# Application of geographical information system (GIS) using artificial neural networks (ANN) for landslide study in Langat Basin, Selangor

To cite this article: Siti Norsakinah Selamat *et al* 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1064** 012052

# Application of geographical information system (GIS) using artificial neural networks (ANN) for landslide study in Langat Basin, Selangor

**Siti Norsakinah Selamat[1], Nuriah Abd Majid[1], Mohd Raihan Taha[2] and Ashraf Osman[3]**

[1] Institute for Environment and Development (LESTARI), Universiti Kebangsaan Malaysia, 43600, Selangor, Malaysia
[2] Department of Civil Engineering, Faculty of Engineering and Build Environment, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia
[3] Department of Engineering, The Palatine Centre, Durham University, Stockton Road, Durham DH1 3LE, UK

E-mail: nuriah@ukm.edu.my

**Abstract**. The landslide was recognized as the most common geologic hazard around the world. The assessment of the relationship landslide conditioning factors is a critical step in managing landslide hazards and risks. Several models have been made to develop the landslide model in recent years. The Artificial Neural Networks (ANN) model was used in this study to develop a landslide model and to identify the most important landslide conditioning factors. Eight conditioning factors, including elevation, slope, aspect, curvature, lithology, soil series, Topographic Wetness Index (TWI), and rainfall, were selected and analyzed using the Geographical Information System (GIS) approach. The multilayer perceptron module and one hidden layer method extracted weighted conditioning factors. The landslide model was validated using the area under the curve (AUC) method. This model validation showed a success rate for training and testing is 0.876, respectively. This study found curvature is the most crucial factor affecting landslide occurrence in the Langat Basin with a 0.213 weight index, followed by rainfall (0.143) and elevation (0.141). Finally, the landslide model can be used as an indicator to identify the most important landslide conditioning factors and assess the relationship between these factors and landslide occurrences.

Keywords: Landslide, Artificial Neural Networks, GIS, Langat Basin

## 1. Introduction

In practically every country, a landslide is a complicated natural hazard that results in property damage and fatalities. Landslide is a geological hazard that has been identified as the most detrimental natural hazards phenomenon [1]. These phenomena are significant geological motions mainly triggered by climate or geophysical factors. A thorough assessment of landslide-prone areas is required to address this problem and significantly reduce the impact of damage resulting from landslides. Therefore, the enormous disaster potential of landslides has drawn some academics to study and measure landslide susceptibilities in recent decades [2, 3].

Landslides are becoming a more common disaster worldwide due to the destruction of forests, uncontrolled development, and rapid urbanization due to population growth [4, 5]. Deforestation on the highlands for development purposes such as roads and homes usually result in landslide phenomena. Landslides in Malaysia are closely related to constructing new development such as roads, highways, dams, and housing areas developed in a hilly area [6]. Landslides are common in Malaysia due to heavy rain, particularly during the monsoon season, caused by the southwest monsoon from May to September and the northeast monsoon from November to March [7]. A landslide incident in the Hulu Langat region in 2011 involving Children's Hidayah Madrasah Al-Taqwa Orphanage was due to heavy rainfall and killed 16 members. This tragedy alerted researchers and governmental agencies to the seriousness of the landslide issue.

The landslide occurrences are exceedingly complex and influenced by various landslide conditioning factors such as topographical, hydrological, geological, and environmental factors [8-11]. In general, the topographical, geological, and hydrological factors are regarded as static since they are relatively stable [12]. On the other hand, environmental factors are usually dynamic and ever-changing in nature. An important phase in landslide study is determining the relationship between landslide conditioning factors and landslides occurrence. Thereby, the factors influencing the occurrence of landslides can be evaluated. Therefore, identifying and analyzing the relationships among the most important factors is an essential process.

In order to reduce the future harm caused by landslide events, this research has focused on developing a landslide model and analyzing the relationships between the conditioning factors that influence landslide occurrences. Therefore, this research aims to develop a landslide model using Artificial Neural Networks (ANN) to investigate the crucial conditioning factors contributing to landslide events in the Langat Basin.

## 2. Study Area

Langat Basin is located at 02º 50' 48" north latitude and 101º 40' 48" east longitude. This study area is one of Selangor's largest river basins, with a total area of about 2,938 km². The Langat Basin has been designated as a transboundary basin shared by three states: Selangor, Negeri Sembilan, and the Federal Territories of Putrajaya. This study area is located in a hilly area with rapid expansion, leading to landslides. The Langat Basin study area is represented in Figure 1.


**Figure 1**. Study area at Langat Basin, Selangor

## 3. Methodology

The purpose of this study is to identify and investigate the relationships between landslide occurrence and landslide conditioning factors using a Geographic Information System (GIS) approach and artificial neural network modelling. The data used in this research was obtained from several sources. The Digital Elevation Model (DEM) from the Department of Survey and Mapping Malaysia (JUPEM), lithology from Department of Mineral and Geosciences Malaysia, soil series from Department of Agriculture, and rainfall from the Department Meteorological Malaysia. In general, a landslide inventory map was created by integrating historical records, field investigations, and satellite photos interpretation. Historical landslides can be identified using high-resolution satellite images [8]. Therefore, landslide inventory in this study was selected using historical landslide events and google earth imagery analysis.

Next is a preparation of the landslide conditioning factors. The selection and preparation of landslide conditioning factors were critical steps in landslide study. Landslide conditioning factors were selected based on the information obtained from the literature connected to the study area. Elevation, slope, curvature, aspect, lithology, soil series, Topographic wetness Index (TWI), and rainfall were considered landslide conditioning factors in this study. These selected conditioning factors were interpreted and analysed using ArcGIS 10.3 software. The layers used a 5 m x 5 m spatial resolution with a rectified skewed orthomorphic (RSO) coordinate system. Lastly is the development of landslide modelling. Machine learning methods have recently become a cornerstone in solving spatial modelling issues in the field of natural hazards [9]. This study used ANN as a machine learning method to investigate landslide issues. ANN was conducted using a multilayer layers perceptron with one hidden layer.

## 4. Results and discussion

Eight landslide conditioning factors were considered in this study, as shown in figure 2. The elevation layer ranged from -79.15 to 1448.25 meters, according to the data. The negative value represents an elevation in the water bodies area, while the positive value represents elevation on land. Langat Basin is dominated by slopes ranging from less than 15º to 74º degrees. The slope was classified into four classes which are <15º, 16º - 25º, 26º to 36º, and > 35º. Curvature can influence landslide movement direction and resistance stress since this conditioning factor controls the speed and convergence of displaced and water flowing down the slope [10]. A negative value represents a concave surface for the curvature conditioning factor. While a zero value represents a flat surface, and the positive value indicates the convex surface. The aspect indicates the change of highest rate in the direction of downslope from a specific cell to the surrounding areas, where it is also considered a slope direction [11]. The direction is identified like a compass for a whole circle direction with 360º. The aspect map was constructed into nine directions such as flat (–1), north (0–22.5 and 337.5–360), northeast (22.5–67.5), east (67.5–112.5), southeast (112.5–157.5), south (157.5–202.5), southwest (202.5–247.5), west (247.5–292.5), and northwest (292.5–337.5).

Lithology is one of the most crucial conditioning elements in landslide study since they are a variety of rocks and soils with distinct inner structures and mineral composition [12]. The lithology map was classified into several classes following the standard classification by the Department of Mineral and Geoscience Malaysia in this study. Lithology at the Langat Basin area was dominated by two classes: clay, silt, peat, minor gravel, and acid intrusives. The soil factor is important element in the research of landslides. The different physical and mechanical features of soil can affect surface water infiltration and groundwater flow [13]. The geological structure of the bedrock or subsurface has a significant impact on soil development [14]. The soil series map was classified into several classes regarding the Department of Agriculture Malaysia. The soil series at the Langat Basin area is dominated by peat, steepland, and Selangor-kangkong.

TWI represents the upslope contributing area and slope to calculate the steady-state wetness and water flow across the region [15]. The index is extensively used in hydrological processes to quantify water flow disposition and topographic impact [16]. The TWI at the Langat Basin area ranges from 0.440 to 22.056. Rainfall is classified as a natural phenomenon where it is regarded as a potential contributor that significantly impacts the occurrence of landslides [17]. The rainfall affected the nature of the soil, such as the decrease in the strength of the soil structure, causing liquefaction of the mass of soil. The rainfall data were obtained from the Department of Meteorological Malaysia. The rainfall map was created using Thiessen polygon analysis, was separated by the station of rainfall.

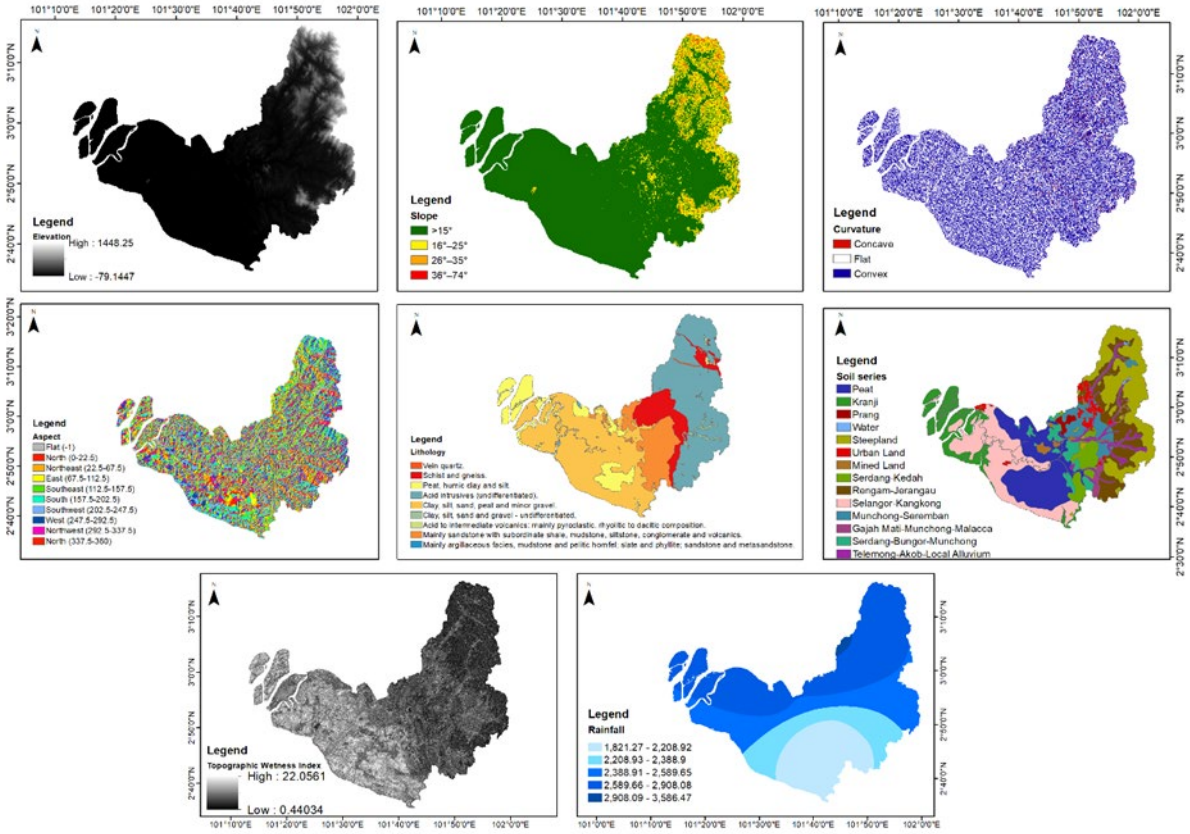


**Figure 2.** Map of conditioning factors

The Artificial Neural Networks approach creates a landslide model. Multilayer Perceptron (MLP) module was used to build up the landslide model in this study. The data were randomly assigned to 70% for training and 30% for testing. The training dataset is used to determine the weights and develop the model. Meanwhile, the testing data identifies errors during the training mode and avoids overtraining. The activation function for the hidden layer in this study was a hyperbolic tangent, while the output layer was the SoftMax function. The landslide model was validated using the area under the curve (AUC) method. This model validation showed a success rate for training and testing is 0.876, respectively. Figure 3 shows the AUC graph for this study. This study used automatic architecture, where the selection was automatic chose three nodes for the hidden layer and two nodes for the output layer to code the dependent variable outcome. In this study, dependent variable outcomes were classified into two classes: one represented landslide event and 0 represented non landslide event, and figure 4 shows the neural network diagram.

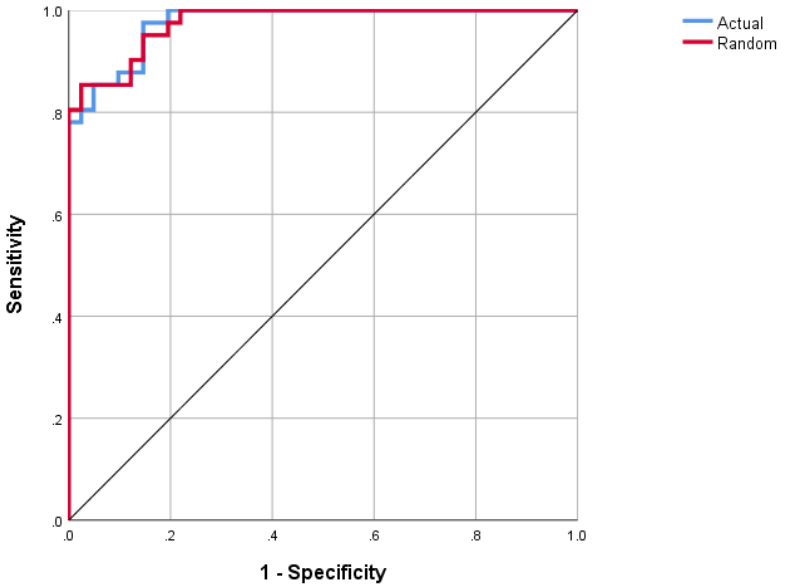


**Figure 3.** Area under the curve (AUC) graph

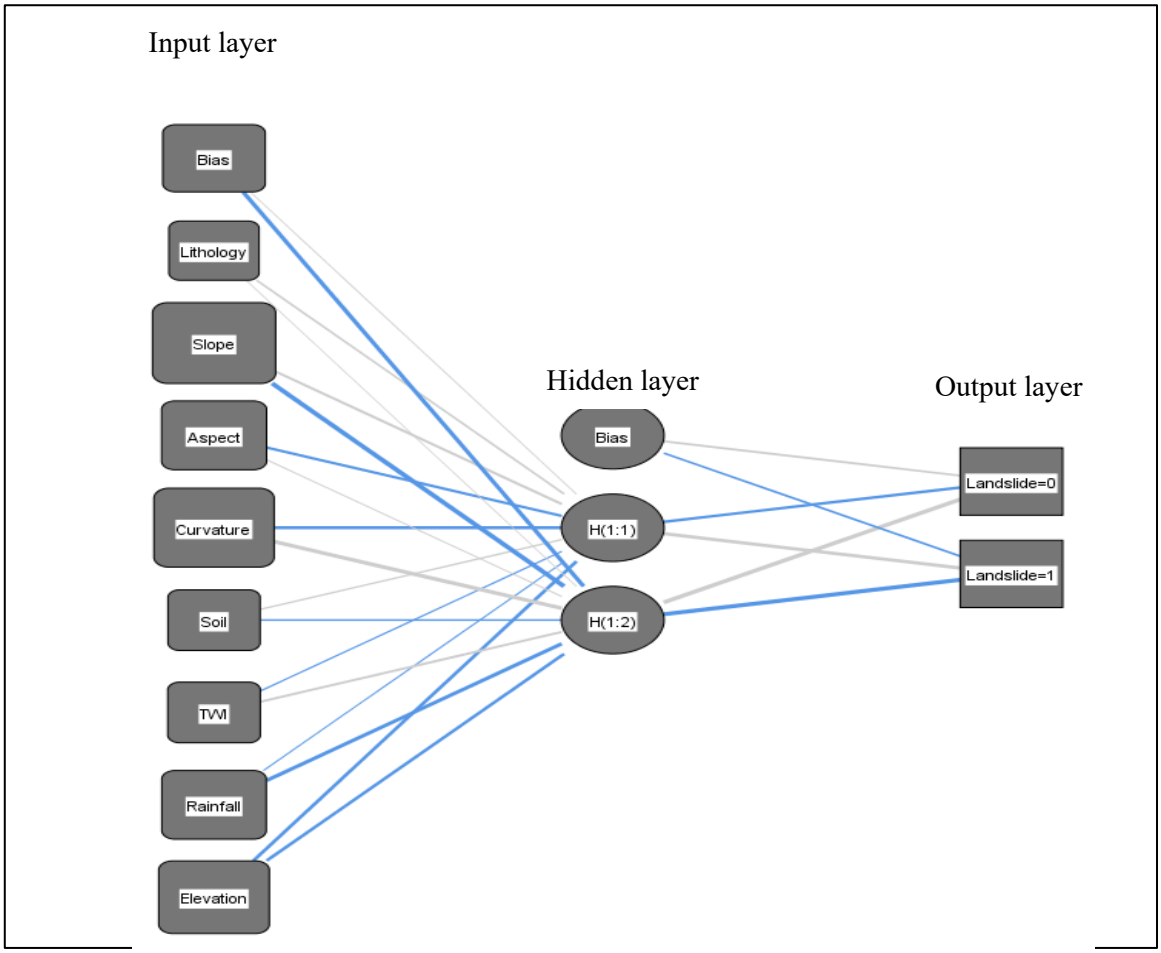


**Figure 4.** Neural networks diagram

An artificial neural network determines various factors' relative importance or weight. This study found that rainfall is the most crucial factor affecting landslide occurrence in the Langat Basin. The weighted index of curvature in this study is 0.213. A negative value represents a concave surface for the curvature factor. A flat surface is represented by a zero number, whereas a positive value represents a convex surface. Curvature can influence landslide movement direction and resistance stress since this conditioning factor controls the speed and convergence of displaced and water flowing down the slope [18]. The second important conditioning factor due to contributing landslide occurrence is rainfall. The weighted index for the rainfall factor is 0.143. Rainfall is classified as a natural phenomenon where it is regarded as a potential contributor that significantly impacts the occurrence of landslides [19]. The third important conditioning factor is an elevation with a 0.141 weight index. The difference between a higher and lower point in the terrain significantly influences the incidence of landslides [12].

## 5. Conclusion

A landslide is a complex geological phenomenon that occurred due to one or more landslide conditioning factors. The purpose of this study was to use machine learning to determine the most important landslide conditioning factors that lead to landslide occurrences in the study area. Artificial neural networks model was used to develop landslide modelling in the Langat Basin, Selangor using an integration of various landslide condition factors such as elevation, slope, curvature, aspect, lithology, soil series, Topographic wetness Index (TWI), and rainfall. Landslide modelling was computed using the Multilayer Perceptron (MLP) module. The area under the curve (AUC) method verifies the landslide model. This model validation resulted in a 0.876 success rate for training and testing. Determining the importance of conditioning factors in landslide study was an important part. The study revealed that the most important conditioning factors according to the accuracy measure of an ANN model are curvature, rainfall, and elevation. The importance of landslide conditioning factors could be used as an indicator to analyse the relationship between landslide conditioning factors and landslide occurrences.

## Acknowledgement

This research is supported by Dana Padanan Antarabangsa (MyPAIR), Natural Environment Research Council (NERC): NEWTON/1/2018/TK01/UKM/2.

## References

[1] Bera A, Mukhopadhyay B P and Das D 2019 Landslide hazard zonation mapping using multi-criteria analysis with the help of GIS techniques: a case study from Eastern Himalayas, Namchi, South Sikkim *Natural Hazards* **96** 935-59

[2] Ji J, Gao Y, Lü Q, Wu Z, Zhang W and Zhang C 2019 China's early warning system progress *Science* **365** 332-

[3] Saha A K, Gupta R P, Sarkar I, Arora M K and Csaplovics E 2005 An approach for GIS-based statistical landslide susceptibility zonation—with a case study in the Himalayas *Landslides* **2** 61-9

[4] Cui Y, Cheng D, Choi C E, Jin W, Lei Y and Kargel J S 2019 The cost of rapid and haphazard urbanization: lessons learned from the Freetown landslide disaster *Landslides* **16** 1167-76

[5] Hasnat G T, Kabir M A and Hossain M A 2018 Major environmental issues and problems of South Asia, particularly Bangladesh *Handbook of environmental materials management* 1-40

[6] Azmi A S M, Salleh W A R W M and Nawawi A H 2013 Cognitive behaviour of residents toward living in landslide prone area: Ulu Klang *Procedia-Social and Behavioral Sciences* **101** 379-93

[7] Abd Majid N, Rainis R and Ibrahim W M M W 2018 Spatial Modeling Various Types of Slope Failure Using Artificial Neural Network (ANN) In Pulau Pinang, Malaysia/Pemodelan

Ruangan Pelbagai Jenis Kegagalan Cerun Menggunakan Rangkaian Saraf Buatan (ANN) Di Pulau Pinang, Malaysia. *Jurnal Teknologi* **80** 135-46

[8] Youssef A M, Pourghasemi H R, Pourtaghi Z S and Al-Katheeri M M 2016 Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia *Landslides* **13** 839-56

[9] Sevgen E, Kocaman S, Nefeslioglu H A and Gokceoglu C 2019 A novel performance assessment approach using photogrammetric techniques for landslide susceptibility mapping with logistic regression, ANN and random forest *Sensors* **19** 3940

[10] Dou J, Yamagishi H, Pourghasemi H R, Yunus A P, Song X, Xu Y and Zhu Z 2015 An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan *Natural Hazards* **78** 1749-76

[11] Kadirhodjaev A, Kadavi P R, Lee C-W and Lee S 2018 Analysis of the relationships between topographic factors and landslide occurrence and their application to landslide susceptibility mapping: a case study of Mingchukur, Uzbekistan *Geosciences Journal* **22** 1053-67

[12] Pham B T, Tien Bui D, Prakash I and Dholakia M B 2017 Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS *Catena* **149** 52-63

[13] Saha A and Saha S 2020 Application of statistical probabilistic methods in landslide susceptibility assessment in Kurseong and its surrounding area of Darjeeling Himalayan, India: RS-GIS approach *Environment, Development and Sustainability* **23** 4453-83

[14] Chawla A, Pasupuleti S, Chawla S, Rao A C S, Sarkar K and Dwivedi R 2019 Landslide Susceptibility Zonation Mapping: A Case Study from Darjeeling District, Eastern Himalayas, India *Journal of the Indian Society of Remote Sensing* **47** 497-511

[15] Pourghasemi H, Pradhan B, Gokceoglu C and Moezzi K D 2013 A comparative assessment of prediction capabilities of Dempster–Shafer and weights-of-evidence models in landslide susceptibility mapping using GIS *Geomatics, Natural Hazards and Risk* **4** 93-118

[16] Lee S, Lee M-J and Jung H-S 2017 Data mining approaches for landslide susceptibility mapping in Umyeonsan, Seoul, South Korea *Applied Sciences* **7** 683

[17] Tsunetaka H 2021 Comparison of the return period for landslide-triggering rainfall events in Japan based on standardization of the rainfall period *Earth Surface Processes and Landforms* **46** 2984-98

[18] Kim S W, Chun K W, Kim M, Catani F, Choi B and Seo J I 2021 Effect of antecedent rainfall conditions and their variations on shallow landslide-triggering rainfall thresholds in South Korea *Landslides* **18** 569-82

[19] Dou J, Tien Bui D, P. Yunus A, Jia K, Song X, Revhaug I, Xia H and Zhu Z 2015 Optimization of causative factors for landslide susceptibility evaluation using remote sensing and GIS data in parts of Niigata, Japan *PloS one* **10** e0133262