China Geology 8 (2025) 691−706

China Geology

Journal homepage: http://chinageology.cgs.cn
https://www.sciencedirect.com/journal/china-geology

Enhancing rock slope stability prediction using random forest machine learning: A case study

Afiqah Ismail[a], Ahmad Safuan A Rashid[a, b,*], Ali Dehghanbanadaki[c, d], Rafiuddin Hakim Roslan[a], Mohd Firdaus Md Dan @ Azlan[e], Abd Wahid Rasib[f], Radzuan Saari[a], Mushairry Mustaffar[a], Azman Kassim[a, b], Rini Asnida Abdullah[a], Khairul Hazman Padil[a], Norbazlan Mohd Yusof[g], Norisam Abd Rahaman[g]

[a] *Faculty of Civil Engineering, Universiti Teknologi Malaysia, Johor Bahru, Johor* 81310, *Malaysia*
[b] *Centre of Tropical Geoengineering, Faculty of Civil Engineering, Universiti Teknologi Malaysia, Johor* 81310, *Malaysia*
[c] *Department of Civil Engineering, Damavand Branch, Islamic Azad University, Damavand* 3971878911, *Iran*
[d] *Research Center of Concrete and Soil, Damavand Branch, Islamic Azad University, Damavand* 3971878911, *Iran*
[e] *Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor* 86400, *Malaysia*
[f] *Faculty of Built Environment and Survey, Universiti Teknologi Malaysia, Johor Bahru, Johor* 81310, *Malaysia*
[g] *Projek Lebuhraya Usahasama Berhad (PLUS) Sdn Bhd, Menara Korporat, Persada PLUS, Persimpangan Bertingkat Subang, KM* 15, *Lebuhraya Baru, Lembah Klang, Petaling Jaya, Selangor* 47301, *Malaysia*

ARTICLE INFO

Article history:
Received 3 August 2023
Received in revised form 3 January 2025
Accepted 20 January 2025
Available online 25 August 2025

Keywords:
Slope stability prediction
Random Forest Algorithm
Remote sensing in Geology
Factor of Safety (FOS)
Geometrical parameters
Rock quality designation (RQD)
Multilayer perceptron (MLP)

ABSTRACT

The prediction of slope stability is a complex nonlinear problem. This paper proposes a new method based on the random forest (RF) algorithm to study the rocky slopes stability. Taking the Bukit Merah, Perak and Twin Peak (Kuala Lumpur) as the study area, the slope characteristics of geometrical parameters are obtained from a multidisciplinary approach (consisting of geological, geotechnical, and remote sensing analyses). 18 factors, including rock strength, rock quality designation (RQD), joint spacing, continuity, openness, roughness, filling, weathering, water seepage, temperature, vegetation index, water index, and orientation, are selected to construct model input variables while the factor of safety (FOS) functions as an output. The area under the curve (AUC) value of the receiver operating characteristic (ROC) curve is obtained with precision and accuracy and used to analyse the predictive model ability. With a large training set and predicted parameters, an area under the ROC curve (the AUC) of 0.95 is achieved. A precision score of 0.88 is obtained, indicating that the model has a low false positive rate and correctly identifies a substantial number of true positives. The findings emphasise the importance of using a variety of terrain characteristics and different approaches to characterise the rock slope.

©2025 China Geology Editorial Office.

1. Introduction

Geotechnical engineers struggle to precisely estimate rock slope stability. The complexity of the physical system and the difficulty of obtaining accurate geotechnical input data are the main causes of this. The analysis must consider rock mass conditions, ground behaviour, and external environmental factors. This includes geological features such as the length, roughness, and infill materials of joints, bedding, and faults behind the rock face, as well as the orientation of the rock slope. Additionally, rainfall, topography, artificial disturbances, and other factors may also affect rock stability. Geospatial data that is accurate and consistent is critical when mapping a rock slope, as it represents essential aspects contributing to slope stability, such as geomorphologic, lithologic, soil, and land use characteristics.

This research explores the mapping of rock slopes at Bukit Merah, Perak, and Twin Peak (Kuala Lumpur) by integrating geological and remote sensing methods. Drone-derived topographical information, geospatial data, and slope orientation are utilised. These properties reflect rock structure at different scales. There may be correlations among them that enable predictive modelling, allowing known properties to help infer others.

First author: *E-mail address*: afiqah.ismail@utm.my (Afiqah Ismail).

* Corresponding author: *E-mail address*: ahmadsafuan@utm.my (Ahmad Safuan A Rashid).

Literary editor: Li-qiong Jia
doi:10.31035/cg2023102
2096-5192/© 2025 China Geology Editorial Office.

Copyright © 2025 Editorial Office of China Geology. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd.
This is an open access article under the CC BY-NC-ND License (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Slope engineering can be considered an uncertain, nonlinear, dynamic, and open complex system due to the structural complexity and discontinuities in the mechanical and physical characteristics of rock slopes, as well as the variability and effect of control factors acting on them (Bui X and Nguyen Q, 2019; Dai C and Wang X, 2021; He XL et al,. 2021; Wu H et al., 2022). Most of these factors demonstrate ambiguity, variability, unpredictability, and complex nonlinear relationships with differing weights of influence on slope stability (Kang F et al., 2017). In other words, it is difficult to represent rock slope stability using a simple mechanical or mathematical model.

In recent decades, geotechnical engineers have utilised various machine learning (ML) models to simulate different geotechnical topics (Amiri ST et al., 2020; Dehghanbanadaki A, 2021; Zhang WG et al., 2023; Lin S et al., 2024). These ML techniques have also been employed to map, estimate, and model slope failures (Lu P and Rosenbaum MS, 2003; Gomez H and Kavzoglu T, 2005; Mohamed T et al., 2012; Tsangaratos P and Benardos A, 2013; Qu JK et al., 2023; Singh P et al., 2023; Wengang Z et al., 2023). In this regard, numerous geotechnical scholars are continually working to develop new predictive models for determining slope stability (Sahin E, 2020; Ye J et al., 2022; Khan AF and Rahman MT, 2022). For example, Nanehkaran YA et al. (2023) focused on the estimation process of the factor of safety (FOS) and provided a comparative analysis using computational intelligence and ML methods. Multilayer perceptron (MLP), decision trees (DT), support vector machines (SVM), and random forest (RF) algorithms were employed to predict the FOS for earth slopes. The models were applied to a dataset of 100 earth slopes from Fars, Isfahan, and Tehran provinces in Iran, which were randomly divided into training and testing datasets. Validation was conducted using Janbu's limit equilibrium analysis method (LEM) and GeoStudio software. MLP achieved the highest accuracy (0.901) and precision (0.90), showing good agreement with LEM results. SVM followed, with an accuracy of 0.873 and a precision of 0.85. The average loss in FOS prediction was lowest for MLP (0.29), followed by SVM (0.41), RF (0.45), and DT (0.62).

In a similar study, aimed to develop a new and highly accurate artificial intelligence model for predicting slope failure in open-pit mines. For this purpose, they combined the M5Rules algorithm with a genetic algorithm (GA) to create a novel hybrid technique named the M5Rules–GA model for slope stability estimation and analysis. The authors modelled 450 slope observations in an open-pit mine in Vietnam using GeoStudio software based on essential parameters, with the factor of safety serving as the model outcome. Additionally, Bui X and Nguyen Q, (2020b) developed other models, such as artificial neural networks (ANN), support vector regression (SVR), and previously introduced models (e.g. FFA-SVR, ANN-PSO, ANN-ICA, ANN-GA, and ANN-ABC) for comparison with the proposed M5Rules–GA model. The evaluation of model performance included determination coefficient, variance accounted for, and root mean square error computation, as well as a general ranking and colour scale analysis. The results of the study confirmed that the proposed M5Rules–GA model is a robust tool for analysing slope stability, outperforming the other investigated models in terms of performance metrics.

In another study, Wang G et al. (2023) selected five quantitative indicators to improve the accuracy of prediction models for slope stability based on 77 field cases. The indicators included slope angle, slope height, internal friction angle, cohesion, and unit weight of rock and soil. Six dimension reduction methods were analysed: Principal components analysis (PCA), kernel PCA, factor analysis (FA), independent component analysis (ICA), non-negative matrix factorisation (NMF), and t-SNE (stochastic neighbour embedding). Seven prediction models were established using classic ML methods, and their reliabilities were examined through random cross-validation. The results indicate that dimension reduction is unnecessary for the data processing of the established prediction models. Random forest (RF), support vector machine (SVM), and k-nearest neighbour (KNN) achieved the best prediction accuracy, exceeding 90%. The decision tree (DT) achieved an accuracy of 86%. Slope height was identified as the most important factor influencing slope stability, while the unit weight of rock and soil was the least significant. RF and SVM models demonstrated the highest accuracy and superiority in slope stability prediction.

In addition to various ML models, random forest (RF) has been widely used in geoscience sectors (Bhattacharya S and Mishra S, 2018; Corcoran JM et al., 2013; Fouedjio F, 2020; Feng R et al., 2021), due to its simplicity, resistance to overfitting or noise, and computational efficiency. Corcoran JM et al. (2013) found that RF produces results that are at least as accurate as those generated by other classifiers, such as support vector machines or k-nearest neighbours. Radar (Du P et al., 2015), geophysics (Reading A et al., 2015), geochemistry (McKay G and Harris JR, 2016), and light detection and ranging (Zhu J and Pierskalla WP, 2016) are examples of the numerous datasets that have been effectively classified using RF.

Recently, many researchers have integrated computer domain knowledge with geotechnical engineering to develop novel approaches for resolving slope issues. In the research presented here, the RF network is used in modelling, to produce a probabilistic result rather than categorising slope failure extents using inputs from combining datasets from different approaches. This study aims to test the capabilities and accuracy of the RF network in predicting factor of safety (FOS), considering the combined influence of geological and geospatially derived information.

2. Methodology

To categorise or predict the value of a target variable, the Random Forest (RF) technique utilises an ensemble of decision trees (DTs) (Breiman L, 2001). A decision tree is a tree-like model composed of three types of nodes: The root node, intermediate nodes, and leaf nodes (Fig. 1). Each node represents an attribute or feature, and the path from the root node through intermediate nodes to a leaf node corresponds to

a specific decision rule. The structure of the entire tree reflects a set of rules derived from the training data. The final prediction is obtained through a majority vote (for classification) or averaging (for regression), aggregating the outputs of all individual decision trees. In the RF model, each decision tree operates independently and concurrently to produce a robust and generalised prediction (Xie HY et al., 2022).

Fig. 2 illustrates the RF algorithm workflow. In this figure, variable D represents the dataset, and y_1, y_2, ⋯, y_k denote the individual outputs (predictions) of each decision tree. The RF regression predictor is then expressed as the average of these outputs:

$$\hat{f}_{rf}^{K}(x) = \frac{1}{K}\sum_{k=1}^{K} T(x). \quad (1)$$

where (x) is the input vector. The bagging method divides the data set into several sample subsets. Every sample subset is trained individually and separately, and to increase the variety of the decision trees (DTs), a randomly chosen number of features from all the features are used in the training process rather than all features (Paudel U et al., 2016). One-third of the available training samples that are not included in the RF trees are used to create training data for each tree using the bagging technique. Furthermore, a different subset, known as out-of-bag (or 'oob'), contains the samples that were not chosen for the k-th tree's training during the bagging phase (Rodriguez-Galiano V et al., 2015). They are employed to verify the prediction accuracy. However, as this is performed internally, cross-validation accuracy for RF is not required.

One of the primary strengths of RF is its ability to handle high-dimensional data, which is common in rock slope stability assessment, where multiple variables are considered. RF can effectively select the most relevant features from a large dataset, reducing the risk of overfitting and improving model interpretability. Additionally, RF provides estimates of feature importance, which can help identify the most critical factors influencing rock slope stability. RF is also robust to noise and outliers, which is essential in rock slope stability prediction, where data quality can be variable. RF's non-parametric nature and ability to handle missing values effectively make it a reliable choice for handling imperfect data. Furthermore, RF can be parallelised for computational efficiency, making it suitable for large datasets. Another significant advantage of RF is its internal cross-validation mechanism through out-of-bag error estimates. This ensures

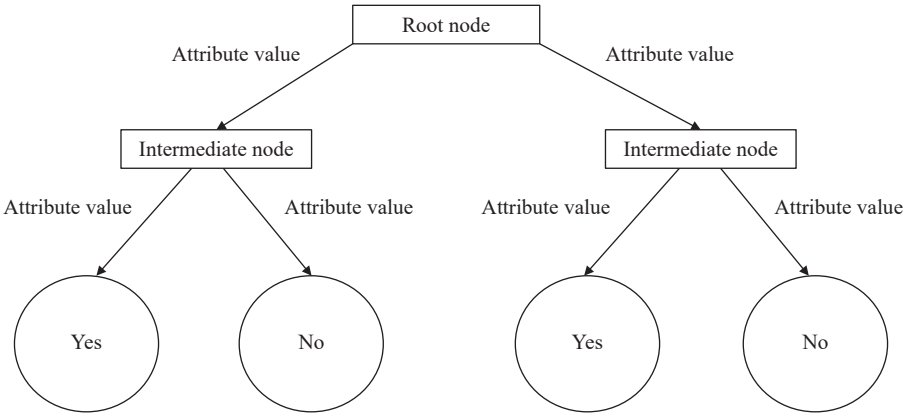

Fig. 1. Schematic of Decision Tree.

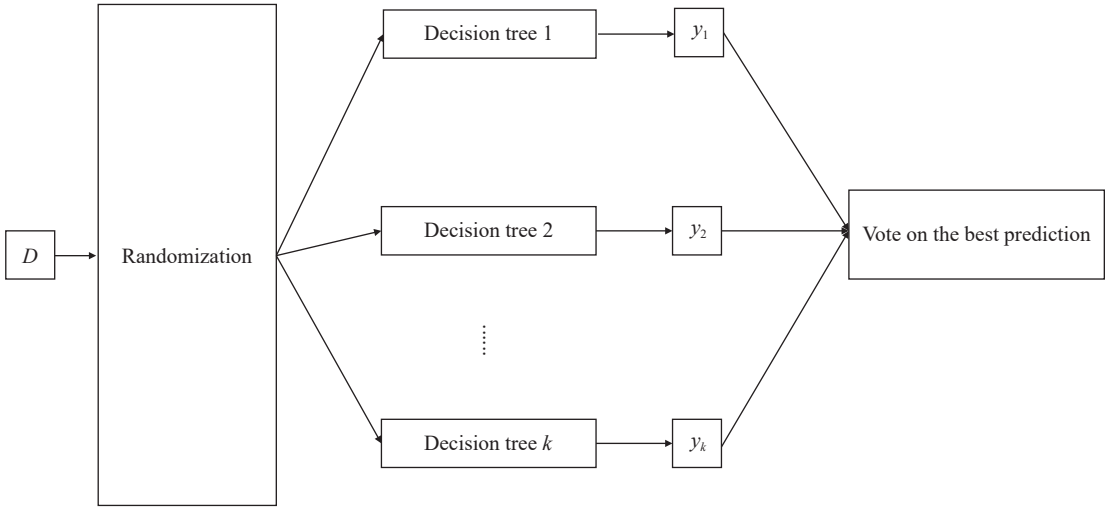

Fig. 2. Random forest algorithm flow.

that the model's accuracy is reliable and not overly optimistic, which is critical in rock slope stability prediction, where the consequences of incorrect predictions can be severe (Salman HA et al., 2024).

While RF has many strengths, it is not without its limitations. One potential weakness of RF is its sensitivity to hyperparameter tuning. However, RF's performance is relatively stable with minimal hyperparameter tuning, making it a low-maintenance option. Additionally, RF can be prone to overfitting if the number of trees is too large, but this can be mitigated through careful tuning of the hyperparameters (Iranzad R and Liu X, 2024).

The RF algorithm is used in modelling, to produce a probabilistic result rather than rock slope classification. According to this study, slope stability refers to an inclined slope's property that allows it to support internal and external forces without shifting. The goal is to predict the safety factor and the slope mass rating with a combination of geological mapping and drone sensor data**.** The first phase of the research comprised data preparation by collecting different types of data from geological mapping and drone sensors. The second phase was random forest prediction analysis.

The RF model used in the present study was built using Pycaret, a recently created Python-based library (Ali M, 2020). This library provides a variety of functions, such as feature selection, hyperparameter tweaking, model ensembling, data preprocessing, training and comparing several ML models. The Google Colab IDE (Bisong E, 2019) platform, which is used by most data scientists and ML practitioners, was used to code this work. This platform offers a convenient cloud-based environment that eliminates the requirement for setting up and managing local development environments. It enables easy accessibility to code execution and provides notable advantages, especially for executing computationally intensive machine learning models. The platform offers high-performance GPUs and the ability to concurrently execute code, making it highly suitable for handling large-scale machine-learning tasks.

3. Study area

3.1. Bukit Merah of Perak

The study area is located in Bukit Merah, Perak, at coordinates 4°59'46.98"N, 100°39'21.39"E (Fig. 3). Bukit Merah lies in northern Perak, approximately 11 kilometres from Ipoh, the capital city of Perak, Malaysia. The rock slope measures 400 m in length and 100 m in height. Generally, Perak's geology is characterised by eruptive masses that form its hills and mountain ranges. The state is divided by three mountain chains into the three plains of Kinta, Larut, and Perak, which run parallel to the coast. Based on the geological map of Kerian, Selama, and Larut Matang Districts, as shown in Fig. 3, three different lithologies underlie the study area: granite, sedimentary rocks, and alluvium (Department of Mineral and Geoscience, 2012). The granite rock in the eastern portion of the study region forms high topography. The majority of Kerian's coastal region, extending to Bukit Merah Lake, is covered by alluvial sediment. According to the geological map, sedimentary rock is located between the granite and alluvial sediment.The sedimentary rock group in the study area is named the Semanggol Formation. The Semanggol Formation (SgF) is a massive mixed sandstone-mud-rich submarine fan complex that formed in an accretionary foredeep basin as a result of the collision of the Sibumasu and Indo-China East Malaya blocks (Hutchinson CS and Tan DNK, 2009). According to Burton CK (1970), the Semanggol Formation in the Baling area is composed of quickly alternating shale, siltstone, and sandstone, with a few bands of chert. Since the Middle and Late Triassic, chert has been seen in lateral deposition with rhythmite and conglomerate elements; radiolarian biostratigraphy established that the chert component is the Semanggol Formation's earliest unit. Sandstone and mudstone (or shale) are found in alternating beds in the middle section, called the rhythmite unit. According to Jasin B (1997), the rhythmite unit is made up of mudstone and turbiditic sandstone interbeds; the sandstone may have turned into siltstone or shale.

In terms of the climatic condition of the study area, Perak is located in a tropical area with an equatorial climate that is often hot, humid, and wet. The entire year is marked by heavy rainfall. In general, the temperature fluctuates between 21°C–27°C. Over 80% humidity is not uncommon. The state's central region averages 5000 mm of rain annually, or roughly 3000 mm annually. The northeast and southwest monsoon seasons alternate across the state, and the transitional months

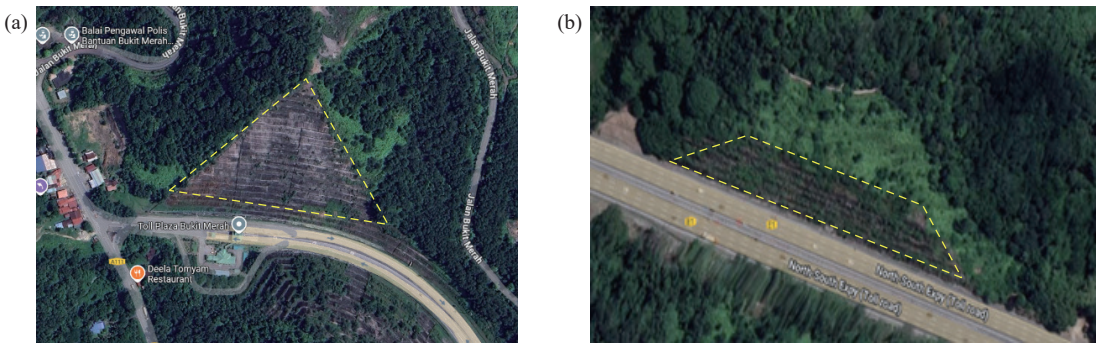

Fig. 3. Satellite image of (a) rock slope located in Bukit Merah, Perak (inside yellow dash) and (b) Twin Peak. Note: rock slope which is along the NKVE highway (inside yellow dash).

for the monsoon season are April and June. The northeast season lasts from November to March, and the southwest season lasts from May to September. Floods are brought on by the northeast monsoon's strong rains, especially in Hulu Perak's upper regions. Little effect of the southwest monsoon is felt in the Kinta Valley. However, the coastal areas of southern Perak occasionally experience thunderstorms, heavy rain, and strong, gusting winds in the predawn and early morning.

3.2. Twin Peak, Kuala Lumpur

The Twin Peak rock slope is located at 3°10′45.07″N, 101°38′26.99″E along the New Klang Valley Expressway (NKVE), a controlled-access highway that traverses the entirety of Selangor and Kuala Lumpur in the Klang Valley of Peninsular Malaysia (Fig 4). Kuala Lumpur is situated at the confluence of the Klang and Gombak Rivers, in the central-western region of Peninsular Malaysia. The Klang Valley, where the city lies, is encircled by hills, with the Titiwangsa Mountains to the east and the Klang Gates Quartz Ridge to the north. The bedrock geology of Kuala Lumpur comprises a variety of rock types (Department of Mineral and Geoscience, 2012). Paleozoic metasedimentary successions primarily underlie the Klang Valley area. Apart from several rock slopes in mildly to moderately weathered schist or metasedimentary formations, most of the slopes in the region consist of cut slopes in Grade I to II granite, which define the mountainous terrain bordering the Klang Valley to the east and west (Abdul Rahim AF et al., 2019).

Triassic granite (ca. 220–198 Ma), part of the tin-bearing Main Range Granite Province of the 'Western Belt' in Peninsular Malaysia, is exposed in the elevated hills surrounding Kuala Lumpur to the east, north, and west (Ghani AA et al., 2013). Granite plutons intrude the entire bedrock sequence of Kuala Lumpur (Fig 4). These lithological units are dissected by a series of WNW–ESE-trending faults, which locally host thick, decameter-scale quartz veins, such as the Klang Gates Quartz Reef within the Kuala Lumpur Fault Zone (Stauffer D, 1968; Shu CK, 1969). The predominant rock type is megacrystic biotite granite, characterised by coarse to extremely coarse grain sizes (Ghani AA, 2009).The Titiwangsa Range in the east and Sumatra Island (Indonesia) in the west protect Kuala Lumpur from strong winds. Kuala Lumpur experiences a tropical rainforest climate that is hot, humid, and sunny, with a lot of rain, particularly from October to March, when the northeast monsoon season occurs. Maximum temperatures range from 32°C–35°C, occasionally rising to 38°C, while minimum temperatures range from 23.4°C–24.6°C and have never dropped below 17.8°C. June through August are normally dry, and the rainfall in Kuala Lumpur typically exceeds 131 mm every

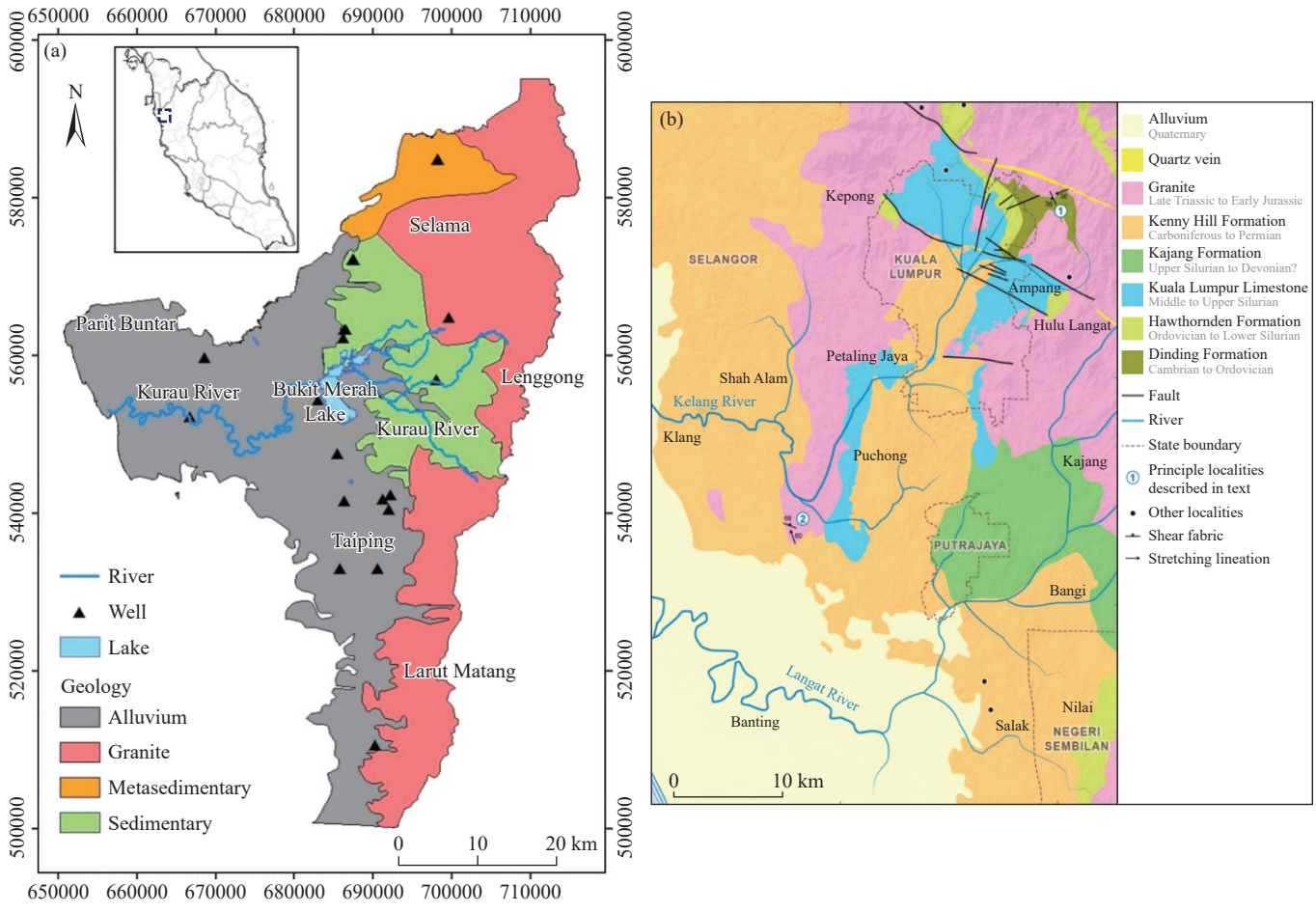

Fig. 4. a−Geological map of Bukit Merah; b−Geological map of Kuala Lumpur (Department of Mineral and Geoscience, 2010).

month, even then. The city typically receives at least 2600 mm of rain yearly.

4. Data preparation

Input datasets, to predict the binary outcomes of rock slope stability ('No' or 'Yes'), consist of different types of variables. Data were collected using two methods: (1) the acquisition of vertical and oblique aerial photographs from an unmanned aerial vehicle (UAV) and (2) conventional rock slope mapping.

4.1. Aerial photographs from a UAV

The rock slope was photographed using a UAV with a DJI Phantom 4 Pro drone, with a 20-megapixel camera attached. Two hundred shots were obtained of top and side views of the rock slope, using reference points to enable the drone to take as many pictures as necessary to cover the area. The images were edited using the Agisoft PhotoScan programme. The software employs the structure from motion (SfM) technique to recognise features in the image for bundle adjustment. To produce three-dimensional surface models, the SfM photogrammetric technique overlaps photos. In addition, ground controls were implemented to guarantee that the photogrammetry procedure produced precise and dependable findings. An orthophoto, a 3D dense point cloud, a 3D model, and a digital surface model represent the photogrammetry process (DSM) results.

The raw data from the drone were stitched together, and layer stacking was utilised to combine two orthophotos. Red, green, and blue (RGB) photos were processed using the Pix4D Mapper Software, while near-infrared (NIR) images were analysed using the Agisoft PhotoScan professional programme. When the stitching was finished, the RGB and NIR orthophotos went through a stacking procedure. To determine the value of the normalised difference water index, the stacking procedure superimposed two separate photos (NDWI). The ArcGIS ArcMap 10.3 software's georeferencing tool was used to stack the RGB and NIR orthophotos for this project. The RGB orthophoto was subjected to the process of geometric image rectification. The output of the data processing produced three different remote sensing maps: the NDVI map, NDWI map and the thermal infrared (IRT) distribution map. All of these maps were used to extract information that may contribute to the rock slope stability. Figs. 5 and 6 show the remote sensing maps at Bukit Merah and Twin Peak.

The normalised difference vegetation index (NDVI) is a vegetation index derived using near-infrared (NIR) imagery. In addition to assessing vegetation health, NDVI can also be applied to evaluate the stability of rock surfaces that are vegetated. Bhandari D et al. (2012) indicated that barren areas —such as those dominated by rock, soil, or sand — typically yield low NDVI values (i.e. ⩽0.1). McFeeters SK (2013) stated that NDVI is a basic yet reliable metric for evaluating ground surface vegetation cover, making it widely applicable in remote sensing applications. NDVI values, typically obtained via satellite-based observations, provide insights into the presence and condition of live green vegetation. In this study, NDVI values were extracted using the Raster Calculator tool in ArcGIS ArcMap 10.3 software. The NDVI is calculated using the following equation:

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + \mathrm{Red})}. \tag{2}$$

The NDWI algorithm was used in this research to indicate the existence of water bodies on the rock surface. McFeeters

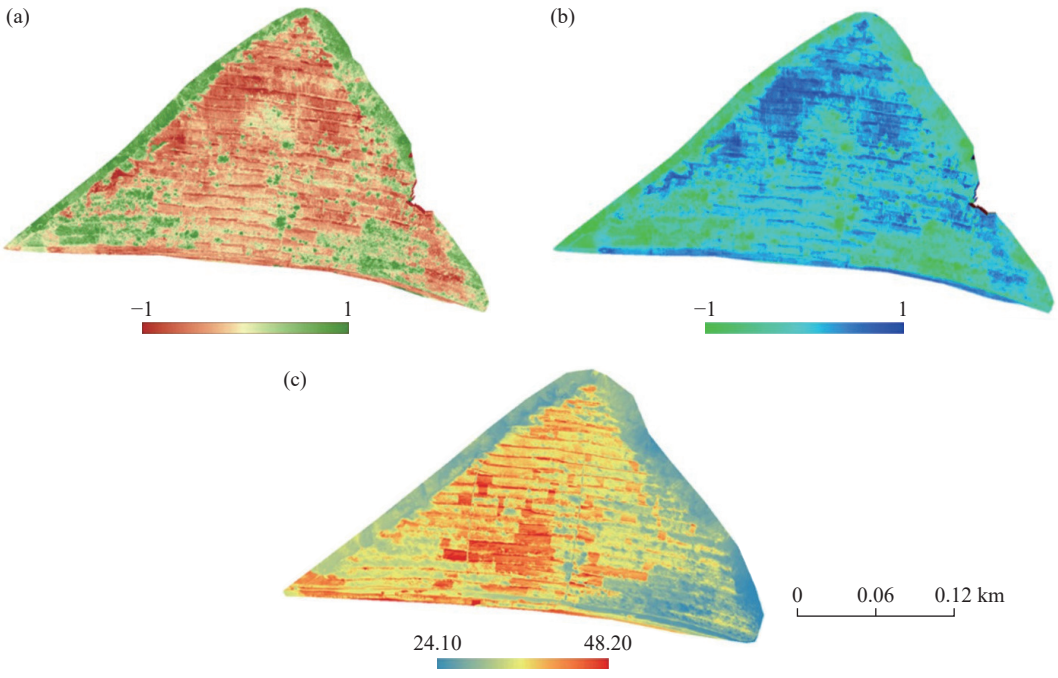

Fig. 5. Distribution maps of NDVI (a), NDWI (b), and IRT(c) at Bukit Merah, Perak.

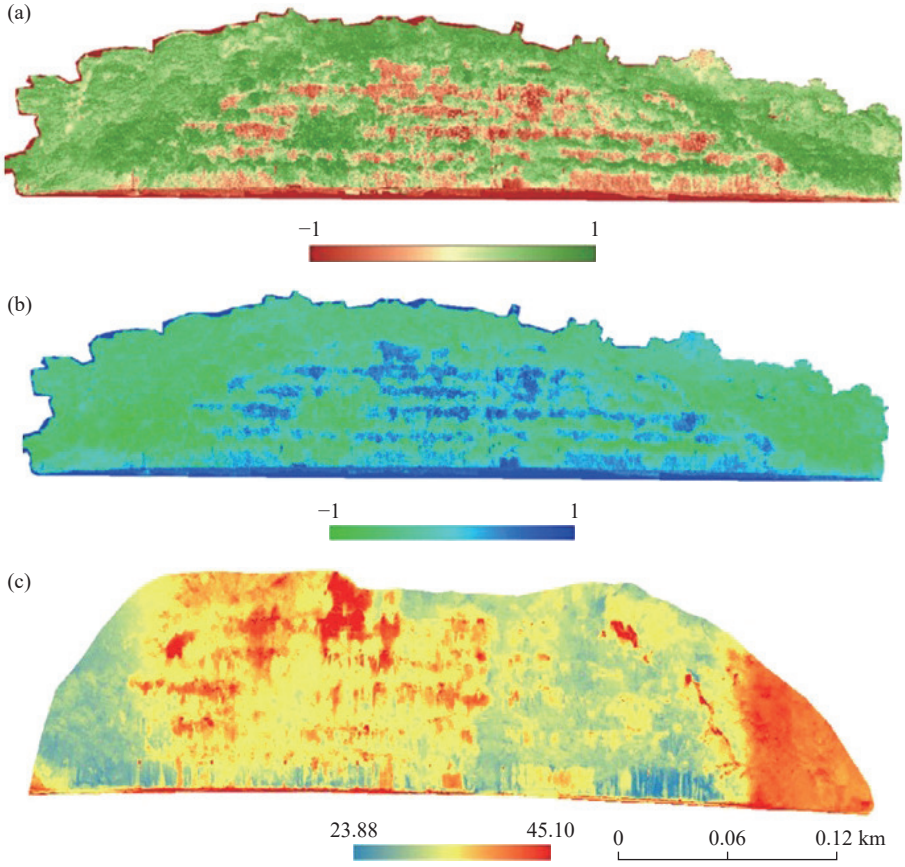

Fig. 6. Distribution maps of NDVI (a), NDWI (b), and IRT (c) at Twin Peak, Kuala Lumpur.

SK (2013) stated that the presence of water on a rock surface is characterised by a larger value (greater than 0) of NDWI. However, a lack of water on the rock surface is indicated by a value less than or equal to zero. The NDWI relies on high reflectance in the NIR range and is generally an excellent proxy for recognising a water surface or the existence of water on the ground surface. The Raster Calculator tool in ArcGIS ArcMap 10.3 was used to derive the NDWI. The equation used to calculate the NDWI is as follows:

$$\mathrm{NDWI} = \frac{(\mathrm{Green} - \mathrm{NIR})}{(\mathrm{Green} + \mathrm{NIR})}. \quad (3)$$

According to Ji L et al. (2009), the ground surface is classified as a water body when the Normalised Difference Water Index (NDWI) is greater than 0 (NDWI > 0), and it is considered non-water when the NDWI is less than or equal to 0 (NDWI $\leqslant$ 0). This aligns with the findings of Lu S et al. (2010), who noted that a negative NDVI value or a value close to zero typically indicates the presence of water. These observations suggest an inverse relationship between NDVI and NDWI: when NDWI values are high, NDVI values tend to be low, and vice versa. In this research, a similar inverse trend was observed between NDVI and NDWI values.

As for the rock surface temperature, the threshold was obtained using an infrared thermography (IRT) image from a FLIR Zenmuse XT sensor designed for thermal imaging purposes. To analyse temporal changes of the rock surface due to different seasons, the data acquisition was carried out in two distinct phases: Phase 1 occurred during the dry season, and phase 2 occurred during the wet season. These two different phases were expected to show the effect of weathering on the rock slope through the differences between the indices and temperature values from the two seasons. The IRT threshold image was used to identify the infrared energy transmitted from an object, change it to the apparent temperature, and display the result as an infrared image. The thermography threshold image depends on the features of the rock slope or the surface thermal conductivity. Rock and concrete features give higher conductivity values than vegetation, which parallels the research finding that rock and concrete surface temperatures are much higher than vegetation.

4.2. Conventional rock slope mapping

The high-density point cloud was utilised to create a detailed 3D surface model of the rock slope. To facilitate a more accurate and efficient analysis, the digitalised 3D rock slope at Bukit Merah and Twin Peak was divided into 30 and 51 panels, respectively (Fig. 7). This division was based on the need to optimise the size of the panels to capture the variability in the rock slope's geological and geomechanical characteristics, while also ensuring that each panel was large enough to be representative of the overall rock slope. Since the entire rock slope is composed of a single lithology, the panels were divided based on size rather than material type.

The division of the rock slope into panels allowed for a

more detailed and accurate analysis of the digital photogrammetry data, which was validated with discontinuity survey and field observation data. The field data collection included measurements of the dip angle of joints, rock slope durability, the presence of filtration, and other discontinuity information. A Brunton compass was used to measure the geographical characteristics of the joint's dip and dip direction. By combining the digital photogrammetry data with field observations, we were able to develop a comprehensive understanding of the rock slope's geological and geomechanical characteristics, which informed the development of the random forest model.

Geological mapping was carried out for each panel through the geostructural domain's definition and the rock mass' geomechanical description. A preliminary geometric report was given for each region, with geostructural domains and principal joint sets. Then, a series of geostructural surveys were performed along scanlines, recording the orientation (dip and dip direction), spacing, persistence, roughness, and general condition (alteration, aperture, filling), in accordance with the method suggested in ISRM (1978). The scanline traversal method was performed at an interval of 100 m for every region. The alignment of the traverse and the locations of both ends of the traverse needed to be determined. The combination of the point cloud and the software application provided the location coordinates for each point, making it easier to define the target area. Feature locations were projected along the strike to the line and the distances recorded. The different parameters taken, along with rock mass encountered, height, length, slope orientation, major joint sets, critical percentage of wedge failure, etc., at those cuttings, were recorded carefully. The kinematic analyses of major joint sets and the slope orientation of cuttings were also carried out by choosing the wedge/planar sliding mode and entering the kinematic properties. All of the parameters shown in Table 1 were used as input data when performing the RF.

5. Results and discussions

5.1. Construction of random forest prediction model

This study estimated the FOS by considering the influence factors of rock mass properties, kinematic analysis, and drone-derived methods. Fig. 8 depicts the stages of the stability of the rock slope prediction model, based on the RF model created in this study. On the basis of the original data, the rock slope parameter's initial index system was developed. A correlation analysis was carried out after ranking each affecting factor's significance. Finally, the RF

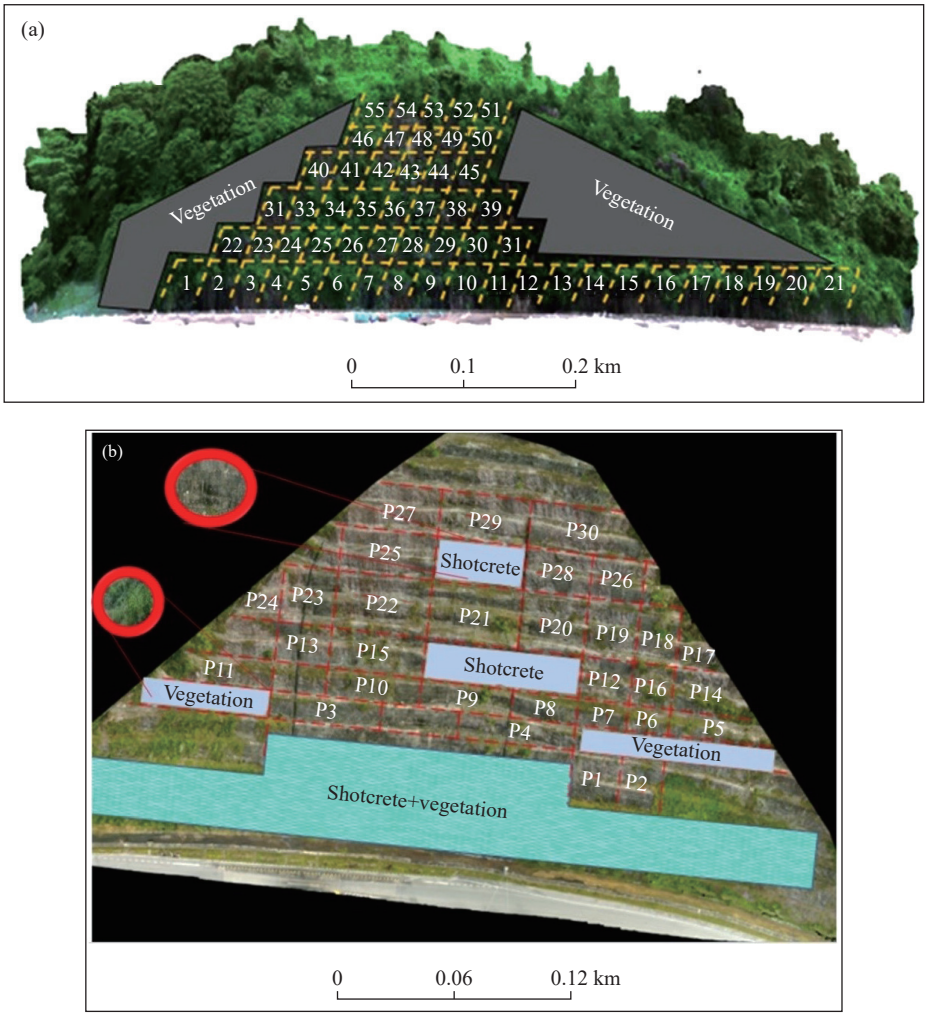

Fig. 7. a−Rock slope of Bukit Merah divided into 30 panels. b−Rock slope of Twin Peak NKVE divided into 51 panels.

Table 1. Input parameters.

Input	Features
i. Rock mass properties	UCS
	RQD
	Spacing
	Persistence
	Aperture
	Roughness
	Infilling
	Weathering
	Orientation of discontinuity factor
	Excavation method
ii. Drone-based remote sensing image (vegetation indices and thermal)	NDVI
	NDWI
	Temperature
iii. kinematic analysis	Number of joints identified
	Number of joint sets
	Critical number of failures/%
	Type of failure (P/T/W)
	Number of critical joints

model performed the regression prediction of the test samples and the accuracy analysis.

18 influencing parameters of the rock slopes were chosen as independent variables from the slope stability study, while FOS was the output-dependent variable. The parameters are rock strength, RQD, joint spacing, continuity, openness, roughness, filling degree, weathering degree, groundwater, NDWI, NDVI, IRT, number of joint sets, critical number of failures, type of failure, and number of critical joints. Tables 2 and 3 show the training set and its output for Bukit Merah, while Tables 4 and 5 show the input and output variables for Twin Peak. The 18 parameters for our study were selected based on a thorough evaluation of factors that significantly impact slope stability. These parameters were chosen to ensure a holistic and multidisciplinary approach, integrating geological, geotechnical, and remote sensing analyses. By combining traditional geotechnical methods with modern remote sensing techniques, we aimed to capture a comprehensive range of influential factors. This multidisciplinary approach is intended to enhance the accuracy and reliability of slope stability predictions.

5.2. Output data: FOS

Slope stability analysis is necessary to prevent slope failure and the FOS is a key consideration. To assess the stability of the slope, the minimal FOS against sliding or shear failure must be calculated. The ratio of the ultimate shear strength divided by the mobilised shear stress during an incipient failure is the FOS for slope stability. The FOS for all panels in the study area focused on planar and wedge failure. Based on the requirement that PLUS Berhad provided, the FOS for an untreated slope should not be less than 1.3, while, for a treated slope, it should not be less than 1.5. The proposed rock slope panel was modelled in Rocscience, Rocplane, and Swedge software to carry out rock slope

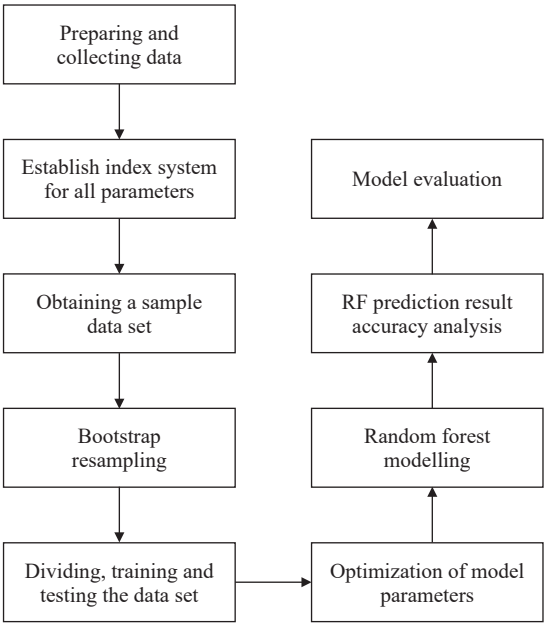

Fig. 8. Random forest machine learning flow in prediction rock slope stability.

stability assessment. The shear strength model used was based on the Barton-Bandis model, where the JRC and JCS values were obtained based on site observations and several studies by the previous author. This failure criterion is more accurate than the Mohr-Coulomb envelope as it reflects real-time shear strength in a changing stress state.

5.3. Machine learning results

Fig. 9 shows the correlation between rock mass properties. In the present study, the slope was considered unstable when the calculated FOS was determined to be below 1. Conversely, the slope was considered stable when the FOS was greater than 1. It should be noted that, in the PyCaret library, several ML methods are trained on the data and ranked based on performance indices. The initial results indicated that LightGBM (Light Gradient Boosting Machine) and RF were the top two models, based on the training data. However, performance indices on the new data demonstrated that RF was the best model. Consequently, RF was used for the final estimation in our results.

To enhance the accuracy of the RF model, a tuning process was performed using the Pycaret library. For RF models, the hyperparameters which were most commonly adjusted were the number of trees in the forest (n_estimators), the maximum depth of each tree (max_depth), the minimum number of samples required to split an internal node (min_samples_split), and the maximum number of features considered for the best split (max_features). To execute the hyperparameter search, the tune_model function provided by PyCaret was utilised, which accepts an RF model as input. This function performed a grid search over the specified hyperparameter space. Each hyperparameter combination was evaluated using cross-validation, a technique that partitions

Table 2. RF prediction model training sample for Bukit Merah.

		Panels														
Categories	Description	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Rock mass properties	UCS/MPa	4	4	4	4	2	2	2	4	4	4	4	4	4	4	4
	RQD	17	17	17	20	20	20	17	13	20	3	8	13	3	17	20
	Discontinuity spacing	15	15	15	15	10	5	10	10	10	8	8	10	8	10	15
	Persistence	2	4	4	2	2	4	4	4	2	2	4	4	2	2	2
	Aperture	4	5	5	4	4	5	5	5	4	4	5	5	4	4	4
	Joint roughness coefficient (JRC)	3	5	5	3	3	5	5	5	3	3	5	5	3	3	3
	Weathering grade	3	5	5	5	2	3	3	5	3	3	5	5	1	3	3
Kinematic analysis	Number of joints identified	9	9	45	12	6	6	7	7	6	8	9	33	25	17	13
	Number of joint set	3	2	3	3	1	3	2	3	3	3	3	3	3	3	1
	Type of failure (P/T/W)	1	1	2	1	1	2	1	2	1	2	t	t	3	3	2
	Critical number of wedge failure plane	1	2	14	0	9	6	3	6	20	4	0	0	13	11	1
	Critical number of topple failure plane	0	0	1	0	0	9	0	3	0	1	1	2	20	2	4
	Critical number of planar failure plane	0	0	0	1	0	0	0	0	0	0	0	0	3	3	0
UAV multispectral and drone	NDVI	0.006	0.178	0.122	0.11	0.11	0.116	0.098	0.118	0.0912	0.08	0.09	0.112	0.058	0.082	0.096
	NDWI	−0.046	0.088	−0.076	−0.078	0.004	−0.044	−0.05	−0.066	−0.042	−0.064	−0.054	−0.029	0.002	−0.056	−0.02
	Temp (Celcius)	34.89	33.25	34.59	32.94	39.2	35.4	40.1	35.7	35.7	40.1	33.6	33.31	34.72	31.68	34.42
		Panels														
Categories	Description	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Rock mass properties	UCS/MPa	4	4	4	4	4	2	7	4	4	4	4	4	4	4	4
	RQD	20	8	8	13	8	3	20	20	8	20	20	20	20	20	20
	Discontinuity spacing	10	8	8	5	10	10	10	10	10	10	15	15	10	15	15
	Persistence	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	Aperture	5	1	1	0	4	1	1	1	1	1	1	1	4	1	1
	Joint roughness coefficient (JRC)	5	3	3	3	3	3	3	3	3	3	3	3	3	3	3
	Weathering grade	5	3	1	3	5	5	5	3	5	3	3	3	3	1	3
Kinematic analysis	Number of joints identified	11	7	8	9	6	23	4	8	7	28	8	10	12	10	6
	Number of joint set	3	3	3	3	4	3	3	1	3	2	3	3	4	2	3
	Type of failure (P/T/W)	2	1	1	1	1	2	1	1	1	3	1	1	2	1	2
	Critical number of wedge failure plane	2	12	6	0	0	3	5	10	4	4	10	3	0	1	1
	Critical number of topple failure plane	1	0	0	3	15	0	0	0	0	6	0	0	2	0	1
	Critical number of planar failure plane	0	0	0	0	0	1	0	0	0	3	0	0	0	0	0
UAV multispectral and drone	NDVI	0.036	0.162	0.031	0.018	0,028	0.031	0.093	0.03	0.082	−0.2	0.01	−0.098	0.032	0.01	0.054
	NDWI	−0.016	−0.02	−0.044	−0.02	−0.042	−0.03	0.022	−0.034	0	0.074	0.092	0.114	0.08	0.068	−0.022
	Temp (Celcius)	30.58	31.6	33.21	33.78	33.04	38.1	37.9	36.1	35.3	34.09	31.15	34.43	33.65	33.7	34.8

Table 3. RF prediction model dependent output for Bukit Merah, Perak.

Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
FOS	7.39	4.62	1.89	1.27	1.08	6.86	1.59	5.91	2.24	6.99			1.39	0.84	6.78	5.09	1.49	1.17			1.44	1.28	5.51	1.38	1.12	4.97	3.91		1.73	1.98

Table 4. RF prediction model training sample for Twin Peak, Kuala Lumpur.

c	Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Rock Mass Properties	UCS/MPa	4	4	4	4	4	4	4	2	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
	RQD	13	20	20	17	20	17	20	17	13	13	8	20	20	20	17	20	20	20	20	15	17	20	20	13
	Discontinuity spacing	10	15	15	10	15	10	15	10	10	10	8	10	15	10	15	10	10	15	15	15	10	10	15	10
	Persistence	10	14.5	10	2	10	15	10	10	2	5	2	2	20	15	15	5	9.5	5	2	15	10	15	10	2
	Aperture	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
	Joint roughness coefficient (JRC)	7	8	9	8	9	7	9	10	7	8	6	7	7	5	6	6	5	5	6	6	8	6	5	7
	Weathering grade	5	5	5	5	5	5	3	3	3	3	5	5	5	5	5	5	5	5	5	5	5	3	3	2
Kinematic analysis	Number of joints identified	10	11	10	10	9	12	11	9	8	8	14	13	13	13	16	16	16	16	16	18	14	10	12	41
	Number of joint set	4	3	3	2	2	3	3	2	2	1	2	1	1	2	3	3	3	3	2	3	3	2	3	2
	Type of failure (P/T/W)	3	2	1	1	2	1	1	1	3	1	1	2	1	1	1	2	6	3	1	1	0	3	3	2
	Critical number of wedge failure plane	18	12	2	2	4	6	0	4	14	12	6	3	0	5	21	33	32	43	4	17	2	15	2	374
	Critical number of planar failure plane	2	1	1	0	1	0	0	1	3	2	0	0	0	0	3	5	3	3	0	1	1	2	1	13
Drone based	NDVI	0.27	0.41	0.36	0.36	0.43	0.45	0.43	0.31	0.39	0.37	0.36	0.31	0.47	0.32	0.32	0.34	0.35	0.34	0.36	0.26	0.45	0.44	0.296	0.43
	NDWI	−0.23	−0.18	−0.22	−0.19	−0.33	−0.31	−0.25	−0.196	−0.12	−0.18	−0.23	−0.16	−0.31	−0.2	−0.24	−0.17	−0.2	−0.19	−0.16	−0.16	−0.23	−0.21	−0.23	−0.3
	TEM (Celcius)	25.9	25.8	27.94	27.96	27.08	26.28	26.18	25.32	27.46	28.2	26.84	26.86	27.04	28.06	27.78	26.28	29.7	28.62	29.56	30.4	28.62	26.7	29.38	29.12

Table 4. RF prediction model training sample for Twin Peak, Kuala Lumpur (continue).

c	Panel	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
Rock mass properties	UCS/MPa	4	4	4	2	2	4	4	4	4	4	4	4	2	2	4	4	4	4	2	2	2	4	4	2
	RQD	20	20	20	17	13	20	8	20	20	20	17	20	13	17	13	20	20	20	13	17	13	20	17	20
	Discontinuity spacing	15	10	10	10	10	10	8	10	10	15	8	10	10	8	10	10	15	15	10	10	10	15	10	10
	Persistence	10	15	10	12	2	4.5	3	10	2	9.5	15	10	9	3	4	1	11.5	16.5	9	8.5	2	10	13.5	10
	Aperture	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Joint roughness coefficient (JRC)	7	7	6	5	6	5	5	7	9	4	7	5	5	4	4	3	5	8	3	4	6	6	5	5
	Weathering grade	1	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
Kinematic analysis	number of joints identified	0	28	29	62	22	41	37	39	34	21	34	37	38	27	19	44	27	26	38	25	27	33	20	37
	Number of joint set	0	3	3	4	2	4	3	4	3	2	3	2	4	4	2	4	4	3	4	3	3	2	3	3
	Type of failure (P/T/W)	2	0	0	4	2	2	2	2	3	1	1	1	4	3	1	5	2	0	5	2	3	2	2	2
	Critical number of wedge failure plane	0	30	180	1233	139	0	379	387	265	0	199	284	230	117	89	593	132	5	390	110	215	137	137	182
	Critical number of planar failure plane	0	0	0	27	8	5	12	14	10	0	5	0	5	6	0	23	6	0	12	10	10	0	0	0
Drone based	NDVI	0.45	0.42	0.34	0.4	0.37	0.35	0.35	0.34	0.36	0.65	0.41	0.298	0.33	0.32	0.28	0.36	0.37	0.38	0.34	0.23	0.28	0.38	0.28	0.27
	NDWI	−0.29	−0.23	−0.2	−0.23	−0.21	−0.21	−0.2	−0.13	−0.196	−0.31	−0.24	−0.17	−0.19	−0.198	−0.21	−0.21	−0.22	−0.24	−0.21	−0.18	−0.15	−0.24	−0.23	−0.19
	TEM/Celcius	27.12	26.82	26.34	26.28	27.6	28.12	26.92	29.24	29.78	26.66	25.98	26.28	27.96	29.2	28.9	28.58	28.7	26.74	26.48	26.26	29.1	28.34	29.84	29.58

Table 5. RF prediction model output for Twin Peak, Kuala Lumpur.

Panel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
FOS (W)	1.26	1.71	1.25	1.67	1.67	1.51		1.47	1.21	3.06	2.93	3.34		1.58	1.46	1.45	0.68	1.38	1.55	1.54	1.52	1.02	1.55	3.12
Panel	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
FOS (W)	1.54	4.78	1.61	1.58	1.26	1.53	1.68	2.84	5.33	None	1.54	1.25	1.15	1.52	2.71	1.86	1.74	1.55	1.42	1.28	1.51	5.47	1.24	1.83

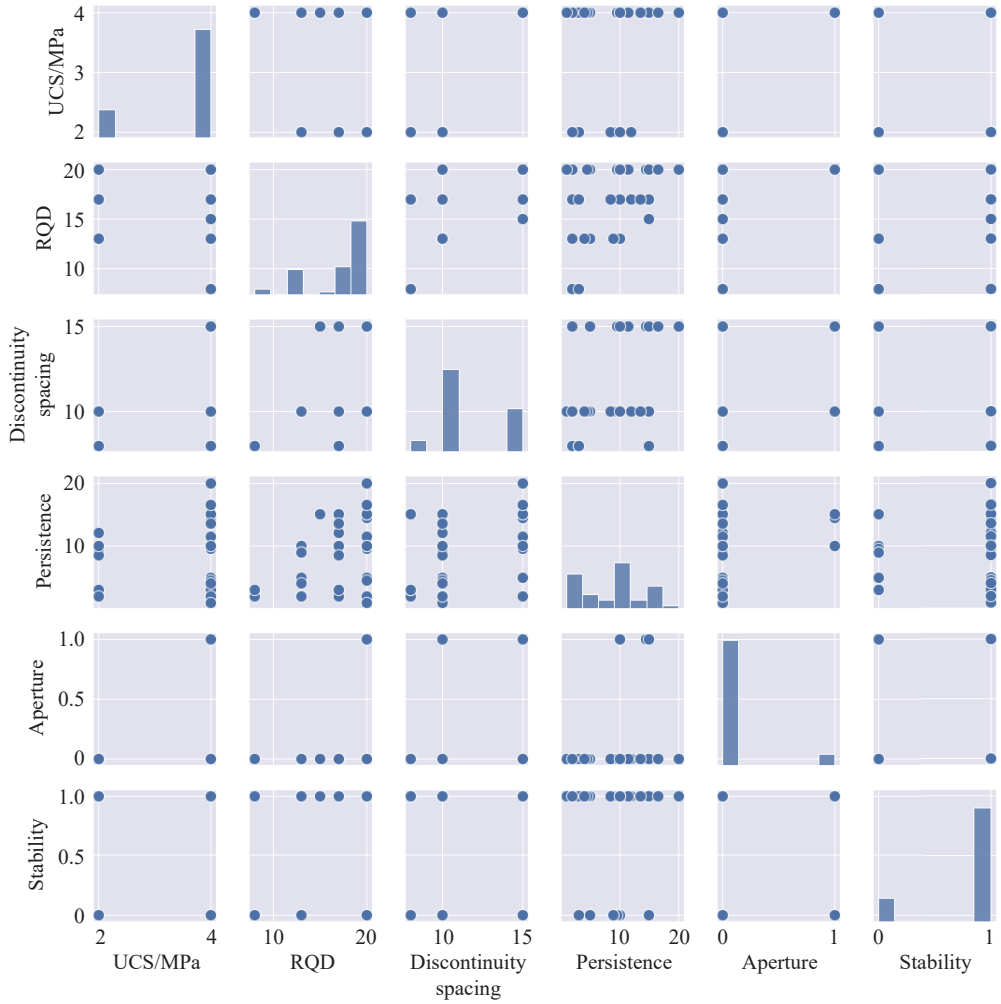

Fig. 9. Correlation between rock mass properties.

the dataset into multiple folds and measures the model's performance on each fold. This iterative process facilitated the estimation of the model's ability to generalise unseen data. The tuning process yielded the results displayed in Table 6, illustrating the characteristics of the optimised model.

Fig. 10 demonstrates the AUC (area under the receiver operating characteristic curve) of RF. The concept of this performance metric refers to evaluating the overall performance of a binary classification model. AUC is used to evaluate the performance of a model when distinguishing between the positive and negative classes. The AUC values range between 0 and 1, where a higher value indicates better model performance. An AUC score of 0.5 suggests random guessing, while an AUC score of 1.0 represents a perfect classifier that can perfectly distinguish between the positive and negative classes. As can be seen, both classes with AUC=0.95 showed good performance in classifying the slopes. In practical terms, an AUC score of 0.95 means that the model has correctly classified 95% of the positive and negative instances in the dataset. This indicates that the model can discriminate between the two classes and can be considered a good classifier. However, it is important to note that the AUC score alone does not provide information on the model's specific performance at different classification thresholds or the quality of the model's predictions for individual instances. Therefore, other metrics, such as precision and F1 score, should also be considered when evaluating the performance of a binary classification model.

Fig. 11 shows the precision-recall (PR) curve obtained for the RF model. This curve illustrates the balance between precision and recall at different classification thresholds in binary classification. Precision measures the proportion of true positives out of all positive predictions, while recall measures the proportion of true positives out of all actual positive instances. A precision score of 0.88 is generally considered high, indicating that the model has a low false positive rate and correctly identifies a substantial number of true positives. However, a high precision score may lead to a lower recall, causing the model to miss some positive instances and increase the false negative rate.

The learning curve in Fig. 12 illustrates the performance of the RF model for the size of the training dataset. This curve is useful for evaluating whether the model is overfitting or underfitting and whether it would benefit from additional training data. The y-axis of the curve represents the performance score of the RF model which, in this case, is the AUC. The x-axis represents the size of the training dataset. The curve shows that, as the size of the training dataset

Table 6. Characteristics of the optimised model.

Parameters	Values
bootstrap	False
ccp_alpha	0.0
class_weight	{}
criterion	'entropy'
max_depth	7
max_features	1.0
max_leaf_nodes	None
max_samples	None
min_impurity_decrease	0.002
min_impurity_split	None
min_samples_leaf	3
min_samples_split	5
min_weight_fraction_leaf	0.0
n_estimators	270
n_jobs	−1
oob_score	False
random_state	8502
warm_start	False

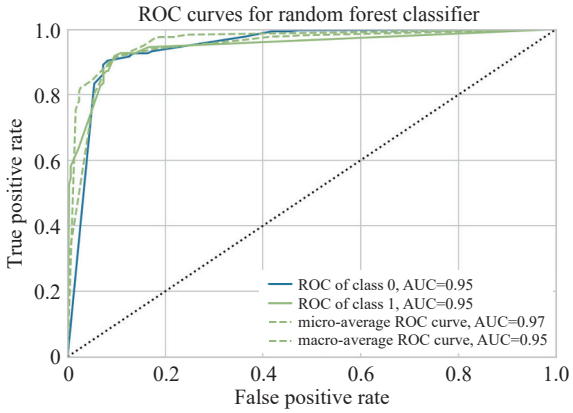

Fig. 10. AUC curve for RF.

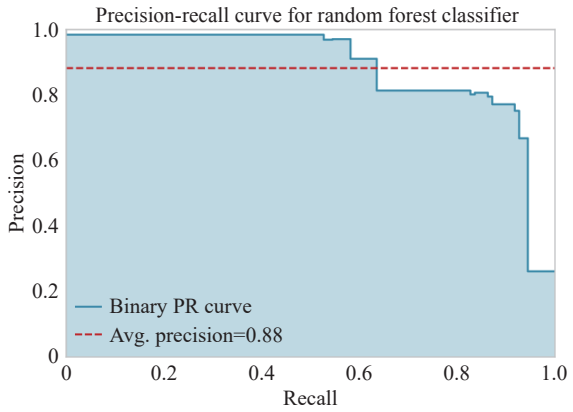

Fig. 11. AUC curve for RF.

increases, the AUC score also increases, indicating that the model is improving. However, the curve also suggests that the model may continue to improve with additional training data, as there appears to be a slight gap between the training and validation scores, indicating some degree of overfitting.

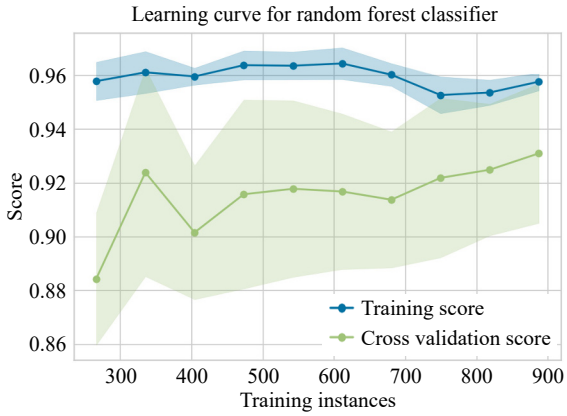

Fig. 12. Learning curve for RF.

6. Research limitations and future directions

The proposed method for slope stability prediction was applied specifically to the Bukit Merah, Perak, and Twin Peak study area in Kuala Lumpur. It is important to validate and assess the generalisability of the method in other geographical locations and different types of rocky slopes. Additionally, the availability and quality of data poses challenges to accurate predictions. Obtaining comprehensive and reliable data for all control factors, such as rock strength, joint spacing, and weathering, requires careful consideration.

Based on these findings, the authors recommend several specific applications: First, the model should be integrated into standard slope assessment protocols, where periodic evaluations using our 18-parameter framework can provide consistent stability assessments. Second, organizations should develop a systematic database incorporating these parameters for all monitored slopes, enabling trend analysis and early detection of stability changes. Third, the model's predictions should be used to establish a risk-based maintenance schedule, where slopes with lower stability predictions receive more frequent inspections and maintenance. Fourth, the findings can guide the optimization of drone flight planning and data collection strategies, particularly for the remote sensing parameters (NDVI, NDWI, temperature). Finally, we recommend using the model's results to develop site-specific threshold values for each parameter, which can serve as triggers for implementing preventive measures. These recommendations aim to transform the model from a theoretical tool into a practical asset for slope management and safety enhancement.

7. Conclusion

Geotechnical engineers encounter formidable challenges when attempting to ensure slope stability, a task compounded by the complex interplay of geological, hydrological, and joint characteristics. To address this, this study proposes leveraging a random forest (RF) ML approach for automated slope stability evaluation, particularly on steep granitic slopes. Considering a comprehensive array of input parameters, including rock mass properties, drone-based remote sensing

imagery (incorporating vegetation indices and thermal data), and kinematic analysis, our findings demonstrate the effectiveness of this approach. The optimisation of the RF model yielded impressive results, with an AUC score of 0.95 and a precision score of 0.88, offering a novel strategy for predicting and managing slope stability with remarkable accuracy. A precision score of 0.88 signifies a low false positive rate, indicating the model's ability to correctly identify a significant number of true positives. While minor errors may arise in establishing correlations between input and output parameters, the model's accuracy is verified through a rigorous comparison of predicted outcomes with actual observations. These results underscore the efficacy of the RF model as a reliable tool for slope stability assessment and management.

CRediT authorship contribution statement

Afiqah Ismail, Ahmad Safuan A. Rashida, Ali Dehghanbanadaki, Rafiuddin Hakim Roslan, Mohd Firdaus Md Danazlan, Abd Wahid Rasi, Radzuan Saari, Mushairry Mustaffar, Azman Kassim, Rini Asnida Abdullah, Khairul Hazman Padil, Norbazlan Mohd Yusof, and Norisam Abd Rahman conceived the presented idea. Afiqah Ismail, Ahmad Safuan A. Rashida, and Ali Dehghanbanadaki designed and performed the experiments. Rafiuddin Hakim Roslan, Mohd Firdaus Md@Danazlan, and Abd Wahid Rasi contributed to data analysis and interpretation. Radzuan Saari, Mushairry Mustaffar, and Azman Kassim contributed to methodology and validation. Rini Asnida Abdullah, Khairul Hazman Padil, Norbazlan Mohd Yusof, and Norisam Abd Rahman contributed to manuscript preparation and review. All authors discussed the results and contributed to the final manuscript.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgement

The authors would like to thank PLUS Sdn Bhd for their support in providing the data and the Universiti Teknologi Malaysia supported this work under UTM Flagship CoE/RG–Coe/RG 5.2: Evaluating Surface PGA with Global Ground Motion Site Response Analyses for the highest seismic activity location in Peninsular Malaysia (Q.J130000.5022.10G47) and Universiti Teknologi Malaysia - Earthquake Hazard Assessment in Peninsular Malaysia Using Probabilistic Seismic Hazard Analysis (PSHA) Method (Q.J130000.21A2.06E9).

References

Abdul Rahim AF, Simon N, Mohamed TR, Md Rafek AG, Serasa AS, Chen YL, Zhang MW, Lee KE, Goh TL. 2019. Probabilistic analysis of potential planar-type rock slope failure of selected Malaysian rock slopes. Bulletin of the Geological Society of Malaysia, 67, 83–90. doi: 10.7186/bgsm67201910.

Ali M. 2020. Pycaret: An open source, low-code machine learning library in Python. Pycaret version, 2.

Amiri ST, Dehghanbanadaki A, Nazir R, Motamedi S. 2020. Unit composite friction coefficient of model pile floated in Kaolin clay reinforced by recycled crushed glass under uplift loading. Transportation Geotechnics, 22, 100313. doi: 10.1016/j.trgeo.2019.100313.

Bhandari AK, Kumar A, Singh GK. 2012. Feature extraction using normalized difference vegetation index (NDVI): A case study of Jabalpur City. Procedia Technology, 6, 612–621. doi: 10.1016/j.protcy.2012.10.074.

Bhattacharya S, Mishra S. 2018. Applications of machine learning for facies and fracture prediction using Bayesian Network Theory and Random Forest: Case studies from the Appalachian basin, USA. Journal of Petroleum Science and Engineering, 170, 1005–1017. doi: 10.1016/j.petrol.2018.06.075.

Bisong E. 2019. Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners. Berkeley, CA, Apress. doi: 10.1007/978-1-4842-4470-8.

Breiman L. 2001. Random forests. Machine Learning, 45(1), 5–32. doi: 10.1023/A:1010933404324.

Bui XN, Nguyen H, Choi Y, Nguyen-Thoi T, Zhou J, Dou J. 2020. Prediction of slope failure in open-pit mines using a novel hybrid artificial intelligence model based on decision tree and evolution algorithm. Scientific Reports, 10(1), 9939. doi: 10.1038/s41598-020-66904-y.

Burton CK. 1970. The palaeotectonic status of the Malay Peninsula. Palaeogeography, Palaeoclimatology, Palaeoecology, 7(1), 51–60. doi: 10.1016/0031-0182(70)90026-X.

Corcoran J, Knight J, Gallant A. 2013. Influence of multi-source and multi-temporal remotely sensed and ancillary data on the accuracy of random forest classification of wetlands in northern Minnesota. Remote Sensing, 5(7), 3212–3238. doi: 10.3390/rs5073212.

Dai C, Li WL, Wang D, Lu HY, Xu Q, Jian J. 2021. Active landslide detection based on sentinel-1 data and InSAR technology in Zhouqu County, Gansu Province, northwest China. Journal of Earth Science, 32(5), 1092–1103. doi: 10.1007/s12583-020-1380-0.

Dehghanbanadaki A. 2021. Intelligent modelling and design of soft soil improved with floating column-like elements as a road subgrade. Transportation Geotechnics, 26, 100428. doi: 10.1016/j.trgeo.2020.100428.

Department of Mineral and Geoscience 2012. Geological Map of Peninsular Malaysia: Modified based on the 8th edition, 1985. Director-General of Minerals and Geoscience Malaysia.

Du PJ, Samat A, Waske B, Liu SC, Li ZH. 2015. Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features. ISPRS Journal of Photogrammetry and Remote Sensing, 105, 38–53. doi: 10.1016/j.isprsjprs.2015.03.002.

Feng RH, Grana D, Balling N. 2021. Imputation of missing well log data by random forest and its uncertainty analysis. Computers & Geosciences, 152, 104763. doi: 10.1016/j.cageo.2021.104763.

Fouedjio F. 2020. Exact conditioning of regression random forest for spatial prediction. Artificial Intelligence in Geosciences, 1, 11–23. doi: 10.1016/j.aiig.2021.01.001.

Ghani A A. 2009. Plutonism, In: Hutchison CS, Tan DNK (Eds.), Geology of Peninsular Malaysia. University of Malaya/Geological Society of Malaysia, Kuala Lumpur, 211–231.

Ghani AA, Searle M, Robb L, Chung SL. 2013. Transitional I S type characteristic in the main range granite, peninsular Malaysia. Journal of Asian Earth Sciences, 76, 225–240. doi: 10.1016/j.jseaes.2013.05.013.

Gómez H, Kavzoglu T. 2005. Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Engineering Geology, 78(1–2), 11–27. doi: 10.1016/j.enggeo.2004.10.004.

He XL, Xu C, Qi WW, Huang YD, Cheng J, Xu XW, Yao Q, Lu YK, Dai BY. 2021. Landslides triggered by the 2020 Qiaojia Mw5.1 earthquake, Yunnan, China: Distribution, influence factors and tectonic significance. Journal of Earth Science, 32(5), 1056–1068. doi: 10.1007/s12583-021-1492-1.

Hutchison CS, Tan DNK (Eds.). 2009. Geology of peninsular Malaysia. Published jointly by the University of Malaya and the Geological Society of Malaysia.

Iranzad R, Liu X. 2024. A review of random forest-based feature selection methods for data science education and applications. International Journal of Data Science and Analytics, 1–15. doi: 10.1007/s41060-024-00509-w.

ISRM. 1978. Suggested methods for the quantitative description of discontinuities in rock masses. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 15(6), 319–368. https://doi.org/10.1016/0148-9062(78)91472-9.

Jasin B. 1997. Permo-Triassic radiolaria from the Semanggolformation, northwest Peninsular Malaysia. Journal of Asian EarthSciences, 15(1), 43−53. doi: 10.1016/s1367-9120(96)00079-x.

Ji L, Zhang L, Wylie B. 2009. Analysis of dynamic thresholds for the normalized difference water index. Photogrammetric Engineering & Remote Sensing, 75(11), 1307–1317. doi: 10.14358/pers.75.11.1307.

Kang F, Xu B, Li JJ, Zhao SZ. 2017. Slope stability evaluation using Gaussian processes with various covariance functions. Applied Soft Computing, 60, 387–396. doi: 10.1016/j.asoc.2017.07.011.

Khan, AF. Rahman MT. 2022. Deep learning approach for landslide susceptibility mapping using remote sensing and GIS: A case study of northern Pakistan. Geomatics, Natural Hazards and Risk, 13(1), 1537–1560. doi: 10.1080/19475705.2022.2094789

Lin S, Liang ZL, Zhao SX, Dong M, Guo HW, Zheng H. 2024. A comprehensive evaluation of ensemble machine learning in geotechnical stability analysis and explainability. International Journal of Mechanics and Materials in Design, 20(2), 331–352. doi: 10.1007/s10999-023-09679-0.

Lu P, Rosenbaum MS. 2003. Artificial neural networks and grey systems for the prediction of slope stability. Natural Hazards, 30(3), 383–398. doi: 10.1023/B:NHAZ.0000007168.00673.27.

Lu SL, Wu BF, Yan NN, Wang H. 2011. Water body mapping method with HJ-1A/B satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 13(3), 428–434. doi: 10.1016/j.jag.2010.09.006.

McFeeters S. 2013. Using the normalized difference water index (NDWI) within a geographic information system to detect swimming pools for mosquito abatement: A practical approach. Remote Sensing, 5(7), 3544–3561. doi: 10.3390/rs5073544.

McKay G, Harris JR. 2016. Comparison of the data-driven random forests model and a knowledge-driven method for mineral prospectivity mapping: A case study for gold deposits around the huritz group and nueltin suite, Nunavut, Canada. Natural Resources Research, 25(2), 125–143. doi: 10.1007/s11053-015-9274-z.

Mohamed T, Kasa A, Taha MR. 2012. Fuzzy logic system for slope stability prediction. International Journal on Advanced Science, Engineering and Information Technology, 2(2), 151. doi: 10.18517/ijaseit.2.2.174.

Nanehkaran YA, Zhu LC, Jin CY, Chen JD, Anwar S, Azarafza M, Derakhshani R. 2023. Comparative analysis for slope stability by using machine learning methods. Applied Sciences, 13(3), 1555. doi: 10.3390/app13031555.

Paudel U, Oguchi T, Hayakawa Y. 2016. Multi-resolution landslide susceptibility analysis using a DEM and random forest. International Journal of Geosciences, 7(5), 726–743. doi: 10.4236/ijg.2016.75056.

Qu JK, Qi SC, Zhou JW. 2023. A three-dimensional limit equilibrium method for slope stability analysis integrating machine learning optimized intercolumn force function. Geological Journal, 58(6), 2438–2453. doi: 10.1002/gj.4770.

Reading AM, Cracknell MJ, Bombardieri D, Chalke T. 2015. Combining machine learning and geophysical inversion for applied geophysics. In ASEG–PESA 24th International Geophysical Conference and Exhibition, Extended Abstracts (ASEG Extended Abstracts, Vol. 2015, p. ab070). Perth, Western Australia. doi: 10.1071/ASEG2015ab070

Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, Chica-Rivas M. 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geology Reviews, 71, 804–818. doi: 10.1016/j.oregeorev.2015.01.001.

Sahin E. 2020. A comparative study of machine learning algorithms for landslide susceptibility mapping: A case study from Turkey. Environmental Earth Sciences, 79(5), 1–17. doi: 10.1007/s12665-020-08847-1.

Salman HA, Kalakech A, Steiti A. 2024. Random forest algorithm overview. Babylonian Journal of Machine Learning, 2024, 69–79. doi: 10.58496/bjml/2024/007.

Singh P, Bardhan A, Han FC, Samui P, Zhang WG. 2023. A critical review of conventional and soft computing methods for slope stability analysis. Modeling Earth Systems and Environment, 9(1), 1–17. doi: 10.1007/s40808-022-01489-1.

Shu YK, 1969. Some NW trending faults in Kuala Lumpur and other areas. Newsletter Geological Society of Malaysia, 17, 1–5.

Stauffer, P. H. 1968. The Kuala Lumpur fault zone: A proposed major strike-slip fault across Malaya. Newsletter Geological Society of Malaysia, 15, 2–4.

Tsangaratos P, Benardos A. 2016. Applying artificial neural networks in slope stability related phenomena. Bulletin of the Geological Society of Greece, 47(4), 1901. doi: 10.12681/bgsg.10945.

Wang GJ, Zhao B, Wu BS, Zhang C, Liu WL. 2023. Intelligent prediction of slope stability based on visual exploratory data analysis of 77 *in situ* cases. International Journal of Mining Science and Technology, 33(1), 47–59. doi: 10.1016/j.ijmst.2022.07.002.

Wen GZ, Han LL, Lin W, Xing Z, Yan MZ. 2023. Machine Learning Algorithms. Application of Machine Learning in Slope Stability Assessment, 13−29.

Wu HS, Chen YL, Lv HY, Xie QH, Chen YG, Gu J. 2022. Stability analysis of rib pillars in highwall mining under dynamic and static loads in open-pit coal mine. International Journal of Coal Science & Technology, 9(1), 38. doi: 10.1007/s40789-022-00504-1.

Xie HY, Dong JJ, Deng Y, Dai YW. 2022. Prediction model of the slope angle of rocky slope stability based on random forest algorithm. Mathematical Problems in Engineering, 2022, 9441411. doi: 10.1155/2022/9441411.

Ye J, Liu Z, Wang Y, Zhang L. 2022. Slope stability prediction using improved random forest and SHAP analysis: A case study in Southwest China. Engineering Geology, 302, 106618. doi: 10.1016/j.enggeo.2022.106618.

Zhang WG, Gu X, Hong L, Han L, Wang L. 2023. Comprehensive review of machine learning in geotechnical reliability analysis: Algorithms, applications and further challenges. Applied Soft Computing, 136, 110066. doi: 10.1016/j.asoc.2023.110066.

Zhu JF, Pierskalla WP. 2016. Applying a weighted random forests method to extract Karst sinkholes from LiDAR data. Journal of Hydrology, 533, 343–352. doi: 10.1016/j.jhydrol.2015.12.012.