



Predicting slope failure with intelligent hybrid modeling of ANFIS with GA and PSO

Jayanti Prabha Bharti^{1,2} · Pijush Samui¹

Received: 1 March 2024 / Accepted: 28 May 2024 / Published online: 18 June 2024
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2024

Abstract

In geotechnical engineering, soil slopes are crucial in various civil engineering projects, including highways, embankments, dams, and excavations. Understanding the behavior of soil slopes is essential for designing stable and safe structures. Combining different soft computing (SC) models can provide more robust slope stability predictions. This paper employs two hybrid computational algorithms to make accurate slope stability predictions. In this research project, the adaptive neuro fuzzy inference system (ANFIS) model is optimized by two novel meta-heuristic optimization algorithms (MOAs): genetic algorithm (GA) and particle swarm optimization (PSO). To this end, slope inputs are taken from a literature survey consisting of 206 input datasets for the training and testing of models. Eleven statistical indices have been evaluated for assessing the performance of proposed hybrid models, along with evaluating rank analysis. ANFIS, ANFIS-GA, and ANFIS-PSO outcomes from the suggested models have R^2 values of 0.6783, 0.7624, 0.7378 during training, 0.6684, 0.8143, and 0.7013 during testing. Also, the ANFIS-GA hybrid model yielded error matrices such as RMSE, MAE, and MSE with values of 0.1217, 0.0912, and 0.0148 in training and 0.12570, 0.0968, and 0.1391 in testing; in contrast, the ANFIS PSO model yielded values of 0.1264, 0.0902, 0.016 in training, and 0.1591, 0.1170, 0.1290 in testing; the ANFIS model yielded values of 0.1345, 0.1127, 0.0172 in training, and 0.1642, 0.1267, 0.1391 in testing. The regression plot was analyzed to compare the predicted value with the actual one. In the present paper, the Metropolis Hastings MCMC sampling method has been introduced to establish the relationship between the inputs, which is slope height (H), slope angle (α), cohesion (c), pore water pressure ratio (R_u), unit weight (γ), angle of internal friction (φ), and output reliability of slopes. A sensitivity analysis was also performed to determine which variable affects the reliability of soil slope more. After that, comparing hybrid models with the ANFIS model notified the engineers and researchers that the model best predicts slope failure for extensive observations.

Keywords Slope failure · ANFIS · Genetic algorithm (GA) · Particle swarm optimization (PSO) · Regression plot · Metropolis Hastings sampling distribution · Sensitivity analysis

1 Introduction

Landslides are frequent geological happenings that significantly harm society and the economy. They cause public and private property damage yearly, totaling hundreds of millions of dollars. According to Massey et al. (2017), these natural disasters are influenced by several variables, including internal variables like slope configuration and soil characteristics, as well as external triggers like rainfall and earthquakes.

Slope stability modeling necessitates assessing the techniques used to control the behavior of the slopes to prevent or lessen landslide damage. Slope stability is traditionally assessed by evaluating the factor of safety (FOS) and choosing an appropriate method design. A FOS based on a proper geotechnical model is needed to evaluate the slope stability. Estimating slope stability involves many factors, and the FOS calculation needs physical information about the geologic materials, geometrical information and its shear-strength parameters, information on pore water pressures, etc. Das et al. (2011) said the slope is considered stable if the FOS is greater than 1.0 and unstable if it is less than 1.0. The limit equilibrium method (LEM) and strength reduction method SRM were described by Cheng et al. (2007a, b) and Nash (1981). For simple homogenous soil slopes, it was

✉ Jayanti Prabha Bharti
jayantib.phd18.ce@nitp.ac.in

¹ Department of Civil Engineering, National Institute of Technology, Patna, India

² Assistant professor, GEC JEHANABAD, Hulasganj, India

found that the safety factor and critical surface determination from these two methods, LEM and SRM, are generally in good agreement except for frictionless soil. However, they were sensitive to the nonlinear solution with frictional material and could not determine their failure surfaces. Griffiths and Fenton (2004) and Luo et al. (2016) used the random finite element and material point methods. Griffiths and Fenton (2004) investigated the probability of slope failure for cohesive soil using the random finite-element method (RFEM) combined with elastoplasticity and random field theory. Also, Luo et al. (2016) analyzed geosynthetic reinforced soil slopes using the shear strength reduction method in combination with the finite element method (FEM). They observed the influence of random soil properties on the probability of failure. However, the predicted safety factor was sensitive to uncertainty regarding soil friction angle. However, the utility of these approaches decreases if large numbers of observations are required to achieve very low probabilities of failure. Sun et al. (2014) and Tschuchnigg et al. (2015a, b) studied the strength reduction method and finite element limit analysis. Sun et al. (2014) explored a new method for the selection and forecasting of mechanics parameters of the rock mass, which tested the integrated use of analogy simulation experiment technology and numerical calculation. According to Li and Xu (2016), the SSPC method assessed the Failure probability of the excavation slopes based on the possible failure mechanisms and modes of the slopes. The results show that the SSPC method exhibited an accuracy of 70 % in the stability assessment of the highway slopes in the Yunnan province. Tschuchnigg et al. (2015a, b) observed that the numerical instabilities of non-associated displacement were minimized using finite element limit analysis with the Davis B procedure. However, it was still tricky in the case of limit theorems of plasticity. The finite difference method for the analysis of soil slopes was studied by Lim et al. (2015). They examined the three-dimensional (3D) slope stability of two-layered undrained clayey slopes using the finite element upper and lower bound limit analysis methods that are finite difference methods, and the results were presented in the form of stability charts that may be convenient tools for civil engineers. However, this method still had many difficulties in handling complex geometry. According to Li et al. (2016) and Reale et al. (2016), the Markov chain model and the numerical limit analysis methods are all LEM techniques based on elasticity and plasticity concepts. They may frequently be used to calculate the uncertainty in slope stability. However, they may have limited predictive capability for extreme events such as rapid landslides or catastrophic failures. Tschuchnigg et al. (2015a, b) said that the assumptions made for slip surfaces and inter-slice forces limit the LEM's accuracy. In contrast, numerical methods call for an effectively fitting core model, which is challenging to develop. Hence, we can conclude that the discussed slope stability method has less accuracy

due to the intricacy of the relationships between the variables influencing slope stability.

Recently, SC approaches have been used successfully to model and classify any problem. They are helpful when the precise scientific tools cannot provide a comprehensive, analytical, and economical solution. Fattahi (2017) studied the advantages of employing SC, which is more precise and specific in achieving tractability and robustness. Furthermore, in recent years, with the development of cheaper personal computers, the intelligence system approaches have been increasingly used in the stability of slope analysis, such as slope stability prediction using fuzzy logic (FL) studied by Saboya et al. (2006) and Ercanoglu and Gokceoglu (2002). McCombie and Wilkinson (2002) used a simple Genetic Algorithm. They applied successfully to the hunt for the critical circle in slope stability analysis, producing results much more quickly than a 'brute force' approach. Kahatadeniya et al. (2009) used an ant colony optimization algorithm, and Cheng et al. (2007a, b) used a PSO algorithm to analyze the soil slope stability.

Additionally, with the advent of more affordable personal computers, intelligence system approaches have been utilized to stabilize slopes. The major drawback of these models are studied by Park and Rilett (1999) that they provide no information on the relative importance of the various parameters like other statistical models. Since the information learned during training is automatically retained in ANNs, it is exceedingly challenging to reasonably interpret the network's overall structure. Additionally, Kecman (2001) noted that ANN has certain natural shortcomings, including delayed convergence time, poorer generalizing ability, getting at local minimums, and over-fitting issues.

The hybrid SC model is frequently used to solve different engineering problems. According to research published by Riahi Madvar et al. (2020), the hybridization approach will increase the accuracy of the ANFIS technique as it modifies and enhances the computational parameters based on a reliable database worldwide. Furthermore, the fuzzy set theory is a valuable tool for managing ambiguity in engineering application decision-making. As a result, fuzzy sets are gaining more and more interest in contemporary data analysis, industrial techniques, detection of patterns, diagnostics, and information technology. Agnihotri et al. (2022) optimize model parameters for predicting flood at the Matijuri gauge station of the Barak River basin, Assam, India. They combined the ANFIS model with Ant Colony Optimization (ACO) for optimization. Bousnina et al. (2023) used the PSO, ANN, and ANFIS models to predict surface quality constant energy during the milling of alloy. They confirmed the efficiency of integrating the PSO model into the ANN neural network by comparing it with the adaptive neuro-fuzzy inference system (ANFIS). Tao et al. (2024) reviewed that nature inspired algorithm are superior than all other

algorithms for river flow modeling. Zhang et al. (2023) optimized the ANFIS model to predict the thermophysical properties of nanofluids. They presented GP-ANFIS, SC-ANFIS, and FCM-ANFIS for predicting specific heat capacity and thermal conductivity of hybrid fluids.

Samantaray et al. (2023) used the AI hybrid method to predict flood discharge collected from four gauging stations of River Brahmani, Odisha, India. They developed a hybridized metaheuristic algorithm, i.e., ANFIS-PSOSMA, that improves the Slime mold algorithm's (SMA) exploration capability by integrating it with particle swarm optimization (PSO) for discharge prediction. They summarized that combining the optimization algorithms with ANFIS enhances performance in modeling monthly flood discharge time series. Mu'azu (2023) investigate the efficacy of teaching–learning-based optimization (TLBO) for tuning two well-known predictive models, namely artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS), applied to the prediction of the factor of safety for a cohesive soil slope-footing system. He investigated that ANN-TLBO was the most robust model, followed by ANFIS-TLBO, ANN, and ANFIS for slope stability prediction. Qin et al. (2022) investigated the SVNN ANFIS model to predict open-pit mine slope stability. The SVNN-ANFIS method provides a new, effective way to evaluate open-pit mine slope stability in an uncertain environment. Nemato-lahi et al. (2018) proposed the PSO FIS (Fuzzy Inference System) and GA FIS models to predict the hydraulic conductivity of saturated soil. They concluded that FIS trained by PSO and FIS trained by GA could be good estimators in filling the gap between available primary soil data and other soil hydraulic characteristics. Moayedi et al. (2020) built 60 different ANN models, 12 hybrid GA-ANFIS models, and 36 PSO-ANFIS models to predict α in driven piles. They used input parameters such as pile diameter, pile length, relative density, embedment ratio (L/D), and both pile end resistance and base resistance at a relatively 10% base settlement from the CPT result, whereas the output was α . They confirmed that the hybridized model is more efficient for the problem above. Yang et al. (2020) also proposed a hybrid algorithm of GA and PSO with an ANFIS model for the computation of ground vibration during rock blasting. They gathered only 86 data for hybrid modeling and concluded that hybrid models are more reliable for evaluating ground vibration. This indicates higher reliability of the optimized GA-ANFIS model in estimating α ratio in driven shafts. The above literature study of hybrid model conclude that intelligent hybrid models give more accurate and reliable results than ANN, ANFIS, and other SC models. Yang et al. (2020) and Moayedi et al. (2020) used 80% of data in the training phase and 20% in testing. Despite the researchers' best efforts, the dataset is frequently too small to determine its repeatability and applicability. The severe flaws in earlier research on the determination of

safety factors for naturally occurring soil slopes are employed techniques having a limited scope of data used, a restricted range of calibration data, a limited scope of applicability, non-generality of acquired data, and imprecise estimations over unseen data. This study attempts to address all of these issues by combining the nature-inspired algorithm with conventional ANFIS to further enhance the process of modeling mechanism in an extended database of reliability index (RI) for prediction of safety of slope.

The present study proposes a comparative analysis of two hybrid intelligent models, ANFIS-GA and ANFIS-PSO with ANFIS, to predict the safety of slopes and determine the most influencing input parameter. Also, in the proposed research, 70% of the data was used for the training phase and 30% in the testing phase. The novelty of the proposed study is to explore the two evolutionary algorithms, GA and PSO, to enhance the performance of the ANFIS model to predict the failure of slopes for large datasets and also assess the effectiveness and precision of applied hybrid models in comparison to ANFIS model for slope stability problem in the field of civil engineering.

2 Physical significance of study

Slope stability investigations are crucial in various fields, including civil engineering, environmental science, and natural hazard assessment. The physical significance of slope stability studies lies in understanding and managing the stability of slopes, which can have significant implications for safety, infrastructure, and the environment as per Bharti et al. (2021). Understanding the stability of slopes helps engineers design structures that can withstand potential slope failures, reducing the risk of accidents and property damage. It contributes to predicting and preventing landslides, which are often triggered by factors such as heavy rainfall, earthquakes, or human activities. Authorities can implement mitigation measures to protect communities and infrastructure by identifying landslide-prone areas. Agriculture, forestry, and mining often involve working on or near slopes. Slope stability studies help manage these activities to minimize environmental impact and optimize resource extraction. Existing infrastructure, such as roads and pipelines, may be affected by changes in slope stability over time. Regular slope stability assessments can guide maintenance and repair efforts to ensure infrastructure's continued functionality and safety. Many SC models are being used and implemented to estimate soil slope stability accurately. In the present research, ANFIS, with its hybrid model ANFIS-GA and ANFIS-PSO, has been deeply studied, elaborated on, and described in subsequent sections of the paper. These hybrid models are assessed through 11 performance parameters. In

this work, we have done sensitivity analysis to evaluate the factors that influence the reliability of soil slopes.

3 Dataset preparation

3.1 Data collection

This section covers developing a data set for modeling soil slope variables. Figure 1 represents a basic soil slope model simulation using GEOSTUDIO software. First, 206 natural soil slope inputs were taken from the literature of Zhou et al. (2019). That literature lists 206 natural soil slopes worldwide with six input variables c , φ , γ , α , H , and R_u , and output FOS. In the present research, FOS has been evaluated for 206 datasets with their six inputs using the Morgenstern price method to predict slope failure accurately. This method is a classic method for analyzing slope stability. It is a LEM used to check the stability pattern of slopes in soil mechanics. Morgenstern and Price developed the technique in the 1960s. In this method, firstly specify the geometric details of the slope, including α , H , and any variations in soil layering, and define the soil properties for each layer in the slope. This includes c , φ , γ , α , H and R_u . Then, potential failure surfaces are identified within the slope. These surfaces are typically assumed or determined based on geological considerations. The slip surface with the lowest FOS has been considered as the critical slip surface. The RI (β_x) for 206 data is computed to analyze the model's reliability.

Practically, RI is computed by the formula

$$\beta_x = \frac{E(F) - 1}{\sigma(F)} \quad (1)$$

where, $\beta_x = \text{RI}$, $E(F) = \text{Expected value of safety factor}$, and $\sigma(F) = \text{Standard deviation}$.

Finally, the data set has been prepared to have six input data: H , α , ϕ , γ , c , and R_u with one output β_x .

3.2 Data preprocessing

Data preprocessing is critical in preparing data for machine learning models, ensuring that the data is clean, relevant, and suitable for training and testing. We must ensure the model is trained on a subset of the data and tested on another independent subset. Before training the model, review the statistical distribution of the target variable in the training and testing phases to ensure that they are representative of each other. Statistical distributions of different variables in the data set are tabulated in Tables 1 and 2.

Before training, we should scale the input and output data between 0 and 1 using the following normalization equation.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (2)$$

X represents the measured value, X_{norm} is the normalized value of the measured parameter, and X_{min} and X_{max} denote the calculated parameters' minimum and maximum values. The available dataset was split into two sections to optimize the network: training and testing. The training data set was used to adjust the network parameters, and the testing data was used to measure the model's performance. In the present research, 30 percent of the dataset was used for testing, and 70 percent was used for training.

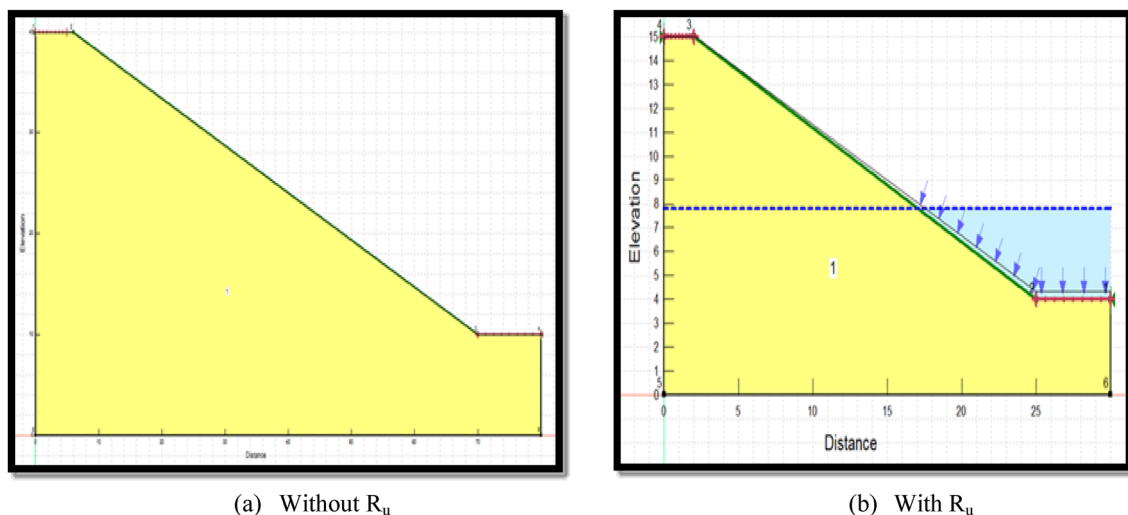


Fig. 1 Soil slope model with and without R_u

Table 1 Statistical distribution of different variables in the testing phase

Statistical parameter	H	α	Γ	c	ϕ	R_u	Actual reliability (Υ)
Mean	0.1624	0.5900	0.5184	0.0924	0.5897	0.1792	0.4575
Standard deviation	0.2028	0.2363	0.2113	0.1420	0.1724	0.2266	0.2930
Median	0.0826	0.5814	0.5181	0.0667	0.6095	0.2266	0.2930
Skewness	1.9550	-0.3807	-0.3472	4.7827	-1.3201	0.6952	0.4420
Lower bound	0.0861	0.3488	0.7565	0.1413	0.7474	0.0000	0.3778
Upperbound	0.5814	0.9070	0.6736	0.1600	0.8081	0.0000	0.2051

Table 2 Statistical distribution of different variables in the training phase

Statistical Parameter	H	α	γ	c	ϕ	R_u	Actual reliability (Υ)
Mean	0.1611	0.5198	0.5274	0.0880	0.5876	0.1988	0.3493
Standard deviation	0.2067	0.2258	0.2075	0.0814	0.1854	0.2311	0.2475
Median	0.0825	0.4651	0.5091	0.0800	0.6076	0.0000	0.2983
Skewness	2.0440	0.1977	-0.2491	2.4033	-0.7491	0.6513	1.1474
Lower bound	0.6330	0.7441	0.7772	0.1166	0.7070	0.4000	0.3444
Upper bound	0.4513	0.4651	0.5181	0.0700	0.4646	0.0000	0.5127

4 Proposed approaches

4.1 Adaptive neuro fuzzy inference system (ANFIS)

ANFIS is a SC system that conjugates the adaptability of neural networks with the interpretability of FL. Jang introduced it in the early 1990s. This hybrid system aims to leverage the strengths of fuzzy systems and neural networks for modeling and control purposes. There are two critical components in this model. The first is a fuzzy inference system (FIS), and the other is a neural network. ANFIS starts with a FIS, which is based on FL principles. FL allows for representing and processing imprecise or uncertain information using linguistic variables and fuzzy rules. This integrates a neural network structure to learn and optimize the parameters of the FIS adaptively. The structure typically consists of five layers, with nodes representing specific functions:

Layer 1 (Input Layer): Nodes represent input variables. Input variables in the present research are H, α , Υ , c, ϕ , and R_u (see Fig. 2)

Layer 2 (Fuzzy Layer): Nodes compute the membership grades for each linguistic term based on the input data.

Layer 3 (Rule Layer): Nodes compute the firing strengths of the fuzzy rule nodes.

Layer 4 (Normalization Layer): Nodes normalize the firing strengths.

Layer 5 (Output Layer): Nodes compute the overall output of soil slope RI (see Fig. 2).

Here, H, α , Υ , c, ϕ , and R_u symbolize the input nodes. $A_1, A_2, B_1, B_2, C_1, C_2, D_1, D_2, E_1, E_2, F_1, F_2$ indicate the linguistic variables, and $\mu_{A_1}, \mu_{A_2}, \mu_{B_1}, \mu_{B_2}, \mu_{C_1}, \mu_{C_2}, \mu_{D_1}, \mu_{D_2}, \mu_{E_1}, \mu_{E_2}, \mu_{F_1}, \mu_{F_2}$ represent the membership function of the proposed node.

4.2 Genetic algorithm (GA)

GA is a robust optimization meta-heuristic search algorithm inspired by the principles of genetics. It is a part of the broader class of evolutionary algorithms. Genetic Algorithms are handy for solving complex optimization problems compared to traditional methods. Researchers have used genetic algorithms as heuristic algorithms to address learning processes and optimization issues in various domains. This algorithm considers the structure as a population, and each population member is referred to as a chromosome, which is a potential solution to the problem. In this algorithm, distinct problem variables function as an individual’s genes, as observed by Mitchell (1998). In this research process, three operators are used to generate the population: selection operator, crossover

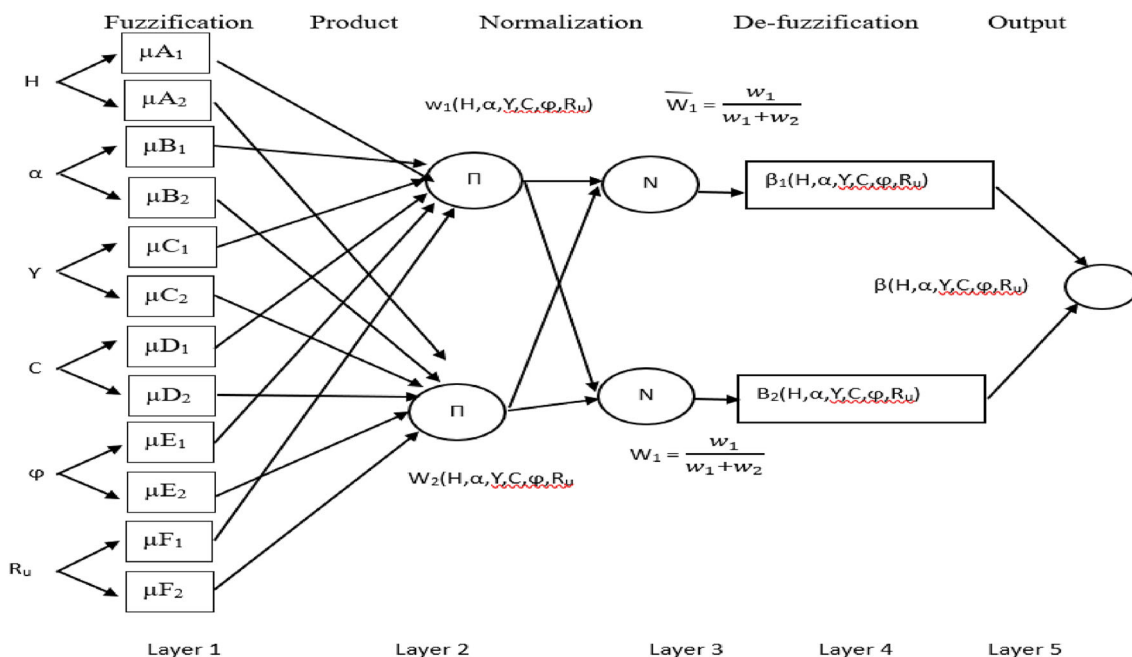


Fig. 2 Different layers used to analyze the ANFIS model

operator, and mutation operator. The selection operator is used to compute the fitness function of each chromosome, and the results are assessed at each level. The best solutions to the problem are the best chromosomes for the present research. The Crossover operator combines the selected chromosomes for better fitness. The crossover operator reported by Saeidian et al. (2016) can be implemented using various techniques, including Uniform, Cycle, N-Point, Tournament, Order, Ranking Selection, and partially mapped. Nourani et al. (2014) applied the mutation operator in a function node (arithmetic operations) or a terminal node (variables and constants). This operator randomly alters some chromosomes’ genetic information, which introduces diversity in the population. These three operators are utilized to create new child chromosomes, regarded as the following generation’s parents. Furthermore, the pattern continues until the termination requirement is satisfied.

4.3 Particle swarm optimization (PSO)

PSO is an optimization technique that draws inspiration from nature and is frequently used to find approximations for optimization and search issues. It was developed in 1995 by James Kennedy and Russell Eberhart and was influenced by the social interactions of fish and birds. This algorithm uses swarm intelligence, where a population of possible solutions, particles, moves through the solution space to explore the ideal solution. Every particle within the swarm symbolizes a possible resolution to the optimization problem. A possible solution is represented by a particle’s location in the

solution space. Each particle has a velocity that determines its movement in the solution space. The velocity is adjusted based on the particle’s previous velocity, its personal best solution, and the best solution found by the entire swarm. Each particle keeps track of its personal best solution and the global best solution obtained by any particle in the swarm. These guide the particle’s movement toward better solutions. A system of mathematical equations is used to update the position and velocity of every particle. The algorithm continues for several iterations or until a termination criterion is met, such as finding a satisfactory solution. Wide-range optimization such problems as feature selection, functional optimization, parameter tuning, and neural network training are all being effectively addressed by PSO. It is relatively simple to implement and has fewer parameters to tune compared to some other optimization algorithms. It is effective in solving optimization problems with high-dimensional and complex search spaces.

4.4 Hybrid methodology with flowchart

4.4.1 Optimization of GA with ANFIS model

ANFIS hybrid modeling starts from dataset preparation with its normalization, as presented in Fig. 3. After normalization, the dataset is split into training and testing. 70% of the data has been used for the training phase, and 30% of the data has been used for the testing phase. After the dataset preprocessing, the ANFIS model’s structure, including the

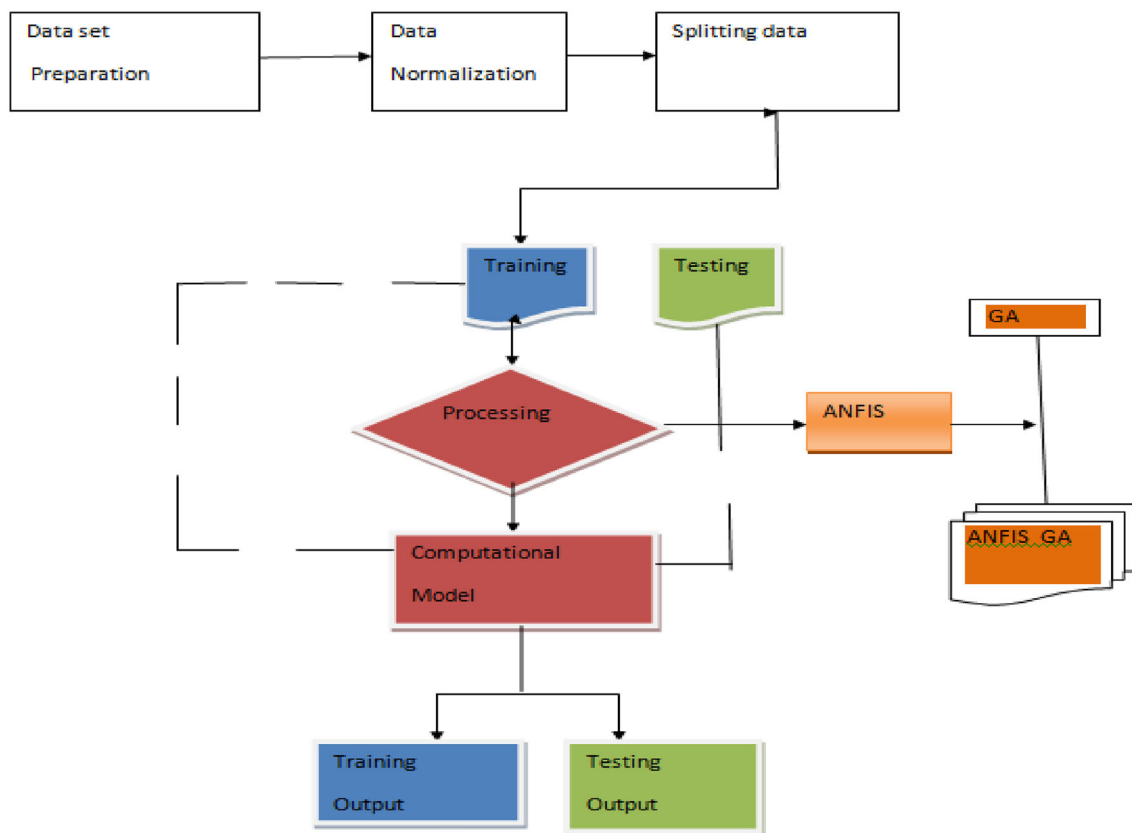


Fig. 3 Flow chart of ANFIS-GA hybrid model

number of fuzzy rules and the types of membership functions, has been defined, and training starts with its training dataset. Implement the Genetic Algorithm in MATLAB. In the next step, with the help of GA algorithm built-in functions provided by MATLAB's Global Optimization Toolbox, all genetic operators, such as selection, crossover, and mutation, have been defined according to the problem requirements. The fitness function has also been described to evaluate the model's performance. The fitness function takes ANFIS model parameters as input, evaluates the model's performance on the validation dataset, and returns a fitness score to be maximized by the genetic algorithm. The following process is to hybridize the ANFIS model with the GA model and integrate the ANFIS model with the genetic algorithm optimization process. Membership functions and rules weight parameters of the ANFIS model have been optimized with GA, which defines the ANFIS parameters' chromosome representation and implements mechanisms for encoding, decoding, and updating the chromosomes during the genetic algorithm optimization process. Integrate the ANFIS model with the genetic algorithm optimization process. Finally, the genetic algorithm will be run to evolve the population of ANFIS parameters over multiple generations, apply genetic operators (selection, crossover, mutation) to generate new

offspring chromosomes and evaluate the performance of the hybrid ANFIS-GA model using performance indices.

4.4.2 Optimization of PSO with ANFIS model with flowchart

In ANFIS PSO hybrid modeling, the dataset is first prepared by creating and loading the data into MATLAB, presented in Fig. 4. Then, the data is normalized and grouped by SSMD. The dataset is split into training (70%) and testing (30%) sets. Further, the ANFIS model was created with MATLAB's Fuzzy Logic Toolbox, and the fuzzy inference system (FIS) structure was defined, including the number of fuzzy rules and types of membership functions. After that, the ANFIS model is being trained using the training dataset. The PSO algorithm has been implemented using MATLAB optimization tools. PSO has optimized the fitness function, combining the ANFIS model with the PSO optimization algorithm. In the next step, the parameters of the ANFIS model have been updated using the PSO algorithm until convergence. The hybrid ANFIS-PSO model is tested using the testing dataset to assess its generalization ability. Depending on the performance of the hybrid model, further optimization and refinement are being done to ANFIS-PSO parameters or adjusting the model architecture to improve performance.

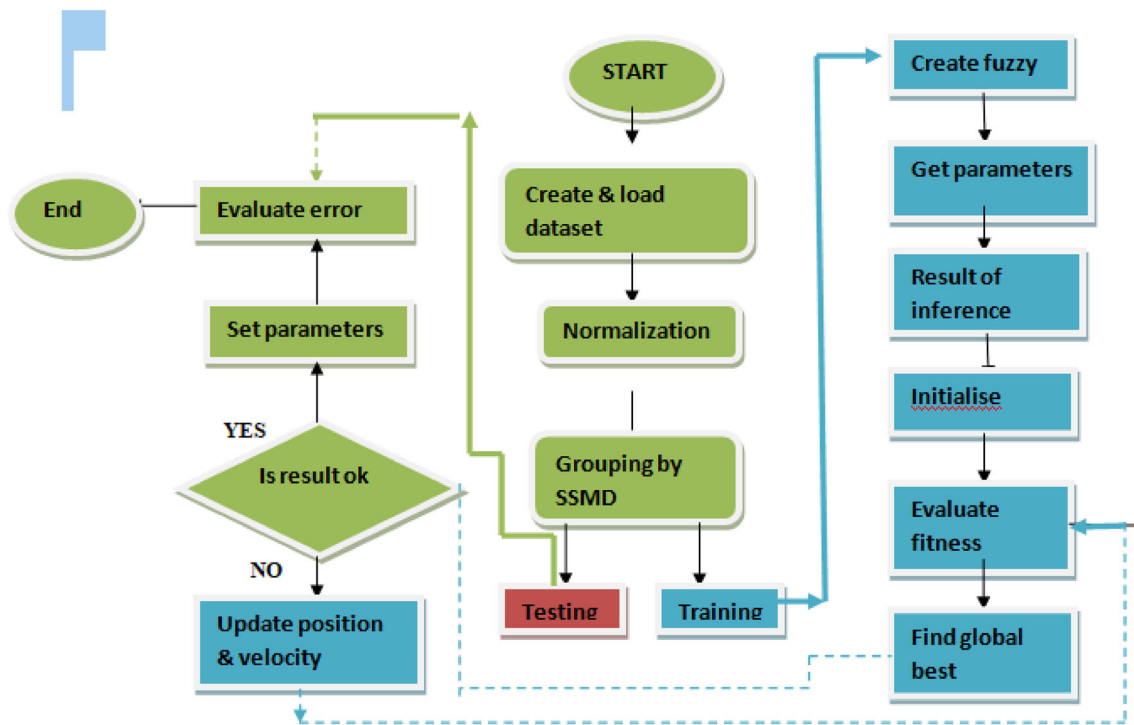


Fig. 4 Flow chart of ANFIS-PSO hybrid model

Finally, the fitness is evaluated, and the global best is obtained after updating position and velocity. If the result is ok, then the parameter has been set, and the error has been computed with the help of the statistics performance index.

5 Statistical performance parameters

5.1 Regression analysis

A statistical evaluation technique, regression analysis, examines the relationship between actual and model-predicted values. Its primary goal is to explore the model and make predictions or inter-causal relationships. These relationships were drawn during the construction and validation of the model phase. The most common type of regression is linear regression, where the relationship between actual and predicted value is modeled as a linear equation. In the present research, regression analysis is made using a scattered plot prepared with the help of PYTHON programming.

5.2 Statistical measures

Several commonly used statistical parameters determine the model's performance. Different parameters have been used to evaluate the present hybrid model's accuracy. These parametric equations are as follows:

$$R^2 = \frac{\sum_{i=1}^N (d_i - d_{mean})^2 - \sum_{i=1}^N (d_i - y_i)^2}{\sum_{i=1}^N (d_i - d_{mean})^2} \quad (3)$$

$$\begin{aligned} \text{VAF (Variance account factor)} \\ = \left(1 - \frac{\text{var}(d_i - y_i)}{\text{var}(d_i)} \times 100 \right) \end{aligned} \quad (4)$$

$$\text{Adj. } R^2 = 1 - \frac{(n-1)}{(n-p-1)} (1 - R^2) \quad (5)$$

$$\text{RMSE (Root mean square error)} = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - y_i)^2} \quad (6)$$

$$\text{PI (Persistence index)} = \text{Adj. } R^2 + (0.01 \times \text{VAF}) - \text{RMSE} \quad (7)$$

$$\begin{aligned} \text{WI (Willmott index)} \\ = 1 - \left[\frac{\sum_{i=1}^N (d_i - y_i)^2}{\sum_{i=1}^N \{|y_i - d_{mean}| + |d_i - d_{mean}|\}^2} \right] \end{aligned} \quad (8)$$

$$\text{MAE (Mean Absolute Error)} = \frac{1}{N} \sum_{i=1}^N |y_i - d_i| \quad (9)$$

$$\begin{aligned} \text{WMAPE (Weighted Mean Absolute Percentage Error)} \\ = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - d_i}{d_i} \right| \times 100 \end{aligned} \quad (10)$$

$$\text{NS (Nash–Sutcliffe Efficiency)} = 1 - \frac{\sum (d_i - y_i)^2}{\sum (d_i - d_{\text{mean}})^2} \quad (11)$$

MAPE (Mean Absolute Percentage Error)

$$= \frac{\sum \left| \frac{d_i - y_i}{d_i} \right| \times 100}{N} \quad (12)$$

$$\text{MSE (Mean Square Error)} = \frac{\sum (d_i - y_i)^2}{N} \quad (13)$$

where d_{mean} is the mean value of the observed data, y_i is the estimated i th value, N is the number of the data in the sample, R^2 is the coefficient of determination, RMSE is the Root-mean-square error, and SD is the standard deviation.

5.3 Rank analysis

According to Zhang et al. (2020), rank evaluation is the easiest method of analysis to determine the model's effectiveness. Rank prediction in the context of a machine learning model typically involves evaluating and comparing the performance of different models or variations of a model to determine their effectiveness in solving a particular problem. It is the most extensively adapted technique for assessing and contrasting the efficiency of developed frameworks. This study employs statistical factors to calculate their score, with their perfect measures serving as a standard. It is based on how many frameworks are being used. The highest value is awarded to the outcomes framework that is best suited, and vice versa. The total value of the score is entirely dependent on the quality of models and their standard value for different parameters applied in both the testing and training stages. One parameter's score may be ranked higher than the others evaluated for model efficiency. When two parameters are scored, they will be ranked equal. As a result, the top-ranked model receives a top score value, while the worst model receives a low score value. A specific predictive model's total score is computed as the sum of total score measures for testing and training data sets.

5.4 Sampling distribution with Metropolis–Hastings

Metropolis–Hastings (MH) is one of the strong Markov Chain Monte Carlo algorithms (MCMC) commonly used for sampling from a probability distribution. Markov chain is constructed for the desired distribution, called equilibrium distribution. It is a powerful statistical technique used to sample complex probability distributions. It is advantageous when direct sampling from the distribution is impractical or impossible. MCMC methods generate a Markov chain that converges to the target distribution, allowing the sampling of

representative values. MCMC methods aim to create samples from a distribution that approximates the target distribution. Convergence is assessed by checking whether the Markov chain has reached a stationary distribution, meaning that samples no longer change significantly. Once convergence is achieved, the samples obtained from the Markov chain are used as an approximation of the target distribution.

5.5 Sensitivity analysis

Sensitivity analysis is a technique used in various research fields such as engineering, economics, finance, and environmental science to evaluate how variations to input variables affect a system's or model's output. It helps understand the robustness and reliability of models and informs decision-making by identifying which parameters significantly influence the results. The primary goal of sensitivity analysis is to quantify and understand how variations or uncertainties in input parameters affect the output of a model or system. Sensitivity Analysis examines the impact of changing one input variable at a time while keeping others constant. It focuses on the immediate vicinity of a particular point in the input space to understand the system's behavior in that region. It quantifies the contribution of each input parameter to the variability in the model output. Sensitivity analysis is applied in various fields, including finance (e.g., assessing the impact of interest rate changes), environmental modeling (e.g., understanding the influence of climate variables), and engineering (e.g., evaluating the sensitivity of a structural model to material properties). Sensitivity analysis is closely related to uncertainty and risk management. By identifying influential factors, decision-makers can focus on mitigating or managing the most critical uncertainties. Sensitivity analysis assumes that input parameters are independent, which may not be accurate in all situations. It also does not account for potential interactions between parameters. It is a valuable tool for decision-makers to gain insights into the behavior of models, systems, or processes, especially when uncertainty and variability are present. It aids in making informed decisions and improving the robustness of models in the face of changing conditions.

SHAP values are one of the powerful tools that represent sensitiveness. SHAP value stands for (Shapley Additive explanations) in which input parameter performance is assessed through the model's output value. SHAP values quantify the contribution of each input parameter by observing the difference between the actual model prediction and the average prediction. The sum of SHAP values for all inputs plus the average model prediction is approximately equal to the exact model prediction for the given slope stability problem. Hence, SHAP values provide a clear interpretation of input variables. Positive SHAP values show a positive contribution to output prediction, while negative SHAP values

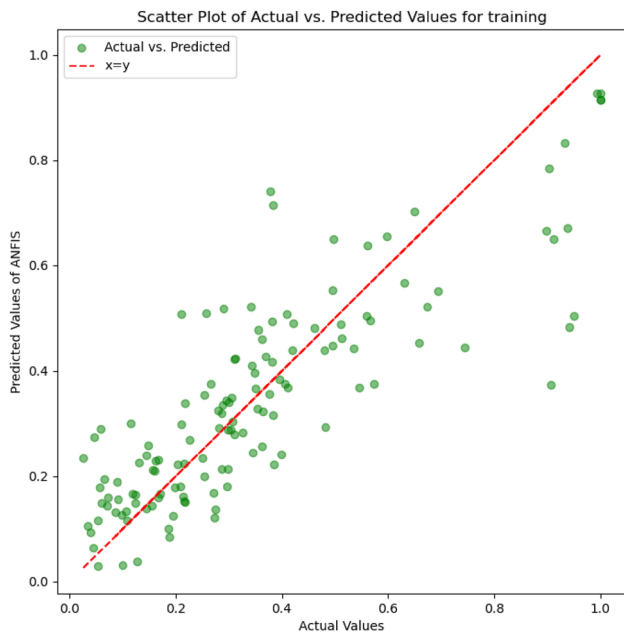


Fig. 5 Regression plot of ANFIS model in training

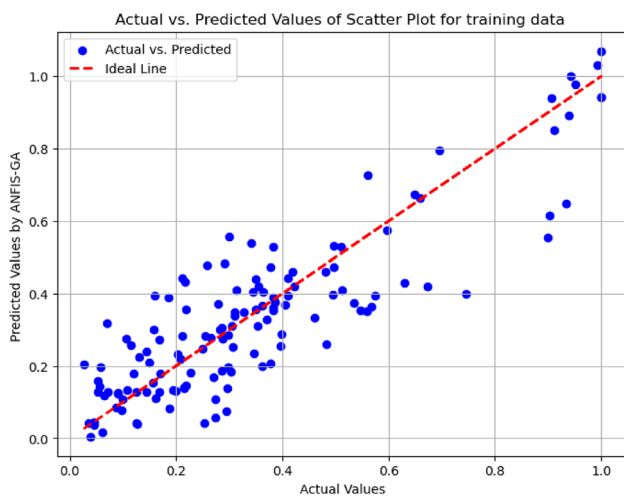


Fig. 6 Regression plot of ANFIS-GA model in training

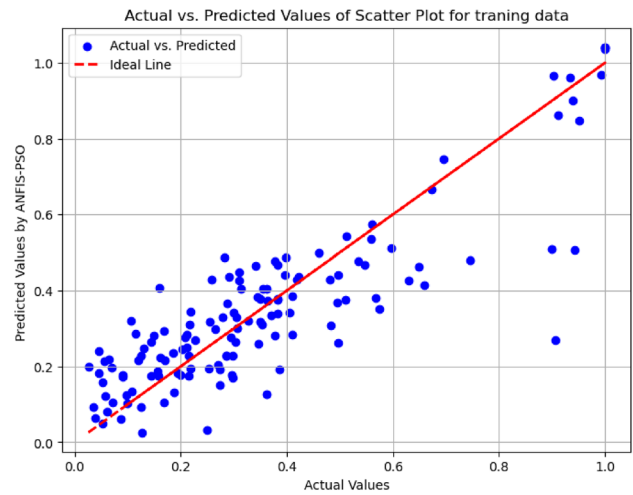


Fig. 7 Regression plot of ANFIS-PSO hybrid model in training

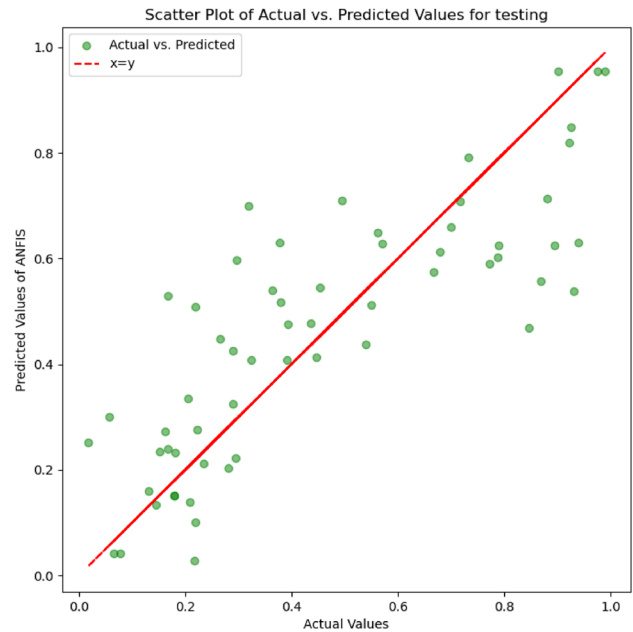


Fig. 8 Regression plot of ANFIS hybrid model in testing

indicate a decreasing contribution to output prediction. In the present research, the SHAP value is represented by a summary plot.

6 Results and discussions

6.1 Regression analysis

In machine learning, a regression plot of training and testing data visually represents how well a model performs in the

training and testing set. This plot helps to evaluate the ability of the meta-heuristic hybrid model. For the regression plot, predicted and actual target values are collected for the training and testing phases. Figures 5, 6, 7 show the regression plot of both the hybrid models, such as ANFIS-GA and ANFIS-PSO along with ANFIS in training and Figs. 8, 9, 10 testing plot. It has been seen that the ANFIS-GA model's predicted value is closer to the actual than the ANFIS-PSO model, and the ANFIS-PSO model is closer to the actual one than the ANFIS model (Table 3).

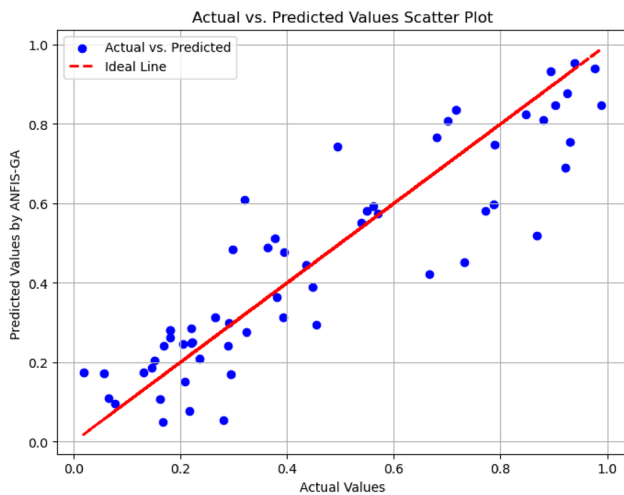


Fig. 9 Regression plot of ANFIS-GA hybrid model in testing

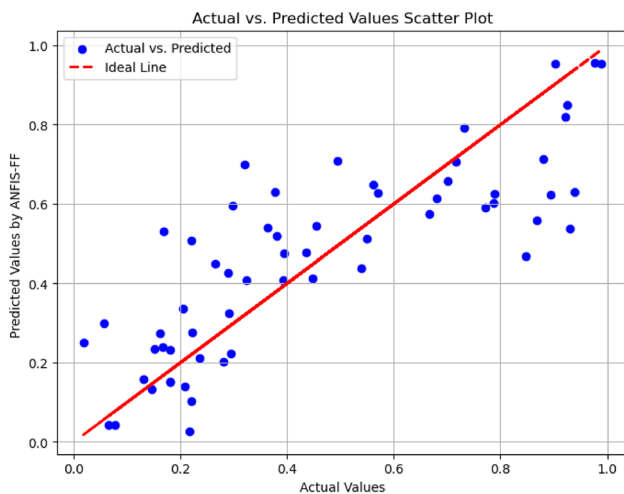


Fig. 10 Regression plot of ANFIS-PSO hybrid model in testing

6.2 Statistical measurement

Some error matrices and some performance matrices evaluate statistical efficiency. It has been seen that the ANFIS-GA hybrid model is a statistically more efficient model.

6.3 Rank prediction

In this section, the ranking of hybrid models is predicted based on the total score obtained through statistical indices ranking in the testing and training phases. Total scores are evaluated by adding scores obtained in testing and training for three models: ANFIS, ANFIS-GA, and ANFIS-PSO. The estimated scores are 24, 64, and 44, respectively. Thus, the highest scores are obtained in the ANFIS-GA model, followed by ANFIS-PSO and ANFIS. ANFIS-GA hybrid model has been proven more effective than ANFIS-PSO, and ANFIS-PSO has been proven more effective than ANFIS (Tables 4, 5).

6.4 Sampling distribution graph

Metropolis Hasting’s empirical distribution shows the sample distribution graph (see Figs. 11, 12, 13, 14, 15, 16).

6.5 Sensitive analysis result

Figure 17 represents a sensitivity bar graph in which the bar length represents the feature’s importance. Figure 18 illustrates the SHAP value of the dependent graph that indicates the influencing behavior of input variables on output. The longer the bar, the more influential the feature is in making predictions. In Fig. 17, the longest bar is of pore water pressure ratio, which means the most influencing parameter

Table 3 Statistical value of different parameters obtained in hybrid models along with ANFIS

Statistical parameter	Training statistical value			Testing statistical value			Ideal value
	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS	ANFIS-GA	ANFIS-PSO	
R ²	00.6783	00.7624	00.7378	00.6684	00.8143	00.7013	1
VAF	68.7643	76.2211	74.0646	64.4476	80.8319	69.7016	100
Adj.R ²	00.7575	00.7575	00.7575	00.7422	00.7422	00.7422	1
PI	01.3328	01.3980	01.3717	01.3046	01.4248	01.2802	2
RMSE	00.1345	00.1217	00.1264	00.1642	00.12570	00.1591	0
WI	00.8986	00.9319	00.9200	00.8764	00.9466	00.9073	1
MAE	00.1127	00.0912	00.0902	00.1267	00.0968	00.1170	0
WMAPE	00.2486	00.2612	00.2583	00.2871	00.2117	00.2556	0
NS	00.6854	00.7565	00.7373	00.6728	00.8128	00.7002	1
MAPE	48.6206	40.0227	45.4622	59.3262	44.3995	55.3007	0
MSE	00.0172	00.0148	00.0160	00.1391	00.1264	00.1290	0

Table 4 Ranking evaluation based on the different parameters of hybrid models with ANFIS

Statistical parameter	Training phase			Testing phase		
	ANFIS	ANFIS-GA	ANFIS-PSO	ANFIS	ANFIS-GA	ANFIS-PSO
R ²	1	3	2	1	3	2
VAE	1	3	2	1	3	2
Adj.R ²	3	3	3	1	1	1
PI	1	3	2	1	3	2
RMSE	1	3	2	1	3	2
WI	1	3	2	1	3	2
MAE	1	3	2	1	3	2
WMAPE	1	3	2	1	3	2
NS	1	3	2	1	3	2
MAPE	1	3	2	1	3	2
MSE	1	3	2	1	3	2
Total score	13	33	23	11	31	21

Table 5 Score and ranking of hybrid models

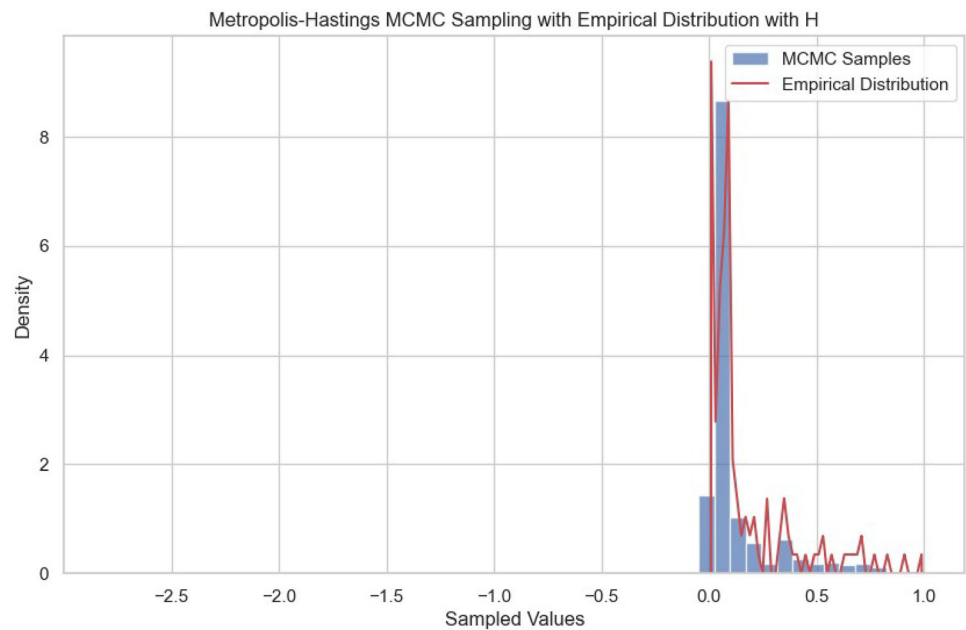
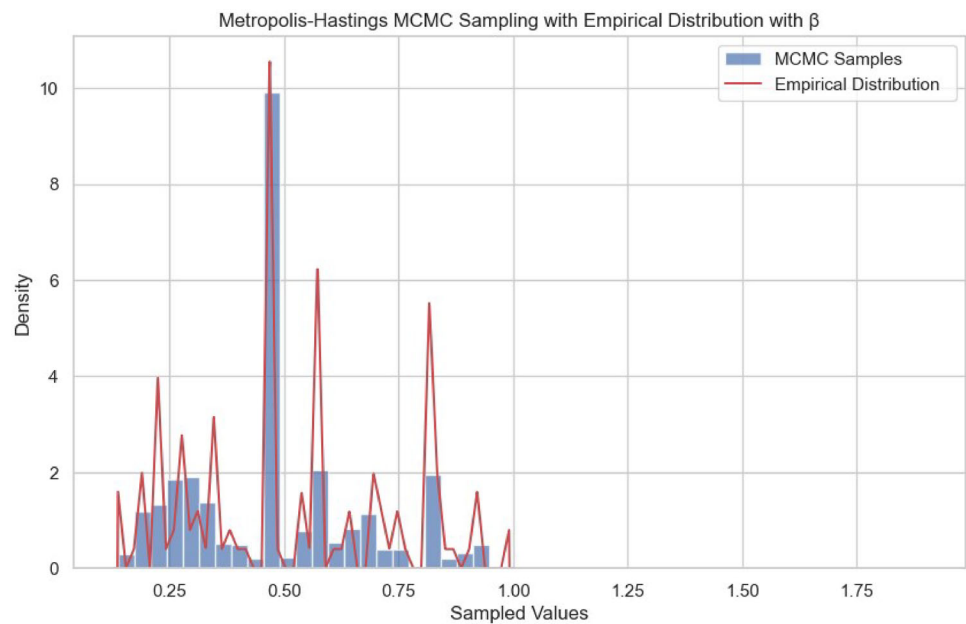
Phase	ANFIS	ANFIS-GA	ANFIS-PSO
Training	13	33	23
Testing	11	31	21
Total score	24	64	44
Ranking	3	1	2

is pore water pressure ratio followed by friction angle, unit weight of soil, angle of soil slope, cohesion, and then last one is the height of slope that means the reliability of soil slope is much affected by pore water pressure. In Fig. 18, the circle's color indicates the direction of the feature's importance on the model's predictions. Here, red indicates a positive impact (increases the prediction), while blue indicates a negative effect (decreases the prediction). The red circle is marked with a positive impact in the pore water pressure variable, representing a high output impact.

7 Conclusions

This paper uses two strong hybrid models, ANFIS-GA and ANFIS-PSO, along with the ANFIS model in the assessment of failure of soil slopes. This paper aims to predict the best model for calculating slope failure in civil engineering. The comparison results showed excellent agreement between the predicted and measured data. For precise and robust results, 11 statistical parameters, including error and performance matrices, have been computed for all three models. The final conclusions of the research are as follows :

- ANFIS, ANFIS-GA, and ANFIS-PSO outcomes from the suggested models have R² values of 0.6783, 0.7624, 0.7378 during training, 0.6684, 0.8143, and 0.7013 during testing. The R² value is closest to the ideal value in the ANFIS-GA model, and it is the strongest model for solving significant data problems related to slope stability, followed by the ANFIS PSO and ANFIS models.
- The ANFIS-GA hybrid model best matches slope stability problems, outperforming the ANFIS-PSO and ANFIS models because its predictive value is more accurate during training and testing.
- The ANFIS-GA hybrid model yielded error matrices such as RMSE, MAE, and MSE with values of 0.1217, 0.0912, and 0.0148 in training and 0.12570, 0.0968 and 0.1391 in testing in contrast, and the ANFIS PSO model yielded values of 0.1264, 0.0902, 0.016 in training, and 0.1591, 0.1170, 0.1290 in testing; the ANFIS model yielded values of 0.1345, 0.1127, 0.0172 in training, and 0.1642, 0.1267, 0.1391 in testing; these results indicate that the ANFIS GA model is more valuable and efficient for assessing slope reliability in today's practice than ANFIS-PSO followed by ANFIS model.
- According to rank analysis, high scores indicate a significant consensus between the actual and expected classes. The combined score for the ANFIS GA hybrid model was 64, the ANFIS PSO was 44, and the ANFIS was 24. The

Fig. 11 Distribution of height of slope**Fig. 12** Distribution of angle of soil slope

ANFIS-GA hybrid model has been ranked highest in the current study, followed by ANFIS-PSO in second place and ANFIS in last place that means ANFIS-GA model is more compatible for solving slope stability problem

- Metropolis–Hastings sampling distribution classifies the soil slopes based on their input variables and reveals the empirical distribution of every input variable.
- Sensitivity analysis shows the impact behavior of different input variables on the reliability status of soil slope. The most influencing parameters are R_u and friction angle, followed by unit wt. of soil slope, slope angle, and slope

height as shown in Figs. 17 and 18 that means the key factor impacting and playing a major part in the occurrence of slope failure is pore water pressure.

- In civil engineering, hybridization of models using meta-heuristic algorithms is an emerging artificial intelligence system. It may improve the model's efficiency for assessing the failure status of any structure.
- The ANFIS-GA hybrid algorithm is the most capable algorithm for predicting the reliability of soil slopes in large data sets.

Fig. 13 Distribution of cohesion of soil slope

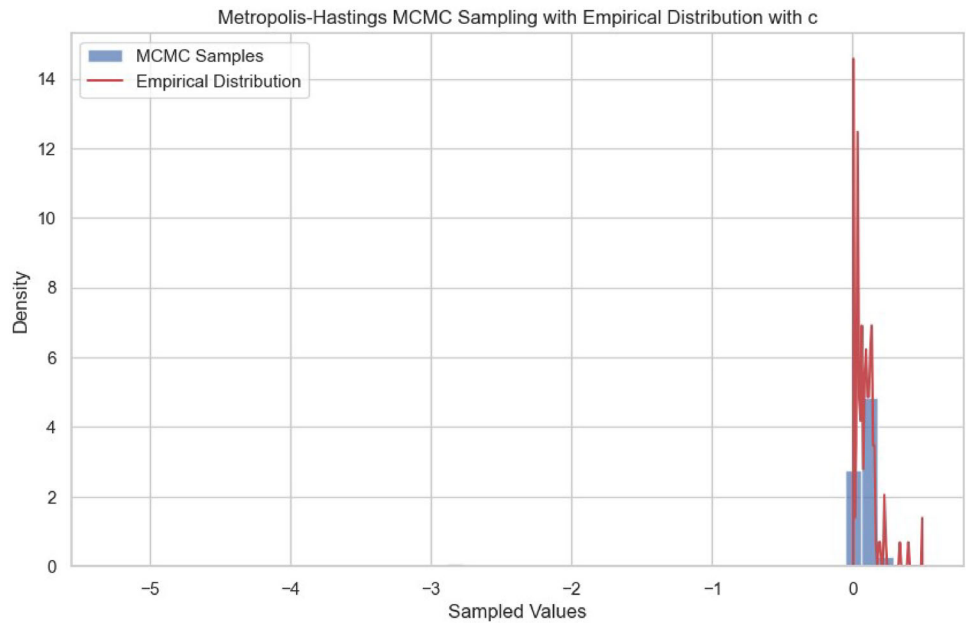
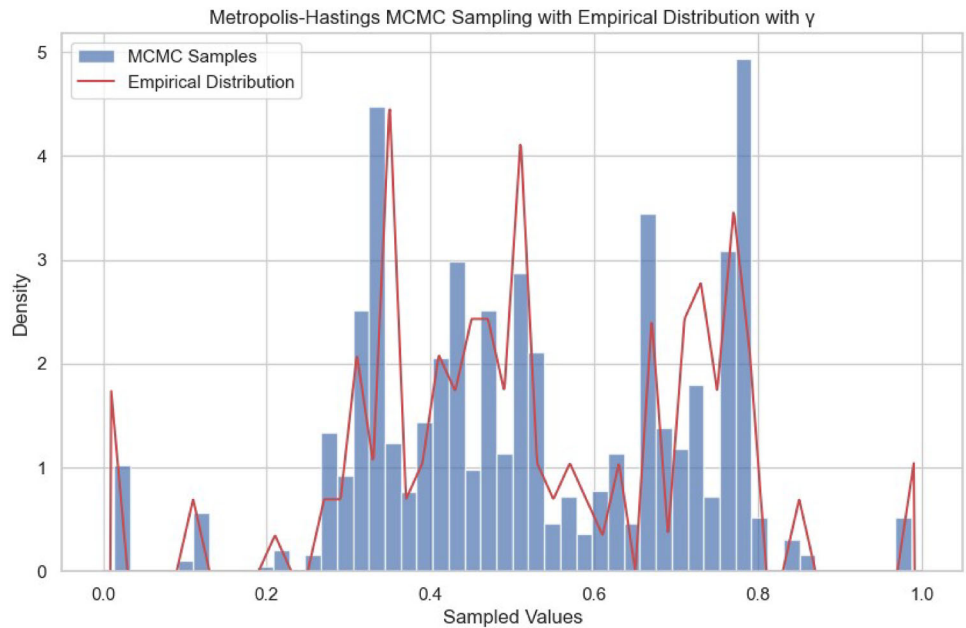


Fig. 14 Distribution of unit wt of soil slope



- Last but not least, hybrid algorithms like ANFIS- GA and ANFIS-PSO stand in a good position in Artificial intelligence for predicting soil slope failure.

8 Research significance

Hybrid optimization of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) presents a significant advancement in AI methods for civil engineering. Civil engineering problems involve complex, nonlinear relationships between various

parameters. Hybrid models such as ANFIS GA and ANFIS PSO can model nonlinear systems and effectively handle this complexity related to civil engineering problems. By integrating hybrid optimization techniques like genetic algorithms particle swarm optimization, ANFIS can refine its parameters to fit the data better, resulting in more accurate predictions and solutions. Hybrid optimization further enhances the data-driven approach of the ANFIS model to extract the most relevant information from the data, leading to improved performance and reliability where large amounts of data are often available from sensors, surveys, and simulations related to civil engineering problems. Civil engineering

Fig. 15 Distribution of friction angle of soil slope

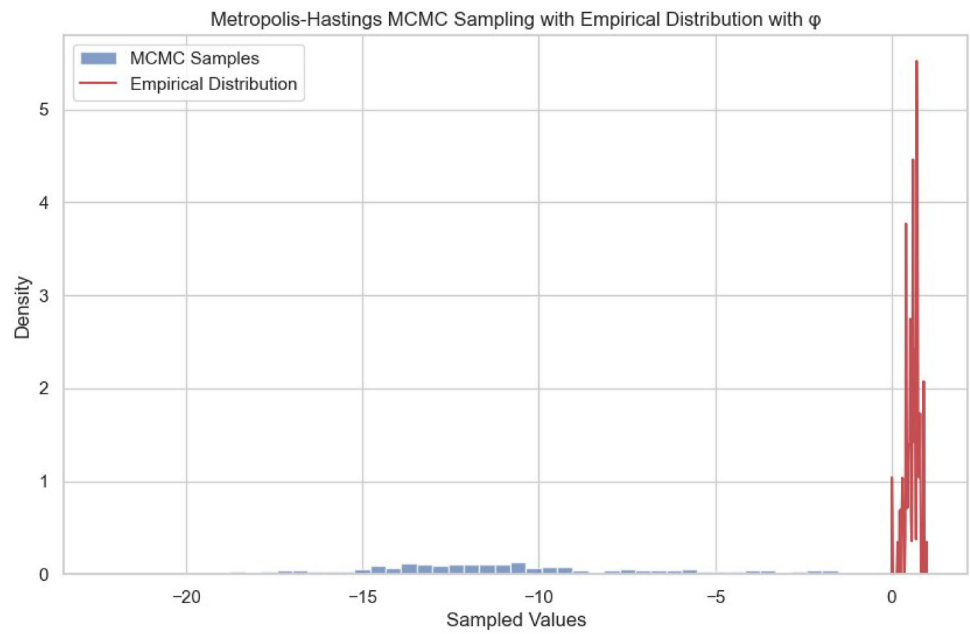
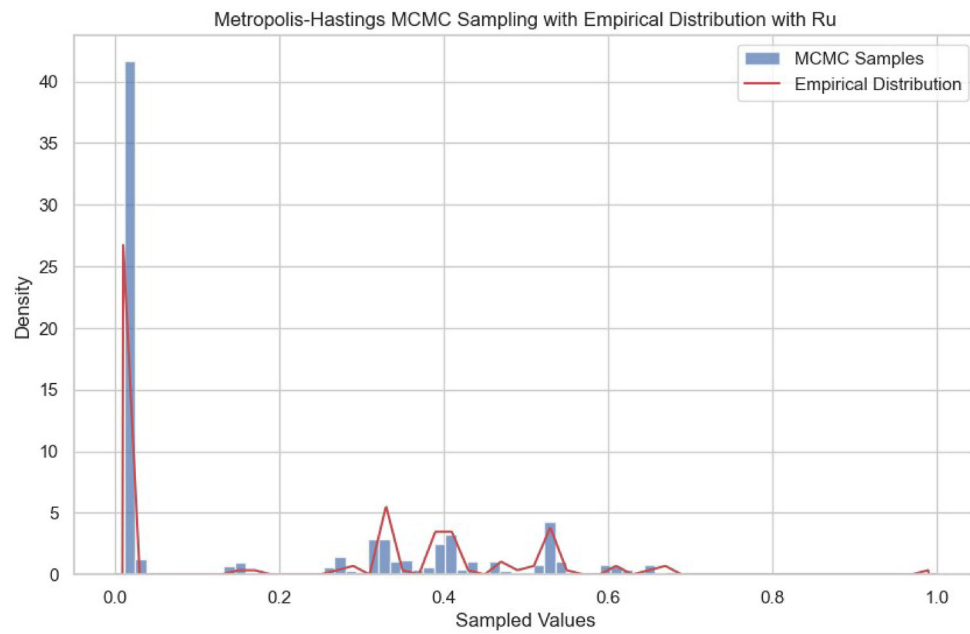


Fig. 16 Distribution of pore water pressure ratio of soil slope



projects usually face uncertainties due to weather variations, material properties, and human error. ANFIS models enhanced with hybrid optimization techniques can adapt to these uncertainties by continuously updating their parameters based on real-time data feedback. This adaptability improves

the robustness of the models and enhances their performance under changing conditions. Therefore, hybrid optimization with the metaheuristic model may be highly recommended in civil engineering.

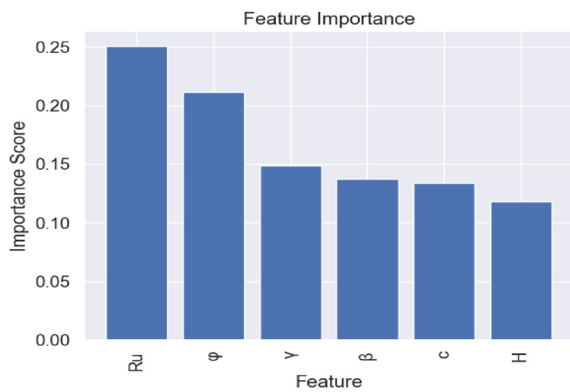


Fig. 17 Feature importance of variables used in this work



Fig. 18 SHAP value of variables used in this work

Author contributions Author Contribution- Jayanti Prabha Bharti; Conceptualization, data collection, formal analysis, investigation, methodology, software validation, visualization and writing original draft. Pijush Samui; Review and supervision “Both the authors read and approved the final manuscript.”

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Agnihotri A, Sahoo, Diwakar M.K (2022) Flood prediction using hybrid ANFIS-ACO model: a case study. In: Inventive computation and information technologies: proceedings of ICICIT 2021, pp 169–180. https://doi.org/10.1007/978-981-16-6723-7_13
- Bharti JP, Mishra P, Moorthy U, Sathishkumar VE, Cho Y, Samui P (2021) Slope stability analysis using Rf, gbm, cart, bt, and xgboost. *Geotech Geol Eng* 39:3741–3752. <https://doi.org/10.1007/s10706-021-01721-2>
- Bousnina K, Hamza A, Yahia NB (2023) An integration of PSO-ANN and ANFIS hybrid models to predict surface quality, cost, and

- energy (QCE) during milling of alloy 2017A. *J Eng Res*. <https://doi.org/10.1016/j.jer.2023.09.016>
- Cheng YM, Lansivaara T, Wei WB (2007) Two-dimensional slope stability analysis by limit equilibrium and strength reduction methods. *Comput Geotech* 34(3):137–150. <https://doi.org/10.1016/j.compgeo.2006.10.011>
- Cheng YM, Li L, Chi S, Wei WB (2007) Particle swarm optimization algorithm for the location of the critical non-circular failure surface in two dimensional slope stability analysis. *Comput Geotech*. <https://doi.org/10.1016/j.compgeo.2006.10.012>
- Das SK, Biswal RK, Sivakugan N, Das B (2011) Classification of slopes and prediction of factor of safety using differential evolution neural networks. *Environ Earth Sci* 64:201–210. <https://doi.org/10.1007/s12665-010-0839-1>
- Ercanoglu M, Gokceoglu C (2002) Assessment of landslide susceptibility for a landslide-prone area north of Yenice, NW Turkey by fuzzy approach. *Environ Geol* 41:720–730. <https://doi.org/10.1007/s00254-001-0454-2>
- Fattahi H (2017) Prediction of slope stability using adaptive neuro-fuzzy inference system based on clustering methods. *J Min Environ* 8(2):163–177. <https://doi.org/10.22044/jme.2016.637>
- Griffiths DV, Fenton GA (2004) Probabilistic slope stability analysis by finite elements. *J Geotech Geoenviron Eng* 130(5):507–518. [https://doi.org/10.1061/\(ASCE\)1090-0241\(2004\)130:5\(507\)](https://doi.org/10.1061/(ASCE)1090-0241(2004)130:5(507))
- Kahatadeniya KS, Nanakorn P, Neaupane KM (2009) Determination of the critical failure surface for slope stability analysis using ant colony optimization. *Eng Geol* 108(1–2):133–141. <https://doi.org/10.1016/j.enggeo.2009.06.010>
- Kecman V (2001) Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. MIT Press, Cambridge
- Li XZ, Xu Q (2016) Application of the SSPC method in the stability assessment of highway rock slopes in the Yunnan province of China. *Bull Eng Geol Environ* 75:551–562. <https://doi.org/10.1007/s10064-015-0792-z>
- Li DQ, Qi XH, Cao ZJ, Tang XS, Phoon KK, Zhou CB (2016) Evaluating slope stability uncertainty using coupled Markov chain. *Comput Geotech* 73:72–82. <https://doi.org/10.1016/j.compgeo.2015.11.021>
- Lim K, Li AJ, Lyamin AV (2015) Three-dimensional slope stability assessment of two-layered undrained clay. *Comput Geotech* 70:1–17. <https://doi.org/10.1016/j.compgeo.2015.07.011>
- Luo N, Bathurst RJ, Javankhosdel S (2016) Probabilistic stability analysis of simple reinforced slopes by finite element method. *Comput Geotech* 77:45–55. <https://doi.org/10.1016/j.compgeo.2016.04.001>
- Massey C, Della Pasqua F, Holden C, Kaiser A, Richards L, Wartman J, McSaveney MJ, Archibald G, Yetton M, Janku L (2017) Rock slope response to strong earthquake shaking. *Landslides* 14:249–268. <https://doi.org/10.1007/s10346-016-0684-8>
- McCombie P, Wilkinson P (2002) The use of the simple genetic algorithm in finding the critical factor of safety in slope stability analysis. *Comput Geotech* 29(8):699–714. [https://doi.org/10.1016/S0266-352X\(02\)00027-7](https://doi.org/10.1016/S0266-352X(02)00027-7)
- Mitchell M (1998) An introduction to genetic algorithms. MIT Press, Cambridge
- Moayed H, Raftari M, Sharifi A, Jusoh WAW, Rashid ASA (2020) Optimization of ANFIS with GA and PSO estimating α ratio in driven piles. *Eng Comput* 36(1):227–238. <https://doi.org/10.1007/s00366-018-00694-w>
- Mu'azu MA (2023) Enhancing slope stability prediction using fuzzy and neural frameworks optimized by metaheuristic science. *Math Geosci* 55(2):263–285. <https://doi.org/10.1007/s11004-022-10029-7>
- Nash DFT (1981) A comparative review of limit equilibrium methods of stability analysis. *Slope stability*, pp 11–76.

- Nematolahi M, Jalali V, Hejazi-Mehrzi M (2018) Predicting saturated hydraulic conductivity using particle swarm optimization and genetic algorithm. *Arab J Geosci*. <https://doi.org/10.1007/s12517-018-3846-2>
- Nourani V, Pradhan B, Ghaffari H, Sharifi SS (2014) Landslide susceptibility mapping at Zonouz Plain, Iran using genetic programming and comparison with frequency ratio, logistic regression, and artificial neural network models. *Nat Hazard* 71(1):523–547. <https://doi.org/10.1007/s11069-013-0932-3>
- Park D, Rilett LR (1999) Forecasting freeway link travel times with a multilayer feed-forward neural network. *Comput Aided Civ Infrastruct Eng* 14(5):357–367. <https://doi.org/10.1111/0885-9507.00154>
- Qin J, Du S, Ye J, Yong R (2022) SVNN-ANFIS approach for stability evaluation of open-pit mine slopes. *Expert Syst Appl* 198:116816. <https://doi.org/10.1016/j.eswa.2022.116816>
- Reale C, Xue J, Gavin K (2016) System reliability of slopes using multimodal optimization. *Géotechnique* 66(5):413–423. <https://doi.org/10.1680/jgeot.15.P.142>
- Riahi Madvar H, Dehghani M, Parmar KS, Nabipour N, Shamshirband S (2020) Improvements in the explicit estimation of pollutant dispersion coefficient in rivers by subset selection of maximum dissimilarity hybridized with ANFIS-firefly algorithm (FFA). *IEEE Access* 8:60314–60337
- Saboya Jr F, GlóriaAlves M, Pinto WD (2006) Assessment of failure susceptibility of soil slopes using fuzzy logic. *Eng Geol* 86(4):211–224. <https://doi.org/10.1016/j.enggeo.2006.05.001>
- Saeidian B, Mesgari MS, Ghodousi M (2016) Evaluation and comparison of genetic algorithm and bees algorithm for location-allocation of earthquake relief centers. *Int J Disaster Risk Reduct* 15:94–107. <https://doi.org/10.1016/j.ijdr.2016.01.002>
- Samantaray S, Sahoo P, Sahoo A, Satapathy DP (2023) Flood discharge prediction using improved ANFIS model combined with hybrid particle swarm optimization and slime mould algorithm. *Environ Sci Pollut Res* 30(35):83845–83872. <https://doi.org/10.1007/s11356-023-27844-y>
- Sun S, Sun H, Wang Y, Wei J, Liu J, Kanungo DP (2014) Effect of the combination characteristics of rock structural plane on the stability of a rock-mass slope. *Bull Eng Geol Environ* 73(4):987–995. <https://doi.org/10.1007/s10064-014-0593-9>
- Tao H, Abba SI, Al-Areeq AM, Tangang F, Samantaray S, Sahoo A, Siqueira HV, Maroufpoor S, Demir V, Bokde ND, Goliatt L (2024) Hybridized artificial intelligence models with nature-inspired algorithms for river flow modeling: a comprehensive review, assessment, and possible future research directions. *Eng Appl Artif Intell* 129:10755. <https://doi.org/10.1016/j.engappai.2023.107559>
- Tschuchnigg F, Schweiger HF, Sloan SW (2015a) Slope stability analysis by means of finite element limit analysis and finite element strength reduction techniques Part II: back analyses of a case history. *Comput Geotech* 70:178–189. <https://doi.org/10.1016/j.compgeo.2015.06.018>
- Tschuchnigg F, Schweiger HF, Sloan SW, Lyamin AV, Raissakis I (2015b) Comparison of finite-element limit analysis and strength reduction techniques. *Géotechnique* 65(4):249–257. <https://doi.org/10.1680/geot.14.P.022>
- Yang H, Hasanipanah M, Tahir MM, Bui DT (2020) Intelligent prediction of blasting-induced ground vibration using ANFIS optimized by GA and PSO. *Nat Resour Res* 29:739–750. <https://doi.org/10.1007/s11053-019-09515-3>
- Zhang X, Yates A, Lin J (2020) A little bit is worse than none: ranking with limited training data. In: *Proceedings of SustainLP Workshop on Simple and Efficient Natural Language Processing*, pp 107–112. <https://doi.org/10.18653/v1/2020.sustainlp-1.14>
- Zhang Z, Al-Bahrani M, Ruhani B, Ghalehsalimi HH, Ilghani NZ, Maleki H, Ahmad N, Nasajpour-Esfahani N, Toghraie D (2023) Optimized ANFIS models based on grid partitioning, subtractive clustering, and fuzzy C-means to precise prediction of thermo-physical properties of hybrid nanofluids. *Chem Eng J* 471:144362. <https://doi.org/10.1016/j.cej.2023.144362>
- Zhou J, Li E, Yang S, Wang M, Shi X, Yao S, Mitri HS (2019) Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. *Saf Sci* 118:505–518. <https://doi.org/10.1016/j.ssci.2019.05.046>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.