INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES
Vol 14, Issue 4, (2024) E-ISSN: 2222-6990

Urban Resilience in the Face of Natural Hazards: Leveraging Machine Learning to Assess Landslide Risk in Kuala Lumpur, Malaysia

Syaidatul Azwani Zulkafli & Nuriah Abd Majid
Institute Environment and Development (LESTARI), Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
Corresponding Author Email: nuriah@ukm.edu.my

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v14-i4/20831 DOI:10.6007/IJARBSS/v14-i4/20831

Published Date: 08 April 2024

Abstract
Landslides represent a significant threat to urban areas globally, causing substantial loss of life, property damage, and infrastructure disruption. The rapid urbanization witnessed in Kuala Lumpur, Malaysia since the 1970s has heightened the susceptibility to landslides, driven by factors such as vegetation removal and slope cutting. This study employs logistic regression (LR), a supervised machine learning technique, to develop a landslide model for Kuala Lumpur. LR is chosen for its simplicity and effectiveness in landslide susceptibility mapping. The methodology involves collecting and pre-processing landslide inventory data, extracting relevant factors from geospatial data, and applying LR to model the relationship between landslides and these factors. The resulting model is validated using an independent landslide dataset, demonstrating a good overall accuracy of 74.1%, with a sensitivity of 84.7% and specificity of 63.5%. The study concludes that LR serves as a valuable tool for landslide hazard assessment and risk management in Kuala Lumpur. The developed model offers guidance for land-use planning and infrastructure development, contributing to Sustainable Development Goal (SDG) 11 by fostering inclusive, safe, resilient, and sustainable cities. By mitigating landslide risk, the model contributes to the protection of lives and livelihoods, promotes sustainable urbanization, and enhances Kuala Lumpur's resilience to natural hazards.
Keywords: Landslide, Geographic Information System (GIS), Logistic Regression (LR), Urban, Malaysia.

Introduction
Natural hazards, particularly landslides appeal as the 3rd most notorious disaster in the world imparting severe damages and necessitating extensive mitigation methods (Zillman, 1999). Landslide impacts can vary depending on their severity. Most of the time, landslide brings crucial fatalities, injuries, and property loss (Mahmud et al., 2013). For years, landslides have affected nearly 4.8 million people, and more than 18,000 deaths were caused by landslides

between 1998 to 2017 (World Health Organization, 2022). According to NASA, there were more than 8,935 landslides that took place around the world with approximately 1,120 cases recorded in Southeast Asia.
Before this, Malaysia had experienced multiple landslides. In regions where steep terrain and urban residential areas coexist as in many parts of Malaysia, the risk is higher for their physical environment and socio-economy. Heavy rainfall and significant ground shaking when integrated with anthropogenic and other spatial interactions were deemed major contributors to landslides (Suzen and Kaya, 2011). Landslides in developing countries, however, have a high and wide range impact due to rapid urbanization trends as it is highly associated with the changes in landscape patterns (Althuwaynee et al., 2015; Saadatkhah et al., 2014). Recently, netizens were left in shock by four deaths following a major landslide in Taman Bukit Permai 2. Unfortunately, that was not the first case involving tragic demise. Malaysia had also experienced a few major landslides back in the day. The infamous Highland Tower collapsed due to significant human errors that claimed 48 lives (Kazmi, et al., 2017; Sardi and Razak, 2019; Zubaidi, et al., 2020). Bukit Antarabangsa held a record as the second most horrific landslide in 2008 at only 1.4 km away from Highland Tower. These horrific events indirectly suggest that landslides have continually occurred in the main cities, such as Kuala Lumpur ever since the development started taking place in the 1970s. On top of that, as Kuala Lumpur achieved a 100% rate of urbanization, this area often involves the modification of natural topography, such as removing vegetation, cutting slopes, and depressions from the land surface which later leads to slope instability. High demand for housing also forced new residences to be built on high terrains, inflating the pressure of the ground due to the heavy materials (Pradhan and Lee, 2010; Azmi et al., 2013). Zulkafli and Abd Majid (2020) stated that a total of 17 landslide events had taken place in 2010 resulting in the highest number of cases that occurred in a year. However, during the final quarter of 2021 and early 2022, there were nine out of 115 locations were categorized as critical (Ahzan, 2022).
Therefore, landslide susceptibility mapping is crucial for urban and land management to prevent such disasters, particularly in high-density populated areas. Landslide susceptibility assessment is a complex process as it requires an in-depth analysis of the spatial relationship between various landslide factors and landslide occurrences (Sujatha and Sridhar, 2021). Based on local terrain, a region's susceptibility to landslides is the probability that a landslide would occur in certain areas (Tebar et al., 2022). Landslide susceptibility can be measured by different factors, such as geological, morphological, or human factors (Table 1) as landslides are influenced by one or more factors and occasionally correspond to one another (Leonardi et al., 2022; Leonardi et al., 2020). According to (Budimir et al., 2015), many studies used factors to predict susceptibility that is relatively stable, namely geology, slope, aspect, and vegetation. These factors modify landscapes over a longer period. Hence, this study will not be using the rainfall measurement factor as it acts on a much shorter time frame.

Table 1
Landslide Factors Classification by Varnes (1984)

Geological	Morphological	Human
Weak or sensitive materials	Tectonic or volcanic uplift	Excavation of slope or its toe
Weathered materials	Glacial rebound	Loading of slope or its crest
Sheared, jointed, or fissured materials	Fluvial, wave, or glacial erosion of slope toe or lateral margins	Drawdown (of reservoirs)
Adversely oriented discontinuity	Subterranean erosion (solution, piping)	Deforestation
	Deposition loading slope or its crest	Irrigation
The contrast in permeability and/or stiffness of materials	Vegetation removal (by fire, drought)	Mining
	Thawing	Artificial vibration
	Freeze-and-thaw weathering Shrink-and-swell weathering	Water leakage from utilities

Although accurately predicting landslide occurrences remains the greatest challenge to this day, landslide susceptibility mapping is still crucial for urban and land management to prevent such disasters, especially in high-density populated areas (Alcantara-Ayala et al., 2017). Therefore, in today's established technologies, machine learning (ML) techniques provide solutions as ML produces predictions, performs clustering, extracts association features, and makes decisions from given information, which are coming to the fore (Ma et al., 2020). ML can be categorized into supervised and unsupervised learning. In supervised learning, classification and regression are two types of problems in data mining where classification problems use an algorithm to accurately assign test data into specific categories, including support vector machines, decision trees, and random forests, while, regression problems use an algorithm to understand the relationship between dependent and independent variables (Delua, 2021). The commonly used regressions in landslide studies are linear regression and logistic regression. On the other hand, unsupervised learning models are used for three main tasks: clustering, association, and dimensionality reduction. These two approaches convey a distinct difference as supervised learning uses labeled input and output data, while an unsupervised learning algorithm does not. The versatility of ML approaches, together with the quantity of landslide-related data acquired over time, has made ML a commonly utilized analytic tool for modeling complex landslide problems by weighting its factors (Ma et al., 2020; Tehrani et al., 2021; Li et al., 2022). Hence, to analyze the relationship between the landslide occurrences and the contributing factors that were later used to develop a landslide susceptibility map, a supervised algorithm, Logistic Regression (LR) is employed in this current study.

LR is one of the most used multivariate analyses in constructing a landslide susceptibility map, which is significant for classifying landslide susceptibility classes (Reichenbach et al., 2018; Sekarlangit et al., 2022). Furthermore, LR is also a supervised learning algorithm that uses a logistic function to map the input variables to categorical dependent variables. LR has been established and proven to be highly reliable for assessing landslide susceptibility, providing larger area coverage (Nwazelibe et al., 2023).

Study area and Materials

Area of Interest, Kuala Lumpur, Malaysia

Malaysia's capital city, Kuala Lumpur is located at 3.1390° N, 101.6869° E covering an area of approximately 243 km^2 with an average elevation of 81.95 m (Figure 1). According to the World Population Review (2020), Kuala Lumpur is divided into districts comprising nearly 8 million population and is predicted to rise to 9 million in another decade. In Kuala Lumpur, the hottest month is March with an average temperature of 28°C (82°F), while the coldest month is January with a temperature of 27°C (81°F) (Holiday Weather, 2022). However, Kuala Lumpur was one of the regions in Malaysia that experienced severe impact when the Air Pollution Index (API) exceeded 200 due to the haze conditions increase in the number of hot spots in 2019 (Annual Report, 2019). Kuala Lumpur is widely known for its diverse landmarks, including Petronas Twin Towers and KL Tower. Kuala Lumpur has become one of the most efficient regions to migrate to over the years as it provides economic opportunities and employment. As of today, internal migration can be seen as a bigger issue because it contributes to the rising population size thus leading to immense urban growth. To accommodate sufficient space for the migrants, urbanization in the Kuala Lumpur area has become a concern to nature as overdevelopment is one of the factors of landslides.

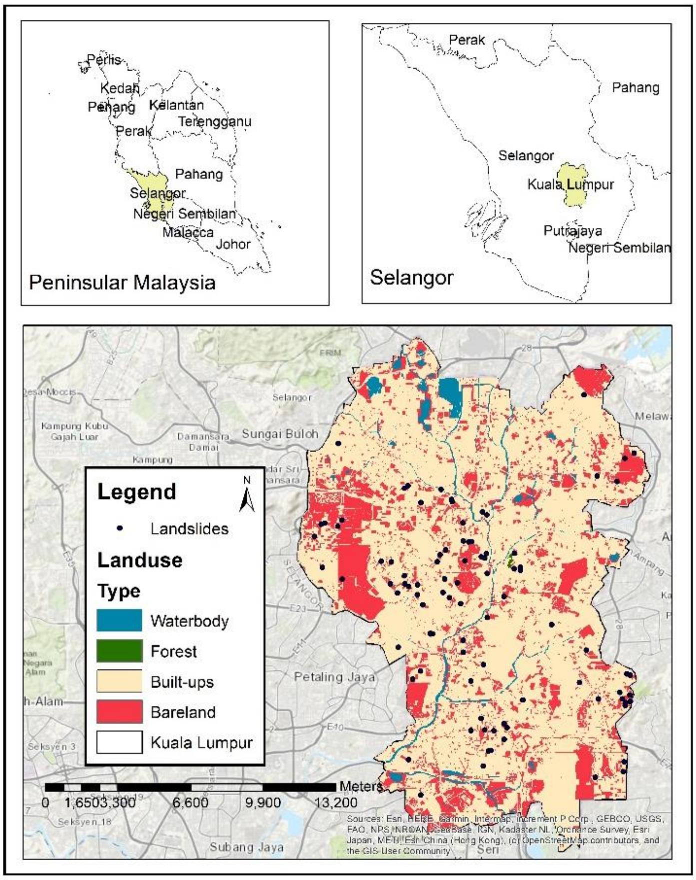


Figure 1. Study area map, Kuala Lumpur, Malaysia.

Data Preparation

A landslide inventory contains the primary input of the location of past and recent landslides. In this study, the inventory was prepared through past studies records, reports, and Google Earth image interpretation which then went through several field surveys for validation across distinct boundaries. A total of 100 landslides were identified across the area, most

concentrated in the Northwest and center parts of Kuala Lumpur. Another 100 random datasets representing non-landslide points were also created using create random points analyst in ArcGIS 10.8.2 software.
In the meantime, there are no guidelines for selecting landslide factors (Ayalew et al., 2005; Yalcin, 2008; Chen et al., 2018). Therefore, landslide causative factors were generally selected based on their influence on landslide occurrences in the selected study area, scientific literature, and data availability (Table 1). In this study, there were eight causative parameters considered in influencing landslides, including distance density factors (roads, rivers, and faults), topography factors (slope angle, slope aspect, and curvature) lithology type, and land use type.

Digital Elevation Model (DEM)
To analyze terrain characteristics, a 30-meter by 30-meter digital elevation model (DEM) was constructed using a contour map obtained from the Department of Survey and Mapping Malaysia. This DEM was generated within ArcGIS software using the 3D Analyst extension. Subsequently, the DEM was utilized to derive the slope angle (Figure 2(a)) and slope aspect (Figure 2(b)) for further analysis. In Malaysia, slopes have been classified into four categories since 2002, constructed by The Department of Minerals and Geoscience Malaysia (JMG) as official guidelines on hillside development (Gue and Wong, 2009). Slopes ranging below 15° were classified as Class I meanwhile Class II slopes were determined between 15° to 25°. On the other hand, Class III was reserved between 25° to 35 °, and any slope angle greater than 35° was classified as Class IV. Thus, this study also classified slopes into the stated classes. The slope angle is a measure of the steepness of a slope. It is calculated as the angle between the slope and a horizontal plane. In the context of landslides, slope angle is an important factor because it can influence the stability of a slope. Landslides are more likely to occur on slopes with steeper angles. This is because the weight of the overlying material is greater on steeper slopes, which can put more stress on the underlying material and make it more likely to fail. The slope aspect is one of the crucial landslide-influencing factors considered in most studies.

Distance to Roads
Roads and highways data were sourced from an open database, comprising polylines spanning across the entirety of Kuala Lumpur. Utilizing ArcGIS 10.8.2 software, an analysis employing Euclidean distance measurement was conducted to generate a thematic map illustrating the proximity of various locations to roads. The resulting map was categorized into five distinct classes using a natural breaks classification method. The distance to roads emerges as a significant factor in this study's landslide analysis, given Kuala Lumpur's intricate network of highways and road infrastructure (as depicted in Figure 2(c)).

Distance to Rivers
The study utilized river network data obtained from the Department of Irrigation and Drainage (DID) in the form of polylines. These polylines represent the branches of the river network and water bodies across Kuala Lumpur, including the Klang River, Gisir River, Kerayong River, Bunus River, Belongkong River, Gombak River, and Jinjang River. The river network data plays a crucial role in landslide susceptibility modeling, as the proximity to rivers and water bodies can influence landslide occurrence (Hervas et al., 2017).To incorporate the river network data into the landslide susceptibility model, an Euclidean distance analysis was performed (Burrough & McDonnell, 1998). This analysis calculates the distance from each

pixel location in the study area to the nearest river or water body. The resulting distance map represents the physical relationship between the terrain and the river network, providing valuable information for assessing landslide susceptibility. The Euclidean distance analysis was chosen for its simplicity and effectiveness in capturing the spatial relationship between the study area and the river network (Figure 2(d). Other distance measures, such as Manhattan distance and geodesic distance, could also be considered depending on the specific characteristics of the study area and the desired level of accuracy (Wang et al., 2018).

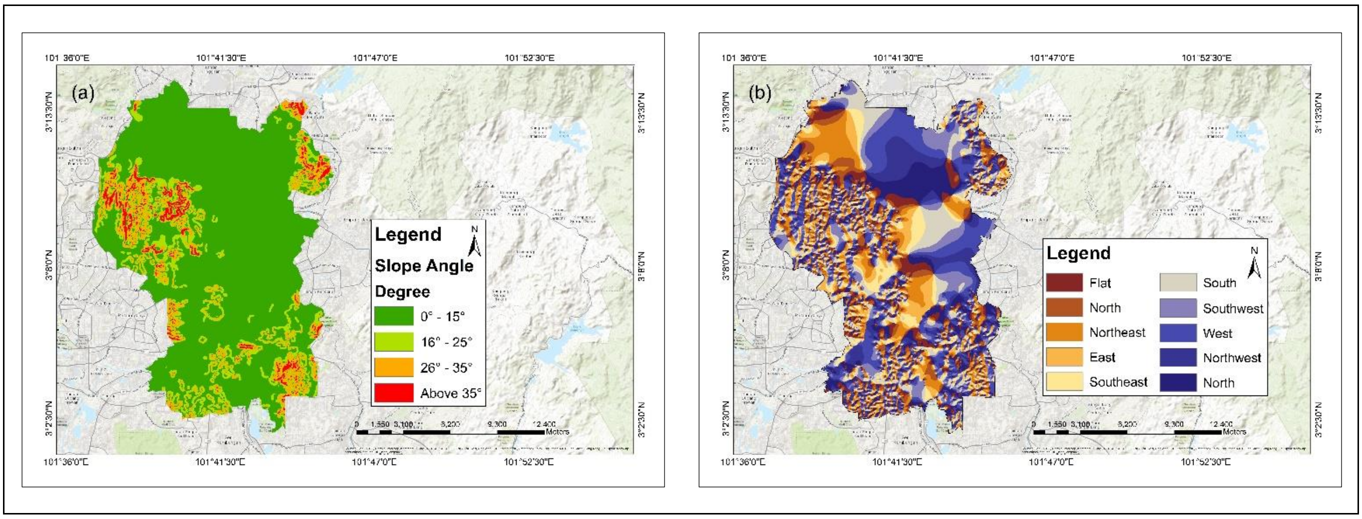
Distance to Faults
Kuala Lumpur's fault lines were extracted from the geological map provided by the Department of Mineral and Geoscience Malaysia. Distance to faults is a measure of how close a location is to a known fault line. Fault lines are areas where the Earth's crust is fractured and can move relative to each other. Earthquakes can occur when these fault lines move, and the closer a location is to a fault line, the more likely it is to experience an earthquake. In the context of Kuala Lumpur, Malaysia, distance to faults is an important factor to consider when assessing earthquake hazards. The city is located in a seismically active region, and there are several active faults in the area. The Kuala Lumpur Fault, for example, is a major strike-slip fault that runs through the city center. The Department of Mineral and Geoscience Malaysia has published a geological map of Peninsular Malaysia that shows the location of known fault lines. This map can be used to calculate the distance to faults for any location in Kuala Lumpur. For example, the distance to the Kuala Lumpur Fault for a location at 3.1450° N, 101.6850° E is approximately 0.0064 kilometers. This means that the location is very close to the fault line and is at high risk of experiencing an earthquake. The Department of Mineral and Geoscience Malaysia also has provided a geology map used to extract lithology types (Figure 2(e)).

Lithology
Figure 2(f), which depicts lithology, it is evident that landslide occurrences in Kuala Lumpur are primarily associated with sandstone, accompanied by subordinate shale, mudstone, siltstone, conglomerate, and volcanic materials. The shale and sandstone, originating from sedimentary rocks within the Kenny Hill formation, have undergone weathering and metamorphosis, resulting in the formation of metasediments such as schist, quartzite, and phyllite (Sanusi et al., 2017). Lithological profiles of the acidic intrusive region, commonly referred to as the Kuala Lumpur granite, exhibit a predominant orientation from the northwest to the southeast, with some areas being surrounded by limestone lithology in the southwest of the study area. Furthermore, similar lithological profiles can be sporadically found in the upper northwest and southeast, with some portions surrounded by limestone to the southwest. Landslide incidents associated with schist and gneiss are less prevalent in the southeast region of Kuala Lumpur. Conversely, a significant proportion of reported landslide incidents occur in areas characterized by limestone lithology, locally known as the Kuala Lumpur Limestone. This limestone terrain is marked by a thin topsoil layer supporting vegetation, underlain by alluvial soil rich in heavy minerals and tin-bearing soil. Extensive quarrying activities on slopes and cliffs over recent decades, primarily for limestone extraction as a primary construction material, have posed significant threats to the stability of the soil layers in this region (Althuwaynee & Pradhan, 2017).

Soil Series
In Kuala Lumpur, there are several types of soil series, including mined land, Munchong Seremban, Rongam Jorangau, Serdang Kedah, Steepland, Telemong-local alluvium, and urban

land. (Figure 2(g).A diverse range of soil series exists, each contributing to the city's unique landscape and environment. One prominent type is mined land, where soil composition has been significantly altered due to past mining activities, resulting in distinct characteristics. Additionally, the Munchong Seremban and Rongam Jorangau soil series are notable, each exhibiting specific profiles and properties that contribute to their identification within the region. Serdang Kedah soil, named after its prevalent area, also plays a role in Kuala Lumpur's soil diversity. Furthermore, areas characterized by steep slopes often have steepland soil, influenced by erosion patterns and nutrient distribution. Telemong-local alluvium, formed from sediment deposition by local rivers, is common in low-lying areas and floodplains. Finally, urban land soil represents areas impacted by human activities, such as construction and pollution, resulting in unique soil characteristics compared to rural environments. Together, these various soil series contribute to the dynamic landscape and environmental diversity of Kuala Lumpur.

Land use
The land use elements in the northwest to northeast areas of Kuala Lumpur exhibit a higher level of safety compared to those in the western and southern regions (Figure 2(h)). Urban areas with dense populations, including residential, commercial, industrial, and utility zones, experienced the highest percentages of land use elements affected by landslides (Althuwaynee & Pradhan, 2017). The proliferation of residential constructions on hilltops has increased significantly due to the scarcity of flat land (Gue & Tan, 2003), which may lead to alterations in water drainage patterns from highland to lowland areas in the future. Furthermore, it is observed that the occurrence of landslides is greatly influenced by the lithological characteristics of the land surface (Dhianaufal et al., 2018).



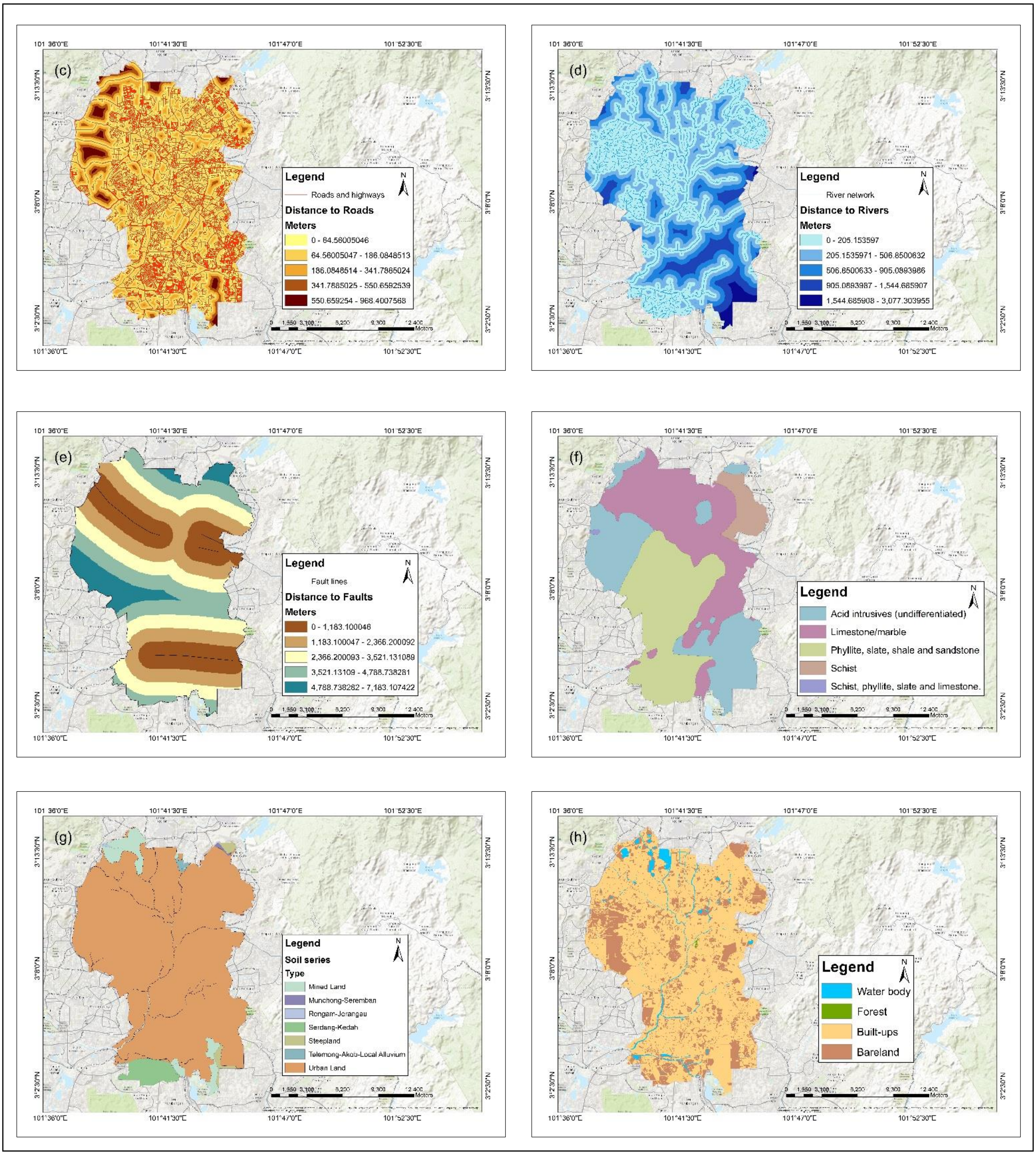


Figure 2. Physical characteristics of landslide factors

Logistic Regression (LR)

Logistic Regression (LR) is a statistical method widely used in various fields, including landslide susceptibility mapping. It's a multivariate analysis technique that establishes a relationship between a binary dependent variable (landslide occurrence or non-occurrence) and a set of independent variables (landslide causative factors). LR is particularly well-suited for landslide

susceptibility mapping due to its ability to handle multiple input factors and generate probability estimates for landslide occurrence.
Mathematical Representation:
LR employs a logistic function to transform a linear combination of independent variables (X1, X2, ..., Xn) into a probability value between 0 and 1. The probability of landslide occurrence (P) is represented by:

$$P = \frac{1}{1 + (exp^{-z})}$$

- P: Probability of landslide occurrence
- Z: Linear combination of causative factors

where P is the probability of landslide occurrence and Z a linear combination of casual X_i factors. Z can be expressed as

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots \beta_n X_n$$

- β0: Constant term
- βi: Coefficient associated with the ith independent variable (Xi)

The coefficients (βi) represent the strength and direction of the relationship between each causative factor and landslide occurrence. Positive coefficients indicate a positive relationship, while negative coefficients imply a negative relationship. In landslide susceptibility mapping, LR is employed to analyze a set of landslide inventory data and a collection of landslide causative factors. The landslide inventory data consists of locations where landslides have occurred, while the causative factors represent various environmental and geological conditions that contribute to landslide susceptibility.
LR is trained using the landslide inventory data to determine the coefficients (βi) for each causative factor. These coefficients represent the relative importance of each factor in influencing landslide occurrence. Once the coefficients are determined, the LR model can be used to predict landslide susceptibility for any location within the study area. Logistic regression (LR) has emerged as a prominent statistical method for landslide susceptibility mapping due to its effectiveness in handling multiple input factors and generating probability estimates of landslide occurrence. LR offers several distinct advantages over other susceptibility mapping techniques, making it a preferred choice among researchers (Akgün and Turkmenoglu, 2004; Pradhan, 2008; An et al., 2010).
One of the primary advantages of LR lies in its interpretability (Akgun and Turkmenoglu 2004). LR provides a direct interpretation of the relationship between landslide causative factors and landslide occurrence through the coefficients (βi) obtained in the model. These coefficients represent the relative importance of each factor in influencing landslide susceptibility, allowing researchers to gain insights into the underlying mechanisms of landslide occurrence (Pradhan, 2008; An et al., 2010).
LR ability to handle multiple independent variables makes it particularly well-suited for landslide susceptibility mapping, as landslide occurrence is often influenced by a complex interplay of environmental and geological factors (Akgun and Turkmenoglu, 2004; Pradhan, 2008; An et al., 2010). LR can effectively analyze a large number of causative factors, including elevation, slope angle, land use, rainfall, and distance to faults, providing a comprehensive

assessment of landslide susceptibility (Akgun and Turkmenoglu, 2004; Pradhan, 2008; An et al., 2010).
Unlike binary classification methods, LR generates probability estimates for landslide occurrence, providing a more nuanced assessment of landslide susceptibility (Pradhan 2008; An et al. 2010). These probability estimates allow researchers to identify areas with varying degrees of landslide risk, enabling more targeted planning and risk mitigation strategies (Pradhan, 2008; An et al., 2010). Logistic regression remains a valuable tool for landslide susceptibility mapping due to its interpretability, ability to handle multiple factors, and probability estimation capabilities. However, it's essential to consider its limitations and carefully evaluate the quality of the data and the choice of causative factors to ensure the robustness of the LR model.

Results
Table 1 presents the results of a logistic regression model aimed at predicting landslide occurrences based on diverse factors. The Classification Table assesses the model's performance comprehensively by comparing observed and predicted values. The model accurately predicted 87 cases of no landslides (0) and correctly identified 116 instances of actual landslides (1). However, it inaccurately predicted 21 cases of landslides in the "No Landslide" category and 50 cases of no landslides in the "Landslide" category. The overall correct prediction percentage stands at 74.1%, underscoring the model's accuracy in evaluating landslide factors. The mentioned cut value of 0.500 serves as a threshold: predicted probabilities exceeding this value are classified as landslides, while those below are labeled as no landslides. This table provides crucial insights into the logistic regression model's effectiveness in predicting landslides, emphasizing the necessity for a nuanced interpretation. The overall percentage, also termed the accuracy rate, gauges the logistic regression model's proficiency in predicting correct outcomes across observations. Calculated by dividing correctly classified observations by the total, the 74.1% overall percentage denotes the model's accurate classification of 74.1% of all landslides and non-landslides. This high accuracy rate suggests the model's reasonable effectiveness in predicting landslides. However, it's imperative to note that the overall percentage isn't the sole performance measure; sensitivity, specificity, and precision are equally crucial. Sensitivity, representing the correctly classified proportion of landslides, is at 84.7%, signifying the model's adept identification of 84.7% of actual landslides. Specificity, reflecting the correctly classified proportion of non-landslides, stands at 63.5%, indicating the model's accurate identification of 63.5% of actual non-landslides. Precision, the proportion of model-classified landslides that are actual landslides, is at 50%, implying that half of the model-classified landslides were indeed landslides. In summary, the logistic regression model appears reasonably effective for predicting landslides. The 74.1% overall accuracy rate is high, and the 84.7% sensitivity underscores the model's proficiency in identifying actual landslides. Nevertheless, the 63.5% specificity suggests a tendency for false positives, potentially causing unnecessary alarm or resource wastage.

Table 1
Logistic Regression Model Percentage on Landslide Factors

Classification Table					
	Observed		Predicted		
			Landslide		Percentage Correct
			0 (No Landslide)	1 (Landslides)	
Step 1	Landslide	0	87	50	63.5
		1	21	116	84.7
	Overall Percentage				**74.1**
a. The cut value is .500					

Table 2 presents a detailed examination of the statistical significance of various factors in the logistic regression model, gauged through the reported significance values (Sig.). These values play a crucial role in determining the influence of each variable on the model's predictive capability and the likelihood of its impact on landslide occurrences. Aspect Value (Asp_Val) and Curvature Value (Curv_Val) exhibit non-significant p-values of 0.292 and 0.328, respectively, suggesting that these variables may not significantly contribute to the model's ability to predict landslides. Conversely, Slope Value (SL_Val) has a notably low p-value of 0.031, indicating statistical significance and implying a substantial impact on the model's predictive power. Forestation Value (F_Val) and Ridge Value (Ri_Val) exhibit higher p-values of 0.994 and 0.240, respectively, suggesting that they may not be statistically significant predictors of landslide occurrences. In contrast, Rock Value (Ro_Val) demonstrates a low p-value of 0.022, signifying its statistical significance and implying a meaningful influence on landslide predictions. The p-values associated with Land Use Value (LU_Val) and its subcategories (LU_Val(1), LU_Val(2), LU_Val(3)) are generally high, with an overall p-value of 0.416. This indicates that, as a categorical variable, land use might not significantly impact the model. Similar observations are made for Soil Value (Soil_Val), with an overall p-value of 0.971, suggesting limited statistical significance. However, specific soil categories, like Soil_Val(3), exhibit a p-value of 1.000, indicating a lack of statistical significance. Lithology Value (Lit_Val) displays varying p-values across its categories, with Lit_Val(1) and Lit_Val(2) showing highly significant values (<.001), indicating their substantial impact on predicting landslides. Rainfall Value (RF_Val) demonstrates a marginal p-value of 0.086, suggesting moderate significance in influencing landslide predictions. The Constant term exhibits a p-value of 1.000, as anticipated, indicating its lack of statistical significance. In summary, the significance values elucidate the relative importance of each factor in the logistic regression model. Variables like Slope Value and Rock Value emerge as statistically significant contributors, while others, such as Forestation Value and certain Land Use and Soil categories, may have limited impact. These findings guide further exploration and prioritization of influential factors in landslide risk assessment and mitigation planning.

=

Table 2
Significant Values of Each Factor

		B	S.E.	**Sig.**	Exp(B)	95% C.I.for EXP(B) Lower	Upper
Step 1[a]	Asp_Val	.001	.001	.292	1.001	.999	1.004
	Curv_Val	-.344	.352	.328	.709	.356	1.412
	SL_Val	.051	.024	**.031**	1.053	1.005	1.103
	F_Val	.000	.000	.994	1.000	1.000	1.000
	Ri_Val	.001	.001	.240	1.001	1.000	1.002
	Ro_Val	-.003	.001	**.022**	.997	.994	1.000
	LU_Val			.416			
	LU_Val(1)	22.077	18281.917	.999	3871942575.409	.000	.
	LU_Val(2)	20.066	18281.917	.999	518187909.341	.000	.
	LU_Val(3)	20.442	18281.917	.999	754478461.636	.000	.
	Soil_Val			.971			
	Soil_Val(1)	1.225	1.943	.528	3.404	.076	153.342
	Soil_Val(2)	.435	1.397	.756	1.544	.100	23.890
	Soil_Val(3)	-18.943	40192.970	1.000	.000	.000	.
	Soil_Val(4)	.004	2.027	.998	1.004	.019	53.393
	Lit_Val			**<.001**			
	Lit_Val(1)	-1.900	.481	<.001	.150	.058	.384
	Lit_Val(2)	.672	.392	.087	1.958	.907	4.225
	Lit_Val(3)	-.785	.525	.135	.456	.163	1.277
	RF_Val	-.004	.002	.086	.996	.992	1.001
	Constant	-11.035	18281.918	1.000	.000		

Discussion

The outcomes of the logistic regression model's significance values offer a nuanced discussion regarding the influential factors in predicting landslide occurrences. Analyzing these values is pivotal for understanding the statistical importance of each variable and its overall impact on the model's performance. Examining the specific results reveals key insights. Firstly, the slope value (SL_Val) emerges as a significant predictor, with a p-value of 0.031, indicating that terrain inclination significantly influences landslide predictions, aligning with geological expectations. Similarly, the rock value (Ro_Val) is identified as a substantial predictor with a low p-value of 0.022, suggesting that geological composition, particularly the presence of rock, plays a crucial role in determining landslide occurrences. On the other hand, curvature value (Curv_Val) and aspect value (Asp_Val) do not exhibit statistical significance with p-values of 0.328 and 0.292, respectively, suggesting that these terrain features may not be strong predictors of landslides in the studied context. Categorical variables like land use value (LU_Val) and soil value (Soil_Val) show relatively high p-values of 0.416 and 0.971, indicating that, as a whole, they may not significantly contribute to the model. Notably, the lack of statistical significance for specific categories within Soil_Val (3) further emphasizes the nuanced impact of individual categories within these variables.

Lithology value (Lit_Val) stands out as an influential factor, particularly Lit_Val(1) and Lit_Val(2), with highly significant p-values of <.001. This suggests that certain lithological characteristics significantly impact landslide occurrences, underscoring the importance of geological composition in landslide prediction. Rainfall value (RF_Val) demonstrates a marginally significant p-value of 0.086, implying that precipitation may play a moderate role in landslide predictions, aligning with established knowledge that heavy rainfall can trigger landslides. Variables such as forestation value (F_Val) and ridge value (Ri_Val) exhibit high p-values of 0.994 and 0.240, respectively, suggesting that these factors may not be statistically significant in predicting landslides in the studied model. Understanding the statistical significance of these factors provides crucial insights for prioritizing interventions and mitigation strategies. Variables like slope and rock value, deemed statistically significant, become focal points for targeted measures. Conversely, factors with less predictive power, such as land use and soil type, may guide resource allocation toward more influential predictors. The nuanced interpretation of significance values enhances the practical applicability of the model's findings in real-world landslide risk assessment and management, allowing for tailored interventions based on the specific characteristics of the study area. Overall, this discussion underscores the importance of considering individual factors' statistical significance in crafting effective strategies for landslide prediction and mitigation.

Conclusion

In conclusion, this study employed a logistic regression model to analyze various factors influencing landslide occurrences. The significance values derived from the model shed light on the relative importance of these factors, offering valuable insights for real-world applications in landslide risk assessment and management. The statistically significant variables, such as Slope Value and Rock Value, provide actionable information for prioritizing interventions. Steeper slopes and areas with certain lithological characteristics demand heightened monitoring and targeted mitigation strategies. These findings empower decision-makers to allocate resources efficiently and implement contextually relevant measures, enhancing the resilience of communities to landslide hazards. Conversely, variables with higher p-values, such as Land Use Value and Soil Value, may have less predictive power in the specific context studied. While these factors should not be disregarded, the study suggests that their influence on landslide occurrences may be limited. This nuanced understanding allows for a more strategic allocation of resources, focusing efforts on factors with greater impact.

The study's findings emphasize the importance of tailoring mitigation strategies to the specific characteristics of the study area. Contextual considerations, guided by statistical analyses, ensure that interventions are not only evidence-based but also responsive to the unique risk factors present. This approach enhances the practical applicability of the model's insights, facilitating their translation into on-the-ground actions. In essence, the significance values obtained through the logistic regression model serve as a roadmap for informed decision-making in landslide risk assessment and management. By leveraging these insights, stakeholders can proactively address key risk factors, ultimately contributing to the development of resilient communities capable of mitigating the impact of landslides. As we navigate the complexities of landslide-prone regions, this study provides a valuable foundation for evidence-based strategies aimed at reducing the vulnerability of communities to landslide hazards.

Acknowledgment
The research was funded by grant no: (FRGS/1/2019/SS07/UKM/02/1), entitled 'Spatial Variation Causal Factors of Slope Failure', and Publishing Incentive UKM for their assistance with the article processing charges enabling the publication of this research.

References

Ahzan, A. (2022). 70 kes tanah runtuh direkodkan di Selangor sehingga kelmarin. Available online: https://selangorkini.my/2022/03/70-kes-tanah-runtuh-direkodkan-di-selangor-sehingga-kelmarin/ [16 Ogos 2022].

Akgun, H., & Turkmenoglu, A. (2004). Landslide susceptibility mapping for the Istanbul Metropolitan Area (Turkey) using logistic regression analysis. *Environmental Geology*, 46(4), 445-453.

Alcantara-Ayala, I., Sassa, K., Mikos, M., Han, Q., Rhyner, J., Takara, K., Nishikawa S., Rouhban B., and Briceno S. (2017). The 4th world landslide forum: landslide research and risk reduction for advancing the culture of living with natural hazards. International Journal of Disaster Risk Science, 8, 498-502.

Althuwaynee, O. F., & Pradhan, B. (2017). Semi-quantitative landslide risk assessment using GIS-based exposure analysis in Kuala Lumpur City. *Geomatics, Natural Hazards and Risk*, 8(2), 706-732. https:// doi.org/10.1080/19475705.2016.1255670.

Althuwaynee, O. F., Pradhan, B., & Ahmad, N. (2015). Estimation of rainfall threshold and its use in landslide hazard mapping of Kuala Lumpur metropolitan and surrounding areas. *Landslides*, 12, 861-875.

An, Q., Liu, J., Zhao, J., & Ma, R. (2010). Landslide susceptibility mapping based on logistic regression and GIS: an example of the Lushan area, China. *Environmental Geology*, 59(2), 391-401.

Ayalew, L., Yamagishi, H., Marui, H., & Kanno, T. (2005). Landslides in Sado Island of Japan: Part II. GIS-based susceptibility mapping with comparisons of results from two methods and verifications. *Engineering geology*, 81(4), 432-445.

Budimir, M. E. A., Atkinson, P. M., & Lewis, H. G. (2015). A systematic review of landslide probability mapping using logistic regression. Landslides, 12, 419-436.

Burrough, P. A., & McDonnell, R. A. (1998). Principles of geographical information systems. Oxford University Press.

Chen, W., Sun, Z., Han, J. (2018). Landslide Susceptibility Modeling Using Integrated Ensemble Weights of Evidence with Logistic Regression and Random Forest Models. *Applied Science,* 9, 171.

Dhianaufal, D., Kristyanto, T. H. W., Indra, T. L., & Syahputra, R. (2018). Fuzzy logic method for landslide susceptibility mapping in volcanic sediment area in Western Bogor. *AIP Conference Proceedings 2023,* Article 020190. https://doi.org/10.1063/1.5064187

Gue, S. S., & Tan, Y. C. (2003, August 19-20). The engineering aspects of hill-site development. [Paper presentation]. Hillside Development–Issues and challenges, Kuala Lumpur, Malaysia.

Hervas, J., Méndez, A., & Blanco, A. (2017). Influence of the proximity to water bodies on landslide occurrence. *Geomorphology*, 287, 112-123.

Holiday Weather. (2022). Kuala Lumpur, Malaysia: Annual Weather Average https://www.holiday-weather.com/kuala_lumpur/averages/ [13 December 2022].

Huqqani, I. A., Tay, L. T., & Mohamad-Saleh, J. (2022). Assessment of Landslide Susceptibility Mapping Using Artificial Bee Colony Algorithm Based on Different Normalizations and

Dimension Reduction Techniques. *Arabian Journal for Science and Engineering*. 47: 7243–7260.

Kazmi, D., Qasim, S., Harahap, I. S. H., & Baharom, S. (2017). Landslide of Highland Towers 1993: a case study of Malaysia. *Innovative Infrastructure Solutions*, 2, 1-9.

Leonardi, G., Palamara, R., & Suraci, F. (2020). A fuzzy methodology to evaluate the landslide risk in road lifelines. T*ransp. Res. Procedia.* 45: 732–739.

Leonardi, G., Palamara, R., Manti, F., & Tufano, A. (2022). GIS-Multicriteria Analysis Using AHP to Evaluate the Landslide Risk in Road Lifelines. *Applied Sciences*. 12(9):4707

Li, H. W. M., Lo, F. L. C., Wong, T. K. C., & Cheung, R. W. M. (2022). Machine learning-powered rainfall-based landslide predictions in Hong Kong—An exploratory study. *Applied Sciences,* 12(12), 6017.

Ma, Z., Mei, G., & Piccialli, F. (2021). Machine learning for landslides prevention: a survey. Neural Computing and Applications, 33(17), 10881-10907.

Mahmud, A. R., Awad, A., & Billa, R. (2013). Landslide Susceptibility Mapping Using Averaged Weightage Score and GIS: A Case Study at Kuala. *Pertanika Journal of Science and Technology,* 21(2): 473–486.

Malaysian Meteorological Department. Annual Report (2019). https://www.met.gov.my/assets/content/penerbitan/pdf/laporantahunan2019.pdf [13 December 2022].

Mandaglio, M. C., Gioffre, D., Pitasi, A., & Moraci, N. (2016). Qualitative Landslide Susceptibility Assessment in Small Areas. *Procedia Engineering, VI Italian Conference of Researches in Geotechnical Engineering*, 158: 440-445.

Nhu, V. H., Mohammadi, A., Shahabi, H., Ahmad, B. B., Al-Ansari, N., Shirzadi, A., ... & Nguyen, H. (2020). Landslide susceptibility mapping using machine learning algorithms and remote sensing data in a tropical environment. *International Journal Of Environmental Research And Public Health*, 17(14), 4933.

Nwazelibe, V. E., Unigwe, C. O., & Egbueri, J. C. (2023). Integration and comparison of algorithmic weight of evidence and logistic regression in landslide susceptibility mapping of the Orumba North erosion-prone region, Nigeria. *Modeling Earth Systems and Environment*, 9(1), 967-986.

Pradhan, P. (2008). Landslide susceptibility mapping using statistical models: a review. *Progress in Physical Geography*, 32(1), 1-59.

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 180, 60-91.

Saadatkhah, N., Kassim, A., & Lee, M. L. (2014). Qualitative and Quantitative LandslideSusceptibility Assessments in HuluKelang area, Malaysia. *Electronic Journal of Geotechnical Engineering*. 19(C): 545-563.

Saadatkhah, N., Kassim, A., & Lee, L. M. (2015). Susceptibility assessment of shallow landslides in Hulu Kelang area, Kuala Lumpur, Malaysia using analytical hierarchy process and frequency ratio. *Geotechnical and Geological Engineering*, 33, 43-57.

Sanusi, M. S. M., Ramli, A. T., Hassan, W. M. S. W., Lee, M. H., Izham, A., Said, M. N., Wagiran, H., & Heryanshah, A. (2017). Assessment of impact of urbanisation on background radiation exposure and human health risk estimation in Kuala Lumpur, Malaysia. *Environment International*, 104, 91-101. https:// doi.org/10.1016/j.envint.2017.01.009

Sardi, M. F., & Razak, K. A. (2019). Assessment of effectiveness of emergency response time during landslide event in Malaysia. *ASM Sci. J*, 12, 1-19.

Sekarlangit, N., Fathani, T. F. Wilopo, W. (2022). Landslide Susceptibility Mapping of Menoreh Mountain Using Logistic Regression. *Journal of Applied Geology*, 7(1): 51-63.

Sujatha, E. R., & Sridhar, V. (2021). Landslide susceptibility analysis: a logistic regression model case study in Coonoor, India. *Hydrology,* 8 (1), 1-18.

Süzen, M. L., & Kaya, B. Ş. (2012). Evaluation of environmental parameters in logistic regression models for landslide susceptibility mapping. *International Journal of Digital Earth,* 5(4), 338-355.

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., & Lacasse, S. (2022). Machine learning and landslide studies: recent advances and applications. *Natural Hazards,* 114(2), 1197-1245.

Trinh, T., Lu, B. T., Le, T. H. T., Nguyen, G. H., Tran, T. V., Nguyen, T. H. V., Nguyen, K. Q., & Nguyen, L. T. (2022). A comparative analysis of weight-based machine learning methods for landslide susceptibility mapping in Ha Giang area. *Big Earth Data*, 1-30.

Wang, F., Wang, Y., & Hong, H. (2018). Assessment of landslide susceptibility in an area with complex terrain based on the integration of multiple information sources and machine learning. *Geocarto International*, 33(10), 1272-1291.

World Health Organization. (2022). Landslides. Available online: https://www.who.int/health-topics/landslides#tab=tab_1 [27 September 2022].

Yalcin, A. (2008). GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations, Turkey. *Catena* 72: 1-12.

Zhang, W. G., Liu, S. L., Wang, L. Q., Samui, P., Chwala, M., & He, Y. W. (2022). Landslide Susceptibility Research Combining Qualitative Analysis and Quantitative Evaluation: A Case Study of Yunyang County in Chongqing, China. *Forests*, 13: 1055.

Zillman, J. (1999). The physical impact of the disaster. In Natural disaster management, ed. J. Ingleton, 320. Leicester: Tudor Rose Holding Ltd.

Zubaidi, A. H., Baharuddin, Z. M., & Mansor, M. (2020). Redevelopment Of Abandoned Highland Towers As Memorial Landscape. Design Ideals Journal, 2(1),36–43.

Zulkafli, S. A., & Abd Majid, N. (2020). Landslides Incidents in Federal Territory of Kuala Lumpur, Malaysia. *Eco. Env. & Cons.*, *27*(3), 990–995.