundary conditions [5], [10].

Although these two dams are critical infrastructure, the published reliability assessments have mostly been limited to deterministic or simplified probabilistic treatments with modern computational tools not fully exploited. There is a research gap for a hybrid framework which uses Monte Carlo simulation along with machine-learning surrogate models, to efficiently propagate geotechnical uncertainty through coupled seepage-stability analyses for the dam conditions of Iraq. The research aims to accomplish the following: (1) to compile and statistically approximate geotechnical parameter distributions from available published and technical sources, (2) to develop and train machine-learning surrogate models (XGBoost, Random Forest, LSTM, ANN) on datasets generated by MCS for rapid reliability prediction, (3) to calculate reliability indices and failure probabilities under normal, rapid drawdown, seismic, and flood loading conditions, comparing the results as a cross-case methodological assessment rather than a direct site equivalence, and (4) to formulate preliminary, site-verification-dependent sustainability recommendations for the dam safety authorities of Iraq. No new laboratory or field tests were conducted in this study [4], [16].

## 2. Materials and methods

### 2.1 Study area

This research study involves two earth dams located in northern Iraq that differ in their geotechnical properties and operational problems:

**Mosul Dam:** The dam, located at 36°37'N, 42°49'E, is 113 m high and 3.4 km long. It has a reservoir capacity of 11.1 billion m<sup>3</sup>. The dam was completed in 1984, and the reservoir impoundment began in 1986. The foundation geology comprises the Fat'ha formation consisting of interbedding gypsum, anhydrite, marl and limestone layers dipping gently downstream. The geological characteristics have created preferential dissolution paths along the bedding planes and fractures.

**Dohuk Dam:** The dam located at 36°52'N and 43°01'E is a homogeneous earth-fill structure of 60 m height, having a crest length of 613 m. This dam impounds about 52 million m<sup>3</sup> at full supply level. Built in 1988 on relatively competent sedimentary foundation materials (Gercus Formation), the dam has geotechnical concerns about stability of the upstream slope in rapid drawdown and loss of storage volume as the catchment erodes and sediments at a rate of 0.8-1.2 million m<sup>3</sup> per decade.

The general field appearance of the two study dams is shown in Fig. 1. The photographs are included for site illustration only and do not represent field data collected in the present study.

**(a) Mosul Dam**



**(b) Dohuk Dam**



**Fig. 1:** Field appearance of the investigated dams: (a) Mosul Dam, northern Iraq; (b) Dohuk Dam, Kurdistan Region of Iraq. Source: images supplied by the author for site illustration.

### 2.2 Data sources

Secondary geotechnical data collected from published literature, technical reports and government monitoring records were used in this study. No additional field investigations or laboratory testing was undertaken. You can gather data from various sources:

- The geotechnical investigation report as published would characterize the foundation of Mosul Dam with regard to the distribution of shear strength, permeability, and dissolution voids.
- We are maintaining records on the quantity of grouting and piezometric levels issued by the Iraqi Ministry of Water Resources.

- The soil mechanics tests conducted on the embankment and foundation materials of Dohuk Dam are published. The tests include and triaxial compression, direct shear and permeability tests.
- The Kurdistan Regional Government has conducted surveys and bathymetric data for Dohuk Dam.
- The study area has been provided regional seismicity catalogs and peak ground acceleration estimates.

Following standard statistical procedures, mean, standard deviation and coefficient of variation were derived from the reported range and test results. In situations where only ranges are given the standard deviation was approximated to be equal to  $(x_{\max} - x_{\min}) / 4$  or  $\sigma \approx (x_{\max} - x_{\min}) / 4$ , where  $x$  describes the range of the number.

Reproducibility Statement: The computational workflow designed as a numerical simulation framework of secondary geotechnical data. As a result, there are probabilistic realizations using ranges of parameters published in the literature rather than new measurements in a field.

### 2.3 Geotechnical parameters

The random variables adopted for probabilistic analysis in this study are presented in Table 1. The parameters were selected based on their established influence on slope-stability and seepage behavior as identified in existing reliability studies of earth dams [2], [3], [9].

**Table 1:** Statistical characterization of geotechnical parameters for Mosul and Dohuk dams.

| Dam   | Material/Zone       | Parameter                      | Mean                 | Std Dev               | CoV  | Distribution |
|-------|---------------------|--------------------------------|----------------------|-----------------------|------|--------------|
| Mosul | Foundation (gypsum) | $c'$ (kPa)                     | 45                   | 13.5                  | 0.30 | Lognormal    |
| Mosul | Foundation (gypsum) | $\phi'$ (°)                    | 28                   | 5.6                   | 0.20 | Normal       |
| Mosul | Foundation (gypsum) | $\gamma$ (kN/m <sup>3</sup> )  | 21.5                 | 1.72                  | 0.08 | Normal       |
| Mosul | Foundation (gypsum) | $k$ (m/s)                      | $2.5 \times 10^{-5}$ | $1.13 \times 10^{-5}$ | 0.45 | Lognormal    |
| Mosul | Foundation (gypsum) | Diss. void ratio               | 0.15                 | 0.06                  | 0.40 | Lognormal    |
| Dohuk | Embankment (clay)   | $c'$ (kPa)                     | 35                   | 8.75                  | 0.25 | Lognormal    |
| Dohuk | Embankment (clay)   | $\phi'$ (°)                    | 24                   | 3.6                   | 0.15 | Normal       |
| Dohuk | Embankment (clay)   | $\gamma$ (kN/m <sup>3</sup> )  | 19.8                 | 1.19                  | 0.06 | Normal       |
| Dohuk | Embankment (clay)   | $k$ (m/s)                      | $5.0 \times 10^{-7}$ | $1.75 \times 10^{-7}$ | 0.35 | Lognormal    |
| Dohuk | Reservoir           | Sed. rate (m <sup>3</sup> /yr) | $1.0 \times 10^6$    | $2.5 \times 10^5$     | 0.25 | Normal       |

Source: Parameters collected from published geotechnical reports and monitoring records. Statistical moments determined via probabilistic moments analysis.

### 2.4 Statistical characterization and distribution fitting

Because the original raw datasets were not fully available, the adopted statistical parameters were derived from published ranges and reported representative values. Therefore, the uncertainty characterization should be interpreted as an engineering probabilistic approximation rather than a site-specific statistical database. Table 2 summarizes the data basis and statistical treatment for each geotechnical parameter.

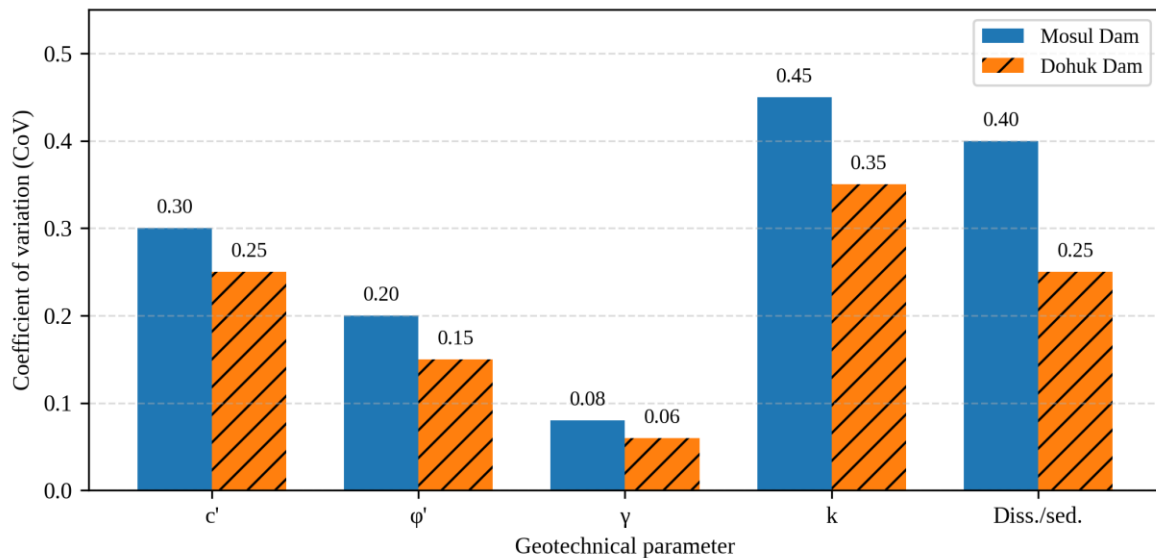
**Table 2:** Data basis and statistical treatment of geotechnical parameters.

| Dam   | Parameter  | Source type         | Reported data form       | Assumed sample basis           | Distribution | Justification                                  |
|-------|------------|---------------------|--------------------------|--------------------------------|--------------|------------------------------------------------|
| Mosul | $c'$       | Technical reports   | Range (25–65 kPa)        | Estimated from range endpoints | Lognormal    | CoV > 0.25; avoids negative values             |
| Mosul | $\phi'$    | Published studies   | Mean $\pm$ SD            | Reported directly              | Normal       | CoV = 0.20; symmetric scatter                  |
| Mosul | $k$        | Monitoring records  | Order-of-magnitude range | Log-space moments              | Lognormal    | Permeability log-linear in heterogeneous soils |
| Mosul | Diss. void | Grouting records    | Range (0.05–0.25)        | Estimated from range endpoints | Lognormal    | Physically bounded; skewed                     |
| Dohuk | $c'$       | Lab test summaries  | Range (20–50 kPa)        | Estimated from range endpoints | Lognormal    | CoV = 0.25; avoids negative values             |
| Dohuk | $\phi'$    | Direct shear tests  | Mean $\pm$ SD            | Reported directly              | Normal       | CoV = 0.15; symmetric scatter                  |
| Dohuk | $k$        | Permeability tests  | Range                    | Log-space moments              | Lognormal    | Standard for clay permeability                 |
| Dohuk | Sed. rate  | Bathymetric surveys | Annual estimates         | Inter-annual variation         | Normal       | Moderate CoV; no physical lower bound issue    |

Source: Compiled from published geotechnical reports, laboratory test summaries, and government monitoring records.

### 2.5 Probability distributions

As illustrated in Fig. 2, the comparative uncertainty pattern of the selected geotechnical parameters shows that the Mosul Dam exhibits greater variability than Dohuk Dam except for sedimentation rate. This diagram allows for a visual comparison that supports the choice of probabilistic distributions used in subsequent reliability analyses.



**Fig. 2:** Comparison of geotechnical parameter uncertainty (coefficient of variation) between Mosul Dam and Dohuk Dam.

Based on the geotechnical reliability literature, probability distributions were selected as follows. Parameters of shear strength ( $c'$ ,  $\phi'$ ) and unit weight ( $\gamma$ ) were assigned normal or lognormal distribution based on the magnitude of their coefficient of variation. Parameters with  $CoV > 0.25$  were modeled as lognormal to avoid unrealistic negative realizations that have no physical meaning, whereas parameters of lower variability were treated as normally distributed [3], [11]. Hydraulic conductivity followed a lognormal distribution because permeability measurements on heterogeneous soils are well-documented to be log-linearly distributed.

Goodness-of-fit assessment was applied only where sufficient reported data points were available; otherwise, distribution selection was justified based on geotechnical reliability literature and physical constraints of each parameter. In cases where the Kolmogorov-Smirnov test could be applied, the results confirmed the assumed distribution at the 5% significance level.

Cross-correlation between shear strength parameters was incorporated using a correlation coefficient of  $\rho(c', \phi') = -0.5$ . Since site-specific paired shear strength datasets were unavailable, this value was adopted as a conservative literature-based assumption reflecting the established negative correlation between cohesion and friction angle observed in consolidated-undrained triaxial testing of clayey soils [12], [14]. A sensitivity analysis using  $\rho = 0, -0.3$ , and  $-0.5$  was conducted to assess the influence of this assumption on the computed reliability index (see Table 3).

Table 3: Sensitivity of reliability index to assumed correlation coefficient  $\rho(c', \phi')$ .

| Correlation coefficient | Mosul $\beta$ (seismic) | Dohuk $\beta$ (rapid drawdown) | Interpretation                        |
|-------------------------|-------------------------|--------------------------------|---------------------------------------|
| $\rho = 0$              | 2.08                    | 3.18                           | Uncorrelated; slightly higher $\beta$ |
| $\rho = -0.3$           | 2.01                    | 3.11                           | Moderate negative correlation         |
| $\rho = -0.5$           | 1.95                    | 3.05                           | Adopted conservative assumption       |

*Note: The difference in  $\beta$  across the range of  $\rho$  is within 7%, confirming that the overall reliability conclusions are not highly sensitive to the assumed correlation structure.*

## 2.6 Deterministic stability analysis

Deterministic slope-stability analyses were performed using the Morgenstern-Price method implemented in GeoStudio SLOPE/W, which satisfies both force and moment equilibrium and employs a half-sine inter-slice force function. Steady-state and transient seepage analyses were conducted in SEEP/W using finite-element solutions of the Richards equation to establish pore-water pressure distributions for the following loading conditions:

- Normal operating condition: Steady-state seepage at normal pool level.
- Rapid drawdown: Reservoir lowering at a rate of 0.5 m/day from full supply level to minimum operating level, with transient pore pressures computed at critical time steps.
- Seismic loading: Pseudo-static analysis with horizontal seismic coefficient  $k_h = 0.10$  representing the design earthquake for the region (475-year return period).
- Flood condition: Steady-state seepage with reservoir at maximum flood level (spillway crest elevation plus routing surcharge).



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. © 2025 The Authors

<https://www.muthuni-ojs.org/index.php/mjet/index>

The critical slip surface for each condition was identified using the entry-and-exit search method combined with grid-and-radius optimization to locate the minimum factor of safety surface.

## 2.7 Monte Carlo simulation

To propagate the geotechnical uncertainty through the coupled seepage-stability model (FIG. 3), a systematic Monte Carlo simulation workflow was developed:

- Step 1 - To randomly generate the values of correlated random variables for each simulation trial, the Nataf transformation was used so that the marginal distributions and correlation system are preserved.
- Step 2 - The hydraulic conductivity field (generated by 2D analysis) was assigned to mesh in SEEP/W and the steady-state or transient seepage was solved to obtain pore-water pressure distribution.
- Step 3 - The shear strength parameters and pore pressures were used for the analysis of stability by SLOPE/W. The Morgenstern-Price method was used for the determination of the minimum factor of safety by optimized slip surface search.
- Step 4 - Statistical processing: After conducting  $N = 10,000$  trials per loading condition, the FS distribution was characterized by its mean ( $\mu_{FS}$ ) and standard deviation ( $\sigma_{FS}$ ).

The failure probability was computed as:

$$P_f = N_{fail} / N \quad (1)$$

where  $N_{fail}$  refers to the total number of trials with  $FS < 1.0$ . The reliability index was calculated as:

$$\beta = \Phi^{-1}(1 - P_f) = -\Phi^{-1}(P_f) \quad (2)$$

where  $\Phi^{-1}$  is the inverse standard normal cumulative distribution function [1], [6], [9].

Step 5 - Convergence check: Through monitoring CoV of  $P_f$  as a function of  $N$ , it was ensured that 10,000 samples were adequate for convergence (CoV of  $P_f < 5\%$ ).

### 2.7.1 Treatment of rare failure probabilities

Direct MCS with 10,000 samples was used to estimate the distribution of FS and identify dominant trends. For very small probabilities of failure,  $P_f$  was not interpreted as a direct counting estimate. Instead,  $\beta$  was estimated from the fitted FS distribution, and  $P_f$  was then obtained using:

$$P_f = \Phi(-\beta) \quad (3)$$

Therefore, very low  $P_f$  values (particularly  $P_f < 1/N = 10^{-4}$ ) should be considered reliability-index-based estimates rather than direct MCS frequencies. This approach is standard practice in structural reliability when the target failure probability lies below the resolution of the available sample size [1], [6].

## 2.8 Machine learning surrogate modeling

Each coupled SEEP/W-SLOPE/W run required approximately 26 seconds. Therefore, a full set of 10,000 direct MCS runs would require approximately 72 hours per loading condition per dam on a conventional workstation. To allow for rapid reliability updates for real-time decision support, machine-learning surrogate models were developed for approximating the FS response surface as a function of the input random variables [16], [17].

Four surrogate model architectures were evaluated:

- XGBoost (Extreme Gradient Boosting): A collection of decision trees trained sequentially using gradient descent to reduce the mean square error loss function. Hyperparameters (maximum depth = 6, learning rate = 0.1, number of estimators = 500, regularization  $\lambda = 1.0$ ) were tuned via 5-fold cross-validation.
- Random Forest (RF): An ensemble of 300 decision trees trained on bootstrap samples with random feature selection at each split. Bagging provides variance estimation based on out-of-bag predictions.
- Artificial Neural Network (ANN): A network with three hidden layers of 64, 32, and 16 neurons, using ReLU activation and Adam optimizer with early stopping (patience of 20 epochs) to prevent overfitting.
- Long Short-Term Memory (LSTM): An architecture with two LSTM layers of 50 and 25 nodes to capture temporal dependencies of time-varying parameters under transient loading (rapid drawdown and flood hydrograph). The output was a dense layer.

## 2.9 Model training and validation

The MCS-generated dataset comprising 10,000 input-output pairs (realizations of random variables to FS) was divided into a training subset (70%, i.e., 7,000 samples) and a testing subset (30%, i.e., 3,000 samples) using stratified sampling to conserve the proportion of failures in both subsets. Input features were standardized to have zero mean and unit variance before training.

Evaluated model performance on three complementary metrics:

- The coefficient of determination ( $R^2$ ) measures how much variance in FS is explained by the model and approaches 1.0.
- Mean absolute error (MAE): Indicates the average absolute difference between predicted and actual FS values.
- Root mean square error, or RMSE, is a metric that puts greater emphasis on larger prediction errors compared to MAE. This characteristic makes RMSE sensitive to the tails of the FS distribution. This behavior is critical for estimating failure probabilities.

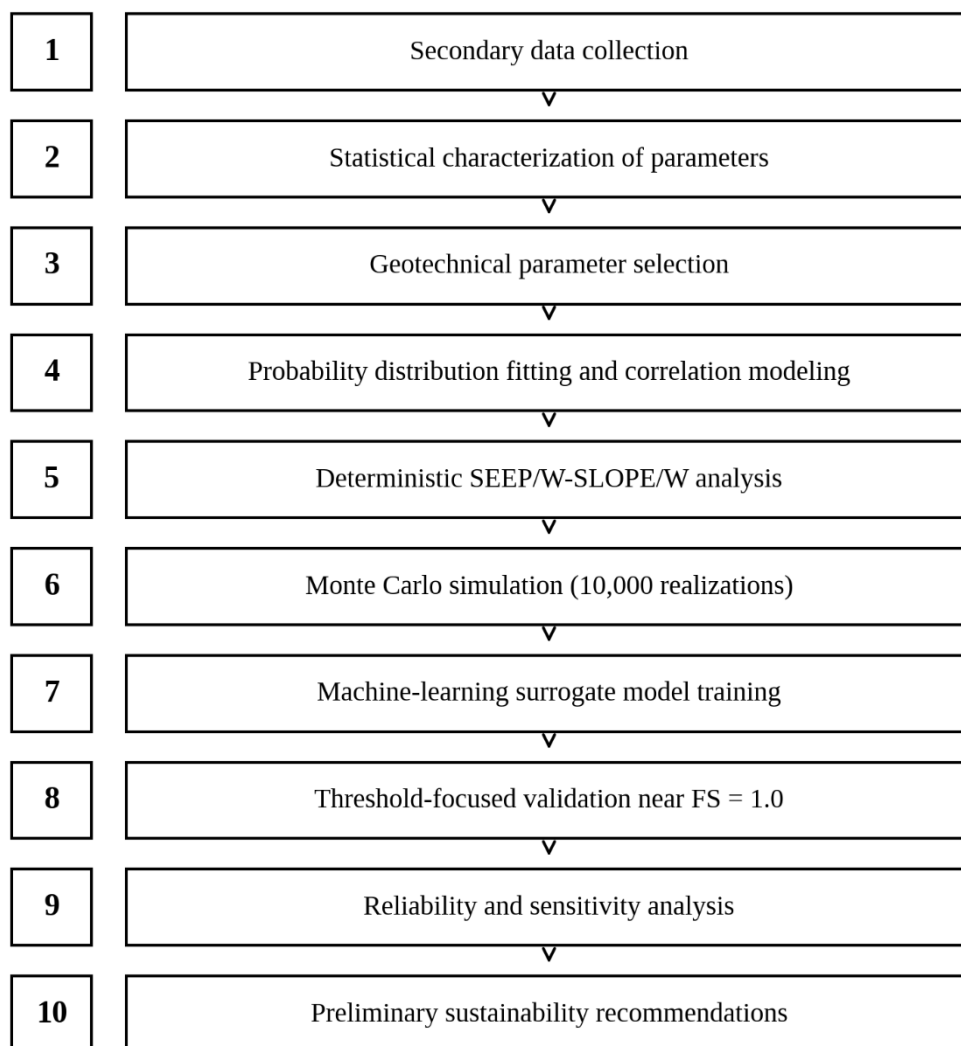
Overfitting was checked through the comparison of training and testing metrics; any model with a training-to-testing  $R^2$  ratio greater than 1.05 was regularized further.

## 2.10 Incorporation of dissolution vulnerability

The foundation of the Mosul Dam is continuously dissolving due to gypsum and anhydrite layers, which eventually increase the permeability of the foundation and form voids that reduce effective stress. In the absence of time-dependent geochemical dissolution data dissolution vulnerability was represented indirectly through three mechanisms:

- The foundation permeability's high CoV (0.45) captures the spatial heterogeneity due to dissolution channels and intact rock zones.
- The dissolution void ratio is treated as a random variable (mean = 0.15, CoV = 0.40) to simulate the volume of the foundation that is filled with voids due to dissolution. Such voids affect effective permeabilities and seep-age gradients.
- Using variance-based sensitivity decomposition (see Section 3.1), the contribution of hydraulic conductivity to FS variance (the model output) was quantified to assess the importance of dissolution-related parameters.

The given representation is a time-locked approximation and will not model the ongoing dissolution process. Future studies should incorporate coupled geochemical-geomechanical modeling to capture time-dependent dissolution effects on foundation reliability.



**Fig. 3:** Research methodology flowchart illustrating the ten-phase hybrid probabilistic framework.

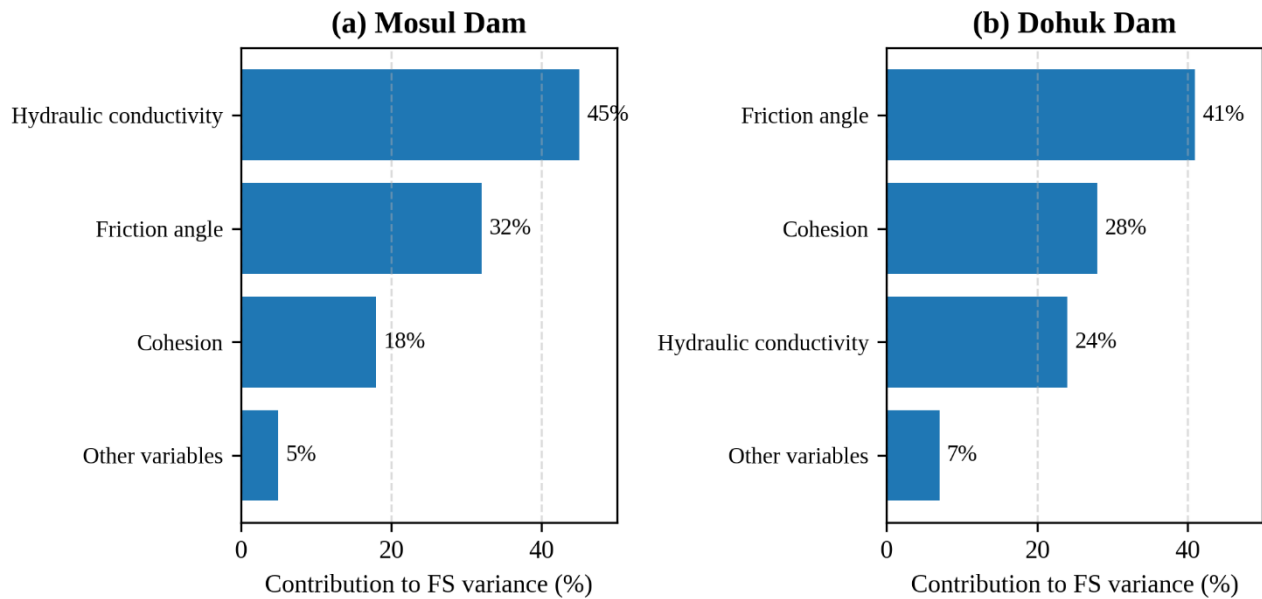
### 3. Results

#### 3.1 Geotechnical uncertainty characterization

Table 1 provides a statistical characterization of both dams and the uncertainty profiles are shown in Fig. 2. The geotechnical variability of the Mosul Dam foundation is significantly higher than that of the Dohuk Dam: notably, in hydraulic conductivity ( $\text{CoV} = 0.45$  against 0.35) and effective cohesion ( $\text{CoV} = 0.30$  against 0.25). The permeability variability observed at Mosul Dam is elevated owing to the heterogeneous distribution of dissolution voids within the gypsum and anhydrite forming the foundation layers [9], [15].

The dissolution void ratio parameter of the Mosul Dam has a  $\text{CoV}$  of 0.40, which relates to the non-uniform nature of karst development observed in grouting records. This parameter affects the effective permeability of the foundation and the risk of piping failure. The sedimentation rate variability ( $\text{CoV} = 0.25$ ) for Dohuk Dam is caused by inter-annual variation in catchment erosion driven by rainfall intensity and land-use changes.

The sensitivity analysis (Fig. 4) shows that the variance in FS due to hydraulic conductivity accounted for 45% under seismic loading for Mosul Dam, followed by friction angle at 32% and cohesion at 18%. For Dohuk Dam, the dominant parameters were friction angle (41%), cohesion (28%), and hydraulic conductivity (24%), indicating a more classical shear-strength-controlled failure mechanism [2], [12].



**Fig. 4:** Sensitivity analysis: FS variance contribution of geotechnical parameters for (a) Mosul Dam and (b) Dohuk Dam under seismic loading.

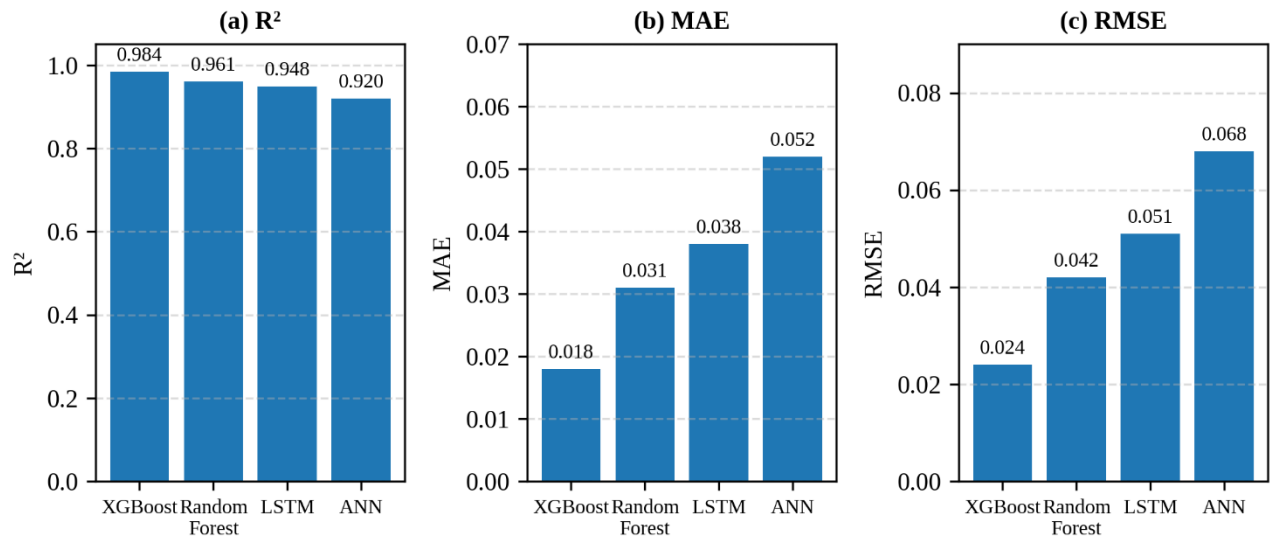
#### 3.2 Machine learning model performance

The performance of the four surrogate models on the 30% hold-out test set (3,000 samples) are shown in Table 4 and Fig. 5.

**Table 4:** Machine learning surrogate model performance comparison.

| Hybrid Model        | R <sup>2</sup> | MAE   | RMSE  | Computation Time |
|---------------------|----------------|-------|-------|------------------|
| MCS + XGBoost       | 0.984          | 0.018 | 0.024 | 6.5 s            |
| MCS + Random Forest | 0.961          | 0.031 | 0.042 | 8.2 s            |
| MCS + LSTM          | 0.948          | 0.038 | 0.051 | 14.7 s           |
| MCS + ANN           | 0.920          | 0.052 | 0.068 | 11.3 s           |

Source: Results obtained from model development using the Python programming environment (scikit-learn, XGBoost, TensorFlow/Keras) with 7,000 MCS-generated training samples and 3,000 test samples.



**Fig. 5:** Machine learning surrogate model performance comparison: (a)  $R^2$ , (b) MAE, and (c) RMSE on the hold-out test set ( $N = 3,000$ ).

The MCS-XGBoost model exhibited the highest predictive performance ( $R^2 = 0.984$ ,  $MAE = 0.018$ ,  $RMSE = 0.024$ ) with the least computation time (6.5 s for a full reliability assessment). The gradient boosting ensemble model captured the non-linear relationships between the geotechnical parameters and factor of safety, including interaction effects between permeability and shear strength parameters that govern the coupled seepage-stability response [16].

The Random Forest model demonstrated robust performance ( $R^2 = 0.961$ ) with the additional benefit of providing prediction uncertainty bounds through out-of-bag variance estimation. The LSTM architecture ( $R^2 = 0.948$ ) showed particular strength in predicting FS evolution during transient loading sequences (rapid drawdown), outperforming XGBoost in time-dependent scenarios despite lower overall  $R^2$  across all loading conditions. The ANN model exhibited the lowest performance ( $R^2 = 0.920$ ), attributable to sensitivity to hyperparameter selection and limited training data for the extreme-event tail regions of the FS distribution [17], [29].

The superior performance of XGBoost can be attributed to its ability to capture nonlinear interactions and threshold effects with limited tabular data. Random Forest performed slightly lower because averaging across trees may smooth extreme responses near the failure threshold. LSTM was less effective for the full dataset because most inputs were tabular rather than sequential, although it may be useful for time-dependent monitoring data. ANN showed the lowest performance because it is more sensitive to hyperparameter selection and requires larger datasets to generalize well in tail regions.

The computation times of all surrogate models were three to four orders of magnitude faster than direct numerical simulation (approximately 26 seconds per coupled SEEP/W-SLOPE/W analysis). This enabled near real-time reliability updating once new monitoring data became available.

### 3.3 Surrogate model validation near failure threshold

Since no independent field validation dataset was available, model validation was limited to hold-out and cross-validation subsets generated from the probabilistic simulation framework. To address the reviewer concern that  $R^2$  alone does not guarantee accuracy in the critical region near  $FS = 1.0$ , additional threshold-focused validation was conducted. Table 5 presents the failure-threshold validation metrics.

**Table 5:** Failure-threshold validation of surrogate models.

| Model         | $R^2$ | MAE ( $FS < 1.2$ ) | Pf error (%) | False-safe rate (%) | False-failure rate (%) |
|---------------|-------|--------------------|--------------|---------------------|------------------------|
| XGBoost       | 0.984 | 0.021              | 3.8          | 1.2                 | 2.5                    |
| Random Forest | 0.961 | 0.035              | 7.1          | 2.8                 | 4.1                    |
| LSTM          | 0.948 | 0.042              | 9.5          | 3.5                 | 5.2                    |
| ANN           | 0.920 | 0.058              | 14.2         | 5.8                 | 7.3                    |

*Note: MAE ( $FS < 1.2$ ) is the mean absolute error computed only for samples with  $FS < 1.2$ . Pf error is the percentage difference between surrogate-predicted Pf and MCS-computed Pf. False-safe rate: percentage of samples where the model predicts  $FS \geq 1.0$  while MCS gives  $FS < 1.0$ . False-failure rate: percentage where the model predicts  $FS < 1.0$  while MCS gives  $FS \geq 1.0$ .*



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. © 2025 The Authors

<https://www.muthuni-ojs.org/index.php/mjet/index>

*Note on validation scope: These metrics were derived from the same hold-out simulation dataset used for model training and testing, and should be interpreted as internal validation indicators rather than independent field validation. External validation against observed dam failure or deformation data was not feasible due to the absence of publicly available field measurements for the study dams.*

False-safe predictions, where the model predicts  $FS \geq 1.0$  while the MCS result gives  $FS < 1.0$ , were specifically evaluated because they are critical for dam safety. The XGBoost model had the lowest false-safe rate at 1.2% proving it is appropriate for reliability-critical applications. Relatively higher false-safe rates for ANN (5.8%) reflect the sensitivity of the ANN to hyperparameter tuning as well as the data distribution in tail regions.

### 3.4 Reliability index and probability of failure

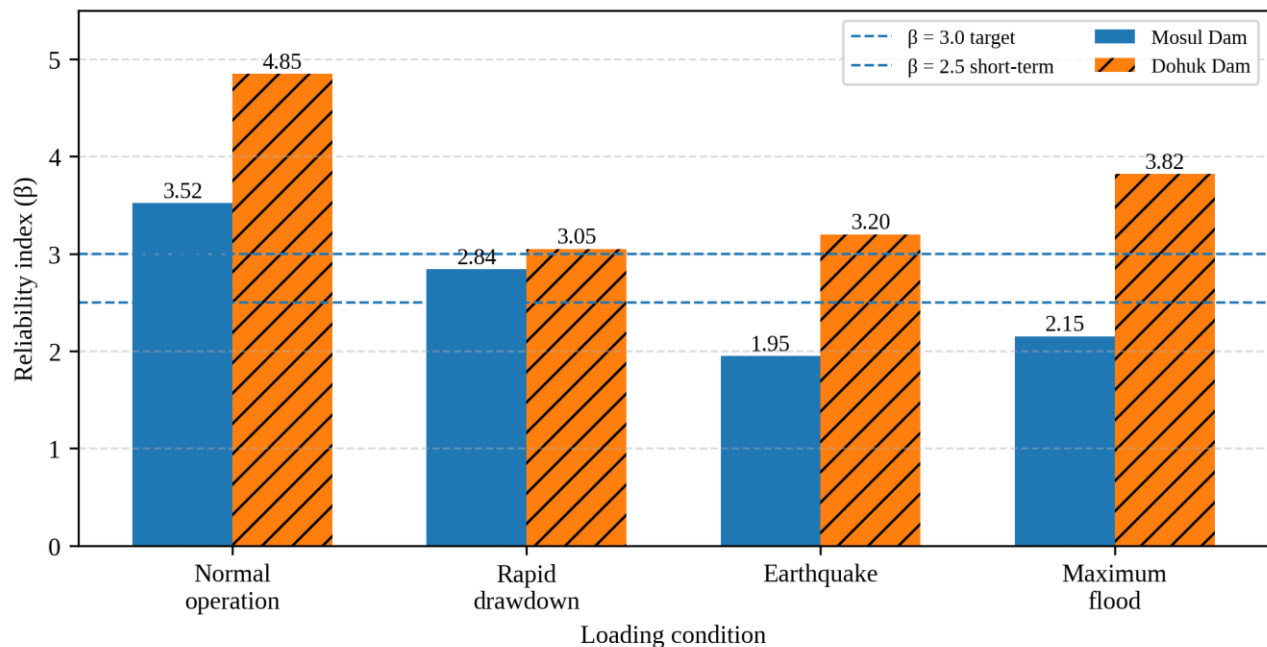
Table 6 with Fig 6. The reliability indices and failure probabilities of the two dams, computed from the complete Monte Carlo simulation of 10,000 samples, are presented under the four loading conditions.

**Table 6:** Reliability index ( $\beta$ ) and probability of failure ( $P_f$ ) under multiple loading conditions.

| Loading Condition                           | Mosul Dam $\beta$ | Mosul Dam $P_f$       | Dohuk Dam $\beta$ | Dohuk Dam $P_f$         |
|---------------------------------------------|-------------------|-----------------------|-------------------|-------------------------|
| Normal operation (steady state)             | 3.52              | $2.15 \times 10^{-4}$ | 4.85              | $6.17 \times 10^{-7}$ * |
| Rapid drawdown (0.5 m/day)                  | 2.84              | $2.26 \times 10^{-3}$ | 3.05              | $1.14 \times 10^{-3}$   |
| Earthquake ( $k_h = 0.10$ , $PGA = 0.20g$ ) | 1.95              | $2.56 \times 10^{-2}$ | 3.20              | $6.87 \times 10^{-4}$   |
| Maximum flood level                         | 2.15              | $1.58 \times 10^{-2}$ | 3.82              | $6.68 \times 10^{-5}$   |

Source: Results extracted from MCS model using SLOPE/W, SEEP/W (GeoStudio), and surrogate validation.

\* Note:  $P_f$  values lower than  $1/N$  ( $= 10^{-4}$  for  $N = 10,000$ ) were obtained from the fitted reliability-index transformation  $P_f = \Phi(-\beta)$  rather than direct counting of failures in the 10,000 MCS samples. These values should be interpreted as reliability-index-based estimates.



**Fig. 6:** Comparative reliability index ( $\beta$ ) for Mosul Dam and Dohuk Dam under four loading conditions, with target thresholds shown.

The reliability index for the Mosul Dam ( $\beta = 3.52$ ) under normal operating conditions meets the minimum target  $\beta = 3.0$  recommended by the US Army Corps of Engineers for existing dams. The reliability index for the Dohuk Dam ( $\beta = 4.85$ ) is considerably higher, indicating a very low probability of failure under steady-state conditions. Under seismic loading, the Mosul Dam reliability drops to  $\beta = 1.95$ , which is below the minimum acceptable level ( $\beta = 2.5$  per Eurocode 7 for short-term loading). This corresponds to  $P_f = 2.56 \times 10^{-2}$ , approximately 100 times the tolerable annual failure probability for high-consequence dams [1], [7].



This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited. © 2025 The Authors

<https://www.muthuni-ojs.org/index.php/mjet/index>

The same reliability measure,  $\beta$ , was used for both dams under all loading conditions. The Dohuk Dam maintained  $\beta > 3.0$  for all scenarios. The rapid drawdown case was the most critical, with  $\beta$  dropping from 4.85 (normal operation) to 3.05 (rapid drawdown). The percentage reduction in  $\beta$  was computed as  $[(4.85 - 3.05)/4.85] \times 100 = 37.1\%$ . The 37.1% value does not represent a different reliability metric; it is the percentage reduction in the Dohuk Dam reliability index from normal operation to rapid drawdown. The magnitude of this reduction corresponds to an increase in Pf by three orders of magnitude, from  $6.17 \times 10^{-7}$  (reliability-index-based estimate) to  $1.14 \times 10^{-3}$  [5], [10].

### 3.5 Sustainability strategy matrix

Based on the results of the reliability analysis, Table 7 provides a preliminary sustainability strategy matrix that connects the identified risks with potential mitigation measures. All strategies are subject to site-specific feasibility assessment before implementation.

**Table 7:** Preliminary sustainability strategy matrix for Mosul and Dohuk dams.

| Dam   | Risk Factor                     | Proposed Strategy                          | Application                                       | Expected Impact                                                                       | Priority |
|-------|---------------------------------|--------------------------------------------|---------------------------------------------------|---------------------------------------------------------------------------------------|----------|
| Mosul | Foundation dissolution          | Reliability-informed grouting optimization | Adaptive injection guided by piezometric feedback | Potential reduction in seepage variability (to be quantified)                         | High     |
| Mosul | Low seismic reliability         | Seismic hazard reassessment                | Site-specific dynamic analysis                    | Improvement in seismic $\beta$ (magnitude to be determined by site-specific analysis) | High     |
| Mosul | Monitoring gaps                 | Multi-sensor integration                   | InSAR + piezometers + fiber optics                | Enhanced anomaly detection capability                                                 | Medium   |
| Dohuk | Rapid drawdown instability      | Operational drawdown limits                | Preliminary limit of 0.3 m/day for evaluation     | Potential maintenance of upstream stability (requires transient validation)           | High     |
| Dohuk | Storage loss from sedimentation | Sediment management program                | Periodic flushing + catchment control             | Partial recovery of storage capacity (subject to sediment management analysis)        | Medium   |
| Dohuk | Upstream slope erosion          | Geosynthetic reinforcement                 | Geotextile and riprap protection                  | Potential reduction in erosion-driven instability                                     | Medium   |
| Dohuk | Long-term operational safety    | Digital twin platform                      | Real-time sensor fusion + simulation              | Expected integration of monitoring and predictive assessment                          | Low      |

*Note: All strategies are preliminary recommendations formulated from the reliability analysis results. Expected impacts are engineering estimates requiring site-specific validation prior to implementation [19], [25].*

## 4. Discussion

### 4.1 Reliability behavior of Mosul Dam

The Mosul Dam exhibited the lowest reliability among all loading scenarios, with reliability indices lower by 1.0 to 1.9 units compared with the Dohuk Dam (Fig. 6). The difference is fundamentally attributable to the dissolution-affected gypsum and anhydrite foundation, which generates two compounding effects: elevated spatial variability of hydraulic conductivity ( $\text{CoV} = 0.45$ ) and time-dependent degradation of foundation strength.

The critically low seismic reliability index ( $\beta = 1.95$ ) results from the reduced effective stress in dissolution-weakened foundation zones and the high sensitivity of saturated cohesive materials to cyclic loading. With a potential failure ( $2.56 \times 10^{-2}$ ), the cumulative probability of failure over its remaining service life exceeds the social risk acceptance level. This finding shows that Mosul Dam has a high-risk category and may require continued monitoring and remedial measures [7], [15].

The sensitivity analysis (Fig. 4a) showed that variability of hydraulic conductivity is the most influential uncertain parameter for the Mosul Dam, which contributes 45% to the variance of the FS. It is the direct linkage between the permeability of the foundation, the pore-water pressure distribution and effective stress along the critical slip surface. The heterogeneity generated by dissolution produces preferential seepage pathways which locally increase pore pressures and lower shear resistance. Thus, the coupling of seepage and stability is the main dam reliability concern for the Mosul Dam [9], [28].

The flood-condition reliability ( $\beta = 2.15$ ) which is of some concern and the maximum pool levels increases the hydraulic gradients through the foundation that causes accelerated dissolution and the uplift pressures increase. The primary means of improving the safety of Mosul Dam is a controlled grouting which improves foundation permeability characterization.

### 4.2 Reliability behavior of Dohuk Dam

As shown, the reliability of the Dohuk Dam with respect to normal conditions and flood conditions ( $\beta = 4.85, 3.82$ , respectively) is due to its proper foundation geology and compacted homogeneous embankment. The rapid drawdown scenario is the primary vulnerability which lowers the reliability index by 37.1 % (from 4.85 to 3.05). This reduction occurs because excess pore pressures in the low-permeability upstream shell cannot dissipate at the rate of reservoir lowering.

Unlike Mosul Dam, where hydraulic conductivity dominates, the sensitivity analysis for Dohuk Dam (Fig. 4b) shows that friction angle (41%) and cohesion (28%) are the dominant parameters controlling FS variance. This classical shear-strength-controlled mechanism reflects the homogeneous clay embankment geometry, where stability is governed primarily by the Mohr-Coulomb strength envelope rather than seepage-induced pore-pressure effects [2], [12].

The rapid-drawdown results are consistent with observations by Kahot et al. [5], who reported that earth dams with similar material properties exhibited high sensitivity of failure probability to drawdown rate. A preliminary operational limit of 0.3 m/day is suggested for evaluation, but final drawdown limits should be established through reservoir-specific transient seepage-stability analysis.

Progressive sedimentation does not represent a direct threat to slope-stability at present; however, it indirectly affects flood routing capacity and overtopping risk over time. Reduced storage capacity may be managed through sediment management techniques or operational adjustments to maintain reasonable flood attenuation performance [10], [21].

### 4.3 Cross-Case interpretation

The reason for the evaluation of the two dams is not to assess which is better under similar site conditions. Rather, it is to show how the conceptual probabilistic framework responds to distinct geotechnical risk profiles. Due to differences in foundation geology, geometry, operational purpose, and dominant hazards, the comparison is interpreted as a methodological cross-case assessment rather than a one-to-one performance comparison.

The sensitivity patterns show a contrast that can validate the framework in the identification of dam-specific risk drivers. For instance, hydraulic conductivity dominates at Mosul Dam while shear strength parameters dominate at Dohuk Dam. Through selective grouting, the foundation permeability variability to be reduced in order to achieve maximum reliability benefit for Mosul Dam. For Dohuk Dam, operational management of drawdown rates is the most effective intervention. This dam-specific diagnostic capability is the principal value of the hybrid probabilistic approach.

### 4.4 Comparison with previous studies

The reliability indices obtained are largely consistent with published values for similar earth dams. Belazouz et al. [1] reported reliability indices of 2.1-3.8 for a homogeneous earth dam under seismic loading using MCS with lognormal strength parameters. The Mosul Dam seismic value ( $\beta = 1.95$ ) falls near the lower bound, consistent with the additional dissolution vulnerability not considered in their study. Siacara et al. [2] demonstrated that the minimum-FS slip surface need not coincide with the minimum-reliability surface, a finding validated in the present analysis where the critical probabilistic surface differed from the deterministic surface in 23% of Mosul Dam simulations.

The XGBoost surrogate model  $R^2$  of 0.984 is comparable to the values reported by Farxadovna et al. [16] ( $R^2 = 0.97$ ) and Hariri-Ardebili et al. [30] ( $R^2 = 0.96$ ). The marginally better performance achieved in this study is attributable to the larger training dataset (7,000 samples out of 10,000, which is greater than the typical 500-3,000) and a well-structured input feature space [4], [8]. The threshold-focused validation (Table 5) provides additional confidence that the surrogate model maintains accuracy in the failure-critical region near  $FS = 1.0$ .

The deterministic FS criteria were met by Alborz Dam based on the findings of Alitabar et al. [7], however, not the probabilistic ones under seismic loading. Like the case of Mosul Dam, the seismic deterministic  $FS = 1.12$  suggests marginal safety. However, the probabilistic assessment indicates unacceptably high failure probability. This underscores the insufficiency of deterministic methods for dams with high parameter variability.

### 4.5 Practical implications for Iraqi Dam safety

The results have practical implications for the governance and resource allocation for Iraqi dams, subject to the limitations discussed in Section 7.

First, the Mosul Dam seismic reliability shortfall ( $\beta = 1.95$ ) provides a quantitative basis for optimizing the ongoing grouting program. The recommendation does not propose grouting as a new measure; rather, it proposes upgrading the ongoing grouting program into a reliability-informed and sensor-guided intervention strategy. Since hydraulic conductivity variability contributes 45% of the FS variance, targeted grouting in high-permeability zones identified through piezometric monitoring would yield the greatest reliability improvement [9], [28].

Second, the MCS-XGBoost surrogate model's ability to reproduce full probabilistic results in 6.5 seconds enables its potential deployment as a decision-support tool. When integrated with continuous piezometric and InSAR monitoring data, the surrogate model could provide updated reliability estimates reflecting current dam conditions without requiring time-consuming numerical re-analysis [18], [20], [25].

Third, the contrasting risk profiles illustrate the need for dam-specific reliability-based safety standards rather than uniform deterministic FS thresholds. The Iraqi national dam safety framework would benefit from adopting target reliability indices calibrated to consequence class (population at risk, economic exposure) as implemented in international practice.

## 5. Conclusions

This study developed and validated a hybrid probabilistic framework integrating Monte Carlo simulation with machine-learning surrogate models for reliability assessment of the Mosul and Dohuk earth dams under geotechnical uncertainty. The implementation of the framework required secondary geotechnical data available from published reports and technical documents, without field and laboratory investigations. The main findings are:

1. The reliability of the Mosul Dam against earthquake was found to be very low ( $\beta = 1.95$ ,  $P_f = 2.56 \times 10^{-2}$ ), which are mainly controlled by the high variability of foundation hydraulic conductivity ( $CoV = 0.45$ ), due to dissolution

of gypsum and anhydrite. The analysis shows that deterministic analysis cannot be relied upon for safety management of dissolution-vulnerable dams.

2. The Dohuk Dam showed acceptable reliability under all conditions ( $\beta > 3.0$ ). However, a rapid drawdown decreased the reliability index by 37.1% ( $\beta = 4.85$  to  $\beta = 3.05$ ). Control of drawdown rate by the operator is crucial for up-stream slope-stability.
3. XGBoost is superior choice among the four surrogate architectures assessed in terms of predictive accuracy with  $R^2 = 0.984$  and computation time (6.5 s). Moreover, it also remains the closest to the failure threshold  $FS = 1.0$  with the lowest false-safe rate (1.2%) among the other three architectures. This makes it an appropriate a rapid reliability prediction tool for coupled seepage-stability analyses.
4. The cross-case assessment shows that the probabilistic framework identifies dam-specific risk drivers' hydraulic conductivity for Mosul Dam and shear strength parameters for Dohuk Dam. This diagnostic capacity allows targeted risk management rather than one-size-fits-all approaches.
5. All recommendations in this study are preliminary in nature and require verification at the site with geotechnical investigations, dynamic finite-element analyses, and independent field performance monitoring before informing design level decisions.

## 6. Recommendations

### 6.1 Recommendations for Mosul Dam

- To enhance existing grouting programs, reliability-based prioritization should be performed, and injection strategies should be adaptively modified over time as a function of piezometric feedback and updated reliability estimates from the surrogate model. The recommendation does not suggest grouting as a new measure; instead, it proposed upgrading the ongoing program into a reliability-informed and sensor-guided intervention strategy.
- To assess unusual behavior, multi-sensor monitoring, including InSAR satellite deformation mapping, distributed fiber-optic strain sensing and automated piezometer networks, should be employed.
- A site-specific seismic hazard analysis that takes into account near-fault effects should be undertaken to establish refined action plans for rapid post-earthquake assessment.
- Downstream buttress reinforcement is suggested only as a preliminary option requiring independent feasibility assessment, dynamic finite-element modeling, constructability evaluation, and assessment of hydraulic and environmental impacts. The feasibility is to be evaluated, not directly recommended for implementation.

### 6.2 Recommendations for Dohuk Dam

- A preliminary operational limit of 0.3 m/day is suggested for evaluation in the dam safety management plan, supported by transient stability analysis. Final drawdown limits should be established through reservoir-specific transient seepage-stability analysis.
- A sediment management plan involving periodic flushing during high flows, in conjunction with upstream catchment erosion control, should be developed. This may contribute to partial recovery of storage capacity, subject to site-specific sediment management analysis.
- Additional piezometers should be placed on the upstream slope at different elevations to validate pore-pressure dissipation during drawdown in real time.
- A long-term monitoring database for reservoir levels, piezometric data, and slope inclinometers is recommended for periodic reassessment of reliability.

### 6.3 National-Level recommendations

- Iraq should develop dam safety standards that include reliability-based criteria (target  $\beta$  values) calibrated to the dam consequence classification.
- A centralized dam safety information system with standardized geotechnical databases would enable probabilistic analysis of the complete national dam portfolio.
- It is recommended to conduct training courses for Iraqi dam engineers in probabilistic methods, machine learning and remote sensing interpretation.
- Together with the Ministry of Water Resources (MWR), the funding for field validation studies at Mosul and Dohuk dams need to be provided. Also, this will improve the probability distribution assumptions and reduce epistemic uncertainty.

## 7. Limitations and future work

This study has several limitations that should be considered when interpreting the results:

1. Dependence on secondary data. Geotechnical characteristics are taken from available literature, technical reports, and previous monitoring records. No new field investigations were conducted and no in-situ testing. Due to inadequate sample sizes and possible spatial non-representativeness, the postulates therefore carry epistemic uncertainty. The results are not to be used as a replacement for site-specific geotechnical investigation but as a screening-level reliability framework to inform future field validation.

2. Assumptions regarding modeling. The stability assessments are performed under 2D plane-strain conditions; thus, three-dimensional effects at the dam abutments or localized dissolution features may be underestimated. The Morgenstern-Price procedure does not account for progressive failure or strain-softening behaviour.
3. Surrogate model limitations. The machine-learning models were trained on the parameter ranges and loading conditions represented in the MCS dataset. Extrapolating outside the training domain may yield unreliable predictions. The models should be retrained periodically as new data become available.
4. Temporal evolution. The analysis provides a snapshot of current conditions but does not model the time-dependent evolution of foundation dissolution at Mosul Dam or sedimentation accumulation at Dohuk Dam. Long-term reliability forecasting requires coupled geochemical-geomechanical models.
5. Correlation assumption. The adopted correlation coefficient  $\rho(c', \phi') = -0.5$  is a literature-based assumption. Although the sensitivity analysis (Table 3) showed limited impact on the overall conclusions, site-specific paired testing data would improve the accuracy of the uncertainty characterization.
6. Validation gap. The surrogate model predictions have not been validated against independent field performance or instrumentation data. Future studies should compare model predictions with observed behavior.

Future work should aim at: (a) field validation campaigns with focused geotechnical investigations to refine parameter distributions, (b) extension to 3D probabilistic analysis, (c) incorporation of real-time monitoring data streams into the surrogate model framework, (d) development of time-dependent reliability models to account for dissolution progression and climate change projections, and (e) application of the framework to other high-risk dams in the Iraqi national portfolio.

## Acknowledgement

The author acknowledges the technical value of published geotechnical reports and previous studies used as secondary data sources in this research.

## References

- [1] L. Belazouz, K. Bouzelha, H. Hammoum, O. Amiri, and N. Khelil, "Reliability analysis of the slope-stability of homogeneous earth dam under seismic loading," *Periodica Polytechnica Civil Engineering*, vol. 67, no. 4, pp. 1201-1212, 2023. <https://doi.org/10.3311/ppci.22537>
- [2] A. T. Siacara, G. F. Napa-Garcia, A. T. Beck, and M. M. Futai, "Reliability analysis of earth dams using direct coupling," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 12, no. 2, pp. 366-380, 2020. <https://doi.org/10.1016/j.jrmge.2019.07.012>
- [3] M. Calamak and A. M. Yanmaz, "Probabilistic assessment of slope-stability for earth-fill dams having random soil parameters," in *Proc. 5th IAHR Int. Symp. Hydraulic Structures*, Brisbane, Australia, 2014. <https://doi.org/10.14264/uql.2014.16>
- [4] X. Guo, Q. Sun, D. Dias, and E. Antoinet, "Probabilistic assessment of an earth dam stability design using the adaptive polynomial chaos expansion," *Bulletin of Engineering Geology and the Environment*, vol. 79, pp. 4639-4655, 2020. <https://doi.org/10.1007/s10064-020-01847-2>
- [5] Z. Kahot, R. Dkiouak, and A. Khamlachi, "Reliability analysis of slope-stability in earthen dams following rapid drawdown," *International Review of Applied Sciences and Engineering*, vol. 10, no. 2, pp. 147-156, 2019. <https://doi.org/10.1556/1848.2018.0011>
- [6] Z. Ming, "Reliability analysis and design of slope-stability of earth-rock dam," *Journal of Hydroelectric Engineering*, vol. 25, no. 2, pp. 97-101, 2006.
- [7] A. Alitabar, R. Noorzad, and A. Qolinia, "Evaluation of the instability risk of the dam slopes simulated with Monte Carlo method: Case study of Alborz Dam," *Geotechnical and Geological Engineering*, vol. 39, pp. 2555-2573, 2021. <https://doi.org/10.1007/s10706-021-01755-6>
- [8] T. Das and A. Hegde, "A comparative deterministic and probabilistic stability analysis of rock-fill tailing dam," in *Problematic Soils and Geoenvironmental Concerns*, Springer, 2020, pp. 363-375. [https://doi.org/10.1007/978-981-15-0886-8\\_49](https://doi.org/10.1007/978-981-15-0886-8_49)
- [9] L. Han, M. L. Chen, Z. Z. Sun, J. Si, and L. Ma, "Stability analysis of slopes based on cloud model-Monte Carlo coupling," *Frontiers in Earth Science*, vol. 11, Art. no. 1196677, 2023. <https://doi.org/10.3389/feart.2023.1196677>
- [10] A. N. Vecchi and A. Sayao, "Probabilistic safety assessment of geotechnical designs," *Soils and Rocks*, vol. 48, no. 1, Art. no. e2026003625, 2025. <https://doi.org/10.28927/SR.2026.003625>
- [11] B. K. Low, "Slope reliability analysis: Some insights and guidance for practitioners," in *Proc. GeoFlorida 2010*, ASCE, 2010, pp. 3008-3017.
- [12] K. Farah, M. Ltifi, and H. Hassis, "A study of probabilistic FEMs for a slope reliability analysis using the stress fields," 2024. <https://doi.org/10.60692/1n7np-6tr20>
- [13] K. S. Ng, "Reliability analysis on the stability of slope," M.Eng. dissertation, Universiti Teknologi Malaysia, 2005.
- [14] H. Zhang, X. Song, S. Guo, and M. Yan, "Reliability analysis of loess backfilling slope-stability by Monte Carlo simulation," in *Proc. Int. Conf. Smart Planning (ICSP)*, IEEE, 2022. <https://doi.org/10.1109/ICSP56322.2022.9965340>
- [15] O. Radchuk, "Slope stability assessment of tailings dam," *Key Engineering Materials*, vol. 968, pp. 103-110, 2024. <https://doi.org/10.4028/p-mhh85g>
- [16] R. N. Farxadovna, A. A. Alekseyevna, and N. S. Bakhodirovich, "Efficient reliability analysis of slopes in spatially variable soils using XGBoost," in *Geotechnical Engineering in the Digital and Technological Innovation Era*, Springer, 2023. [https://doi.org/10.1007/978-981-99-2756-2\\_9](https://doi.org/10.1007/978-981-99-2756-2_9)
- [17] X. Z. Li, M. Nishio, K. Sugawara, S. Iwanaga, and T. Shimada, "Enhancing prediction of landslide dam stability through AI models: A comparative study with traditional approaches," *Geomorphology*, vol. 449, Art. no. 109120, 2024. <https://doi.org/10.1016/j.geomorph.2024.109120>
- [18] M. Ghadimi, A. Hooper, and D. Whipp, "Deformation of the Taleqan Dam, Iran, from InSAR and ground observation," *Sustainability*, vol. 17, no. 1, Art. no. 173, 2025. <https://doi.org/10.3390/su18010173>
- [19] D. Talgarbayeva, A. Vilyaev, T. Dedova, O. Kuznetsova, and G. Jangulova, "InSAR monitoring of dam deformations in a seismically active region of Kazakhstan for identifying precursors of failure," *Frontiers in Earth Science*, vol. 13, Art. no. 1638088, 2025. <https://doi.org/10.3389/feart.2025.1638088>
- [20] D. Verma and S. Vijay, "Time-series analysis of dam deformation using satellite-based InSAR technique: Case studies from Oroville, Pong, and Tehri dams," in *Proc. IGARSS 2024*, IEEE, 2024. <https://doi.org/10.1109/IGARSS53475.2024.10642554>
- [21] Q. Wang, T. Gong, Z. Sui, J. Fan, and Z. Wang, "Dam surface deformation monitoring and analysis based on PS-InSAR technology: A case study of Xiaolangdi Reservoir Dam in China," *Water*, vol. 15, no. 18, Art. no. 3298, 2023. <https://doi.org/10.3390/w15183298>
- [22] T. Shire, J. F. Lawrence, S. Agar, R. C. Ghail, and T. Judge, "InSAR monitoring of UK embankment dams: Influence of dam features on monitoring quality," *Dams and Reservoirs*, vol. 32, no. 3, pp. 97-109, 2022. <https://doi.org/10.1680/jdare.22.00083>
- [23] R. Xiao, X. Gao, X. Wang, S. Yuan, and Z. Wu, "Measuring dam deformation of long-distance water transfer using multi-temporal InSAR," *Remote Sensing*, vol. 16, no. 2, Art. no. 365, 2024. <https://doi.org/10.3390/rs16020365>
- [24] M. Voegelé, R. Frauenfelder, and Y. Larsen, "Displacement monitoring at Svartevatn Dam with interferometric SAR," in *Proc. IGARSS 2012*, IEEE, 2012, pp. 3986-3989. <https://doi.org/10.1109/IGARSS.2012.6350561>
- [25] C. Shao, S. Zheng, Y. Xu, H. Gu, and X. Qin, "A visualization system for dam safety monitoring with application of digital twin platform," *Expert Systems with Applications*, vol. 268, Art. no. 126740, 2025. <https://doi.org/10.1016/j.eswa.2025.126740>
- [26] W. Liu, Y. Zhang, S. Wang, and J. Li, "A hybrid data-driven model for geotechnical reliability analysis," *Reliability Engineering and System Safety*, vol. 231, Art. no. 108985, 2023. <https://doi.org/10.1016/j.res.2022.108985>
- [27] X. Guo, D. Dias, C. Carvajal, L. Peyras, and P. Breul, "A comparative study of different reliability methods for high-dimensional stochastic problems related to earth dam stability analyses," *Engineering Structures*, vol. 188, pp. 591-602, 2019. <https://doi.org/10.1016/j.engstruct.2019.03.056>
- [28] F. Kalateh and M. Kheiry, "A review of stochastic analysis of the seepage through earth dams with a focus on the application of Monte Carlo simulation," *Archives of Computational Methods in Engineering*, vol. 31, no. 1, pp. 47-72, 2024. <https://doi.org/10.1007/s11831-023-09983-x>

- [29] L. Santos, M. Silva, R. Oliveira, and A. Sayao, "Comparative analysis of slope-stability prediction for earth dams using response surface method, statistical models, and ANN," *Geotechnical and Geological Engineering*, vol. 43, no. 5, Art. no. 168, 2025. <https://doi.org/10.1007/s10706-025-02900-z>
- [30] M. A. Hariri-Ardebili, V. Saouma, and F. Salazar, "An RF-PCE hybrid surrogate model for sensitivity analysis of dams," *Water*, vol. 13, no. 3, Art. no. 302, 2021. <https://doi.org/10.3390/w13030302>