PAPER • OPEN ACCESS

Assessment on Recent Landslide Susceptibility Mapping Methods: A Review

IOP Conf. Series: Earth and Environmental Science **971** (2022) 012032 doi:10.1088/1755-1315/971/1/012032

Assessment on Recent Landslide Susceptibility Mapping Methods: A Review

W A A Manan[1*], A S A Rashid[1], M Z A Abdul Rahman[2] and M F A Khanan[2]
[1] Department of Geotechnics and Transportation, School of Civil Engineering, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia;
[2] Department of Geoinformatics, Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, Johor Bahru 81310, Malaysia
*Corresponding author: wardatunahmar@graduate.utm.my

Abstract. Landslide is a destructive natural hazard that causes severe property loss and loss of lives. Numerous researchers have developed landslide susceptibility maps in order to forecast its occurrence, particularly in hill-site development. Various quantitative approaches are used in landslide susceptibility map production, which can be classified into three categories; statistical data mining, machine learning and deterministic approach. In this paper, we choose two regular models in each category, which are Weight of Evidence (WoE) and Frequency Ratio (FR), Artificial Neutral Networks (ANN) and Support Vector Machines (SVM), Shallow Landsliding Stability Model (SHALSTAB) and YonSei-Slope (YS-Slope). Discussion and assessment on these models are based on relevant literature.

1. Introduction

Landslide is a natural disaster that causes extensive property damages and occasionally results in loss of life, especially in a hilly area. World Health Organisation (WHO) reported that landslides affected an estimated 4.8 million people and caused more than 18 000 deaths between 1998-2017. While in the United States, Japan, Italy, and India estimated the annual losses due to landslides disaster, at least USD1 billion for each country [1]. Due to its complexity, uncontrollability, and destructiveness, landslide poses a formidable challenge to us. In view of this, numerous studies have been conducted to forecast the occurrence of landslides in a specific location. Landslide prediction and mitigation techniques are essential for minimising loss of life and property.

Landslide susceptibility mapping (LSM) is an essential and widely used procedure for the spatial prediction of landslides in many countries. Typically, it is used to map a region's landslide hazard, and it can be classified as qualitative or quantitative [2]. [3], propose that a landslide susceptibility map can be performed either qualitatively or quantitatively depending on the availability and accuracy of data, required outcome and scale of analysis. The qualitative method is a relatively subjective approach that demonstrates landslide's susceptibility levels through descriptive expressions based on expert judgement ([4]- [5].

On the other hand, quantitative models quantify the relationship between slope instability and other controlling factors numerically. Generally, the quantitative analyses can be divided into statistical data mining and deterministic approach [4], which are frequently used in previous landslide susceptibility studies [6]-[7]. The deterministic approach mainly measures the factors of safety, while the statistical method mainly focuses on the correlation of data between the historical landslide distribution and the controlling parameters [8]. In the last three decades, machine learning approaches have been applied in various fields and landslides susceptibility analysis due to their superior performance and ability to deal with complex and varying data.

IOP Conf. Series: Earth and Environmental Science **971** (2022) 012032 doi:10.1088/1755-1315/971/1/012032

Therefore, this article comprehensively reviews statistical data mining, deterministic and machine learning approaches. Two representative models from each category which Weight of Evidence (WoE), Frequency Ratio (FR), Artificial Neutral Networks (ANN), Support Vector Machines (SVM), Shallow Landsliding Stability Model (SHALSTAB) and Stability Index Mapping (SINMAP) were discussed in details to analyse and compare the predictive power of landslide susceptibility analysis.

2. Landslide susceptibility mapping (LSM)

Susceptibility to landslides is a term that refers to the likelihood or probability of a landslide occurring in a given area. Thus, landslide susceptibility maps have become a popular method of illustrating where future landslides are likely to occur based on landslide factors. Landslide susceptibility can be assessed in various ways, depending on the available data [2]. However, a landslide susceptibility map can't be produced with limited data, but the most necessary depending on the quality of the data [9] [10].

Based on the previous case studies, identified factors that influence the landslide include slope angle, curvature, aspect, soil type, geology, distance to a river, and drainage. Numerous researchers have agreed that the slope angle and aspect are the most critical variables in landslide spatial analysis ([6], [11]-[12]. The possible explanation for this is that the slope angle is the most significant slope stability analysis parameter with direct effects on shear strength [13].

Susceptibility is a term that refers to the degree to which future slope movements are able to influence terrain [14]. [4] have suggested that it is necessary to answer basic questions such as where what, and how the landslide occurred to define the landslide susceptibility in the studied area. The landslide susceptibility map classifies slope conditions as stable or unstable, indicating the likelihood of a landslide. Besides, it can also be considered an expression of relative hazard, total landslide density of likely frequency [15].

3. The statistical data mining approach

Statistical data mining techniques are applied in LSM to reduce subjectivity in using qualitative approaches. The statistical approach uses GIS tools to assess the spatial distribution of existing landslides with the spatial distribution of various causative factors [4]. Generally, it can be classified into two types: bivariate and multivariate. This article reviews the two models in bivariate analysis: weight of evidence (WoE) and Frequency Ratio (FR).

Bivariate analysis analyses two variables that determine the correlation between a dependent variable and an independent variable [1]. Various types of bivariate models are most commonly applied in previous studies to produce landslide susceptibility mappings. They are Frequency Ratio (FR), Fuzzy Logic (FL), Weight of Evidence (WoE), Statistical Index (Si), Weighted Overlay Model (WOM) and Weighting Factor (Wf). [16]-[17]. The area under the curve (AUC) validates a model's accuracy throughout the model development. Devkota [18] have classified AUC value in four conditions; less than 0.5 is considered as no ability to evaluate, between 0.7 and 0.8 indicates acceptable interpretability, between 0.8 and 0.9 is excellent, and outstanding if it is more than 1.0. AUROC curves are among the most widely used tools in models for evaluating and assessing landslide susceptibility according to the works of literature [19]. Based on Chen *et al.* [20], the AUC can be calculated as in the below equation (1).

$$AUC = \frac{(\sum TP + \sum TN)}{(P + N)} \tag{1}$$

Where P is the total number of landslides, N is the total number of non-landslides, TP is the true positive and TN is the true negative.

3.1 *Weight of Evidence (WoE)*

According to Bayes' theorem, the basis of Weight of Evidence (WoE) is conditional probability. WoE method has been applied to identify the spatial association between the location of a landslide and a set of causal factors [21]. Most of the previous studies [16] [17][22] [23] have been conducted using WoE by focussing on the theoretical background and application. Weight of Evidence (WoE) is derived from the density of observed landslide locations in each class of parameters, referred to as ‘‘evidence," such as slope, lithology, and aspect [24].

The spatial correlation between the causal factor's class and the occurrence of a landslide can be determined using the magnitude of the weight contrast. The landslide susceptibility index (LSI) is calculated by adding the contrasts of all the parameters factors. The greater the landslide susceptibility index values are, the higher the probability of landslide to occur.

3.2 *Frequency Ratio Model*

The frequency ratio (FR) approach has been utilised to identify the correlation between landslide distribution prone areas and each causal factor that induces landslides [22]. According to Pradhan *et al.* [25], the following formula in equation (2) has been used to compute the FR value based on the correlation between the variables equation.

$$FR = P_{LO}/P_{LF} \tag{2}$$

P_{LO} is the percentage of landslide-prone area in each factor subcategory and P_{LF} is the percentage of each factor subcategory that influences the landslide.

The landslide susceptibility index (LSI) has been calculated using the following equation (3) by adding all of the FR values. According to [23], the higher the LSI value is, produced, the greater the tendency will be for a landslide to occur.

$$LSI = \sum_{i=1}^{n} FR_i \tag{3}$$

Where n value is refer to the number of influencing factors.

4. Machine learning approach

Machine learning (ML) is the computer programming that used past experience or samples to optimise the performance criterion by learning from the sample [26]. Numerous studies in the different geotechnical fields had used machine learning as an effectively geotechnical application tool in their study [27]–[37]. Machine learning has also gained popularity in producing landslide susceptibility mapping (LSM) as it has evolved, and various types of ML algorithms have been used to produce LSM [6] [50]-[75].

LSM that produced using ML used the main concept: the correlation between a set of landslide conditioning factors and a past data set of the landslide event were evaluated based on the ML algorithm. This was done to determine their spatial connection to predict the potential landslide event in the study area. B. Pradhan *et al.* [65] has clarified that spatial prediction of a landslide is more efficient to be conducted using ML approach than other methods, for instance, analytic method or expert’s opinion method. This article then presents further discussions on two types of ML frequently used in producing LSM. They are Artificial Neural Network (ANN) and Support Vector Machines (SVM).

4.1 *Artificial Neuro Network*

The artificial neuron network (ANN) is a well-organised mathematical model that is widely used in many fields and this model was created based on the mechanism of the human brain to solve the problem or replicate the human learning process represented by computerised programming [66]–[68]. The ANN has been used as modelling of prediction method in the geotechnical area, particularly in landslide study, to develop good quality LSM or landslide hazard maps (LHM) [80][81]. The ANN has been effectively used to assess landslide susceptibly in many studies. The selection of the ANN as the assessment tool for landslide susceptibility is due to the ability of this model to learn and analyse a set of landslide causal factors based on a non-linear numerical model [77] [81] [82].

Furthermore, ANN can work on the problem that the statical method is unable to resolve due to the limitation of the theory or data [68]. Besides, the ANN model does not require a specific parameter to develop and assess the landslide susceptibility since it is independent and able to manage vague and uncertain data, making this model suitable to be performed using continuous, categorical and binary data without breaching any assumptions [79] [82].

There are various types of ANN, and this article only focuses on the feed-forward multilayer perceptron's (MLP) with backpropagation learning algorithms as most previous literature used this type of ANN modelling. The structure of the ANN-MLP model consists of three main components: the input, hidden, and output layer. The neuron in the input layer is only connected to the hidden layer as well as the neuron in the hidden layer is connected to the output layer, which means that interconnection of the neuron to others layer does not exist [72] [73] [74].

Figure 1 shows an example of a simple ANN-MLP architecture design that is commonly used in LSM. The development of ANN model is commonly assessed with the area under the ROC curve to determine its performance and efficiency. Hosmer and Lemeshow [75] describe the criteria for rejecting or accepting the area under the curve value (AUC) for the assessed model. If the AUC of the model assessment is less than 0.50, it is considered as fail, between 0.70 and 0.80 is acceptable, between 0.80 and 0.90 is excellent, and between 0.90 and 1.00 is outstanding.

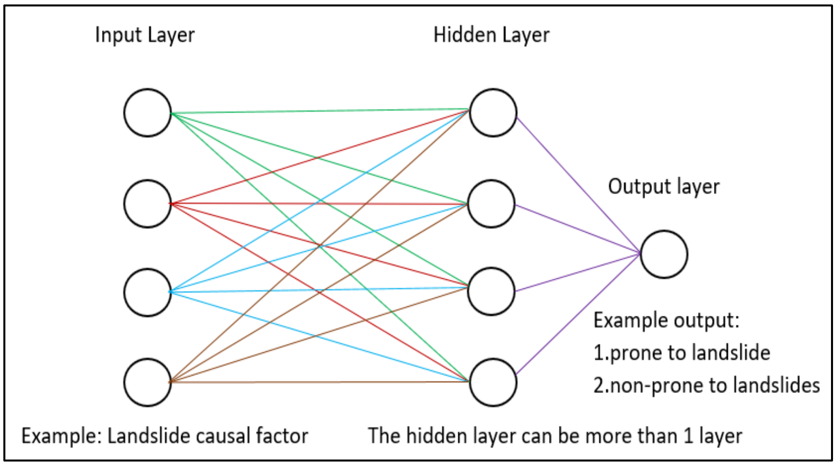

Figure 1. ANN-MLP architecture design.

4.2 *Support Vector Machines (SVM)*

Many studies have been conducted to examine the performance of SVM to produce LSMs either as a single method, hybrid or to compare it with other approaches [66]–[69] [76-89]. Saro *et al.* [90] claim that SVM produces high performance with a minimum setting requirement, making it one of the common ML approaches used to produce LSM. While SVM is categorised as a supervised machine learning method that complements the structural risk minimisation principle [82] [90].

Nuriah *et al.* [78] highlight that SVM can search for an optimal separating hyperplane to map the dimensional future space based on original input space. In other words, the goal of the SVM model during developing LSM is to pursue the optimal separating hyperplane that can differentiate between landslide susceptible area and landslide non-susceptible area. Pradhan and Lee [46] remark that the SVM model is developed based on two key principles: the optimal classification hyperplane and the use of a kernel function.

The basic principle of SVM can be described in figure 2. Dots and squares represent two classes of different data. H1 and H2 are located parallel to the H line, which is defined as their classification line. Those lines move over the data points closest to the classification line, which is the support vectors that create a gap between that two-line known as classification margin [80]. For a given hyperplane, the margin is considered to have the shortest distance from each class's data vector, indicating that the accuracy of classification by the SVM model increases toward the increase of the classification margin [41]. The kernel function role in SVM is to facilitate the input data transformation to the higher dimensional space for it to be linearly classified, as being illustrated in figure 2(b). Thus, the type of kernel function used in SVM has a major influence on the SVM performance, and its selection is crucial [81].

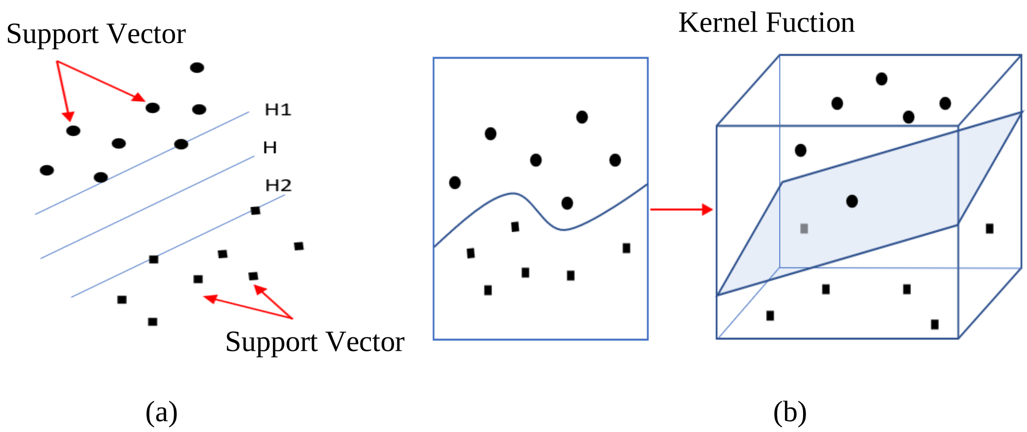

Figure 2. The principle of SVM (a) Dimensional hyperplane differentiating the two classes by the maximum gap; (b) non-separable case and the slack variables [76]

4.3 *Performance of ANN and SVM*

Many previous studies have used ANN as the sole model to produce LSMs. Yao *et al.* [65] conducted a study in Vaz Watershed, Iran, to evaluate the performance of the ANN model with the two separate algorithms, which was in Multi-Layered Perceptron (MLP) and Radial Base Function ((RBF) in producing LSMs. This study was conducted based on nine causal factors and 136 landslide inventory data. The study showed that both MLP and RBF produced an excellent performance, which was more than 80 per cent, but MLP performance was slightly higher in training and validating datasets. Another study by Lo and Leung [39] used only 34 landslides inventory data with 5 causal factors to produce LSMs using MLP and Probabilistic Neural Network (PNN). Again, in this study, the MLP model showed a higher performance compare to PNN. However, the model's performance in this study was slightly poor but still accepted as the MLP and PNN produce 73 per cent and 68 per cent validation performance, respectively. The reduction in the performance of the ANN-MLP model by Lo and Leung [39] compared to Yao *et al.* [65] may reflect the influence landslide causative factor that was used. Moreover, the study by Lo and Luang [39] validated only with past landslide inventory data, while the study by Zare et al. (2013) used the area under curve (UAC) method to validate the model.

Besides, Ortiz and Martínez-Graña (2018) conducted a study at Capitanejo, Colombia, choosing 14 landslide causal factors and 54 landslide inventory data using MLP-ANN model to develop LSMs. This study found that the MLP-ANN model produced an outstanding LSMs performance as the AUC value for tanning and the validation data set were 98.8 per cent and 92.86 per cent. Other researchers [58] [50] [71] also used MLP-ANN model to develop LSMs in the various study area and found that the model produced an excellent performance as the assessment of the LSMs using the AUC method produced a result of more than 80 per cent both for training and validation data set. However, the study by Lee *et al.* [82] and Hong *et al.* [70] used the same approach to produce a slightly lower performance, which was between 71 and 76 per cent. This may be due to the different validation methods used in this study, which may affect their performance. The LSMs provided by Sharir *et al.*[82] only validated the data set through ANN modelling, while Hong *et al.* [70] validated using the scoring wizard method.

Su *et al.* [83] used SVM as the only model in their study at Wencheng, China. In his study, 354 landslide data was used, with 7 landslide causative and 2 triggering factors being considered. The uniqueness of this study was that the data set of this model did not divide training and validation set as in the usual practice in the ML method. As an alternative, fivefold cross-validation was used to remove the impact of the production on the training dataset. This study found that the LSMs produced an excellent performance with an AUC value of 96 per cent.

Another study by Saro *et al.* [90] also used SVM only to produce LSM but in two different study areas: Pyeong Chang and Inje areas of Gangwon Province, Korea. The selection of landslide causative factor remained the same for both areas except slope length only considered at Inje while land used only considered at Pyeong Chang. The result showed that the LSMs for the Peong Chang area had slightly higher performance than Inje, with the AUC value of 81.36 per cent and 77.39 per cent respectively. The study area may be subject to the change in a spatial distribution corresponding to an event in that area, leading to slight differences in the model performance. The sensitivity analysis was conducted in this study to determine the high impact causative factor to the LSMs. The study found that factors that related to the topography of the area such as SPI, TWI, slope, aspect, and slope length had a strong influence on the LSMs performance.

W. Chen *et al.* [89] and Feizizadeh *et al.* [95] conducted a study to compare the performance of SVM with the four different kernel functions which were LN,PL, RBF and SIG to produce LSMs. The result of both studies showed that the SVM-SIG model had the lowest performance and SVM-PL had the second-highest performance for both studies. The main difference between those two studies was the SVM-RBF produced the most outstanding performance of AUC value in a study by W. Chen *et al.* [89] while in the study Feizizadeh *et al.* [95] the SVM-LN had the highest performance. The difference in the model performance in those studies may be due to the difference in a causative factor employed in the study, which may affect the effectiveness of four different kernel functions of SVM in producing LSMs. Almost all the SVM models that are discussed above have produced an AUC value of more than 80 per cent which indicates that the SVM model has a good performance in producing LSMs.

5. Deterministic approach

A deterministic approach to landslide susceptibility mapping can be identified in geotechnical fields. The deterministic approach includes the interaction between hydrology, topography, soil properties, and in some cases, vegetation to understand and predict the location and timing of occurrences of landslides. Mostly, models generated from the deterministic approach generally consist of a methodology that combines the hydrological model for analysing the pore-water pressure and the infinite slope-stability model for computing its safety factors such as Stability Index Mapping, SINMAP [85], Transient Rainfall Infiltration and Grid-based Regional Slope-stability Model, TRIGRS [86], Shallow Slope Stability Model, SHALSTAB [87] and YonSei-Slope, YS-Slope Model [88].

The fundamental of these models is the combination of hydrological and geotechnical approaches based on slope stability analysis. The discrepancies between models are mostly due to the different models used in failure criterion interpretation and hydro-geotechnical consideration. The YS-Slope and TRIGRS models, for example, estimated the Factor of Safety (FS) in unsaturated soil conditions, but SHALSTAB and SINMAP used saturated soil conditions. Although both the YS-Slope and the TRIGRS incorporate water infiltration into the subsurface, the YS-Slope Model is more advanced in addressing groundwater recharge conditions, which can result in deep seated failure in geotechnical assessments.

5.1 *Stability index mapping (SINMAP)*

SINMAP was established by Pack *et al.* [85] to provide terrain stability mapping tools that could complement stability mapping methods currently that were practised in Columbia at the time. Theoretically, this method is applicable for a shallow translational landslide. This model improved the stability mapping methods using Grid-based DEMs configured free extension to a GIS platform so that analysis is easy to use and widely available. In this model, some parameters are derived from the Digital

IOP Conf. Series: Earth and Environmental Science **971** (2022) 012032 doi:10.1088/1755-1315/971/1/012032

Elevation Model (DEM). These can be slope angle, flow direction, and catchment area. In contrast, the remaining parameter, especially geotechnical data, needs to be set manually or it can straight away use the default values proposed by the authors. This model is decent in which the cohesion value either for soil or root strength is considered in the infinite slope stability model and can be set to zero for the cohesionless situation. Simultaneously, this model incorporates parameter uncertainty through uniform probability distributions with lower and upper bounds on uncertain parameters.

The SINMAP model calculates the stability index (SI), which refer to the probability of a specific location in terms of stability, assuming uniform distributions of the parameters over these uncertainty ranges in each grid cell of the map. The value ranges between 0 (most unstable) and 1 (least unstable). Table 1 and 2 shows the range of parameters and SI based on stability classes. The SI is a dimensionless form of the infinite slope stability model, which calculates the factor of safety assuming the wet and dry density is equally based on Hammond *et al.* [89]. :

$$\mathrm{SI} = \frac{C_\mathrm{r} + C_\mathrm{s} + \cos^2\theta[\rho_\mathrm{s} g(D - D_\mathrm{w}) + (\rho_\mathrm{s} g - \rho_\mathrm{w} g)D_\mathrm{w}]\tan\phi}{D\rho_\mathrm{s} g \sin\theta\cos\theta}, \tag{4}$$

Where Cr is root cohesion, Cs is soil cohesion, h is slope angle, qs is the wet soil density, qw is the density of water, D is the vertical soil depth, Dw is the vertical height of the water table within the soil layer, and / is the internal friction angle of the soil.

Table 1. Parameters and ranges.

Parameters	Range
Transmissivity/effective recharge	2000 to 3000 m
Dimensionless cohesion	0 to 0.25
Internal friction angle	30 to 45°
Soil density.	2000 kg/m^3

Table 2. Stability Index (SI) based on stability classes.

Condition	Class	Prediction slope Zone
SI > 1.5	1	Stable
1.5 > SI > 1.25	2	Moderate
1.25 > SI > 1.0	3	Quasy-stable
1.0 > SI > 0.5	4	Lower threshold
0.5 > SI > 0.0	5	Upper threshold
0.0 > SI	6	Defended

5.2 *Yonsei Slope Model, YS-Slope*

YS-Slope model is a GIS-based physical landslide prediction model developed by Kim *et. al.* [88]. This model considers rainfall infiltration depth, recharge, and groundwater flows, whereas the SINMAP model did not include this component in their model. This model was developed using GIS-based raster data and input parameters. It consists of internal and external parameters related to the slope's stability, as shown in Figure 3. This model improved on the hydro-geotechnical approach by incorporating unsaturated soil parameters such as the soil-water characteristic curve and matric suction, as well as internal friction angle and cohesion value is obtained from soil strength properties. The infinite slope failure model was improved in order to compute the factor of safety (FS) of rainfall-induced landslides by taking into account the other influential component in increasing soil strength and interception loss due to plant growth, as shown in the equation below:

$$FS = \frac{(c_s' + c_r') + [(\gamma_{sat} - \gamma_w)(Z_w + D_w) + \gamma_t \cdot (D - Z_w - D_w) + q_0]cos^2\beta \cdot \tan \emptyset'}{[\gamma_{sat} \cdot (Z_w + D_w) + \gamma_t \cdot (D - Z_w - D_w) + q_0] \cdot \sin\beta \cdot \cos\beta} \quad (5)$$

Where c_s' is soil cohesion, c_r' is the shear strength increase by root reinforcement, γ_{sat} is the saturated unit weight of soils, γ_w is the unit weight of water, γ_t is the total unit weight of soils, q_0 is the forest tree load, Z_w is wetting front depth, D_w is groundwater table from the bedrock, D is the thickness of dry soil, $\emptyset'$ is internal friction angle of soil and β is slope inclination.

The modified Green-Ampt model by Mein and Larson [91] was used in hydrological evaluation to analyse rainfall infiltration, and the wetting front depth was estimated based on this analysis. The cumulative infiltration and runoff are computed in transient rainfall-infiltration analysis by comparing the infiltration capacity with the rainfall intensity.

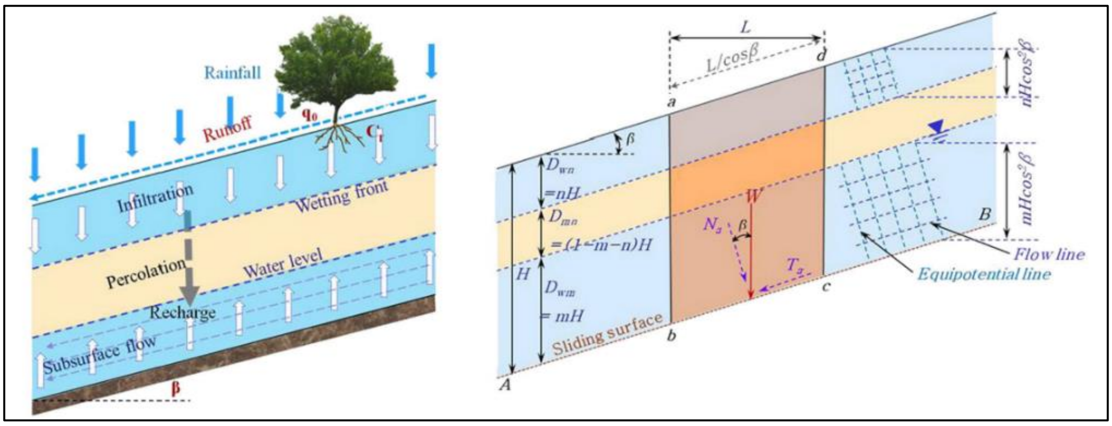

Figure 3. Infinite slope plane for slope stability analysis in YS-Slope model.

5.3 *Performance of SINMAP and YS Slope*

The YS-Slope model has purposed to analyse the rainfall-induced landslides in the spatial based predicting model. Jeong *et al.* [92] used this model to analyse rainfall-induced landslides and found that these procedures can be applied to predicting landslides and debris flows. furthermore, Hong *et al.* [93] propose a numerical method to simulate landslides and debris flows by combining rainfall infiltration analysis and the YS-Slope model in their study. While Samseong *et al.* [92] employed the YS-slope Model and prediction approach to developing a unified prediction method to examine the susceptibility map of landslides at Umyeonsan Mountain in Korea.

Similar to the SINMAP model, the YS-Slope model also can be used for shallow landslide studies. However, YS-Slope can be further used for deep-seated landslides prediction. Jeong *et al.* [92] prove that YS-slope can predict shallow landslides and deep-seated landslides and reasonably agree with historical landslides. Besides, the researcher confirmed with ROC analysis and showed that relatively high accuracy compared to other predictions.

Duc Long and Dung [94] used the YS-Slope model by incorporating the unsaturated soil properties to develop the LSM. The assessment of landslide susceptibility is based on the contributions related to spatial data such as elevation, slope angle and groundwater table spatial distribution map and input data corresponding to unsaturated permeability function, rainfall intensity, soil-water characteristic curve (SWCC). They found that incorporating the unsaturated soil characteristic becomes more realistic in predicting the landslide susceptibility area in terms of factor of safety.

Furthermore, Kim *et al.* [95] investigate the effectiveness of YS-Slope by coupling the effects of hydro-mechanical processes on the infiltration behaviour of unsaturated soils. A few sets of infiltration analyses for a soil column were carried out under varied soil conditions, and the results were compared using the software GEO-SLOPE [96]. The research reveals that the volume change of the soil influences the transient seepage analyses in deformable soils. It is relevant that whenever the matric suction on a slope changes due to rainfall infiltration, consequently changes the effective stress that directly relates to seepage processes. Therefore, the infiltration rate of the YS-Slope is slower than the Geo-slope due to permeability function based on the porosity that altered the matric suction distribution. Thus, the coupled effects of hydro-mechanical behaviour of soils considered in the YS-slope model has a positive effect on unsaturated soil stability, especially during rainfall.

6. Conclusion

Statistical data mining, machine learning, and deterministic approach are widely used in LSM. The statistical approach is the simplest way to compute LSM compared to machine learning and the deterministic approach. Machine learning needs knowledge of computer programming, while deterministic approaches are more toward needing geotechnical input. The advantages and disadvantages of each approach were listed in Table 3 and summaries of statistical, machine learning, and deterministic approaches were discussed in Table 4. Based on previous research works, each approach can yield accurate results in determining the landslide-prone area based on qualified data representing the condition of the study area, scale, and method of data acquisition. In other words, descriptive, reliable and easily accessible input parameters should be selected as the parameter used in a study. Due to this condition, it is necessary to map the landslide-prone areas by forecasting and specifying the region where reliable and accurate landslide susceptibility maps can be helpful for land planners, decision-makers, and risk assessment personnel.

Table 3. The advantages and disadvantages of each approach.

Type of Approach	Advantages	Disadvantages
Statistical Approach	• Simple and easily implemented in GIS using remote sensing inputs. • Minimum data required	• loss of data quality and accuracy with the oversimplification of thematic input data, and loss of data sensitivity in forced individual analysis of causative factors • basically in linear based, so they were inefficient in capturing the nonlinearity and complexity of landslide phenomena.
Machine Learning Approach	• Can map out the detailed occurrence of past landslides • To collect sufficient information on the variables that are the occurrence of landslides	• Difficult to prepare data and need a longer time. • Just susceptibility assessment • Not readily be extrapolated to the neighbouring area
Deterministic Approach	• Able to quantitatively calculate the safety factor. • Externally available models can be used instead of spending time implementing model algorithms in GIS. • Encourage the detailed investigation and analysis of geotechnical parameters.	• The data requirements can be prohibitive, and it is usually difficult to procure the necessary input data to be used in the models effectively.

Table 4. Summaries of statistical, machine learning, and deterministic approaches.

Weight of Evident (WoE) and Frequency Ratio (FR)	Artificial Neural Network (ANN) and Support Vector Machine (SVM)	SINMAP and YS-Slope
• Statistical approaches generally require the collection of a large amount of data. They are particularly useful for medium-scale (regional) mapping. • is used to validate a model's accuracy throughout the model development • WoE appears to be efficient but has limitations - WoE model assessment had a lower performance validation result.	• ANN and SVM can yield accurate results in determining landslide-prone • SVM model outperformed other models of ML in most of the study • unique features in SVM - more efficient such as robustness to noise, non-linear decision boundaries, easily implementable probabilistic outcome, and an inherent ability to deal with high dimensional classification problems.	• these models have covered a shallow translational landslide controlled by groundwater convergence • SINMAP model greatly limits the results, which should only be treated as a general susceptibility indicator, not a true safety factor (FS) value. • The failure criterion used to derive the FS equation in YS-Slope consider unsaturation condition, while SINMAP focus on the saturated condition

Acknowledgments
This research proposal was sponsored by the UTM High Impact Research Grant awarded by UTM R&D FUND Research (Development of UTM Landslide Susceptibility Map -Q.J130000.3651.03M35)

IOP Conf. Series: Earth and Environmental Science **971** (2022) 012032 doi:10.1088/1755-1315/971/1/012032

7. References

[1] H Gao, P S Fam, H C Low, L T Tay, and H Lateh 2019 An overview and comparison on recent landslide susceptibility mapping methods Disaster Adv. **12** 46–64

[2] F Guzzetti, A Carrara, M Cardinali and P Reichenbach 1999 Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy *Geomorphology* **31** 181-216

[3] M Ciurleo, L Cascini and M Calvello 2017 A comparison of statistical and deterministic methods for shallow landslide susceptibility zoning in clayey soils *Eng. Geol.* **223** 71–81

[4] P Aleotti and R Chowdhury 1999 Landslide hazard assessment: Summary review and new perspectives *Bull. Eng. Geol. Environ.* **58** 21–44

[5] M S Sulaiman, A Nazaruddin, N M Salleh, R Z Abidin, N D Miniandi and A H Yusoff 2019 Landslide occurrences in Malaysia based on soil series and lithology factors *Int. J. Adv. Sci. Technol.* **28** 1–26.

[6] A D Regmi *et al.* 2014 Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya *Arab. J. Geosci.* **7** 725–742

[7] L Borrelli, M Ciurleo and G Gullà 2018 Shallow landslide susceptibility assessment in granitic rocks using GIS-based statistical methods: the contribution of the weathering grade map *Landslides* **15** 1127–1142

[8] N Saadatkhah, A Kassim and L M. Lee 2014 Qualitative and quantitative landslide susceptibility assessments in Hulu Kelang area, Malaysia *Electron. J. Geotech. Eng.* **19** 545–563

[9] R J Fabbri, A G Chung and C F Cendrero A 2003 Is Prediction of Future Landslides Possible with a GIS? *Nat Hazards* **30** 487–499

[10] J Coe, J Godt, R Baum, R Bucknam and J Michael 2004 Landslide susceptibility from topography in Guatemala *Landslides Eval. Stab. Terrain Eval. Stabilisation* **1** 69–78

[11] M Teixeira, C Bateira, F Marques and B Vieira 2015 Physically based shallow translational landslide susceptibility analysis in Tibo catchment, NW of Portugal *Landslides*

[12] B C Vieira, N F Fernandes, O Augusto Filho, T D Martins and D R Montgomery 2018 Assessing shallow landslide hazards using the TRIGRS and SHALSTAB models, Serra do Mar, Brazil *Environ. Earth Sci*. **77** 260-274

[13] Q Q Wang, Y Guo, W Li, J He and Z Wu 2019 Predictive modeling of landslide hazards in Wen County, northwestern China based on information value, weights-of-evidence, and certainty factor *Geomatics, Nat. Hazards Risk* **10** 820–835

[14] P Reichenbach, M Rossi, B D Malamud, M Mihir, and F Guzzetti 2016 A review of statistically-based landslide susceptibility models *Earth-Science Rev.* **180** 60–91[15] J Chacón, C Irigaray, T Fernández and R El Hamdouni 2006 Engineering geology maps: Landslides and geographical information systems *Bull. Eng. Geol. Environ.* **65** 341–411

[16] V Vakhshoori and M Zare 2016 Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods *Geomatics, Nat. Hazards Risk* **7** 1731–1752

[17] S Pasang and P Kubíček 2020 Landslide susceptibility mapping using statistical methods along the asian highway, bhutan *Geosci.* **10** 1–26

[18] K C Devkota *2013* Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling-Narayanghat road section in Nepal Himalaya *Nat. Hazards* **65** 135–165

[19] J L Zêzere, S Pereira, R Melo, S C Oliveira and R A C Garcia2017 Mapping landslide susceptibility using data-driven methods *Sci. Total Environ.* **589** 250–267

[20] W Chen 2018 Landslide susceptibility modeling based on GIS and Novel Bagging-based Kernel logistic regression *Appl. Sci.* **8** 12

[21] T Kavzoglu, E K Sahin, and I Colkesen 2014 Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression *Landslides* **11** 425–439

[22] E Nohani 2019 Landslide susceptibility mapping using different GIS-Based bivariate models *Water (Switzerland)* **11** 1–22
[23] A Ozdemir and T Altural 2013 A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan mountains, SW Turkey *J. Asian Earth Sci.* **64** 180–197
[24] I Armaş 2012 Weights of evidence method for landslide susceptibility mapping. Prahova Subcarpathians, Romania *Nat. Hazards* **60** 937–950
[25] B Pradhan, L Saro, and M F Buchroithner 2010 A GIS-based back-propagation neural network model and its cross-application and validation for landslide susceptibility analyses *Comput. Environ. Urban Syst.* **34** 216–235
[26] R. T. Sataloff, M. M. Johns, and K. M. Kost 2010 Introduction To Machine learning, 2nd Edition *The MIT Press*
[27] M K Lo and Y F Leung 2019 Bayesian updating of subsurface spatial variability for improved prediction of braced excavation response *Can. Geotech. J.* **56** 1169–1183
[28] I. Papaioannou and D Straub 2017 Learning soil parameters and updating geotechnical reliability estimates under spatial variability – theory and application to shallow foundations *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **11** 116–128
[29] P Xiong 2021 Towards advancing the earthquake forecasting by machine learning of satellite data *Sci. Total Environ.* **771** 145256
[30] D J Lary, A H Alavi, A H Gandomi and A L Walker 2016 Machine learning in geosciences and remote sensing *Geosci. Front.* **7** 3–10
[31] W W Zhang, A T C Goh, Y Zhang, Y Chen, and Y Xiao 2015 Assessment of soil liquefaction based on capacity energy concept and multivariate adaptive regression splines *Eng. Geol.* **188** 29– 37
[32] X Liang, H He and Y Zhang 2019 Optimization design of micro-piles in landslide safety protection based on machine learning *Saf. Sci.* **118** 861–867
[33] M Zhang, M Li, J Zhang, L Liu, and H Li 2019 Onset detection of ultrasonic signals for the testing of concrete foundation piles by coupled continuous wavelet transform and machine learning algorithms *Adv. Eng. Informatics* **43** 10-15
[34] M Mittal, S C Satapathy, V Pal, B Agarwal, L M Goyal and P Parwekar 2021 Prediction of coefficient of consolidation in soil using machine learning techniques *Microprocess. Microsyst.* **82** 103830
[35] Y Bouslihim, A Rochdi, N El, and A Paaza 2021 Machine learning approaches for the prediction of soil aggregate stability *Heliyon* 7
[36] L Zhang, L Yang, T Ma, F Shen, Y Cai and C Zhou 2020 A self-training semi-supervised machine learning method for predictive mapping of soil classes with limited sample data *Geoderma*, **384** 114809
[37] R Jena 2020 Earthquake hazard and risk assessment using machine learning approaches at Palu, Indonesia *Sci. Total Environ.* **749** 141582
[38] S Lee, J H Ryu, J S Won and H J Park 2004 Determination and application of the weights for landslide susceptibility mapping using an artificial neural network *Eng. Geol.* **71** 289–302
[39] L Ermini, F Catani and N Casagli 2005 Artificial Neural Networks applied to landslide susceptibility assessment *Geomorphology* **66** 327-343
[40] A M S Pradhan and Y T Kim 2016 Landslide susceptibility mapping of Phewa catchment using multilayer perceptron artificial neural network *Nepal J. Environ. Sci.* **4** 1–9
[41] S Lee, M J Lee and H S Jung 2017 Data mining approaches for landslide susceptibility mapping in Umyeonsan, Seoul, South Korea *Appl. Sci.* 7
[42] S A Salleh, A S A Abd Rahman, A N Othman and W M N Wan Mohd 2018 Comparative study of landslides susceptibility mapping methods: Multi-Criteria Decision Making (MCDM) and Artificial Neural Network (ANN) *IOP Conf. Ser. Earth Environ. Sci.***117** 1–6
[43] L Yu, Y Cao, C Zhou, Y Wang and Z Huo 2019 Landslide susceptibility mapping combining information gain ratio and support vector machines: A case study fromWushan Segment in the Three Gorges Reservoir Area, China *Appl. Sci.* 9

[44] D H Lee, Y T Kim and S R Lee 2020 Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various non-linear activation functions *Remote Sens.*12

[45] W Chen, X Chen, J Peng, M Panahi and S Lee 2021 Landslide susceptibility modeling based on ANFIS with teaching-learning-based optimization and Satin bowerbird optimizer *Geosci. Front. **12*** 93–107

[46] X Yao, L G Tham and F C Dai 2008 Landslide susceptibility mapping based on Support Vector Machine: A case study on natural slopes of Hong Kong, China *Geomorphology* **101** 572–582

[47] M Marjanović, M Kovačević, B Bajat and V Voženílek 2011 Landslide susceptibility assessment using SVM machine learning algorithm *Eng. Geol.* **123** 225–234

[48] D Tien Bui, B Pradhan, O Lofman and I Revhaug 2012 Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and nave bayes models *Math. Probl. Eng.*

[49] B Pradhan 2013 A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS *Comput. Geosci.* **51** 350–365

[50] S. Lee 2007 Landslide susceptibility mapping using an artificial neural network in the Gangneung are, Korea *Int. J. Remote Sens.* **21** 4763–4783

[51] N Micheletti 2014 Machine Learning Feature Selection Methods for Landslide Susceptibility Mapping *Math. Geosci.* **46** 33–57

[52] H Hong, B Pradhan, C Xu and D Tien Bui 2015 Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines *Catena* **133** 266–281

[53] B T Pham, B Pradhan, D Tien Bui, I Prakash and M B Dholakia 2016 A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India) *Environ. Model. Softw.* **84** 240–250

[54] A Shirzadi 2017 A comparative study between popular statistical and machine learning methods for simulating volume of landslides *Catena* **157** 213–226

[55] H J Oh, P R Kadavi, C W Lee, and S Lee 2018 Evaluation of landslide susceptibility mapping by evidential belief function, logistic regression and support vector machine models *Geomatics, Nat. Hazards Risk* **9** 1053–1070

[56] O Ghorbanzadeh, T Blaschke, K Gholamnia, S R Meena, D Tiede and J Aryal 2019 Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection *Remote Sens.* 11

[57] B Peethambaran, R Anbalagan, D P Kanungo, A Goswami and K V Shihabudheen 2020 A comparative evaluation of supervised machine learning algorithms for township level landslide susceptibility zonation in parts of Indian Himalayas *Catena* **195** 104751

[58] B Pradhan and S Lee 2009 Landslide risk analysis using artificial neural network model focussing on different training sites *Int. J. Phys. Sci.* **4** 1–15

[59] B Pradhan and S Lee 2010 Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia *Landslides* **7** 13–30

[60] M. Khamehchiyana, P. Abdolmaleki, and B. Rakei 2011 Landslide susceptibility mapping using backpropagation neural networks and logistic regression: The Sephidargole case study, Semnan, Iran*Geomech. Geoengin* **3** 237–250

[61] H. C. Quan and B. G. Lee 2012 GIS-based landslide susceptibility mapping using analytic hierarchy process and artificial neural network in Jeju (Korea) *KSCE J. Civ. Eng.* **7** 1258–1266

[62] S Pascale 2013 Landslide susceptibility mapping using artificial neural network in the urban area of Senise and San Costantino Albanese (Basilicata, Southern Italy) *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **7974** 473–488

[63] M S Alkhasawneh, U K Ngah, L T Tay, N A Mat Isa and M S Al-Batah 2014 Modeling and testing landslide hazard using decision tree *J. Appl. Math.*

[64] M S Al-Batah, M S Alkhasawneh, L T Tay, U K Ngah, H Hj Lateh and N A Mat Isa 2015 Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer

Perceptron *Math. Probl. Eng.*
[65] M Zare, H R Pourghasemi, M Vafakhah and B Pradhan 2013 Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms *Arab. J. Geosci.* **6** 2873–2888
[66] A Abbaszadeh Shahri, J Spross, F Johansson and S Larsson 2019 Landslide susceptibility hazard map in southwest Sweden using artificial neural network *Catena* 183
[67] H J Oh and S Lee 2017 Shallow landslide susceptibility modeling using the data mining models artificial neural network and boosted tree *Appl. Sci.* **7** 1–14
[68] M Conforti, S Pascale, G Robustelli and F Sdao 2014 Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy) *Catena* **113** 236–250
[69] X Hu, H Zhang, H Mei, D Xiao, Y Li and M Li 2020 Landslide susceptibility mapping using the stacking ensemble machine learning method in lushui, southwest China *Appl. Sci.* 10
[70] A M Nuriah, R Ruslan and W I Wan Mohd Muhiyuddin 2018 Spatial Modeling Various Types Of Slope Failure Using Artificial Neural Network (ANN) In Pulau Pinang, Malaysia *J. Teknol. Sciences Eng. 804* **1** 135–146
[71] A R Rasyid, U Hasanuddin, N P Bhandary, and R Yatabe 2018 GIS based Landslide Susceptibility Mapping using Artificial Neural Network (ANN) Model in South Sulawesi Province, Indonesia *Int. J. Earth Sci. Eng.* **11** 107–111
[72] H Moayedi, M Mehrabi, M Mosallanezhad, A S A Rashid and B Pradhan 2019 Modification of landslide susceptibility mapping using optimized PSO-ANN technique *Eng. Comput.* **35** 967–984
[73] S Benchelha 2019 Landslide susceptibility mapping in the municipality of Oudka, northern Morocco: A comparison between logistic regression and artificial neural networks models *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* **42** 41–49
[74] B Peethambaran, R Anbalagan, D P Kanungo, A Goswami and K V Shihabudheen 2020 A comparative evaluation of supervised machine learning algorithms for township level landslide susceptibility zonation in parts of Indian Himalayas *Catena* 195
[75] D. W. Hosmer and S. Lemeshow 2000 *Applied logistic Resgression*, vol. 2nd edition
[76] F Ren, X Wu, K Zhang and R Niu 2015 Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the Shuping landslide in the Three Gorges, China *Environ. Earth Sci.* **73** 4791–4804
[77] D Tien Bui, T A Tuan, H Klempe, B Pradhan and I Revhaug 2016 Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree *Landslides* **13** 361–378
[78] W Chen, H Chai, Z Zhao, Q Wang and H. Hong 2016 Landslide susceptibility mapping based on GIS and support vector machine models for the Qianyang County, China *Environ. Earth Sci.* **75** 1–13
[79] M Abedini, B Ghasemian, A Shirzadi and D T Bui 2019 A comparative study of support vector machine and logistic model tree classifiers for shallow landslide susceptibility modeling *Environ. Earth Sci.* **78** 1–15
[80] Y Huang and L Zhao 2018 Review on landslide susceptibility mapping using support vector machines *Catena* **165** 520–529
[81] C Xu, F Dai, X Xu and Y H. Lee 2012 GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China *Geomorphology* **145–146** 70–80
[82] K Sharir, R Roslee, L K Ern, and N Simon 2017 Landslide Factors and susceptibility mapping on natural and artificial slopes in Kundasang, Sabah *Sains Malaysiana* **46** 1531–1540
[83] C Su, L Wang, X Wang, Z Huang, and X Zhang 2015 Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine *Nat. Hazards* **76** 1759–1779
[84] B Feizizadeh, M S Roodposhti, T Blaschke and J Aryal 2017 Comparing GIS-based support

vector machine kernel functions for landslide susceptibility mapping *Arab. J. Geosci.* 10
[85] R. T. Pack, D. G. Tarboton, and C. N. Goodwin 2016 Terrain stability mapping with SINMAP, technical description and users guide for version 1.0
[86] R. L. Baum, W. Z. Savage, and J. W. Godt 2008 TRIGRS — A Fortran Program for Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Analysis, Version 2.0 *U.S. Geol. Surv.* **78** 2008–1159
[87] D. R. Montgomery and W. E. Dietrich 1994 A physically based model for the topographic control on shallow landsliding *Water Resour. Res.* **4** 1153–1171
[88] J. Kim, K. Lee, S. Jeong, and G. Kim 2014 GIS-based prediction method of landslide susceptibility using a rainfall infiltration-groundwater flow model *Eng. Geol.* 63–78
[89] C Hammond, D Hall, S Miller and P Swetik 1992 Level I Stability Analysis (LISA) documentation for version 2.0 *Gen. Tech. Rep. - US Dep. Agric. For. Serv.*
[90] S Lee, S M Hong and H S Jung 2017 A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea *Sustain.* **9** 15–19
[91] Mein R.G and C. Larson 1973 Modeling Infiltration during a Steady Rain *Water Resour. Res.* 384–394
[92] S Jeong, J Park, M Hong, J Kim and K. Lee 2019 Unified Prediction Method for the Rainfall-Induced Landslides and Debris Flows *Springer* 28.
[93] M Hong, S Jeong, and J Kim 2020 A combined method for modeling the triggering and propagation of debris flows *Landslides* **17** 805–824
[94] P Duc Long and N T Dung 2020 Susceptibility assessment of slope failures in Singapore using GIS- based prediction models
[95] Y Kim, S Jeong and J Kim 2016 Coupled infiltration model of unsaturated porous media for steady rainfall *Soils Found.* **56** 1071–1081
[96] GEO-SLOPE 2012 User’s Manual. Geo-Slope International Ltd, Canada