International Journal of Design & Nature and Ecodynamics
Vol. 20, No. 2, February, 2025, pp. 349-357
Journal homepage: http://iieta.org/journals/ijdne

Comparative Analysis of Local and Transferred ANN Models in Landslide Susceptibility Prediction in a Tropical Region

Nur Hisyam Ramli[1,2], Siti Noor Linda Taib[1*], Norazzlina M. Sa'don[1], Dayangku Salma Awang Ismail[1], Raudhah Ahmadi[1], Imtiyaz Akbar Najar[1], Nazeri Abdul Rahman[3], Rosmina Ahmad Bustami[4], Tarmiji Masron[2]

[1] Department of Civil Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan 94300, Malaysia
[2] Centre for Spatially Integrated Digital Humanities (CSIDH), Faculty of Social Sciences and Humanities, Universiti Malaysia Sarawak, Kota Samarahan 94300, Malaysia
[3] Department of Chemical Engineering and Energy Sustainability, Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan 94300, Malaysia
[4] UNIMAS Water Centre (UWC), Faculty of Engineering, Universiti Malaysia Sarawak, Kota Samarahan 94300, Malaysia

Corresponding Author Email: tlinda@unimas.my

https://doi.org/10.18280/ijdne.200212

Received: 29 December 2024
Revised: 17 January 2025
Accepted: 23 January 2025
Available online: 28 February 2025

Keywords:
slope failures, data sciences, multilayer perceptron, remote sensing, model transfer

ABSTRACT

Landslides are a common form of natural disaster in the tropics due to heavy rainfall in the wet season. Due to the hazards that come with landslides, determining the susceptibility of an area is of utmost importance. Currently, this is done through an ML-based approach. However, some areas may lack the required data. Thus, this study focused on comparing the impact of a transferred ML model from a comprehensive data region to a localized model. This was done by developing an ANN model trained on data from Western Sarawak and comparing it to the localized model in the west coast of Sabah and Selangor. The transferred ANN model results were acceptable, with recall scores of 0.89 and 0.86 for the west coast of Sabah and Selangor, respectively, while the localized models both achieved a recall score of 1. AUC scores were also comparable, at 0.988 and 0.995 for the west coast of Sabah and Selangor, respectively, while the localized models both achieved an AUC of 1. For the LSMs, in both target areas, the transferred ANN model predictions were heavily skewed in comparison to the localised model. It is recommended that future studies test the transferability in other tropical regions beyond Southeast Asia.

1. INTRODUCTION

In a tropical country such as Malaysia, landslide cases are commonplace with an average of 28 cases reported annually [1]. Landslides in Malaysia occur during the wet season, as abundant rainfall is the primary triggering factor that rapidly reduces the structural rigidity of slope-forming material to the point of failure [2]. The failed slopes would then move downward and pose a significant threat to anything in their path [3, 4]. Thus, to mitigate the hazards associated with the moving mass, it is of utmost importance to determine the landslide susceptibility of the affected areas [5, 6]. Conventionally, the landslide risks of a slope are verified through methods such as finite element method or locally developed Slope Assessment Systems such as the Slope Management and Risk Tracking System developed by the Public Works Department in Malaysia [7, 8]. However, the aforementioned methods are not practical for determining landslide risks in large areas of interest due to their dependency on soil physical properties that require laboratory testing for verification [9].

In recent years, quantitative approaches through Machine Learning (ML) models were deemed to be a viable alternative to the conventional methods of landslide risks assessment due to the models being capable of producing results that are comparable to the conventional methods whilst having lesser restrictions on the data requirements [10]. One of the primary sources of data for landslide prediction ML models in contemporary studies are Remote Sensing (RS) data providers [11]. RS provides an easily accessible data platform for ML models development purposes with many providers supporting researches by supplying data that are free to use [12]. Nevertheless, in certain areas, there is a possibility that the requirements could not be met, where the RS data providers do not have the required data. This condition is widely present in the target variable required by the ML models, which is the historical landslide points, due to high terrain deformation rates for purposes such as road construction [13]. This will lead to a false positive classification, as the terrain was changed intentionally without any landslide cases being present.

Historical landslide point scarcity can cause major problems for ML model performance, such as unjustified biases due to a small data pool with similar values [14]. One of the main

features of a supervised ML based approach in landslide prediction is its transferability, which can be used to predict landslides in areas outside of the original sample area, given that the same features are present [15]. Many of the transferred ML models, particularly a neural network-based approach, have shown favourable results when it comes to predicting landslides in the target regions [16]. Furthermore, a recent study has shown that a transferred ML model could accurately predict the target area's landslide susceptibility with an Area Under the Curve score of 0.942 [17]. However, the study area was limited to the Himalayan, where there is a possibility when applying the same methods in the tropics would not yield the same results, due to the difference in climatical and geological properties [18, 19].

Thus, this study aims to compare the performance of a transferred ML model and a local ML model for landslide prediction in the tropical regions of Malaysia. The transfer ML model originated from Western Sarawak, with the targets being the West coast region of Sabah (Target 1), and areas in the borders of Selangor (Target 2). The ML model of choice was an Artificial Neural Network (ANN), which has been determined to be highly transferable in previous studies. The transfer region was used to develop the Landslide Susceptibility Map (LSM) of each target region, whereas the target ANN models were only used to develop their own LSMs. The success of this study will determine the applicability of the transferred ML model for landslide predictions in areas with scarce landslide points throughout the tropical regions of Southeast Asia.

2. MATERIALS AND METHODS

2.1 Study area

Sabah and Sarawak exist on the same island of Borneo [20, 21], whilst Selangor is situated in Peninsular Malaysia. Target 1 is located approximately 824.97 km away from Western Sarawak, and Target 2 is located 1009.89 km away from the region. Although seemingly far apart from each other, each one of the study areas exhibits the same dominant type of landslide, which is shallow rapid landslides. This is due to the shared primary landslide triggering factor which is heavy rainfall [22]. The three regions share the same wet season, which is the Northeast monsoon season that generally occur from November to February annually [23]. According to the rainfall intensity data provided by the Climate Hazards Center InfraRed Precipitation with Station Data (CHIRPS) during the Northeast monsoon season of 2022-2023 in mm per five days (pentad), the transfer area receives the most rainfall, with a minimum of 65.753 mm/pentad, a mean of 85.167 mm/pentad, and a maximum of 121.266 mm/pentad followed by Target 1 and Target 2, as seen in Table 1 [24].

Table 1. Rainfall data statistic in each study area in mm/pentad

Statistics	Region		
	Transfer	Target 1	Target 2
Minimum	65.753	27.881	33.536
Mean	85.167	52.864	43.769
Maximum	121.266	91.889	53.566

Geologically, the conditions in each study area varied vastly from one to the other according to the Harmonised World Soil Database version 2.0 [25]. The only constant geological properties that was present throughout the study areas was Technosols with Geocode 7001 (Table 2). Between the Transfer region and Target 1, the other similarities were 4461, 4463, and 4498 (Table 2, Table 3). This is likely since both areas are located within the same island of Borneo, where some geological properties can be similar [26]. Furthermore, there were no similarities between the Transfer region and Target 1, to Target 2, excluding 7001. Unlike Target 1, Target 2 is in Peninsular Malaysia, which explains the difference between the geological properties of the region to the study areas in Borneo.

Economically, the gross domestic product (GDP) of Target 2 is 145,659 USD, which greatly outweighs the entire state of Transfer region, and Target 1 with a GDP of 31,636 USD, and 18,489 USD respectively [20, 27]. Furthermore, in terms of population density, the density of Transfer region and Target 1 combined could not surpass the population density of two cities in Target 2 that exceeds 10,000 people per square kilometre (km^2), in contrast to Transfer region and Target 1 with a population density of 5,393 people per km^2 and 1,703 people per km^2 respectively [28]. As anthropogenic variables have been determined to influence landslides, it was important to include it in this study [29].

Table 2. Geological property type in each study area

Geological Properties Code (Geocode) of Study Areas		
Transfer	Target 1	Target 2
4461	4461	4284
4463	4463	4324
4476	4498	4464
4498	4536	4527
4550	7001	4552
4562		7001
7001		

Table 3. Geocode dominant composition

Geocode	Dominant Composition (%)
4284	70: Haplic Acrisols
4324	80: Euctric Gleysols
4461	40: Haplic Acrisols, 20: Dystric Cambisols, 20: Humic Acrisols
4463	50: Haplic Acrisols, 20: Luvsiols, 20: Ferric Acrisols
4464	40: Haplic Acrisols, 20: Haplic Nitisols, 20: Ferric Acrisols
4476	80: Ferralic Cambisols
4498	40: Dystric Gleysols, 20: Umbric Gleysols, 20: Gleyic Cambisols
4527	60: Thionic Fluvisols, 30: Umbric Gleysols, 10: Dystric Gleysols
4536	30: Haplic Luvisols, 20: Chromic Luvisols, 20: Haplic Ferralsols
4550	80: Histosols. 10: Umbric Gleysols, 10: Dystric Fluvisols
4552	80: Histosols, 10: Umbric Gleysols, 10: Thionic Fluvisols
4562	60: Albic Arenasols, 30: Haplic Podzols
7001	Technosols

Elevation and slope angle are two topographical variables that frequently coincide with landslides, and have been deployed in multiple landslide susceptibility studies through ML models [30, 31]. With the data provided by the National Space Agency (NASA), Shuttle Radar Topography Mission (SRTM), the elevation and slope angle of the study areas were

evaluated [32]. In terms of elevation, Target 1 was determined to have a maximum elevation of 3973m, a mean of 530.988m, and a minimum of -12m. Target 2 has a maximum elevation of 1262m, a mean of 112.277m, and minimum of -35m. The transfer region has a maximum elevation of 1262m, a mean of 86.010m, and a minimum of -72m. Thus, the Transfer region is more similar in terms of elevation to Target 2, instead of Target 1. The minimum slope angle in each study area were constant with 0°. In Target 1, the maximum slope angle was 85.132°, with a mean of 17.282°. In Target 2, the maximum slope angle was 74.688°, with a mean of 7.429°. In the Transfer region, the maximum slope angle was 75.650°, with a mean of 8.089°. In terms of slope angle, the Transfer region has more similarities to Target 2, in contrast to Target 1. Thus, in can be concluded that topographically, the Transfer region is shares more in similarity to Target 2 in comparison to Target 1. The location of the study areas is depicted in Figure 1.

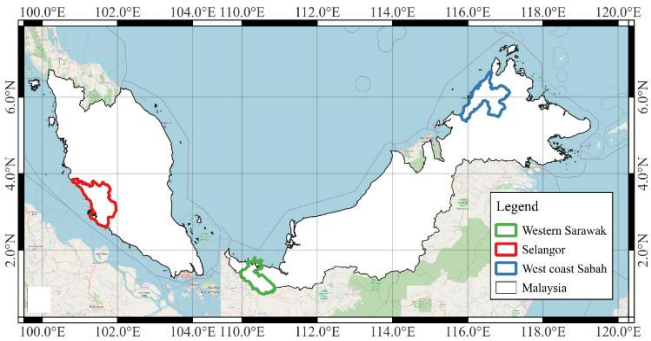

Figure 1. Location of study area

2.2 Methodology

This study was conducted in four primary phases (Figure 2).

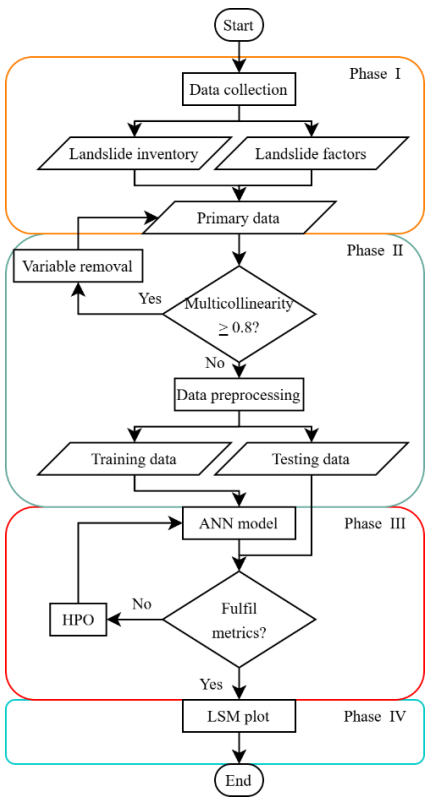

Figure 2. Study flowchart

The first phase was conducted to collect the developmental data required by the ANN models. The second phase was to check for multicollinearity amongst the variables in the primary dataset. The third phase was the ANN model developmental and evaluation phase where hyperparameters optimisation (HPO) and metrics evaluation was conducted. The final phase was to plot the LSMs.

2.2.1 Data collections and primary data preparation

To develop the ANN models, both landslides inventory and landslides factors were required. The data was detrimental for the ANN model as it is a supervised approach to predict landslide susceptibility.

Landslide inventory consists of landslides and non-landslides points. These variables are important at it helps the ANN model to make a distinction between the high risk and low risk areas in conjunction with the landslide factors [33]. For the target areas, the landslide points data were provided by NASA through the Global Landslide Catalogue for landslide triggered by rainfall [34]. As for the Transfer region, the landslide points data were provided by the local authorities. The non-landslide points on the other hand were obtained by randomly sampling points with slope angle less than or equal to 5° [7]. For the Transfer region, there were 53 landslides points, and 977 non-landslides points. For Target 1, there were 34 landslide points, and 199 non-landslide points. For Target 2, there were 41 landslide points, and 154 non-landslide points.

Landslide factors on the other hand are variables that causes landslides to occur which can be divided into triggering factors and conditioning factors. It was crucial to ensure that all of the variables selected are present in all study areas to enable the ANN model transfer, thus geological variables were not included [35]. The triggering variable in this study was rainfall, where the data was provided by CHIRPS for the average rainfall in mm pentad during the wet season. Rainfall is a significant triggering variable of landslides especially in the tropics where it is abundant during the wet season. It is responsible for the increase in pore water pressure that can greatly destabilise a slope forming material stability [36].

For the conditioning factors, it was represented through aspect, curvature, distance from roads (DFR), elevation, Normalised Difference Vegetation Index (NDVI), and slope angle.

The aspect data for the study areas were provided by NASA SRTM's processed from the digital elevation model (DEM) in Google's Earth Engine (GEE) [32]. Aspect refers to orientation of a slope face ranging from 0° North circling clockwise back to its origin. Different aspect has different levels of landslide susceptibility depending on the area of interest. This is due to the localised climate conditions of each aspect, that causes the weathering rate for the slope forming materials to be different from one to the other [37]. The difference in weathering conditions will ultimately lead to differing landslide susceptibility rate.

The curvature data was obtained by conducting a terrain analysis on the DEM through the System for Automated Geoscientific Terrain Analyses extension in Quantum Geographical Information System (QGIS) [38]. Curvature ranges from concave to flat to convex as it goes from the negative values to 0 to positive values [39]. As curvature goes from the positive to negative, the susceptibility towards landslide increases due to the tendency of ponding on concave curvatures, which allows rainfall to better penetrate the surface

soil [30].

DFR in this study was a measure of anthropogenic activity as not only does it correlate with the disturbance of the original slope to make way for the construction, but human settlement has tendency to be nearby a road for easier access [40]. The expansion of human settlements increases the area of impervious surface, which leads to an increase in runoff that could trigger landslide [41]. Thus, being near a road generally increases the likelihood of its surroundings to be more susceptible to landslides. The DFR data was obtained by querying road data using the Quick Open Stream Map service in QGIS, before converting it into a proximity map using QGIS's proximity buffer function [42].

The elevation data was supplied by NASA SRTM through GEE. As elevation increases the likelihood of landslides also increases due to the influence it has on the other landslide conditioning factors [43]. Furthermore, shallow landslides have been observed to occur more in areas with variations in the elevation [44].

NDVI is a measure of vegetation density, where areas with high NDVI tends to have lower susceptibility towards landslides due the support of roots which reinforce the slope forming materials [45]. Furthermore, areas with a low NDVI values are known to be more susceptible to landslides as not only does it lacks the root support, it also increases the rate of rainfall infiltration due to the lack of surface cover, which leads to saturated soils that are more likely to experience slope failure [46]. The NDVI data of this study was supplied by the United States Geological Survey Land Satellite [47].

Slope angle is an important variable to consider in a landslide susceptibility study. As slope angle increases, the stability of the slope forming material decreases making it more prone to landslide in comparison to areas with lower slope angle [48]. Furthermore, slope angle influences landslide motion characteristics with an increase in slope angle leading to a greater runout [49]. The slope angle data in this study was obtained from NASA SRTM through GEE.

After the landslide inventories, and landslide factors have been collected, the primary data was prepared. It was done by sampling the values of each landslide factors at each landslide and non-landslide points using the raster sampling tool in QGIS [50, 51].

2.2.2 Multicollinearity assessment and data preparation

Multicollinearity is not a favourable condition when it comes to developing an ML model. Input variables with multicollinearity amongst each other exceeding 0.8 is highly correlated [52]. When two input variables are highly correlated, it introduces redundancy issues to the ANN model, as it considers two input variables as a single variable, thus increasing the variable influence on the output unjustifiably [53]. In this study, the multicollinearity was accessed on the primary data through Microsoft's Excel data analysis plugin for correlation, where if a correlation amongst two input variables exceeds 0.8, one of it were to be removed from the dataset [54]. After no multicollinearity issue has been determined, the primary data undergoes data preprocessing. In this study, the landslide factors were processed through min-max scaling (Eq. (1)). Min-max scaling ensures the landslide factors to be scaled from 0 to 1, to avoid issues where a single variable would overpower the other variables that have smaller values, such as the values of elevation in comparison to slope angle [55, 56].

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where, $\hat{x}$ is the scaled x value, x_{min} is the minimum value of x, x_{max} is the maximum value of x, and x is the current value of x.

2.2.3 ANN model development and evaluation

The ANN models were developed in R through the "NeuralNet" package, that allows hyperparameter optimisation (HPO) such as architecture, maximum steps, and learning algorithm [57]. The training data was used to develop the ANN model whereas the testing was used to evaluate its predictive performance based on never-before-seen data. The evaluation metrics for the ANN models were Area Under the Receiver Operating Curve (AUC) and recall. AUC provides information on how well the ANN model distinguishes landslide cases and non-landslide cases based on the landslide factors [58]. It is obtained by evaluating the area under the curve for true positive rate (TPR) versus the false positive rate (FPR). Recall on the other hand is TPR, in a ML development where the 1s (landslide) data pool is at an imbalance to the 0s (non-landslide) as it is smaller, it is more important for the ML model to be able to predict more of 1s accordingly (Eq. (2)) [59]. Both AUC and Recall minimum score were set at 0.8, where if the scores were lower, HPO would be conducted. The importance of each landslide factors on landslide occurrence in the ANN models were determined through Garson's algorithm from the "NeuralNetTools" package in R.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

where, TP is true positive, and FN is false negative.

2.2.4 LSM plotting

To plot the LSM for each study area, the landslide factors data was converted into a numerical data frame using the "terra" package in R [60]. Then, the numerical data frame underwent the same min-max scaling process as the training data and the testing data. The ANN model was then deployed to plot the LSM where each target areas landslide susceptibility was predicted by its local ANN model and the Transfer region ANN model. The LSMs were important as it visualised the sensitivity of the ANN models to the variable importance [5].

3. RESULTS AND DISCUSSIONS

3.1 Multicollinearity assessment

The multicollinearity score for each study area primary data can be seen in Figure 3. In the primary data for the Transfer region, the most highly correlated landslide factors were slope angle and elevation with a correlation score of 0.70 (Figure 3(a)). From the correlation of landslide factors to landslide occurrence (fail), slope angle and elevation correlated highly with landslide occurrence with a correlation score of 0.88, and 0.58 respectively.

In Target 1's primary data, the same pattern was observed for the most highly correlated landslide factors, which were slope angle and elevation with a correlation score for 0.72 (Figure 3(b)). The same pattern persists for the correlation between landslide factors to landslide occurrences with a

correlation score between slope angle and landslide occurrence of 0.96.

As for Target 2's primary data, the most highly correlated landslide factors were rainfall and elevation (Figure 3(c)). This shows that in Target 2 region, as elevation increases, the rainfall intensity increases. Contradicting with the common conception of as DFR decreases, the likelihood of landslide increases, the correlation between DFR and landslide occurrences were in the negatives [30]. This could be explained by observing the correlation between DFR and slope angle, which was determined to have a score of -0.20, indicating that most of the roads in Target 2 are in areas with gentle slopes.

(a)	aspect	curvature	DFR	elevation	ndvi	rain	slope	fail
aspect	1.00							
curvature	-0.03	1.00						
dfr	-0.04	0.01	1.00					
elevation	-0.06	0.14	0.16	1.00				
ndvi	0.04	-0.02	0.24	0.26	1.00			
rain	-0.01	0.00	-0.16	0.06	-0.26	1.00		
slope	0.00	0.01	0.04	0.70	0.17	0.13	1.00	
fail	-0.03	0.05	0.02	0.58	0.15	0.15	0.88	1.00

(b)	aspect	curvature	DFR	elevation	ndvi	rain	slope	fail
aspect	1.00							
curvature	-0.03	1.00						
dfr	0.00	0.04	1.00					
elevation	0.01	0.06	0.63	1.00				
ndvi	-0.01	-0.05	-0.05	-0.17	1.00			
rain	-0.08	0.06	-0.02	-0.16	0.22	1.00		
slope	0.17	-0.03	0.44	0.72	-0.01	-0.27	1.00	
fail	0.18	-0.03	0.33	0.66	-0.01	-0.29	0.96	1.00

(c)	aspect	curvature	DFR	elevation	ndvi	rain	slope	fail
aspect	1.00							
curvature	-0.08	1.00						
dfr	0.02	-0.02	1.00					
elevation	-0.17	0.17	-0.17	1.00				
ndvi	0.18	0.10	0.39	-0.21	1.00			
rain	-0.06	0.01	0.19	0.59	-0.16	1.00		
slope	0.15	-0.10	-0.20	0.51	-0.07	0.28	1.00	
fail	0.11	-0.06	-0.24	0.55	-0.12	0.28	0.89	1.00

Figure 3. Multicollinearity score of (a) Transfer region, (b) Target 1, and (c) Target 2

3.2 HPO and evaluation

It was determined that the ANN models would only converge with the stepmax set at 1E+8. The final hyperparameters for all ANN models were 4 neurons in a single hidden layer, a learning rate of 0.001, a backpropagation learning algorithm, and a stepmax of 1E+8. In its own region, the Transfer region ANN model has achieved a recall score of 1.0 based on its own testing data. The results of the ANN model transfer and local ANN model prediction in terms of recall can be seen in Table 4. In Target 1, the local ANN model outperformed the transferred ANN model by 0.11 based on the recall score, where the local ANN model achieved a recall score 1.00 whereas the Transfer ANN model achieved a recall score of 0.89. In Target 2, the local ANN model outperformed the transferred ANN model by a recall score difference of 0.14, where the local ANN model achieved a recall score of 1.00, whereas the transferred ANN model achieved a recall score of 0.86.

As for the AUC score, both localised ANN model performed better locally in comparison to the transferred ANN models with Target 1 scoring an AUC of 1 while the Transfer model scored 0.988, and Target 2 ANN model scoring an AUC of 1 while the Transfer ANN model scored 0.995 (Table 4). However, in this study recall scores were more important as compared to the AUC score due to the lower number of samples for landslide points [61]. Thus, it can be concluded that the Transfer ANN model performed better in Target 1 than in Target 2. This can be explained by the multicollinearity matrix, which shows that the transfer region more closely resembled the primary data of Target 1 than that of Target 2 in terms of correlation.

Table 4. ANN models recall and AUC score

	Region			
	Target 1		Target 2	
ANN	Local	Transfer	Local	Transfer
Recall	1.00	0.89	1.00	0.86
AUC	1.00	0.988	1.00	0.995

3.3 Garson's variable importance and LSMs

The results of the Garson's algorithm on the importance of each landslide factors for respective ANN models can be seen in Table 5. The variable importance determines the influence that each landslide factor has on the landslide occurrence. In all three ANN models, the slope angle was the most influential landside factor.

Table 5. Garson's variable importance

Factors	ANN Model		
	Transfer	Target 1	Target 2
Aspect	0.053	0.162	0.114
Curvature	0.092	0.171	0.048
DFR	0.060	0.143	0.166
Elevation	0.113	0.058	0.125
NDVI	0.070	0.053	0.110
Rain	0.105	0.151	0.103
Slope	0.506	0.260	0.331

Based on the table, anthropogenic factor represented through NDVI and DFR is prevalent in Target 2 as compared to the other study areas. Indicating that, human activities played a significant role in landslide occurrence in Target 2. This was not observed in Transfer and Target 1. Although Target 1 multicollinearity matrix showed that slope angle has the highest correlation with landslide occurrence in the region, based on the Garson's variable importance analysis it was relatively low with a score of 0.260.

The resulting LSM from respective ANN models in each target area can be seen in Figure 4. It was observed that the distribution of landslide susceptibility risk was more linear when the localised ANN model was used compared to the transferred ANN model. Analysis the statistics for the LSMs of Target 1, it was revealed that the mean and the median susceptibility predicted by the local ANN model and the transferred ANN model were 0.312 and 0.133, and 0.285 and 0.006 respectively. It shows application of both ANN models, the results were positively skewed, however, when using the Transfer ANN model on Target 1, the skew was heavier. The

heavier skew correlated with the weightage of the Transfer ANN model, where the slope angle was the most prevalent factor. In comparison, Target 1 had a lower weight for slope angle (0.260) than the Transfer ANN model (0.506).

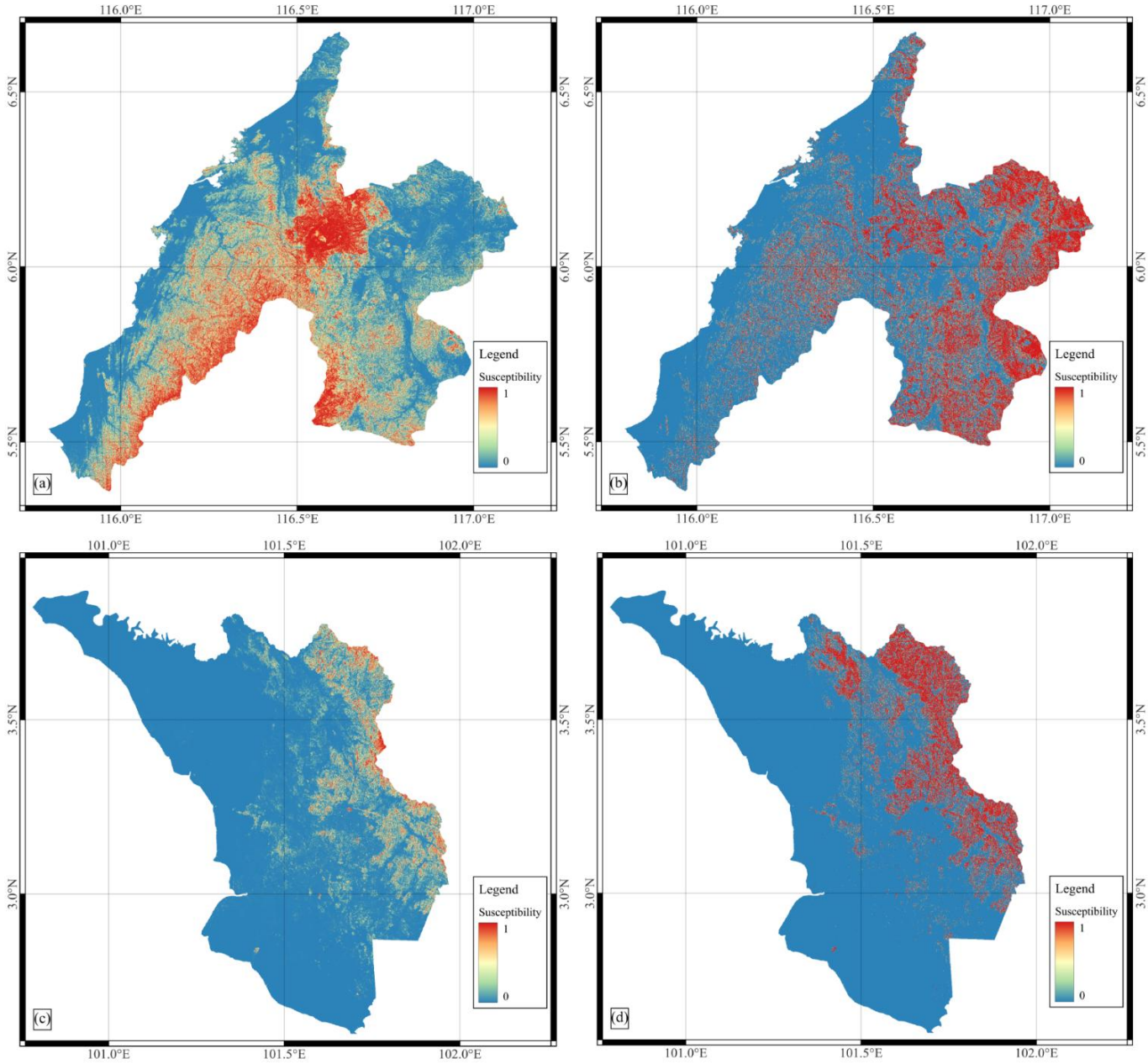

Figure 4. LSM of (a) Target 1 predicted by local ANN model, (b) Target 1 predicted by Transfer ANN model, (c) Target 2 predicted by local ANN model, and (d) LSM of Target 2 predicted by Transfer ANN model

4. CONCLUSIONS

This study aimed to conduct a comparative analysis of local and transferred ANN models in landslide susceptibility prediction in a tropical region. There were similarities and differences between the transfer region and Target 1 and Target 2. In all three regions, the slope angle was determined to have the most influence on landslide occurrence. The geological condition was the primary difference for all of the study areas.

For Target 2 statistics, the local ANN model achieved a mean of 0.068 and a median of 0.007 for landslide susceptibility. When deploying the Transfer ANN model on Target 2, the mean was 0.136 whereas the median was 0.0001. Both ANN models showed that landslide predictions were skewed to the right; however, a stronger skew was observed again in the Transfer ANN model predictions. This can be attributed to differences in variable importance between the target ANN models and the Transfer ANN model.

In Target 1, the Transfer ANN model achieved a recall of 0.89 and an AUC score of 0.988, in comparison to the local ANN model, which achieved a recall and AUC score of 1, respectively. In Target 2, the Transfer ANN model has achieved a recall of 0.86 and an AUC score of 0.995, in comparison to the local ANN model with a recall and AUC score of 1 respectively. The difference was explained through differences in variable importance for each region, as determined by Garson's algorithm. For the LSMs, the Transfer ANN model predictions in each target areas were heavier skewed to the right in comparison to the local ANN models

predictions. This study showed that the transfer was successful, as indicated by high performance metrics on the test data. The transferred ANN model indicated that in the tropical region, a region with scarce landslide related data can be predicted by a transferred ML model developed with a more comprehensive dataset. However, its applicability is limited to a preliminary assessment of landslide susceptibility, as the generalization of variable importance differs. To further improve performance, it is suggested that more data be obtained from local agencies or to combine datasets, as observed in studies on the Himalayas. It is recommended that future research explore the transferability of ML models to different tropical regions outside of Southeast Asia.

ACKNOWLEDGMENT

The authors acknowledge Ministry of Higher Education Malaysia, Fundamental Research Grant Scheme, No.: FRGS/1/2022/TK06/UNIMAS/01/1 and Universiti Malaysia Sarawak for supporting this project.

REFERENCES

[1] Akter, A., Noor, M.J.M.M., Goto, M., Khanam, S., Parvez, A., Rasheduzzaman, M. (2019). Landslide disaster in Malaysia: An overview. International Journal of Innovative Research & Development, 8(6): 292-302. https://doi.org/10.24940/ijird/2019/v8/i6/jun19058

[2] Rosly, M.H., Mohamad, H.M., Bolong, N., Harith, N.S.H. (2022). An overview: Relationship of geological condition and rainfall with landslide events at East Malaysia. Trends in Sciences, 19(8): 3464-3464. https://doi.org/10.48048/tis.2022.3464

[3] Achour, Y., Pourghasemi, H.R. (2020). How do machine learning techniques help in increasing accuracy of landslide susceptibility maps? Geoscience Frontiers, 11(3): 871-883. https://doi.org/10.1016/j.gsf.2019.10.001

[4] Najar, I.A., Ahmadi, R., Khalik, Y.K.A., Taib, S.N.L., Sutan, N.B.M., Ramli, N.H.B. (2023). Soil suffusion under the dual threat of rainfall and seismic vibration. International Journal of Design & Nature and Ecodynamics, 18(4): 849-860. https://doi.org/10.18280/ijdne.180411

[5] Vincent, S., Pathan, S., Benitez, S.R.G. (2022). Landslides in goa: A weight of evidence (WoE) approach for mapping. In 2022 International Conference on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC), pp. 154-160. https://doi.org/10.1109/ICMACC54824.2022.10093638

[6] Ben, S.K., Ahmadi, R., Bustami, R.A., Najar, I.A. (2023). Numerical investigation of seismic hazard and risk of murum hydro-electric dam (MHEP). In International Conference on Dam Safety Management and Engineering, pp. 855-869. https://doi.org/10.1007/978-981-99-3708-0_60

[7] Singh, H., Huat, B.B., Jamaludin, S. (2008). Slope assessment systems: A review and evaluation of current techniques used for cut slopes in the mountainous terrain of West Malaysia. Electronic Journal of Geotechnical Engineering, 13: 1-24.

[8] Ahmad, A.B., Ahmadi, R., Najar, I.A., Abidin, A.S.Z. (2024). Comprehension of energy-based methods for investigating soil suffusion uncertainties. International Journal of Design & Nature and Ecodynamics, 19(3): 733-743. https://doi.org/10.18280/ijdne.190303

[9] Alqawasmeh, H. (2023). Comparative assessment of various artificial neural network techniques for estimating the safety factor of road embankments. Geomate Journal, 24(103): 42-51. https://doi.org/10.21660/2023.103.3532

[10] Anis, Z., Wissem, G., Vali, V., Smida, H., Mohamed Essghaier, G. (2019). GIS-based landslide susceptibility mapping using bivariate statistical methods in North-western Tunisia. Open Geosciences, 11(1): 708-726. https://doi.org/10.1515/geo-2019-0056

[11] Rudra, R.R., Sarkar, S.K. (2023). Artificial neural network for flood susceptibility mapping in Bangladesh. Heliyon, 9(6): e16459. https://doi.org/10.1016/j.heliyon.2023.e16459

[12] Nugroho, F.S., Danoedoro, P., Arjasakusuma, S., Candra, D.S., Bayanuddin, A.A., Samodra, G. (2021). Assessment of Sentinel-1 and Sentinel-2 data for landslides identification using google earth engine. In 2021 7th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR), pp. 1-6. https://doi.org/10.1109/APSAR52370.2021.9688356

[13] Yarmohammad Touski, M., Veiskarami, M., Dehghani, M. (2019). Interferometric point target analysis (IPTA) for landslide monitoring. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42: 1079-1083. https://doi.org/10.5194/isprs-archives-xlii-4-w18-1079-2019

[14] Huang, F., Mao, D., Jiang, S. H., Zhou, C., Fan, X., Zeng, Z., Catani, F., Yu, C., Chang, Z., Huang, J., Jiang, B., Li, Y. (2024). Uncertainties in landslide susceptibility prediction modeling: a review on the incompleteness of landslide inventory and its influence rules. Geoscience Frontiers, 15(6): 101886. https://doi.org/10.1016/j.gsf.2024.101886

[15] Kritikos, T., Robinson, T.R., Davies, T.R. (2015). Regional coseismic landslide hazard assessment without historical landslide inventories: A new approach. Journal of Geophysical Research Earth Surface, 120(4): 711-729. https://doi.org/10.1002/2014jf003224

[16] Asadi, A., Baise, L. G., Chatterjee, S., Koch, M., Moaveni, B. (2024). Regional landslide mapping model developed by a deep transfer learning framework using post-event optical imagery. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 18(1): 186-210. https://doi.org/10.1080/17499518.2024.2316265

[17] Singh, A., Dhiman, N., Niraj, K.C., Shukla, D.P. (2024). Ensembled transfer learning approach for error reduction in landslide susceptibility mapping of the data scare region. Scientific Reports, 14(1): 29060. https://doi.org/10.1038/s41598-024-76541-4

[18] Depicker, A., Govers, G., Jacobs, L., Vanmaercke, M., Uwihirwe, J., Campforts, B., Kubwimana, D., Mateso, J.M., Bibentyo, T.M., Nahimana, L., Smets, B., Dewitte, O. (2024). Mobilization rates of landslides in a changing tropical environment: 60-year record over a large region of the East African Rift. Geomorphology, 454: 109156. https://doi.org/10.1016/j.geomorph.2024.109156

[19] Huggel, C., Khabarov, N., Korup, O., Obersteiner, M.,

Clague, J.J., Stead, D. (2012). Physical impacts of climate change on landslide occurrence and related adaptation. In Landslides, J. J. Clague and D. Stead, Eds. Cambridge: Cambridge University Press, pp. 121-133. https://doi.org/10.1017/cbo9780511740367.012
[20] Najar, I.A., Ahmadi, R.B., Jamian, M.A.H., Hamza, H., Ahmad, A., Sin, C.H. (2022). Site-specific ground response analysis using the geotechnical dataset in moderate seismicity region. International Journal of Mechanics, 16(1): 37-45. https://doi.org/10.46300/9104.2022.16.5
[21] Najar, I.A., Ahmadi, R., Khalik, Y.K.A., Mohamad, N.Z., Jamian, M.A.H., Najar, N.A. (2022). A framework of systematic land use vulnerability modeling based on seismic microzonation: A case study of miri district of Sarawak, Malaysia. International Journal of Design & Nature and Ecodynamics, 17(5): 669-677. https://doi.org/10.18280/ijdne.170504
[22] Rahman, H.A., Mapjabil, J. (2017). Landslides disaster in Malaysia: An overview. Health, 8(1): 58-71.
[23] Mohd Anip, M.H. (2024). Permulaan monsun timur laut 2024/2025.
[24] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. Scientific Data, 2(1): 1-21. https://doi.org/10.1038/sdata.2015.66
[25] FAO and IIASA. (2023). Harmonized world soil database version 2.0. FAO, International Institute for Applied Systems Analysis (IIASA), Rome and Laxenburg. https://doi.org/10.4060/cc3823en
[26] Gray, J.M., Bishop, T.F., Wilford, J.R. (2016). Lithology and soil relationships for soil modelling and mapping. Catena, 147: 429-440. https://doi.org/10.1016/j.catena.2016.07.045
[27] Department of Statistics Malaysia. (2025). Gross Domestic Product by State. OpenDOSM.
[28] Department of Statistics Malaysia. (2023). Learn more about your area today! OpenDOSM.
[29] Caves, B. (2021). Leveraging on multidisciplinary expertise for landslide disaster risk reduction and management: A case study of a limestone hill rockfall hazard assessment. Sains Malaysiana, 50(8): 2179-2191. https://doi.org/10.17576/jsm-2021-5008-04
[30] Youssef, A.M., Pourghasemi, H.R. (2021). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geoscience Frontiers, 12(2): 639-655. https://doi.org/10.1016/j.gsf.2020.05.010
[31] Nur, B.A. (2018). Landslides susceptibility mapping at Gunung Ciremai National Park. In E3S Web of Conferences, 31: 12010. https://doi.org/10.1051/e3sconf/20183112010
[32] Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D. (2007). The shuttle radar topography mission. Reviews of Geophysics, 45(2): 1-33. https://doi.org/10.1029/2005RG000183
[33] Badola, S., Mishra, V.N., Parkash, S., Pandey, M. (2023). Rule-based fuzzy inference system for landslide susceptibility mapping along national highway 7 in Garhwal Himalayas, India. Quaternary Science Advances, 11: 100093. https://doi.org/10.1016/j.qsa.2023.100093
[34] Kirschbaum, D., Stanley, T., Zhou, Y. (2015). Spatial and temporal analysis of a global landslide catalog. Geomorphology, 249: 4-15. https://doi.org/10.1016/j.geomorph.2015.03.016
[35] Avand, M., Kuriqi, A., Khazaei, M., Ghorbanzadeh, O. (2022). DEM resolution effects on machine learning performance for flood probability mapping. Journal of Hydro-Environment Research, 40: 1-16. https://doi.org/10.1016/j.jher.2021.10.002
[36] Zhang, Z., Zeng, R., Meng, X., Zhao, S., Wang, S., Ma, J., Wang, H. (2023). Effects of changes in soil properties caused by progressive infiltration of rainwater on rainfall-induced landslides. Catena, 233: 107475. https://doi.org/10.1016/j.catena.2023.107475
[37] Guo, Y., Ma, C. (2023). Elucidating the role of soil hydraulic properties on aspect-dependent landslide initiation. Hydrology and Earth System Sciences, 27(8): 1667-1682. https://doi.org/10.5194/hess-27-1667-2023
[38] Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J. (2015). System for automated geoscientific analyses (SAGA) v. 2.1. 4. Geoscientific model development, 8(7): 1991-2007. https://doi.org/10.5194/gmd-8-1991-2015
[39] Ahmed, I., Pan, N.D., Debnath, J., Bhowmik, M., Bhattacharjee, S. (2024). Flood hazard zonation using GIS-based multi-parametric Analytical Hierarchy Process. Geosystems and Geoenvironment, 3(2): 100250. https://doi.org/10.1016/j.geogeo.2023.100250
[40] Singha, C., Swain, K.C., Moghimi, A., Foroughnia, F., Swain, S.K. (2024). Integrating geospatial, remote sensing, and machine learning for climate-induced forest fire susceptibility mapping in Similipal Tiger Reserve, India. Forest Ecology and Management, 555: 121729. https://doi.org/10.1016/j.foreco.2024.121729
[41] Yang, C., Wang, J., Li, S., Xiong, R., Li, X., Gao, L., Guo, X., Ma, C., Xiong, H., Qiu, Y. (2024). Landslide susceptibility assessment and future prediction with land use change and urbanization towards sustainable development: The case of the li river valley in Yongding, China. Sustainability, 16(11): 4416. https://doi.org/10.3390/su16114416
[42] Trimaille, E. (2023). Download OSM data thanks to the Overpass API. You can also open local OSM or PBF files. A special parser, on top of OGR, is used to let you see all OSM keys available.
[43] Chung, C.J.F., Fabbri, A.G. (2003). Validation of spatial prediction models for landslide hazard mapping. Natural Hazards, 30: 451-472. https://doi.org/10.1023/b:nhaz.0000007172.62651.2b
[44] Dias, H.C., Gramani, M.F., Grohmann, C.H., Bateira, C., Vieira, B.C. (2021). Statistical-based shallow landslide susceptibility assessment for a tropical environment: A case study in the southeastern Brazilian coast. Natural Hazards, 108(1): 205-223. https://doi.org/10.1007/s11069-021-04676-y
[45] Doan, V.L., Nguyen, B.Q.V., Pham, H.T., Nguyen, C.C., Nguyen, C.T. (2023). Effect of time-variant NDVI on landside susceptibility: A case study in Quang Ngai province, Vietnam. Open Geosciences, 15(1): 20220550.

https://doi.org/10.1515/geo-2022-0550
[46] Hoa, P.V., Tuan, N.Q., Hong, P.V., Thao, G.T.P., Binh, N.A. (2023). GIS-based modeling of landslide susceptibility zonation by integrating the frequency ratio and objective-subjective weighting approach: a case study in a tropical monsoon climate region. Frontiers in Environmental Science, 11: 1175567. https://doi.org/10.3389/fenvs.2023.1175567
[47] Aziz, N.F.A., Yaacob, N., Yusof, A.L., Omar, H. (2025). Statistical analysis for forest fire factors using geography information system (GIS) and remote sensing imagery. Journal of Advanced Research in Applied Sciences and Engineering Technology, 45(2): 177-190. https://doi.org/10.37934/araset.45.2.177190
[48] Donnarumma, A., Revellino, P., Grelle, G., Guadagno, F.M. (2013). Slope angle as indicator parameter of landslide susceptibility in a geologically complex area. Landslide Science and Practice: Volume 1: Landslide Inventory and Susceptibility and Hazard Zoning, 425-433. https://doi.org/10.1007/978-3-642-31325-7_56
[49] Wu, Y.B., Duan, Z., Peng, J.B., Zhang, Q. (2022). Influences of slope angle on propagation and deposition of laboratory landslides. Earth Surface Dynamics Discussions, 2022: 1-34. https://doi.org/10.5194/esurf-2022-38
[50] Janzandr. (2020). Version: [2196] One Click Raster Stacking 1.1.0. QGIS.
[51] Najar, I.A., Ahmadi, R., Amuda, A.G., Mourad, R., Bendary, N.E., Ismail, Ismial, I., Nabilah, A.B., Tang, S. (2025). Advancing soil-structure interaction (SSI): A comprehensive review of current practices, challenges, and future directions. Journal of Infrastructure Preservation and Resilience, 6(5): 1-25. https://doi.org/10.1186/s43065-025-00118-2
[52] Iban, M.C., Sekertekin, A. (2022). Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin provinces, Turkey. Ecological Informatics, 69: 101647. https://doi.org/10.1016/j.ecoinf.2022.101647
[53] Islam, A.R.M.T., Talukdar, S., Mahato, S., Kundu, S., Eibek, K.U., Pham, Q.B., Kuriqi, A., Linh, N.T.T. (2021). Flood susceptibility modelling using advanced ensemble machine learning models. Geoscience Frontiers, 12(3): 101075. https://doi.org/10.1016/j.gsf.2020.09.006
[54] UMW. (2015). Microsoft excel manual. University of Mary, Washington.
[55] Mohammed, M.A. (2024). Effect of using numerical data scaling on supervised machine learning performance. Global Libyan Journal, 67: 1-21. https://doi.org/10.37376/glj.vi67.5903
[56] Najar, I.A., Ahmadi, R., Amuda, A.G., Mourad, R., Bendary, N.E., Tang, S., Ahmad, A., Khalik, Y.K.A., Haris, M.N. (2025). Laboratory Investigation of Soil Suffusion through Particle Size Distribution and Hydraulic Conductivity Analysis, 48(2): 1-9. https://doi.org/10.1007/s40996-025-017
[57] Fritsch, S., Guenther, F., Guenther, M.F. (2019). Package 'neuralnet'. Training of Neural Networks, 2: 30.
[58] Chowdhury, M.S., Rahman, M.N., Sheikh, M.S., Sayeid, M.A., Mahmud, K.H., Hafsa, B. (2024). GIS-based landslide susceptibility mapping using logistic regression, random forest and decision and regression tree models in Chattogram District, Bangladesh. Heliyon, 10(1): e23424. https://doi.org/10.1016/j.heliyon.2023.e23424
[59] Al-Ruzouq, R., Shanableh, A., Jena, R., Gibril, M.B.A., Hammouri, N.A., Lamghari, F. (2024). Flood susceptibility mapping using a novel integration of multi-temporal sentinel-1 data and eXtreme deep learning model. Geoscience Frontiers, 15(3): 101780. https://doi.org/10.1016/j.gsf.2024.101780
[60] Bivand, R., Hijmans, R.J., Pebesma, E., Sumner, M.D. (2024). Package 'terra,'. CRAN.
[61] Dai, L., Zhu, M., He, Z., He, Y., Zheng, Z., Zhou, G., Wang, C., Ren, J., Tang, H., Liu, Q., Huang, F., Li, Z., Li, M., Wang, Z., Li, M., Jiang, L. (2021). Landslide risk classification based on ensemble machine learning. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, pp. 3924-3927. https://doi.org/10.1109/IGARSS47720.2021.9553034