Journal of Advanced Research in Computing and Applications

Journal homepage:
https://karyailham.com.my/index.php/arca/index
ISSN: 2462-1927

Transferability of an Artificial Neural Network Model for Landslide Susceptibility Prediction: Extending from Western Sarawak to Tuaran and Penampang, Sabah

Nur Hisyam Ramli[1,*], Siti Noor Linda Taib[1], Norazzlina M. Sa'don[1], Raudhah Ahmadi[1], Dayangku Salma Awang Ismail[1], Imtiyaz Akbar Najar[1],Tarmiji Masron[2], Nazeri Abdul Rahman[3]

[1] Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia
[2] Faculty of Social Sciences and Humanities Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia
[3] Department of Chemical Engineering and Energy Sustainability, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia

ARTICLE INFO

Article history:
Received 11 October 2025
Received in revised form 19 November 2025
Accepted 25 November 2025
Available online 8 December 2025

Keywords:
Landslide; Borneo; ANN; GIS

ABSTRACT

Landslides have been determined to be one of the most damaging forms of natural disasters as well as common geological hazards around the world that causes significant damage to the ecosystems, and settlements. Landslide susceptibility assessment is a crucial step in understanding the risk of landslides for an area to conduct a successful mitigation effort. Cotemporary research has deemed a machine learning based approach to be highly accurate in determining the susceptibility of an area towards landslide, and a suitable prediction function for the landslide susceptibility map development. Given the dependency of machine learning models on the availability of data used to develop its predictive capabilities, the deployment of the method may not be suitable in regions that are lacking or do not have the required data, especially historical landslide points. In this study, the Artificial Neural Network model was developed with data from Western Sarawak, located 1158 km away from the target region which were Tuaran and Penampang, Sabah, that was lacking in the amount of available historical landslide data. To enable the transfer, the type of variables in both regions were kept uniform, which were aspect, curvature, elevation, land use and land cover, rainfall intensity, slope angle, and topographic wetness index. In its own region, the model was determined to be highly accurate for predicting landslide susceptibility with a training success rate, and prediction success rate of 100%. Transferring the model to the target region reveals the process was a success, with a prediction success rate of 98%, and a precision of 100%.

1. Introduction

The movement of slope forming materials is known as landslides, and it has been determined to be one of the most widely occurring form of natural disasters [1]. The frequency of landslide occurrences is expected to increase in the coming years due to the rapid expansion of settlements,

* *Corresponding author.*
E-mail address: 22020251@siswa.unimas.my

https://doi.org/10.37934/arca.41.1.2435

and anthropogenically induced global warming [2]. The scale of landslides, and its' onset may vary from one case to another, however, all of it contributes to degradation of the victims lives, and the overall economic system of the country [3]. Thus, determining the susceptibility of a location to landslide is a crucial step in planning the mitigation procedures [4].

The term of a landslide susceptibility is used convey the likelihood that a landslide had occurred or has the potential to occur determined through the severity of the landslide causative variables that are present in an area [5]. Uncovering the level of landslide susceptibility for a region of interest is not a straightforward task, as landslides are complex natural phenomena that is influenced by numerous variables, each at a different rate [6]. Several approaches to the problem have been conducted throughout the years be it though a qualitive analysis, quantitative analysis, or both [7].

Cotemporary approach in determining landslide susceptibility, and other form of natural disasters are done through machine learning models to which provide a data-driven result, thus avoiding the potentiality of unjustified biases [8]. The advancement of data science, computing hardware, and availability of remote sensing data, have enabled researchers to deepen their understanding spatially on the occurrence of landslides, as well as developing the landslide susceptibility map (LSM) for their study areas [9].

The most crucial aspect of machine learning based approach for landslide susceptibility assessment is data. Machine learning models regardless of purpose requires data for training, and subsequently evaluating the model's accuracy [10]. The accuracy of the models is highly dependent on the accuracy of the training data, as it is used to develop the predictive capabilities of the model [11]. However, in the certain cases, the data required may not be available or lacking, which would inhibit the use of machine learning model, or decreases its prediction accuracy.

Thus, this study was conducted to evaluate whether the usage of a machine learning model that was developed with a large data pool could be use in predicting the landslide susceptibility of a region which has a small data pool. The machine learning model of choice was an Artificial Neural Network (ANN), as it was determined to have the best chance of success due to its adaptative capabilities [12]. The transferrable ANN model was developed using data from Western Sarawak, and the target region was Tuaran, and Penampang district of Sabah. This study is very important due to increasing rate of urbanisation throughout the Malaysia where both region is located, which have been identified to be the major cause of landslide in the country [13].

1.1 Area of Study

Tuaran and Penampang are two neighbouring districts located in the state of Sabah, Malaysia with an area of roughly 496.52 km^2, and 1583.64 km^2 respectively as shown in Figure 1. Penampang is located within the Western region of Sabah while Tuaran is situated in the Northeastern region of the state. The Crocker Range has the most influence on the terrain of Tuaran, with its high elevation, and steep slopes, making most of the landslide consisting of hills, and mountains with low lying regions on the coastal areas [14]. Being neighbouring districts, the terrain of Penampang is also influenced by the Crocker range, with 70% of the district has hilly to mountainous terrain, whilst the rest consists of low-lying areas concentrated mostly along the coastline [15]. Spatially, the target region is located roughly 1158 km from the source region, albeit both regions are on the same island of Borneo. It was confirmed that there are only 15 historical landslide location data that are freely available for the public through the National Space Agency (NASA) Global Landslide Catalogue (GLC), a 68% difference from the recorded historical landslide points in Western Sarawak made available by the Department of Minerals and Geosciences (JMG) for this study [16].

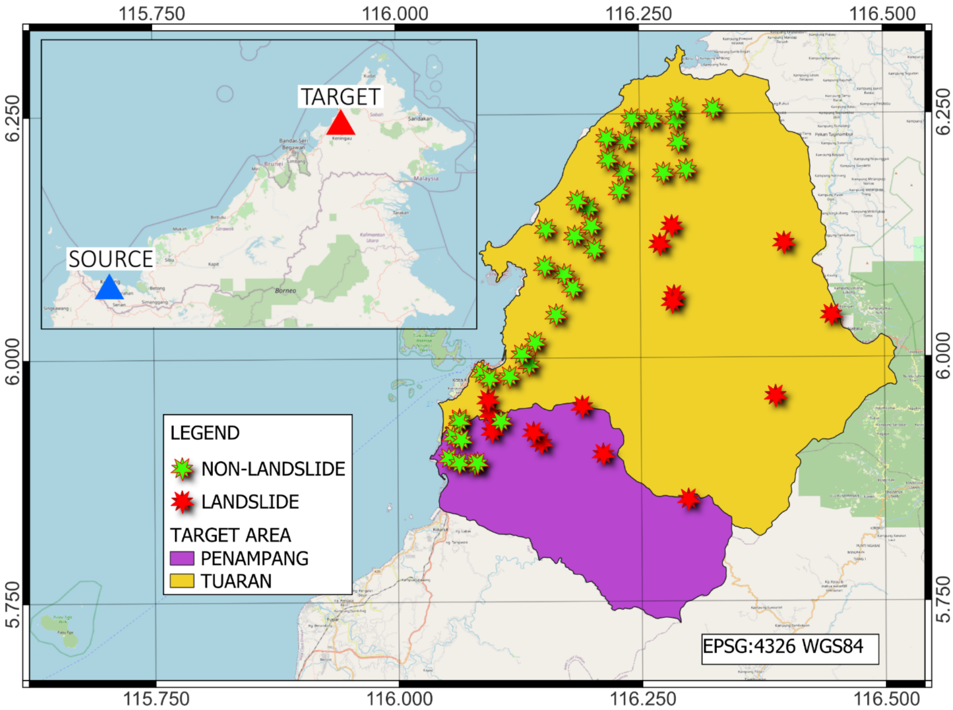

Fig. 1. Study area map

2. Methodology

The main goal of this study was to determine the transferability of an ANN model landslide susceptibility prediction. Initially, the ANN model was trained and tested with data from Western Sarawak. The ANN was developed to predict landslide susceptibility based on aspect, curvature, elevation, land use and land cover (LULC), rainfall intensity from November to December, slope angle, and Topographic Wetness Index (TWI). The ANN model architecture was developed with a single hidden layer consisting of 8 neurons as illustrated in Figure 2a, and learns through the backpropagation algorithm for the weightage adjustments [17].

Backpropagation was selected due to several reasons, which were [18,19]:

- The ability of the algorithm to learn from the training data and adapt accordingly to the network's parameters to understand the relationship between the layers.
- Automation of relevant features learning. As the training datasets consist of a wide range of data from different variables, the backpropagation was able to determine the influence of each variable on the outcome.
- Provided that the study has similar variables to the training data, backpropagation can adapt to the specific conditions of the new study area, making it suitable for studying the transferability of a machine learning model.

The ANN model accuracy was evaluated through the Area Under the Curve (AUC) method which compares the amount of correct prediction to the total amount of prediction, with 1 being the highest score [20]. It was determined that the ANN model AUC for both training success rate (Figure 2b), and prediction success rate (Figure 2c) was 1.

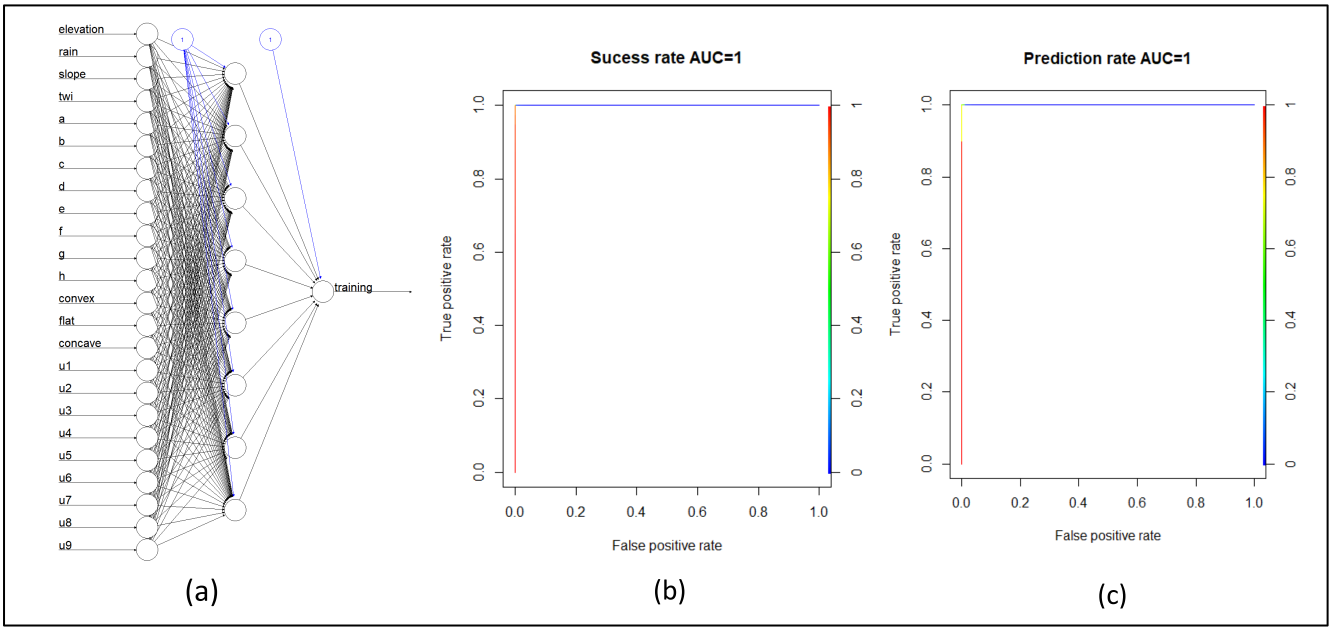

Fig. 2. (a) ANN model architecture, (b) AUC for training data, (c) AUC for testing data

The same AUC method was used to evaluate the ANN model accuracy in predicting the landslide susceptibility of the target region, with added evaluation method of precision that measure the quality of the prediction [21]. AUC requires True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) predictions (see Eq. (1)). Precision on the other hand only requires TP, and FP (see Eq. (2)).

$$AUC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Data that is required from the target area is managed in the same way as the training data and testing data of the ANN model, thus, the same data-preprocessing procedures, and the same input-output variables must be present in both source regions, as well as the target region. In this study, the historical landslide locations were obtained from NASA's GLC, which have underwent geo-referencing through Google Earth Pro as the location given may deviate from the actual point location [22]. Non-landslide locations were determined by selecting areas with slope angle lower than 5° [23].

Data-processing was done based on the type of data for the landslide causative variables, which were categorical and numerical. The categorical variables in this study were aspect, curvature, LULC, and TWI, whilst the numerical variables were elevation, slope angle, and TWI. The categorical variables have underwent data-preprocessing through one-hot encoding, which classifies the data based on predefined groups, as the data was better present in the form of categories and cannot be clearly defined through numerical values [24]. Numerical variables have undergone data-preprocessing through min-max scaling, where the numerical values were scaled to be from 0 to 1, A_s is the scaled value, A is the unscaled value, A_{min} is the lowest value of the variable, and A_{max} is the maximum value of the variable (see Eq. (3)). This scaling processing was conducted to avoid overpowering issues caused by the large differences in values for the numerical variables [25].

$$A_s = \frac{A - A_{min}}{A_{max} - A_{min}} \tag{3}$$

The landslide causative variables map shows the distribution of each variable's values (Figure 3a to Figure 3g).

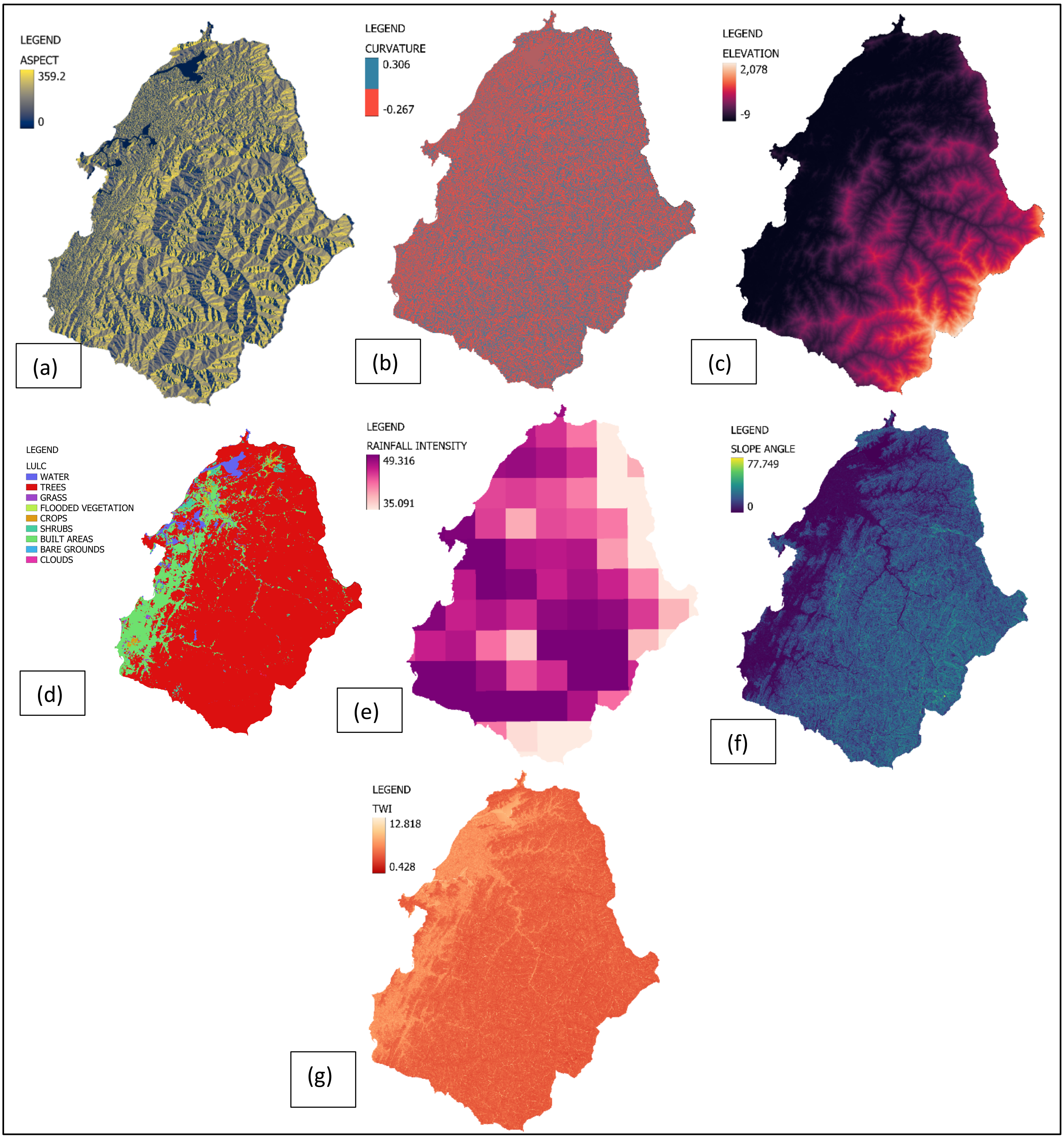

Fig. 3. (a) Aspect, (b) Curvature, (c) Elevation, (d) LULC, (e) Rainfall intensity, (f) Slope angle, (g) TWI

Aspect refers to the orientation of a given slope face. Aspect does not directly causes landslide to occur, instead, different aspect have different climatic conditions such as the amount of sun exposure, and rainfall received, that determine the degree of weathering and type of vegetation [26].

The aspect data was classified into 8 categories from a to h, with a range of 45° respectively. The aspect Geographical Information System (GIS) raster file was obtained by using the System for Automated Geoscientific Analyses (SAGA) extension in Quantum Geographical Information System (QGIS) from the elevation raster [27].

Curvature represents the shape of the slope itself, from concave to flat to convex. Curvature is responsible for the accumulation or redirection of surface runoff from rainfall, a major landslide triggering factor [28]. Concave curvature has been determined to have a higher susceptibility towards landslide in comparison to flat, and convex curvature due to its capability to accumulate water [29]. The curvature raster in this study was obtained through SAGA in QGIS, where the negative score represents concave curvature, 0 score for flat curvature, and a positive score for convex curvature [30].

Elevation shows the vertical distance between a point to the vertical datum, and areas located in the higher elevation are deemed as being more susceptible to landslides compared to areas with lower elevation as the ruggedness of the terrain increases with the elevation [31]. The elevation raster in this study was obtained from NASA's Advanced Land Observing Satellite (ALOS) through the Alaska State Facility (ASF) [32]. The elevation of the target region ranges from -9m to 2078m.

LULC in landslide susceptibility assessment is a measure of the anthropogenic activity present in the study area. In Malaysia, areas with high human activity experience more landslide cases compared to undisturbed areas [33]. The LULC raster file was obtained from the Environmental Systems Research Institute, Inc. (ESRI) [34]. 9 distinct LULC were present in the study area, and as LULC is a form of categorical data, it was categorized, were u1 for water, u2 for trees, u3 for grass, u4 for flooded vegetation, u5 for crops, u6 for shrubs, u7 for built areas, u8 for bare grounds, and u9 for clouds.

Rainfall is a major triggering factor of landslide in a tropical country such as Malaysia, as it penetrates the ground, loosening the soil, and if the intensity exceeds a certain threshold, a landslide would occur [35]. The rainfall raster file was obtained from the Climate Hazard Centre (CHC) with a unit of mm per pentad (five days), which shows that the rainfall intensity for the target region from November to December ranges from 35.091 to 49.316 [36].

Slope angle has been observed to be the main conditioning factor of landslides amongst the landslide causative variables. Region with steep slopes has a higher susceptibility towards landslide in comparison to region with a gentler slope, as the shear stress on the slope forming material increases [37]. The slope angle raster file was obtained from SAGA in QGIS, which shows that the slope angle in the target region ranges from 0° to 77.749° [27].

TWI is a dimensionless index, that shows the location of catchment areas based on where the flow would be accumulated given the properties of terrain [38]. Areas with higher accumulation of water tends to have a higher susceptibility towards landslides as it promotes the alteration of the slope forming material [39]. The TWI raster file was obtained from QGIS through the raster calculator function, and it shows that the TWI for the target area ranges from 0.428 to 12.818 [40].

The values of each landslide causative variables at each landslide points, and non-landslide points were extracted through QGIS raster sampling tool, which was used to determine the accuracy of the source ANN model to predict landslide susceptibility in the target area after data-preprocessing have been applied. However, unlike the source ANN model training data, Pearson's Correlation Analysis (PCA), was not applied to the target data, as it was not used to develop the ANN model predictive capabilities. PCA is an important process to identify whether there is multicollinearity present in the training data. High multicollinearity exceeds PCA coefficient of 0.7, and causes the ANN model prediction to degrade in term of accuracy as redundant input variables is supplied in the training process [41].

3. Results and Discussion

3.1 Source ANN Model

After the data for the source ANN model have been prepared, it was checked for multicollinearity through PCA to avoid any issues caused by multicollinearity. The result of the analysis can be seen in Table 1, with the highest correlation for input-to-input variables was determined to be between elevation and slope angle for the source area with a PCA coefficient of 0.68. The highest collinearity between the input-to-output of landslide however, was not a detrimental evaluation for the training data, instead it is an overview of the relationship between the input variables to the output variables, where for the source region, slope angle has the highest correlation with landslide, indicating that as slope angle increases, the landslide susceptibility also increases [42]. It was concluded that no significant multicollinearity was present in the training data, thus, any potential issues caused by multicollinearity was avoided.

Table 1
PCA matrix of source ANN Model

	Aspect	Curvature	Elevation	LULC	Rainfall	Slope	TWI	Landslide
Aspect	1.00							
Curvature	-0.01	1.00						
Elevation	-0.09	0.04	1.00					
LULC	-0.02	0.04	-0.19	1.00				
Rainfall	-0.03	0.00	0.00	0.01	1.00			
Slope	-0.09	0.07	0.68	-0.17	0.06	1.00		
TWI	0.06	-0.19	-0.32	-0.07	-0.07	-0.38	1.00	
Landslide	-0.07	0.08	0.50	-0.14	0.06	0.86	-0.33	1.00

As for the accuracy of the source ANN model, the training success rate and prediction success rate AUC of 1, indicates a highly accurate ANN model in predicting the landslide susceptibility of the source region. The training success rate evaluation was conducted to determine the overall success rate of the training phase, while the prediction success rate evaluation was conducted to determine the accuracy of the ANN model in predicting landslide susceptibility based on never-before-seen data, which was the testing data.

3.2 Accuracy and Precision of Source ANN Model in Predicting the Landslide Susceptibility of the Target Area

After the values of each landslide causative variables have been extracted from each corresponding raster files using the landslide points, and non-landslide points in QGIS, the data was used to evaluate the source ANN model accuracy and precision in predicting the landslide susceptibility of the target area after the same data-preprocessing measures conducted on the source training data was also conducted on the target data.

The results of the prediction can be seen in Figure 4, where the similarity threshold was set as 80%. In total, there were 34 non-landslide points, and 15 landslide points in the target region. The source ANN model has accurately predicted all 34 non-landslide points as non-landslide cases, and out of the total 15 landslide points, 14 were accurately predicted. Resulting AUC from the source ANN model in predicting the landslide susceptibility for the target area was 0.98.

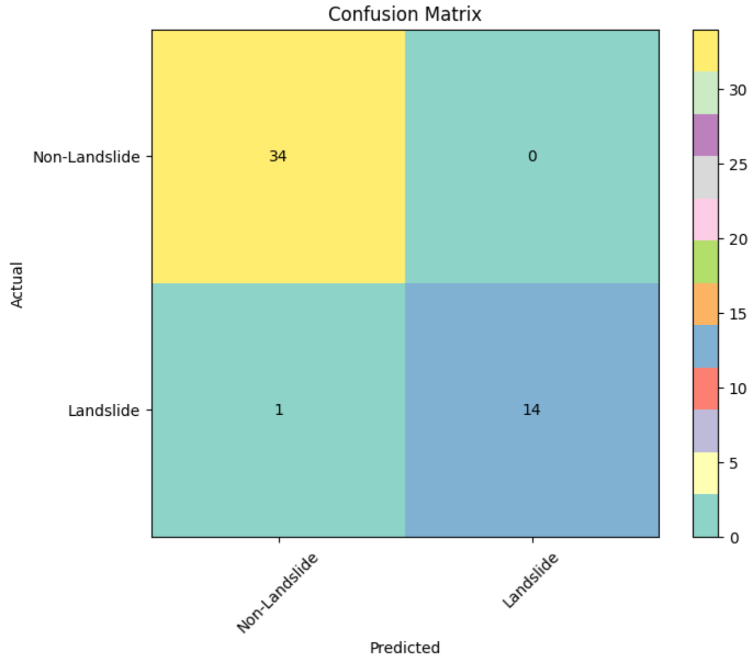

Fig. 4. Prediction outcome

The results from the prediction and the actual values can be seen in Table 2. Precision of the source ANN model in predicting the landslide susceptibility of the target area was determined to be at 0.98.

Table 2
Excerpt from Prediction

Prediction	Actual	Prediction	Actual	Prediction	Actual	Prediction	Actual
1.00	1.00	0.87	1.00	0.06	0.00	0.00	0.00
1.00	1.00	0.85	1.00	0.04	0.00	0.00	0.00
0.97	1.00	0.76	1.00	0.03	0.00	0.00	0.00
0.97	1.00	0.38	0.00	0.03	0.00	0.00	0.00
0.95	1.00	0.35	0.00	0.02	0.00	0.00	0.00
0.93	1.00	0.26	0.00	0.01	0.00	0.00	0.00
0.92	1.00	0.22	0.00	0.00	0.00	0.00	0.00
0.91	1.00	0.19	0.00	0.00	0.00	0.00	0.00
0.91	1.00	0.16	0.00	0.00	0.00	0.00	0.00
0.90	1.00	0.15	0.00	0.00	0.00	0.00	0.00
0.90	1.00	0.11	0.00	0.00	0.00	0.00	0.00
0.88	1.00	0.10	0.00	0.00	0.00	0.00	0.00

3.3 Target Region LSM

After the ANN model was determined to have reached or exceed the satisfactory accuracy of 0.8 through AUC, and a precision threshold of 0.8, the LSM was developed [43]. This was done by converting the raster files into a numerical data frame where the same data pre-processing measures done on the source training data for the ANN model was done on the numerical data frame. The resulting LSM (Figure 5a) shows that most of the highly susceptible areas are concentrated in the highly elevated areas with steep slopes (Figure 5b).

The 3-dimensional (3D) map has shown that elevation and slope angle play a significant role in the landslide susceptibility of the target region, thus, correlating with the source region PCA coefficient of input-to-output variables. As for the low susceptibility areas, most of the areas are concentrated in the low-lying region alongside the coastline, except for a few hilly areas.

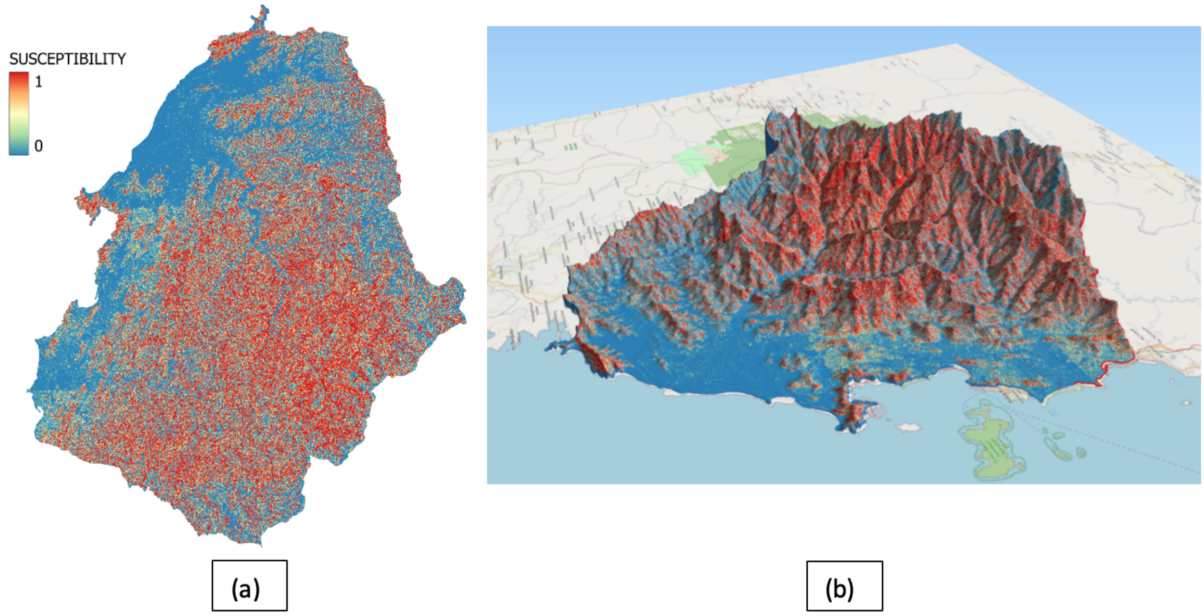

Fig. 5. (a) Target region LSM (b) 3 Dimensional LSM of target region

The reduction of accuracy in the transfer may have resulted from the difference in percentage of the elevation, and slope angle from the source region to the target region shown in Table 3. The percentage error was carried out to the LSM, as the ANN model was used as the predictive function. The 2% error from the AUC of 0.98, is an indicative difference of elevation with a percentage difference of 37.57%, and slope angle with a percentage difference of 2.14%. This was determined by observing the PCA coefficient of the variables, which were 0.5, and 0.86 respectively for elevation and slope angle. Thus, the error may have mostly originated from the elevation. However, given the low percentage of error in term of inaccurate predictions, the LSM can still be used as a preliminary assessment of landslide susceptibility for the target region [17]. Thus, it can be concluded that the transfer process was a success.

Table 3
Difference of source and target region elevation and slope angle

	Source			Target				
Variables	Min	Max	Range	Min	Max	Range	% Range Difference	Output PCA
Elevation	-37.00	1266.00	1303	-9	2078	2087	37.57	0.50
Slope	0	76.09	76.09	0	77.75	77.75	2.14	0.86

4. Conclusions

Landslides are a common form of natural disasters which occur everywhere around the globe. Given its complexity as it is influenced by numerous variables each at a different rate, contemporary research has opted to apply machine learning based approaches to solve the issue. However, the application of machine learning in landslide susceptibility assessment is constrained by the availability of data, as it requires data to both train the model, as well as evaluate its prediction accuracy. In this study, the target region of Tuaran and Penampang of Sabah lack the amount of data required to develop the machine learning model of choice, which was an Artificial Neural Network. Thus, the source region of Western Sarawak data was used to develop the machine learning model. The landslide causative variables were kept constant to facilitate the transfer, which were aspect, curvature, elevation, land use and land cover, rainfall intensity, slope angle, and Topographic Wetness Index, as the variables were present in both source regions, as well as the target region. The machine learning model was evaluated to be very accurate in its own source area with a training success rate, and prediction success rate of 100% for both evaluations. In the target area, the machine learning model was evaluated to have a 98% prediction success rate, with a precision of 100%. Thus, it was concluded that the machine learning model was suitable for the landslide susceptibility map development given the accuracy and precision. The resulting landslide susceptibility map can be used as preliminary assessment in the landslide mitigation effort for the target area.

Acknowledgement
The authors acknowledge Ministry of Higher Education Malaysia, Fundamental Research Grant Scheme, FRGS/1/2022/TK06/UNIMAS/01/1 and Universiti Malaysia Sarawak for supporting this project. This research has also received publication funding from Sarawak Energy Berhad.

References
[1] Zainuddin, Zahir, Muhammad Niswar, and Hafsah Nirwana. "Landslide Susceptibility Zoning Using Fuzzy Logic Algorithm & GIS: Study Case In Majene, West Sulawesi." In *2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, pp. 183-188. IEEE, 2023. https://doi.org/10.1109/IC3INA60834.2023.10285732
[2] P. S. V. Rama Krishna *et al.*, *Guidelines on integrating climate change projection into landslide risk assessment & mapping*. Jakarta: Association of Southeast Asian Nations (ASEAN), 2021. [Online]. Available: https://asean.org/book/guideline-on-integrating-climate-change-projection-into-landslide-risk-assessments-and-mapping-at-the-river-basin-level/
[3] Ma, Zhengjing, Gang Mei, and Francesco Piccialli. "Machine learning for landslides prevention: a survey." *Neural Computing and Applications* 33, no. 17 (2021): 10881-10907. https://doi.org/10.1007/s00521-020-05529-8
[4] Taherynia, Mohammad Hossein, Mojtaba Mohammadi, and Rasoul Ajalloeian. "Assessment of slope instability and risk analysis of road cut slopes in Lashotor Pass, Iran." *Journal of Geological Research* 2014, no. 1 (2014): 763598. https://doi.org/10.1155/2014/763598
[5] El Fahchouch, A. N., L. Ait Brahim, and O. Raji. "Susceptibility assessment of landslides: A comparison of two GIS-based methods in the region of Al Hoceima (Eastern Rif, Morocco)." In *MATEC Web of Conferences*, vol. 149, p. 02055. EDP Sciences, 2018. https://doi.org/10.1051/matecconf/201814902055
[6] Chen, Wei, Jianbing Peng, Haoyuan Hong, Himan Shahabi, Biswajeet Pradhan, Junzhi Liu, A-Xing Zhu, Xiangjun Pei, and Zhao Duan. "Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China." *Science of the total environment* 626 (2018): 1121-1135. https://doi.org/10.1016/j.scitotenv.2018.01.124
[7] Huat, Bujang BK, and S. Jamaludin. "Evaluation of slope assessment system in predicting landslides along roads underlain by granitic formation." *American Journal of Environmental Sciences* 1, no. 2 (2005): 90-96. https://doi.org/10.3844/ajessp.2005.90.96
[8] Wang, Nan, Hongyan Zhang, Ashok Dahal, Weiming Cheng, Min Zhao, and Luigi Lombardo. "On the use of

explainable AI for susceptibility modeling: Examining the spatial pattern of SHAP values." *Geoscience Frontiers* 15, no. 4 (2024): 101800. https://doi.org/10.1016/j.gsf.2024.101800

[9] Chowdhury, Md Sharafat, Md Naimur Rahman, Md Sujon Sheikh, Md Abu Sayeid, Khandakar Hasan Mahmud, and Bibi Hafsa. "GIS-based landslide susceptibility mapping using logistic regression, random forest and decision and regression tree models in Chattogram District, Bangladesh." *Heliyon* 10, no. 1 (2024). https://doi.org/10.1016/j.heliyon.2023.e23424

[10] Sh, Husien, Reem M. El-taweel, KhloodA Alrefaey, Ahmed Labena, Irene Samy Fahim, Lobna A. Said, and Ahmed G. Radwan. "Enhanced removal of crystal violet using rawfava bean peels, its chemically activated carbon compared with commercial activated carbon." *Case Studies in Chemical and Environmental Engineering* 9 (2024): 100534. https://doi.org/10.1016/j.cscee.2023.100534

[11] Djukem, Wamba Danny Love, Anika Braun, Armand Sylvain Ludovic Wouatong, Christian Guedjeo, Katrin Dohmen, Pierre Wotchoko, Tomas Manuel Fernandez-Steeger, and Hans-Balder Havenith. "Effect of soil geomechanical properties and geo-environmental factors on landslide predisposition at Mount Oku, Cameroon." *International journal of environmental research and public health* 17, no. 18 (2020): 6795. https://doi.org/10.3390/ijerph17186795

[12] R. Souissi, A. Al Bitar, and M. Zribi, "Accuracy and transferability of artificial neural networks in predicting in situ root-zone soil moisture for various regions across the globe," *Water (Switzerland)*, vol. 12, no. 11, pp. 1–20, Nov. 2020. https://doi.org/10.3390/w12113109

[13] Akter, Ayesha, M. J. M. M. Noor, Masa Goto, Shamsunnahar Khanam, Ahmed Parvez, and Md Rasheduzzaman. "Landslide disaster in Malaysia: An overview." *International Journal of Innovative Research and Development* 8, no. 6 (2019): 292-302. https://doi.org/10.24940/ijird/2019/v8/i6/JUN19058

[14] Red Crescent Malaysia, "CHAPTER TUARAN." Accessed: Jun. 10, 2024. [Online]. Available: https://pbsmmct.tripod.com/Latar.html

[15] Penampang District Office, "Background of Penampang District Office." Accessed: Jun. 10, 2024. [Online]. Available: https://pdpenampang.sabah.gov.my/page.php?menu=INFO KORPORAT&sub=LATAR BELAKANG&page=text_pic

[16] Kirschbaum, Dalia, Thomas Stanley, and Yaping Zhou. "Spatial and temporal analysis of a global landslide catalog." *Geomorphology* 249 (2015): 4-15. https://doi.org/10.1016/j.geomorph.2015.03.016

[17] Islam, Abu Reza Md Towfiqul, Swapan Talukdar, Susanta Mahato, Sonali Kundu, Kutub Uddin Eibek, Quoc Bao Pham, Alban Kuriqi, and Nguyen Thi Thuy Linh. "Flood susceptibility modelling using advanced ensemble machine learning models." *Geoscience Frontiers* 12, no. 3 (2021): 101075. https://doi.org/10.1016/j.gsf.2020.09.006

[18] Nelli, Fabio. *Python data analytics: Data analysis and science using PANDAs, Matplotlib and the Python Programming Language*. Apress, 2015.

[19] Swamynathan, Manohar. *Mastering machine learning with python in six steps: A practical implementation guide to predictive data analytics using python*. Manohar Swamynathan, 2017. https://doi.org/10.1007/978-1-4842-2866-1

[20] Li, Yao, Frank Badu Osei, Tangao Hu, and Alfred Stein. "Urban flood susceptibility mapping based on social media data in Chengdu city, China." *Sustainable Cities and Society* 88 (2023): 104307. https://doi.org/10.1016/j.scs.2022.104307

[21] Shahabi, H., M. Rahimzad, O. Ghorbanzadeh, S. T. Piralilou, T. Blaschke, S. Homayouni, and P. Ghamisi. "Rapid mapping of landslides from Sentinel-2 data using unsupervised deep learning." In *2022 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS)*, pp. 17-20. IEEE, 2022. https://doi.org/10.1109/M2GARSS52314.2022.9840273

[22] Gómez, Derly, Edwin F. García, and Edier Aristizábal. "Spatial and temporal landslide distributions using global and open landslide databases." *Natural Hazards* 117, no. 1 (2023): 25-55. https://doi.org/10.1007/s11069-023-05848-8

[23] Jamaludin, Suhaimi, Bujang BK Huat, and Husaini Omar. "Evaluation And Development Of Cutâ€"Slope Assessment Systems For Peninsular Malaysia In Predicting Landslides In Granitic Formation." *Jurnal Teknologi (Sciences & Engineering)* (2006): 31â-46.

[24] Ul Haq, Ikram, Iqbal Gondal, Peter Vamplew, and Simon Brown. "Categorical features transformation with compact one-hot encoder for fraud detection in distributed environment." In *Australasian Conference on Data Mining*, pp. 69-80. Singapore: Springer Singapore, 2018. https://doi.org/10.1007/978-981-13-6661-1_6

[25] Sharma, Vinod. "A study on data scaling methods for machine learning." *International Journal for Global Academic & Scientific Research* 1, no. 1 (2022): 31-42. https://doi.org/10.55938/ijgasr.v1i1.4

[26] Guo, Yanglin, and Chao Ma. "Elucidating the role of soil hydraulic properties on aspect-dependent landslide initiation." *Hydrology and Earth System Sciences* 27, no. 8 (2023): 1667-1682. https://doi.org/10.5194/hess-27-1667-2023

[27] Conrad, Olaf, Benjamin Bechtel, Michael Bock, Helge Dietrich, Elke Fischer, Lars Gerlitz, Jan Wehberg, Volker Wichmann, and Jürgen Böhner. "System for automated geoscientific analyses (SAGA) v. 2.1. 4." *Geoscientific model development* 8, no. 7 (2015): 1991-2007. https://doi.org/10.5194/gmd-8-1991-2015

[28] Ahmed, Istak, Nibedita Das Pan, Jatan Debnath, Moujuri Bhowmik, and Shaswati Bhattacharjee. "Flood hazard zonation using GIS-based multi-parametric Analytical Hierarchy Process." *Geosystems and Geoenvironment* 3, no. 2 (2024): 100250. https://doi.org/10.1016/j.geogeo.2023.100250

[29] Moragues, Silvana, María Gabriela Lenzano, Pilar Jeanneret, Verónica Gil, and Esteban Lannutti. "Landslide susceptibility mapping in the Northern part of Los Glaciares National Park, Southern Patagonia, Argentina using remote sensing, GIS and frequency ratio model." *Quaternary Science Advances* 13 (2024): 100146. https://doi.org/10.1016/j.qsa.2023.100146

[30] Wood, Jo. "Geomorphometry in landserf." *Developments in soil science* 33 (2009): 333-349. https://doi.org/10.1016/S0166-2481(08)00014-7

[31] Naceur, Hassan Ait, Hazem Ghassan Abdo, Brahim Igmoullan, Mustapha Namous, Hussein Almohamad, Ahmed Abdullah Al Dughairi, and Motrih Al-Mutiry. "Performance assessment of the landslide susceptibility modelling using the support vector machine, radial basis function network, and weight of evidence models in the N'fis river basin, Morocco." *Geoscience Letters* 9, no. 1 (2022): 39. https://doi.org/10.1186/s40562-022-00249-4

[32] J. Laurencelle, T. Logan, and R. Gens, "ALOS PALSAR - Radiometric Terrain Correction," JAXA/METI ALOS PALSAR. Accessed: Oct. 24, 2024. [Online]. Available: https://asf.alaska.edu/datasets/daac/alos-palsar/

[33] H. A. Rahman and J. Mapjabil, "Landslides disaster in Malaysia: An overview," *Heal. Environ. J.*, vol. 8, no. 1, pp. 58–71, 2017, [Online]. Available: http://psasir.upm.edu.my/id/eprint/54691/

[34] Karra, Krishna, Caitlin Kontgis, Zoe Statman-Weil, Joseph C. Mazzariello, Mark Mathis, and Steven P. Brumby. "Global land use/land cover with Sentinel 2 and deep learning." In *2021 IEEE international geoscience and remote sensing symposium IGARSS*, pp. 4704-4707. IEEE, 2021. https://doi.org/10.1109/IGARSS47720.2021.9553499

[35] Hashim, Kamarulbahrin Bin, and Henry Litong Among. "Geological investigation on Ruan Changkul landslide." (2003).

[36] Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak et al. "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes." *Scientific data* 2, no. 1 (2015): 1-21. https://doi.org/10.1038/sdata.2015.66

[37] Khajehzadeh, Mohammad, Mohd Raihan Taha, Suraparb Keawsawasvong, Hamidreza Mirzaei, and Mohammadreza Jebeli. "An effective artificial intelligence approach for slope stability evaluation." *Ieee Access* 10 (2022): 5660-5671. https://doi.org/10.1109/ACCESS.2022.3141432

[38] Chowdhury, Md Sharafat. "Flash flood susceptibility mapping of north-east depression of Bangladesh using different GIS based bivariate statistical models." *Watershed Ecology and the Environment* 6 (2024): 26-40. https://doi.org/10.1016/j.wsee.2023.12.002

[39] Cobos-Mora, Sandra Lucia, Victor Rodriguez-Galiano, and Aracely Lima. "Analysis of landslide explicative factors and susceptibility mapping in an andean context: The case of Azuay province (Ecuador)." *Heliyon* 9, no. 9 (2023). https://doi.org/10.1016/j.heliyon.2023.e20170

[40] QGIS Development Team, "QGIS desktop 3.22 user guide," 2023. [Online]. Available: https://docs.qgis.org/3.22/pdf/en/QGIS-3.22-DesktopUserGuide-en.pdf

[41] Schober, Patrick, Christa Boer, and Lothar A. Schwarte. "Correlation coefficients: appropriate use and interpretation." *Anesthesia & analgesia* 126, no. 5 (2018): 1763-1768. https://doi.org/10.1213/ANE.0000000000002864

[42] Nnanwuba, Uzodigwe Emmanuel, Shengwu Qin, Oluwafemi Adewole Adeyeye, Ndichie Chinemelu Cosmas, Jingyu Yao, Shuangshuang Qiao, Sun Jingbo, and Ekene Mathew Egwuonwu. "Prediction of spatial likelihood of shallow landslide using GIS-based machine learning in awgu, southeast/Nigeria." *Sustainability* 14, no. 19 (2022): 12000. https://doi.org/10.3390/su141912000

[43] Maharjan, Menuka, Sachin Timilsina, Santosh Ayer, Bikram Singh, Bikram Manandhar, and Amir Sedhain. "Flood susceptibility assessment using machine learning approach in the Mohana-Khutiya River of Nepal." *Natural Hazards Research* 4, no. 1 (2024): 32-45. https://doi.org/10.1016/j.nhres.2024.01.001