مجلة
مركز البحوث الجغرافية والكارتوجرافية

مجلة علمية محكمة تصدر عن
مركز البحوث الجغرافية والكارتوجرافية
كلية الآداب - جامعة المنوفية

الترقيم الدولي الموحد للطباعة: 2357-0091
الترقيم الدولي الموحد الإلكتروني: 2735-5284

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

مجلة مركز البحوث الجغرافية والكارتوجرافية
بكلية الآداب – جامعة المنوفية
مجلة علمية مُحَكَّمَة

تقييم قابلية حدوث الانهيارات الأرضية في حوض وادي ضلع (السعودية) من خلال دمج أساليب نظم المعلومات الجغرافية والاستشعار عن بعد والإحصاء ثنائي المتغيرات ونهج التعلم الآلي الاصطناعي

إعداد

د/ وليد شكري عبد الحميد يوسف
قسم الجغرافيا ونظم المعلومات الجغرافية، كلية الآداب، جامعة أسيوط

د/ نرمين أحمد شكري
قسم الجغرافيا، كلية الآداب، جامعة القاهرة

أ.م.د/ أحمد علي أحمد علي
قسم الجغرافيا ونظم المعلومات الجغرافية، كلية الآداب، جامعة أسيوط

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

Egyptian Knowledge Bank
بنك المعرفة المصري

مجلة مركز البحوث الجغرافية والكارتوجرافية

بكلية الآداب – جامعة المنوفية

مجلة علمية مُحَكَّمَة

موقع المجلة علي بنك المعرفة المصرى: https://mkgc.journals.ekb.eg/

الترقيم الدولي الموحد للطباعة: ٢٣٥٧-٠٠٩١
الترقيم الدولي الموحد الإلكتروني: ٢٧٣٥-٥٢٨٤

تتكون هيئة تحكيم إصدارات المجلة من السادة الأساتذة المحكمين من داخل وخارج اللجنة العلمية الدائمة لترقية الأساتذة والأساتذة المساعدين في جميع التخصصات الجغرافية

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

Article:

An Assessment of Landslide Susceptibility in Wadi Dil'ah Basin (KSA) by Integrating GIS, RS, Bivariate Statistics and Artificial Machine Learning Approaches

Dr. Waleed S. Yousuf
Department of Geography and GIS, Faculty of Arts, Assiut University, Assiut, Egypt.

Dr. Nermin A. Shoukry
Department of Geography, Faculty of Arts, Cairo University, Giza, Egypt.

Dr. Ahmed A. Ali
Department of Geography and GIS, Faculty of Arts, Assiut University, Assiut, Egypt.

ABSTRACT:

Landslide is a natural hazard that causes numerous casualties and property losses worldwide. This study aims to produce landslide susceptibility maps by integrating GIS, Remote Sensing, Machine Learning Artificial Neural Networks (ML-ANN) and Bivariate Statistical {Frequency Ratio (FR), Shannon Entropy (SE)} approaches in Wadi Dil'ah basin in the southwestern part of Saudi Arabia. A total number of 137 landslide sites were identified using high-resolution satellite images, historical records, and field surveys. An equal number of non-landslide sites (areas with a slope angle less than 2°) were selected and divided into two groups; 70% were used for model training and 30% for model validation. Eighteen landslide-related factor layers were selected and prepared to be examined, including: Altitude, Slope-angle, Slope-aspect, Slope-length, Topographic Position Index (TPI), Terrain Ruggedness Index (TRI), Landform patterns, General, Plan, and Profile Curvature, Lithology, Distance from Fault, Topographic Wetness Index (TWI), Distance from Stream, Rainfall, Land Use/Land Cover (LULC), Normalized Difference Vegetation Index (NDVI), and Distance from Road. Moreover, the Variance Inflation Factors (VIF) and Tolerance Level (TOL) indices were generated to detect and measure multicollinearity assessment to avoid strong correlations among the factors. The relationships between the landslide-related factors and the landslide inventory map were calculated by using ANN, FR, and SE models. The Receiver Operating Characteristic (ROC) Curve and the Area Under the Curve (AUC) were applied to assess the model performance. The results of the VIF and TOL indices indicated no multicollinearity among the selected factors. The AUC values for the training rates were 0.966, 0.955, and 0.953, while the testing rates were 0.983, 0.993, and 0.971 for the examined ANN, FR, and SE models, respectively. The resultant landslide susceptibility maps (LSMs) were divided into five categories: very low, low, moderate, high, and very high. The classification was accomplished by using the Natural Breaks (Jenks) tool. The percentage for each landslide susceptibility category was calculated. The research final results revealed that the performance of the ANN model was better than the FR and SE models. Therefore, the ANN model is recommended as a suitable approach for applying landslide susceptibility maps in the mountainous region. It is hoped that the findings of this study will assist decision-makers and researchers in mitigating landslides and understanding their dynamics.

Keywords: *Landslide susceptibility mapping (LSM) · Artificial neural networks (ANN) · Frequency ratio (FR) Shannon entropy (SE) · Landslides. GIS. Saudi Arabia.*

I. Introduction

Landslides are among the most destructive natural environmental hazards and pose a significant risk to human life in mountainous regions (Ali et al. 2021). They are recognized as a threat to human life, human activities, and various economic, social, and political aspects, as well as natural resources (Bista, 2022; Wang et al. 2015). Landslides are responsible for substantial losses in lives and property and hinder social development (Cui et al. 2019; Khan et al. 2021). Additionally, they cause damage to infrastructure, agricultural lands, and urban areas in various mountainous regions worldwide (Li et al. 2020; Mondini et al. 2021).

The United Nations Development Program (UNDP) states that landslides rank second among the most common geological hazards in the world, causing significant financial losses annually (Pham et al. 2020). Approximately 66 million people live in areas highly susceptible to landslides, and 17% of casualties in these regions are attributed to landslides, which are recurrent disasters in mountainous areas (Achu et al., 2022a, 2022b). The economic losses have amounted to about USD of 10.8 billion from 1990 to 2020. These losses are expected to increase in the future due to the growing urban expansion, economic development, and unusually high levels of rainfall caused by climate change (Saha et al., 2021; Jakob, 2022; Li et al., 2022; Naceur et al., 2022). However, these damages and losses can be mitigated through effective planning and management (Rajakumar et al., 2007).

Landslides are defined as the downward movement of rock masses and debris down slopes (Cruden et al., 1996), and occur due to natural phenomena or human activities (Cruden et al., 1991; Shano et al., 2021). These causes include rainfall, earthquakes, groundwater level changes, tectonic movements, the formation and erosion of water channels, accelerated severe slope erosion, road construction, deforestation, and mineral extraction (Gomez et al., 2023). These factors cause a rapid increase in the stress borne by slope materials and a decrease in shear strength, surpassing what is known as the triggering threshold, an indicator of slope instability and the occurrence of landslides (Cardinali et al., 2002).

Additionally, human activities have a direct impact on the landscape, causing changes in slopes due to urban expansion into areas with unstable slopes affected by past landslides. Excavation and filling activities, as well as road construction, are significant factors contributing to landslides. These activities alter the surface slope during and after urban development in these areas (Farhan, 2002).

The southwestern region is considered one of the most important gateways to the Kingdom of Saudi Arabia due to its strategic location, attracting significant attention from the Saudi government. This region is characterized by its rugged mountainous terrain, notably the towering Arabian Shield Mountains, which constitute more than 70% of its area and receive a substantial amount of rainfall annually. Consequently, numerous landslides have been recorded, triggered by various mechanisms such as rainfall, earthquakes, and human activities (Youssef et al. 2022).

Therefore, landslides are among the most common phenomena in the southwestern part of the Kingdom of Saudi Arabia, particularly along Asir Highlands (Sarawat mountain range). This region is characterized by its steep scarps, and the catchment areas are particularly susceptible to landslides due to intense rainstorms.

The pace of development in this region has accelerated significantly, with the establishment of numerous urban areas and infrastructure roads, escarpment roads, tunnels, and highways throughout the mountainous areas. Given the geological characteristics of this region, where numerous tectonic movements occur, many of these rock formations become steep and structurally weak, leading to frequent landslides and mass movements such as landslides and debris flows (Sidle et al. 2018). This poses serious threats to urban areas and the infrastructure that supports transportation.

Landslide Susceptibility Assessment (LSA) is an important assessing measurement in geological hazard research (Merghadi et al. 2020; Azarafza et al. 2021). It is considered as a significant value for studying the distribution of landslide probabilities and understanding the relationship between landslides and the environmental factors that cause them. Previously, there have been few attempts to address landslide susceptibility in Saudi Arabia. However, with the increasing rate of landslides in recent years, it is essential to raise awareness of the landslide problem and work on reducing their impacts and/or preventing them to some extent through the use of predictive models for landslide occurrence and assessing the factors that would be contributed to them.

Remote sensing (RS) and geographic information systems (GIS) techniques are effective and beneficial in mapping landslide susceptibility. By utilizing these technologies, suitable and unsuitable areas for developmental activities can be identified. Recently, with advancements in various software programs based on Machine Learning Algorithms (ML) and Data Management (DM), numerous studies have been conducted to map landslide susceptibility using machine learning algorithms, including Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Random Forest (RF), Logistic Regression (LR), and Support Vector Machines (SVM), along with other statistical models such as Frequency Ratio (FR) and Shannon Entropy (SE), with the assistance of GIS and remote sensing (Vayadande et al. 2024; Ganesh et al. 2022; He et al. 2023; Selamat et al. 2022; 2023; Masruroh, 2023; Youssef, 2023).

However, these models produced varying results with different types of data according to the selective study area by different researchers. The significance of this study lies in the ability to map landslide susceptibility in the basin of Wadi Dil'ah area of Asir in the southwestern (SW) region of Saudi Arabia, along Abha-Jazan Road, relying on the selection of effective conditioning factors (Anis et al. 2019; Dam et al. 2022) and based on remote sensing and geographic information systems that integrated with machine learning algorithms (ANN) and statistical models (FR - SE). It is known that Abha-Jazan Road (Al-Dil'ah Pass) has suffered damage and disruptions due to landslide occurrences, especially following heavy rainstorms, particularly during the rainy season when sudden floods lead to rock falls along cracks or create debris flows of sedimentary materials accumulated along drainage networks. Therefore, the results of this study can assist planners and decision-makers in identifying areas prone to landslides to mitigate their risks in the study area.

1. Description of the Study Area:

Wadi Dil'ah basin (the study area) is located in Asir region in the southwestern part of Saudi Arabia, south of Abha City. It extends between the latitudes of 18° 12' 25" N and 17° 54' 11" N, and the longitudes of 42° 25' 44" E and 42° 35' 25" E (Fig. 1 left). Wadi Dil'ah basin is approximately 33.817 km long and 12.378 km wide, covering an area of 283.2 km². The study area falls within the Arabian Shield. Wadi Dil'ah basin is one of the sub-basins of Etwid Wadi in Asir region. Wadi Dil'ah flows from the Sarawat Mountains towards Tihamah plain in a southwest direction and continues in this direction until it meets the lower Etwid Wadi. Moreover, Wadi Dil'ah basin is characterized by its rugged topography, especially in its upper basin, with the highest elevation peak reaching 2,661 meters above sea level, while the lowest elevation being 415 meters above sea level.

Wadi Dil'ah takes on a triangular shape, with a Form Factor of 0.21 (Horton, 1945). Several important roads pass through the study area, including Dali'ah escarpment, which connects various villages within the region (Fig. 1 right) and links to other major cities such as Abha, Jizan, and El-Drab. The slope angles range from 0 to 66.3 degrees. The climate of the study area is moderate, with summer temperatures not exceeding 30 degrees Celsius, while winter temperatures tend to be cooler, reaching around five degrees Celsius in the highlands. Wadi Dil'ah is one of the most significant sub-valleys of Wadi Etwid, rich in the amounts of water flowing through it or stored within it. It receives rainfall in the form of intense storms between March and May, with average monthly precipitation of 29.5 mm in March, 46.5 mm in April, and 64 mm in May. It is worth noting that the southwestern mountainous region where the study area is located has experienced numerous unprecedented rainstorms across the Saudi kingdom, with these rains increasing in intensity, duration, and frequency, resulting in various forms of widespread destruction (Abu Abdullah et al. 2020).

2. Data Sources and Types

Several sources were used to produce different types of data. This data includes information from historical records, reports from the Civil Defense Department, the Ministry of Transport, and agencies responsible for road maintenance in Asir region, along with field studies (data collected over different periods from the local population between 2013 and 2019).

Data sources also included high-resolution satellite images (GeoEye images with a spatial resolution of 0.5 meters, obtained from King Abdul-Aziz City for Science and Technology (KACST), and Google Earth Professional images with an approximate spatial resolution of one meter) and Landsat 8 (OLI) images with a spatial resolution of 30 meters (Landsat 8 satellite image Path 167 Row 048 Scene Identifier LC08_L2SP_167048_20230410, Acquisition Date: 10-04-2023, obtained from the United States Geological Survey USGS earth explorer website (https:// earth explorer. usgs. gov/).

This aided in identifying land use/land cover patterns. The study also utilized Band 5 (Near Infrared) and Band 4 (Red Band) to derive the Normalized Difference Vegetation Index (NDVI) values for the study area. The data included the ALOS-PALSAR digital elevation model with a resolution of 12.5 meters (https://search.asf.alaska.edu/#/), which was useful in deriving various parameters,

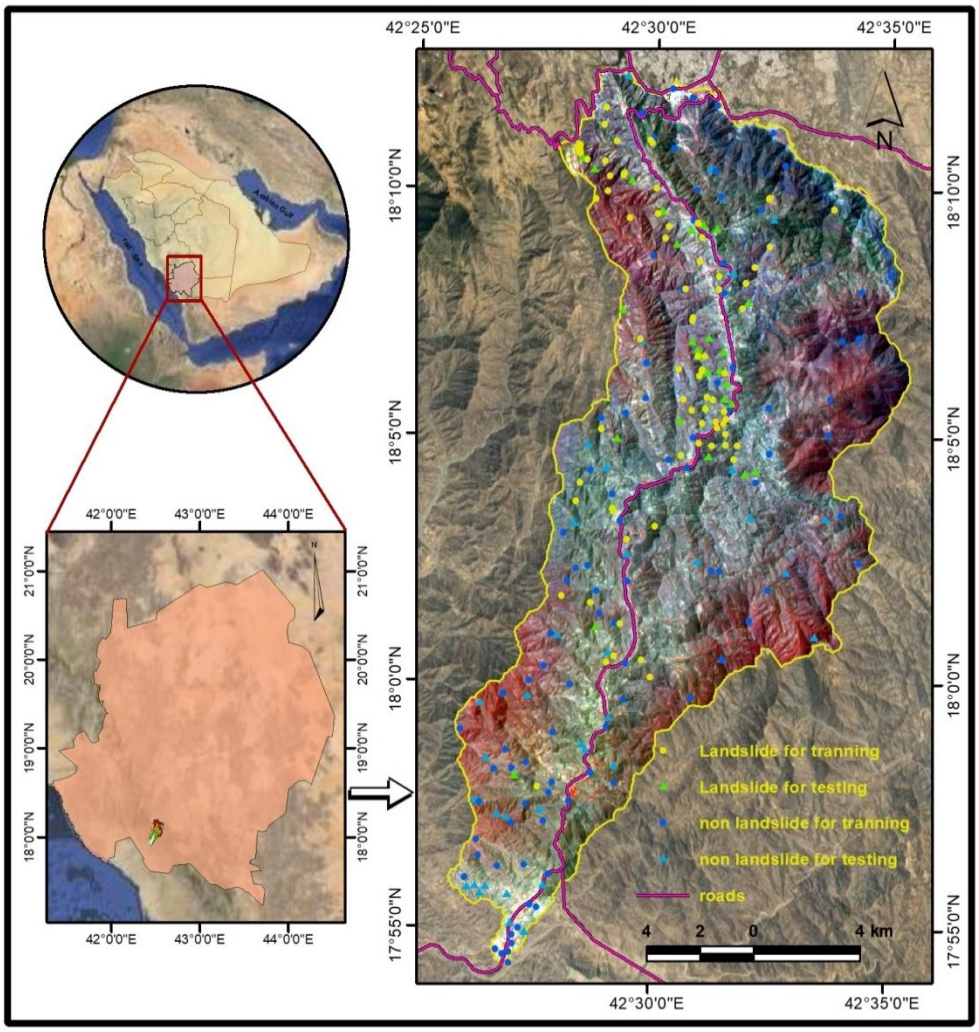

Fig. 1 Study area map with Landsat-8 false-color composite image RGB (5,4,2).

including elevation, slope angle, slope aspect, slope length, Topographic Position Index (TPI), Terrain Ruggedness Index (TRI), Terrain Wetness Index (TWI), landform types, general, plan, and profile curvature, as well as the drainage network. Geological and topographic maps were scanned and then georeferenced within ArcMap interface in ArcGIS software. Following this, a database was constructed and built within the Arc Catalog environment. This was followed by the digitization process, which converted the maps from their paper format to digital format while establishing a uniform datum and projection system, specifically WGS_1984_UTM_Zone_38N. Geological formations and faults were digitized from the geological maps at a scale of 1:250,000 issued by the Ministry of Petroleum and Mineral Resources, specifically Wadi Baysh (GM-77) and Abha (GM-75) quadrangle geological maps, which were obtained from the Saudi Geological Survey database.

The road network in the study area was also digitized from the topographic map at a scale of 1:250,000 published by the Aerial Survey (A.S) Department of the Ministry of Petroleum and Mineral Resources, particularly Sabya Sheet (NE36-9) and Abha Sheet (NE36-5). Additionally, road data were downloaded from the website (https://www.openstreetmap.org/export) to update the road network derived from the topographic maps. Distance layers to roads, faults, and waterways were derived in

raster format using the Euclidean Distance Tool in ArcGIS. Furthermore, additional data was derived from the annual surface information report issued by the National Center for Meteorology on average rainfall from 2012 to 2021 at Abha, Al-Soudah, Tamniah, and Al-Drab stations were used to derive a raster format layer for rainfall distribution. Finally, some landslide locations were surveyed using a Global Positioning System (GPS) in the study area.

II. Literature Review

The literature review of this research provides the necessary background of the variation of some essential studies of landslides that were carried out by many researchers and analysts that utilized different methods and techniques in various geographically mountainous areas and regions. Moreover, it sheds light on the various statistical and technical combination models that could be used and applied in this new field of study. World widely, it appears that about 1.3% of destructive natural disasters occurred and existed from landslide, with Asia accounting for about 54% of this phenomena (Khaliq, et al., 2022; Shahabi, et al., 2023). However, it is directed toward some selective researches that were applied generally in Asia and particularly in Saudi Arabia.

In 2015, a study of delineating landslide susceptibility (LS) was applied by the geologist "Ahmed Youssef" in Ar-Rayth mountainous area in Jizan (KSA). He used an Analytical Hierarchy Process (AHP) that infused with both Frequency Ratio (FR) and Logistic Regression (LR) models. A land inventory map was constructed of all landslides locations based on many data sources along with the proposed causative factors' weights in the study area. The study ends with producing three susceptibility maps indicating the preference usage of LR model for landslides studies (Youssef, A., 2015). Additionally, another study of the former author in the same year is accomplished to map landslide susceptibility in another different mountainous area in Jizan (KSA). It was conducted by using GIS-based frequency ratio and index of entropy models. The study area was delineated as "Al-Hasher" area that is located NE Jizan City. In this paper, the (FR) and Index of Entropy statistical models were experienced along with the aid of GIS tools and remote sensing data. Landslide Susceptibility Maps (LSMs) were produced by achieving nearly a value of 0.7 in both models, respectively (Youssef, A., et al., 2015a). Moreover, Youssef and three other researchers assessed landslide susceptibility in Wadi Jawrah basin that is located in the Jizan region, SW KSA. He utilized different types of data sources that depends more on geological structures and variables for the selected influence factors. In this study, he concentrated on evaluating two bivariate statistical approaches, the (FR) and Weights-of-Evidence (WoE) to produce and assess the final LSMs. His final results revealed that both experimented models produced reasonable accuracy (Youssef, A., et al., 2015b).

In 2019, an assessment study of LS in Mazandran Province (Iran) is accomplished by using an integration method of two statistical models. The first is FR and the second is Random Forest (RF) that were infused in satellite ASTER and SRTM DEMs data. The main advantage of this research methodology is the capability of determining the relative importance of effective factors and enlighten the spatial relationship between these factors and landslides locations (Arabameri, 2019).

Another Landslide Susceptibility Mapping (LSM) study in the northern part of Iran to compare four bivariate statistical models: FR, SE, WoE, and Evidential Belief Function (EBF). The research ends with a final result of recommending the usage of WoE model in the study area that achieved the highest AUC value indicating the highest accuracy (Nohani, 2019).

In 2021, a geological/geophysics study of LSA and its impacts on the urban expansion of Makkah Al-Mukarramah (KSA). The data were processed and analyzed using the Horizontal-Vertical Spectral Ratio (HVSR) method to gain the resonance frequency and H/V amplification factor. The study ends with recommending the usage of microtremor measurements as an effective technique in locating the sites that suffer from landslides (Abdelrahman, et al., 2021). In a study area of Al-Fawar basin (Syria), an assessment of LSM zonation was applied. The study used bivariate statistics of FR model and Statistical Index (SI) model with utilizing 13 dominant factors (Abdo, 2022).

In 2023, a study to produce a LSM for Al-Hada Mountainous terrain of Makkah province in Saudi Arabia, a usage of GIS and RS tools only were a dominant research methodology without infusing any statistical approach. A GIS-based weighted overlay analysis along with a remotely sensed data were applied. Selective eight raster format layers work as conditioning factors are processed. The study lacked of measuring an accuracy assessment for the produced final map and it is not working for regions with small-scale landslides (Alharbi and El-Sorogy, 2023). Another study to produce a LS evaluation that based on remotely sensed, geological, and seismological data to be infused with moicrotremor measurements of Al-Taif urban area in Saudi Arabia. The study offered a recommendation stating that the microtremor measurements give a thorough method for assessing landslides. It expedites landslide analyses and lowers the initial expenses of numerical computations with a significantly high accuracy (Abdelrahman, et al., 2023).

Furthermore, a recent study offers a new methodological approach to generate LSMs by assessing the efficiency of applying three conventional Machine Learning Algorithms (MLAs) including RF, Decision Tree (DT), and Support Vector Machine (SVM), utilizing 14 influential factors. The study area is located in Iran's western Kurdistan province and resulted of showing that DT, RF, and SVM have respective prediction rates of 0.94, 0.82, and 0.75 (Shahabi, et al., 2023). Finally, another recent study offers a new methodological approach to handle the LS Mapping in different perspective. The study area is located in Chattogram (Bangladesh). Its incorporated GIS-based machine learning algorithms of Logistic Regression (LR) with RF and Decision & Regression Tree (DRT) models. Sixteen landslide conditioning factors were determined and experimented. Three LSMs for the three models were produced and accuracy assessed to obtain a final accuracy of LR, RF, and DRT models were 0.94, 0.91, and 0.95, respectively (Chowdhury, S., et al., 2024).

III. Research Methodology

The mentioned above material of data sources with their diverse types are significantly contributed to apply intelligent algorithm methods such as Machine Learning-Artificial Neural Networks (ML-ANN) algorithms and Bivariate Statistical

models in mapping landslide susceptibility. The overall methodology of the study is presented in (Fig. 2). Methodologically, the study followed several key stages:

1. Data Sources and Data Input: Identification of various data sources used to create the landslide susceptibility map.
2. Selection of Landslide Conditioning and Indicator Factors: Choosing the factors contributing to landslide occurrences.
3. Creating Landslide Inventory Map: Selection of random samples with 70% of training and 30% of testing samples of landslide sites.
4. Testing Multicollinearity: Assessing Multicollinearity before building the predictive models.
5. Modeling Techniques: Application of the suggested modeling techniques (FR, SE, and ML-ANN) to create landslide susceptibility maps (LSMs).
6. Verification and Testing: Conducting ROC-AUC tests to evaluate the accuracy and performance of the suggested predictive models.
7. Data Outputs and Final Results.

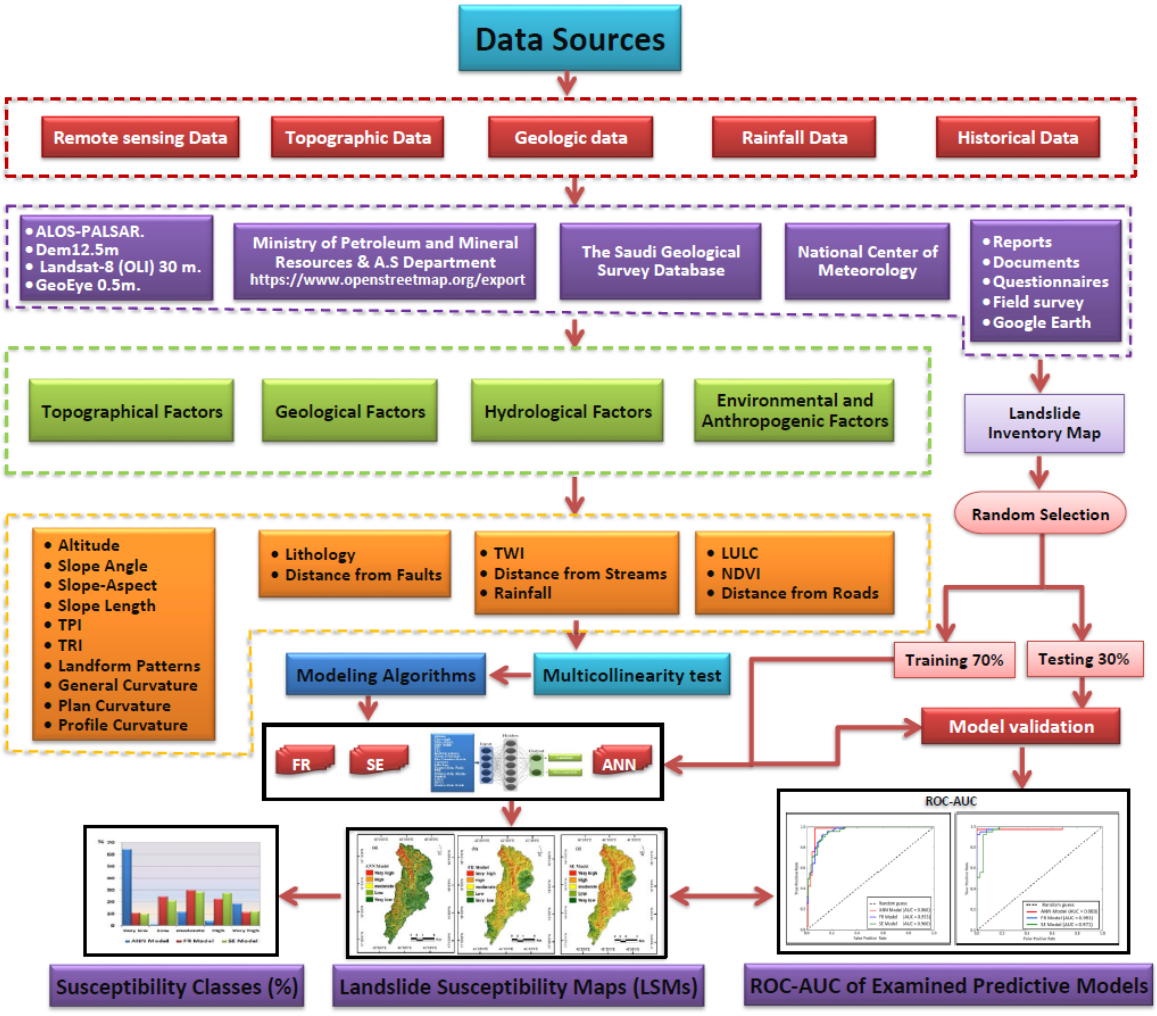

Fig. 2 Research schematic flowchart showing proposed methodology to predict the final outputs of Land Susceptibility Maps (LSMs) of the study area.

IV. Processing and Analyses

1. Landslide Inventory Map

The preparation of landslide inventory data is a crucial step in the landslide modeling process to build an accurate and effective prediction model for landslides (Lee et al. 2020). This is because there is an assumption that past events have a strong influence on the future (Zhou et al. 2018). Thus, landslide inventory maps can provide useful information regarding the locations of past landslides and may also identify areas where future landslides are likely to occur (He et al. 2023).

With the assistance of remote sensing technologies and field investigations, inventory maps can be evaluated more effectively. Based on a comprehensive analysis of various datasets, including historical documents, field surveys, interviews with some local residents in the study area, and the interpretation of high-resolution satellite imagery, a landslide location map was created.

In this study, the data for spatial prediction was classified into two categories: the first category is landslide locations, for which 137 total landslide sites were identified for each of training and testing samples (back to Fig. 1 right) to be used in building the predictive model for landslide susceptibility (Phong et al. 2020; Azarafza et al. 2021). The second category is non-landslide locations, where an equal number of sites that have not experienced landslides (areas with a slope angle of less than 2°) were identified.

Using ArcGIS Desktop software ver10.5 and the Geostatistical Analyst extension, the data was randomly divided into training datasets 70% of landslides (96 locations of landslides and 96 locations of non-landslides) to build the landslide susceptibility models. The remaining 30% (41 locations of landslides and 41 locations of non-landslides) were used as testing data to evaluate the model (back to Fig. 1 right).

The training and testing datasets were converted into a raster format. The landslide and non-landslide locations were coded with the numbers 1 and 0, respectively (Chen et al. 2019). Finally, the training dataset was overlaid with the conditioning factors that had been prepared in advance to extract the descriptive values for each factor.

2. Selection of Landslide Conditioning Factors

Landslides occur as a result of the influence of a set of selected factors, which serve as inputs for creating susceptibility maps for any region. In this study, eighteen landslide conditioning factors were extracted, categorized into topographical, geological, hydrological, environmental and anthropogenic factors, to evaluate the spatial prediction of landslides in the study area.

The selected factors include several continuous variables (altitude, slope angle, slope aspect, slope length, Topographic Position Index (TPI), Terrain Ruggedness Index (TRI), Topographic Wetness Index (TWI), general curvature, plane curvature, profile curvature, lithology, rainfall, NDVI), while others are categorical variables (LULC, distance from streams, distance from roads, distance from faults, geomorphological landform patterns) (Table 1; Fig. 3, 4, and 5). Below is a description of these factors on details:

Table 1: Spatial relationship between each conditioning factor and landslide occurrence using FR and SE models

Conditioning factors	Classes	Class pixels	Pixels (%)	landslide	Landslide pixel %	FR	$RF_i=P_{ij}$	$W_{j,FR}$	P_j	$W_{j,SE}$
Altitude	< 500	4042	0.20	0	0.00	0.000	0.000	2.917	0.000	0.037
	500 - 1000	128693	7.10	41	29.90	4.214	0.663		-0.361	
	1000 - 1500	479399	26.50	34	24.80	0.938	0.148		-0.346	
	1500 - 2000	748666	41.30	54	39.40	0.954	0.150		-0.367	
	2000 - 2500	430853	23.80	8	5.80	0.246	0.039		-0.166	
	> 2500 (m)	20582	1.10	0	0.00	0.000	0.000		0.000	
Slope angle	0 - 2 (degree)	15182	0.80	0	0.00	0.000	0.000	4.174	0.000	0.070
	2 - 5	32233	1.80	0	0.00	0.000	0.000		0.000	
	5 - 10	91325	5.00	0	0.00	0.000	0.000		0.000	
	10 - 18	273172	15.10	0	0.00	0.000	0.000		0.000	
	18 - 30	800952	44.20	6	4.40	0.099	0.003		-0.137	
	30 - 45	578802	31.90	77	56.20	1.760	0.048		-0.324	
	45 - 67	20569	1.10	54	39.40	34.728	0.949		-0.367	
Slope-Aspect	Flat	6485	0.40	0	0.00	0.000	0.000	1.523	0.000	0.027
	North	180717	10.00	6	4.30	0.844	0.101		-0.157	
	Northeast	201093	11.10	21	15.30	1.381	0.164		-0.287	
	East	231694	12.80	51	37.20	2.912	0.346		-0.368	
	Southeast	240987	13.30	26	19.00	1.427	0.170		-0.315	
	South	243265	13.40	5	3.60	0.272	0.032		-0.121	
	Southwest	296872	16.40	10	7.30	0.446	0.053		-0.191	
	West	236572	13.10	12	8.80	0.671	0.080		-0.213	
	Northwest	174550	9.60	6	4.40	0.455	0.054		-0.137	
slope length	0 - 5 (m)	1634013	90.20	91	66.40	0.737	0.048	1.000	-0.272	0.046
	5 - 15	84681	4.70	19	13.90	2.968	0.195		-0.274	
	15 - 30	66551	3.70	20	14.60	3.975	0.261		-0.281	
	30 - 45	23836	1.30	6	4.40	3.330	0.219		-0.137	
	> 45	3154	0.20	1	0.70	4.194	0.276		-0.036	
TPI	< -50	358142	19.80	3	2.20	0.111	0.018	1.977	-0.084	0.016
	-50 :-10	497711	27.50	15	10.90	0.399	0.065		-0.242	
	-10 : 30	437014	24.10	33	24.10	0.999	0.164		-0.343	
	30 : 70	305721	16.90	40	29.20	1.731	0.284		-0.359	
	>70	213647	11.80	46	33.60	2.848	0.468		-0.366	
TRI	< 0.04	1493161	82.40	34	24.80	0.301	0.010	2.573	-0.346	0.040
	0.04 - 0.08	280600	15.50	75	54.70	3.536	0.115		-0.330	
	0.08 - 0.12	35384	2.00	24	17.50	8.972	0.291		-0.305	
	0.12 - 0.16	2929	0.20	4	2.90	18.065	0.585		-0.103	
	> 0.16	161	0.00	0	0.00	0.000	0.000		0.000	
landform patterns	Pit	53621	3.00	0	0.00	0.000	0.000	2.189	0.000	0.078
	valley	207513	11.50	0	0.00	0.000	0.000		0.000	
	footslope	831	0.00	0	0.00	0.000	0.000		0.000	
	shoulder	339	0.00	0	0.00	0.000	0.000		0.000	
	hollow	354788	19.60	0	0.00	0.000	0.000		0.000	
	spur	320962	17.70	5	3.60	0.206	0.020		-0.121	
	Peak	10459	0.60	4	2.90	5.059	0.498		-0.103	
	Ridge	118520	6.50	28	20.40	3.125	0.307		-0.325	
	Slope	745202	41.10	100	73.00	1.775	0.175		-0.230	

Table 1: Continued.

Conditioning factors	Classes	Class pixels	Pixels (%)	landslide	Landslide pixel %	FR	$RF_i=P_{ij}$	$W_{j,FR}$	P_j	$W_{j,SE}$
General Curvature	Flat	346415	19.10	10	7.30	0.382	0.142		-0.191	
	Convex	724946	40.00	95	69.30	1.733	0.645	2.212	-0.254	0.035
	Concave	740874	40.90	32	23.40	0.571	0.213		-0.340	
Plan Curvature	Flat	146883	8.10	0	0.00	0.000	0.000		0.000	
	Convex	848390	46.80	85	62.00	1.325	0.612	2.689	-0.296	0.048
	Concave	816962	45.10	52	38.00	0.842	0.388		-0.368	
Profile Curvature	Flat	107663	5.90	0	0.00	0.000	0.000		0.000	
	Convex	881894	48.70	38	27.70	0.570	0.264	3.238	-0.356	0.056
	Concave	822678	45.40	99	72.30	1.592	0.736		-0.235	
Lithology	Sabya Group (sa)	265542	14.70	0	0.00	0.000	0.000		0.000	
	Jiddah Group (jt)	7892	0.40	0	0.00	0.000	0.000		0.000	
	Baish Group (ba)	155800	8.60	0	0.00	0.000	0.000	4.398	0.000	0.122
	Bahah Group (bt)	1358039	74.90	137	100.00	1.334	1.000		0.000	
	Sabya Group (sa)	24962	1.40	0	0.00	0.000	0.000		0.000	
Distance from Faults	0 - 750 (m)	686984	37.90	112	81.80	2.157	0.705		0.000	
	750 - 1500	486738	26.90	20	14.60	0.544	0.178		0.000	
	1500 - 2250	301717	16.60	2	1.50	0.088	0.029		-0.036	
	2250 - 3000	183312	10.10	2	1.50	0.144	0.047	3.100	-0.062	0.084
	3000 - 3750	103700	5.70	1	0.70	0.128	0.042		-0.062	
	3750 - 4500	44669	2.50	0	0.00	0.000	0.000		-0.281	
	> 4500	5115	0.30	0	0.00	0.000	0.000		-0.165	
TWI	< 4	314494	17.40	105	76.60	4.416	0.919		-0.204	
	45447	1082456	59.70	32	23.40	0.391	0.081		-0.340	
	45510	305013	16.80	0	0.00	0.000	0.000	4.040	0.000	0.080
	45573	60650	3.30	0	0.00	0.000	0.000		0.000	
	> 10	49622	2.70	0	0.00	0.000	0.000		0.000	
Distance from Streams	0 - 100	561955	31.00	28	20.40	0.659	0.119			
	100 - 200	457820	25.30	44	32.10	1.271	0.230		0.000	
	200 - 300	355886	19.60	36	26.30	1.338	0.242		-0.062	
	300 - 400	234292	12.90	21	15.30	1.186	0.214	1.064	-0.137	0.026
	400 - 500	141133	7.80	6	4.40	0.562	0.102		-0.287	
	500 - 600	51540	2.80	2	1.50	0.513	0.093		-0.351	
	> 600	9609	0.50	0	0.00	0.000	0.000		-0.365	
Rainfall	200 - 250 (mm)	17812	1.00	0	0.00	0.000	0.000		0.000	
	250 - 300	254437	14.00	9	6.60	0.468	0.176	2.007	-0.179	0.044
	300 - 400	783645	43.20	72	52.60	1.215	0.456		-0.338	
	400 - 500	756341	41.70	56	40.90	0.979	0.368		-0.366	
LULC	Built up area	61223	3.40	1	0.70	0.216	0.087		-0.036	
	Roads	89753	5.00	4	2.90	0.590	0.238	2.292	-0.103	0.087
	Vegetation	555154	30.60	8	5.80	0.191	0.077		-0.166	
	Bare ground	1106105	61.00	124	90.50	1.483	0.598		-0.090	
NDVI	-0.0214 - 0.0717	377596	20.80	35	25.50	1.226	0.320		-0.349	
	0.0717 - 0.0906	637563	35.20	69	50.40	1.432	0.373		-0.345	
	0.0906 - 0.1109	424058	23.40	30	21.90	0.936	0.244	1.534	-0.333	0.037
	0.1109 - 0.1365	285094	15.70	2	1.50	0.093	0.024		-0.062	
	0.1365 - 0.3229	87924	4.90	1	0.70	0.150	0.039		-0.036	
distance from Roads	0 - 1000 (m)	547384	30.20	91	66.40	2.199	0.586		0.000	
	1000 - 2000	437370	24.10	35	25.50	1.059	0.282		0.000	
	2000 - 3000	338536	18.70	8	5.80	0.313	0.083		0.000	
	3000 - 4000	216512	11.90	3	2.20	0.183	0.049	2.576	-0.084	0.067
	4000 - 5000	126535	7.00	0	0.00	0.000	0.000		-0.166	
	5000 - 6000	89032	4.90	0	0.00	0.000	0.000		-0.349	
	> 6000	56866	3.10	0	0.00	0.000	0.000		-0.272	

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

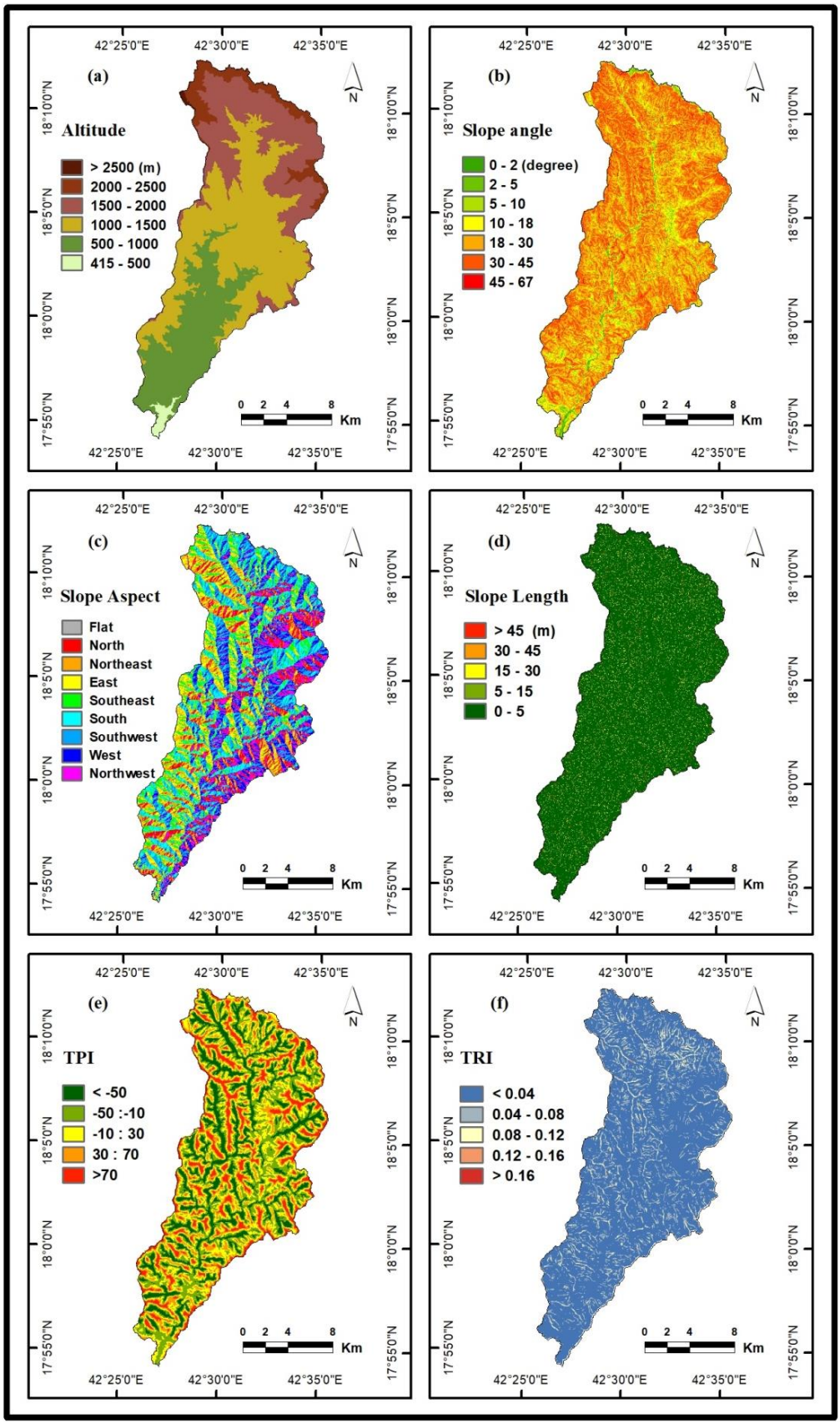

Fig. 3. Landslide conditioning (Topographical) factors: (a) Altitude, (b) Slope angle, (c) Slope-Aspect, (d) Slope Length (LS), (e) Topographic Position Index (TPI), (f) Terrain Ruggedness Index (TRI).

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

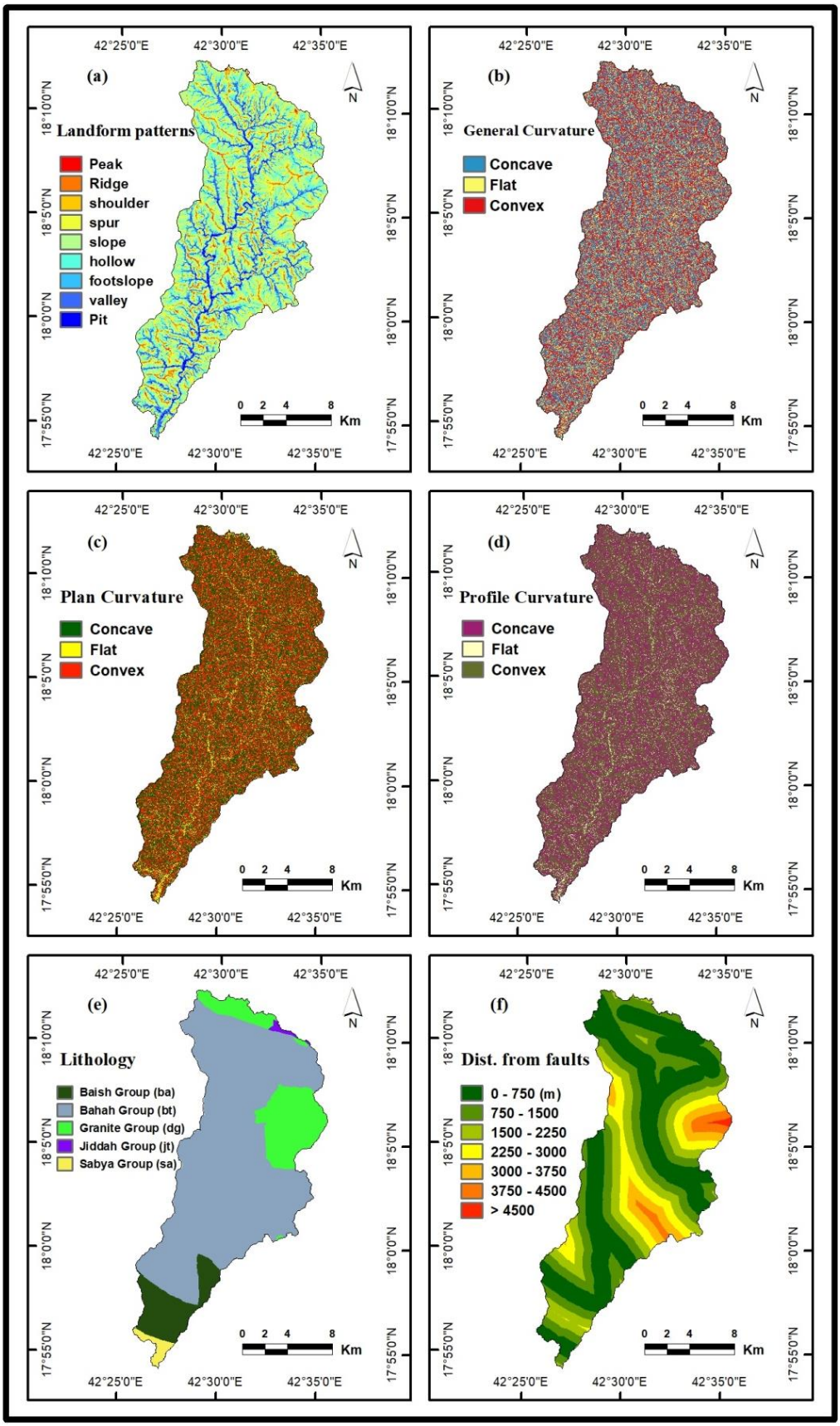

Fig. 4. Landslide conditioning (Topographical & Geological) factors: (a) landform Patterns, (b) General Curvature, (c) Plan Curvature, (d) Profile Curvature, (e) Lithology, (f) Distance from Fault.

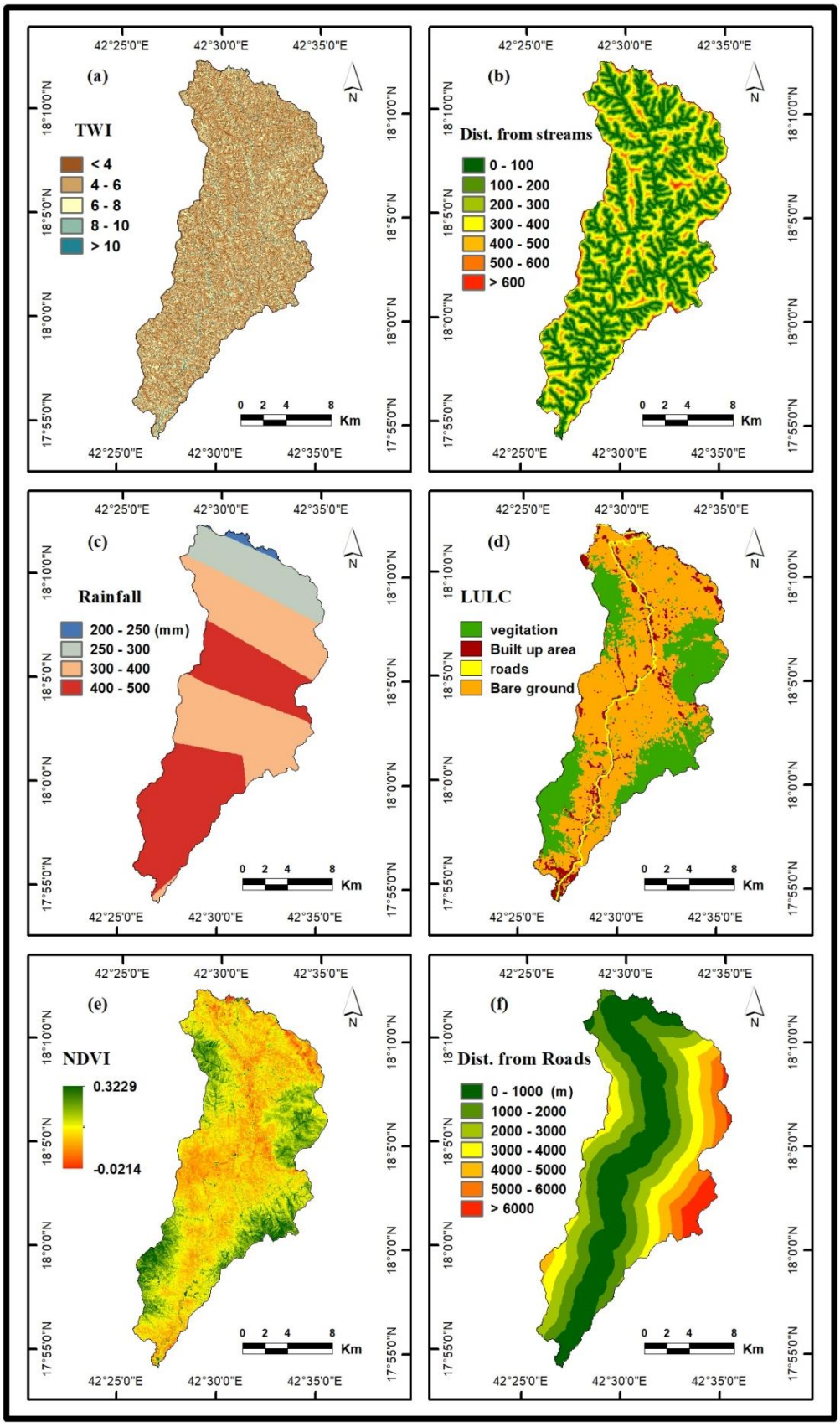

Fig. 5. Landslide conditioning (Hydrological & Environmental) factors: (a) Topographic Wetness Index (TWI), (b) Distance from streams, (c) Rainfall, (d) Land use/land cover (LULC), (e) Normalized Difference Vegetation Index (NDVI), (f) Distance from roads.

2.1. Topographical Factors

2.1.1. Altitude

Altitude is one of many topographical factors that influence slope stability and, consequently, leads to landslides (Hong et al., 2018; Feizizadeh et al. 2014). In this study, altitude values ranged from 415 to 2661 meters above sea level. This altitude was categorized into six classes: (415 - 500), (500 - 1000), (1000 - 1500), (1500 - 2000), (2000 - 2500) meters, and (> 2500 m), as illustrated in (back to Fig. 3a).

2.1.2. Slope Angle

The slope angle is one of the most influential factors in landslide assessments (Hong et al., 2018; Nguyen et al., 2019), with landslides occurring more frequently on steeper slopes (Poudel et al., 2016). Areas susceptible to landslides are characterized by steep slopes that contribute to the instability of the underlying rock and soil (Jebur et al., 2014). Additionally, slope plays a vital role in subsurface flow and affects soil moisture, which is directly related to landslide occurrence. In this study, slope degree values ranged from 0 to 66.3 degrees, and the slopes were classified into seven categories according to Young's classification (Young, A., 1972) (back to Fig. 3b).

2.1.3. Slope Aspect

The slope aspect refers to the direction of maximum change in value from each cell to its neighbors (Al-Najjar et al., 2019). It is also one of the most important factors influencing landslide occurrence due to the varying moisture levels across different aspects (Pham et al., 2018). The slope aspect affects various processes that have direct and indirect impacts on landslides, including wind direction, rainfall, sunlight exposure, hydrological processes, evaporation, transpiration, soil moisture concentration, and vegetation cover (Devkota et al., 2013). Furthermore, the slope aspect significantly influences the distribution of landslide types. In this study, a raster layer indicating the slope aspect of the study area was derived from a 12.5-meter Digital Elevation Model (DEM) and classified into nine categories with values: (-1), (0–22.5 degrees, 337.5–360 degrees), (22.5–67.5 degrees), (67.5–112.5 degrees), (112.5–157.5 degrees), (157.5–202.5 degrees), (202.5–247.5 degrees), (247.5–292.5 degrees), and (292.5–337.5 degrees) representing flat, north (N), northeast (NE), east (E), southeast (SE), south (S), southwest (SW), west (W), and northwest (NW), respectively (back to Fig. 3c).

2.1.4. Slope Length (LS)

The slope length; sometimes Length of Slope (LS) is also one important topographical factors influencing landslide susceptibility. Slope length, in conjunction with slope angle, affects soil loss and hydrological processes in mountainous areas (Pourghasemi and Rahmati, 2018). In the current study, the LS factor was derived from the Digital Elevation Model (DEM) using the SAGA software according to the following equation (Eq. 1) (Moore and Burch, 1986). In this study, the LS values range from 0 to 312.27 meters and were classified into five categories as follows: (>45), (30-45), (15-30), (5-15), and (0-5) (back to Fig. 3d).

$$LS = \left(\frac{A_s}{22.13}\right)^{0.4}\left(\frac{\mathrm{Sin}\beta}{0.0896}\right)^{1.3} \quad \textbf{(Eq. 1)}$$

where, A_s (m^2) is specific catchment area and β is in degree.

2.1.5. Topographic Position Index (TPI)

Topography is defined as specific geomorphological features on the earth's surface, ranging from large-scale features such as plains and mountain ranges to secondary features like hills and valleys (Jenness, 2010). Weiss and Jenness introduced a new application in Geographic Information Systems for the automated classification of topography known as the Topographic Position Index (TPI) (Weiss, 2001; Jenness, 2006). This index determines the position of a point within the context of its surrounding topography, indicating whether the point is in a valley, on a ridge, or on a slope.

Since the establishment of the Topographic Position Index (TPI) by Weiss and Jenness, the TPI has been used at various levels to classify landscapes into both slope position categories: (valley, lower slope, flat slope, typical slope, model slope, upper slope, ridge, hill) and landform categories: (channels/gorges, drainage, upland drainage, high drainage, shallow valleys, U-shaped valleys, plains, canyons / deeply incised streams, mid-slopes, open slopes, upper slopes, open slopes, local hills, hills in valleys, middle hills, small hills in the plain, plains, mountain peaks , local ridges, mid-slope ridges and high ridges) (Jenness, 2007). The Topographic Position Index (TPI) provides useful information about the landscape's topographic features that are used to assess landslide risks. The Topographic Position Index (TPI) can be calculated by using the following equation (Eq. 2):

$$TPI_t = z_0 - \frac{\sum_I - n\, z_n}{n} \quad \textbf{(Eq. 2)}$$

Where;

z0 = elevation of the model point under evaluation Zn = elevation of grid within the local window

n = the total number of surrounding points employed in the evaluation.

SAGA GIS was used to calculate the TPI, and the values of the index ranged from (-157.64 to 208.95). These values were classified into five categories as follows: (< -50), (-50 to -10), (-10 to 30), (30 to 70), and (> 70). (back to Fig. 3e)

2.1.6. The Terrain Ruggedness Index (TRI)

It is used to describe the terrain as either smooth or rugged, as well as the local variation in slopes or surface curvature (Dahal et al. 2008). TRI is also defined as the difference in elevation between adjacent pixels (Al-Najjar and Pradhan 2021). Terrain Ruggedness Index was calculated by using the following equation (Eq. 3):

$$TRI = \sum_{\alpha,\beta} \frac{|Z_{i+\alpha,j+\beta} - Z_{i,j}|}{8} \quad \text{(Eq. 3)}$$

Where is the elevation of the cell, and α and β are the indices representing the cell and its eight neighboring cells (Riley et al. 1999).

The TRI values were derived using SAGA GIS, and the TRI was divided into five categories as follows: (0-0.04), (0.04-0.08), (0.08-0.12), (0.12-0.16), and (>0.16) (back to Fig. 3f). The geomorphology of the area affects the occurrence of landslides, and it is taken into account in many sensitivity studies such as monitoring wildfires, landslide susceptibility, or hydrology (Pham et al., 2019).

2.1.7. The Geomorphological Landform Patterns

Recently, the landform patterns are classified using an innovative method introduced by Jasiewicz and Stepiski (2013) known as Geomorphic; this algorithm classifies landscapes by categorizing 498 unique geomorphological patterns based on elevation differences within the area surrounding the target cell (Fig. 6). In this study, the geomorphic algorithm integrated into SAGA GIS was used to classify landform patterns into 10 different landform types, outputting them in raster format as follows: Flat, Peak, Ridge, Shoulder, Spur, Slope, Hollow, Foot slope, Valley, and Pit (Depression) (Jasiewicz and Stepiski, 2013) (back to Fig. 4a).

2.1.8. Curvature (General, Plan, and Profile)

General, Plan, and Profile Curvature are fundamental terrain variables that have a direct impact on the acceleration and deceleration of surface runoff and the deposition of materials by managing the speed of material movement on a slope (Xiao et al., 2019; Aghdam et al., 2016). This, in turn, affects the occurrence of landslides (Pham et al., 2018).

In this study, the layers for General, Plan, and Profile Curvature were derived from the DEM using ArcGIS Ver10.5 software. The curvatures were classified into three categories: negative curvature (Concave) (<-0.05), zero curvature (Flat) (-0.05 to 0.05), and positive curvature (Convex) (>0.05) (Nohani et al., 2019). The values for General Curvature ranged from (-23.04 to 17.92), Plan Curvature from (-10.2336 to 10.6918), and Profile Curvature from (-10.5108 to 12.9354) (back to Fig. 4b, c & d).

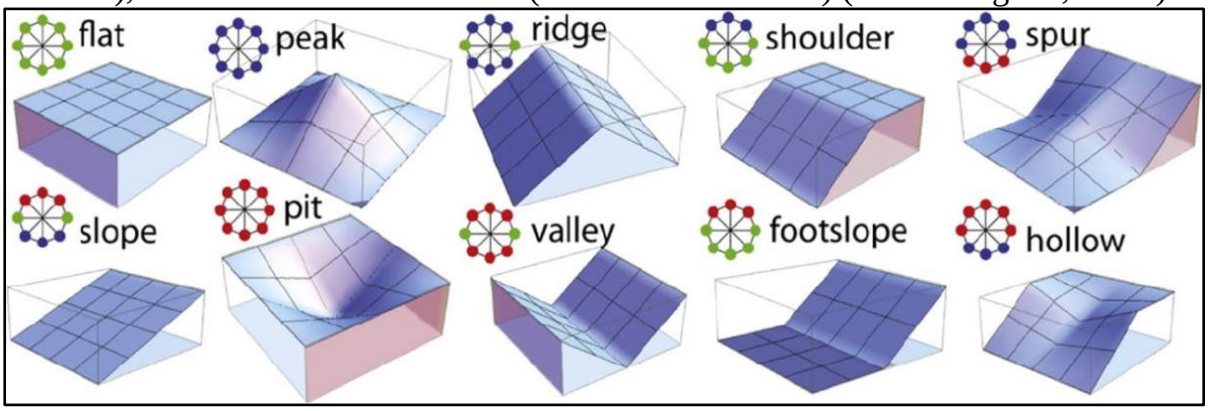

Fig. 6. A Geomorphic line of sight concept. Symbolic 3D morphologies and their corresponding Geomorphic (ternary patterns) for the 10 most common patterns found in nature. Figure after Jasiewicz and Stepinski (2013).

2.2 Geological Factors

2.2.1. Lithology

Lithology is the primary factor that directly influences the development of landslides in a specific area (Abedini et al., 2019; Tian et al., 2019) and contributes to the mapping of landslide susceptibility. Variations in lithology significantly impact different types of geographical hazards, such as landslides (Mekonnen et al., 2022). These geological units differ in physical and mechanical properties, including type, strength, weathering degree, durability, density, and permeability (Shano et al., 2021). Many researchers have used lithology in landslide studies (Dang et al., 2019; Panchal & Shrivastava, 2020, 2021, 2022). However, it requires understanding the nature of rock types to identify those that allow water storage in pores, which act as landslide factors (Jennifer et al., 2021). In the current study, five geological groups were identified as follows (back to Fig. 4e):

a. Sabya Group (SA): it consists mainly of Quartz-bearing meta-sedimentary rocks consisting of quartzite, quartz pebble conglomerate, argillite, limestone, and graywacke; widely converted to micaceous schist.
b. Baish group (BA): which consists of Greenstone, tholeiitic meta basalt (local pillow structures), and minor. metagraywacke, meta chert, and marble.
c. Bahah group (BT): it consists mainly of Biotite-quartz schist, phyllite, and calcareous metagraywacke.
d. Jiddah Group: (JT): which consists of pillow lava, flow breccia, tuff, dacite tuff, interbedded subordinate, often lava carbonaceous conglomeratic gray wacke and phyllite.
e. Granite group (DG): which consists Biotite monzogranite with diorite, gabbro, Foliated uniform body of biotite granodiorite and monzogranite.

2.2.2. Distance from Faults

Geological variables such as lineaments and faults play a crucial role in the occurrence of landslides and are considered when analyzing landslide susceptibility because they affect rock stability and encourage landslides (Arabameri et al., 2020). Faults have a significant impact on landslides and slope stability because rocks near these structures are often fractured, broken, and weathered, resulting in much lower engineering properties compared to intact rocks (Miller et al., 2009).

In this study, faults were extracted from the geological map (scale: 1: 250,000). The distance to faults was calculated using the Euclidean distance function in ArcGIS 10.5. The maximum distance to faults was found to be 4934 meters. The distances from faults were classified into seven categories: (0 - 750), (750 - 1500), (1500 - 2250), (2250 - 3000), (3000 - 3750), (3750 - 4500), (> 4500) (back to Fig. 4f).

2.3. Hydrological Factors

2.3.1. Topographic Wetness Index (TWI)

The Topographic Wetness Index (TWI) is a function of both the slope and the contributing area per unit orthogonal to the flow direction. It is an indicator that identifies water-saturated areas resulting from runoff under topographic conditions (Mitra et al., 2022). Low TWI values are typically associated with steep slopes and

efficient drainage systems, while high TWI values indicate high surface saturation with gentle slope angles. Increased surface saturation raises the risk of landslides and reduces the shear strength of materials on slopes (Zhao and Chen, 2020).

In our current study, TWI values were calculated using SAGA GIS with the following equation (Eq. 4) (Roy and Saha, 2019). The TWI values ranged from 1.75689 to 23.8445 and were classified into five categories: (< 4), (4 - 6), (6 - 8), (8 - 10), (> 10) (back to Fig. 5a).

$$TWI = \ln \frac{a}{\tan \beta} \quad \textbf{(Eq. 4)}$$

where a is the confluence area per contour length, $\tan\beta$ is the local slope. In addition, *TWI* increases with the increase of runoff accumulation.

2.3.2. Distance from Streams

Previous studies have indicated a relationship between drainage density and factors such as climate, soil, slope, and geological composition (Nohani et al., 2019). Terrain with high drainage density and thin soil layers is typically prone to shallow landslides (Paul and Bhowmik, 2016). Additionally, many researches stated that about 65% of landslides occurred near the first-order drainage network within a distance that range from zero to 40 meters. Proximity to watercourses is a crucial factor in mapping landslide susceptibility and has been widely used in landslide susceptibility studies (Abedi & Feizizadeh, 2021; Moragues et al., 2021; Shano et al., 2021).

In this study, the drainage network was derived from the ALOS-PALSAR digital elevation model using ArcGIS 10.5 with the Hydrology extension. The distance from major streams was calculated using the Euclidian distance function in ArcGIS 10.5, with the maximum distance to streams being 814.9 meters. The distances from streams were classified into seven categories: (0 - 100 meters), (100 - 200 meters), (200 - 300 meters), (300 - 400 meters), (400 - 500 meters), (500 - 600 meters), (> 600 meters) (back to Fig. 5b).

2.3.3. Rainfall

Rainfall is another significant factor influencing the occurrence of landslides due to its direct impact on the stability of surface slopes in a given area (Jennifer et al., 2021). Moreover, heavy rainfall, whether short-term or long-term, controls surface runoff and activates pore water pressure in the soil, leading to soil weakening and destabilization of the terrain (Jennifer et al., 2021). Numerous researchers have incorporated rainfall data in landslide studies, demonstrating it as a crucial factor in landslide occurrences (Gheshlaghi and Feizizadeh, 2021; Hong et al., 2018; Jennifer et al., 2021).

In this study, a raster layer representing rainfall distribution was derived using ArcGIS version 10.5 and the spatial interpolation tool. (IDW) Interpolates a raster surface from points using an inverse distance weighted (IDW) technique (back to Fig. 5c).

2.4. Environmental and Anthropogenic Factors

2.4.1. Land Use/Land Cover (LULC)

Human activities such as road construction, urban expansion, and infrastructure development significantly impact landslides as they alter land use and land cover patterns. These activities often require cutting or excavating slopes, which can lead to slope instability (Xiao et al., 2019). Land use also influences water infiltration rates and, ultimately, the potential for surface accumulation in the soil (Shu et al., 2019). Pham et al. (2016) noted that landslides are generally less likely to occur in forested areas compared to barren or sparsely vegetated regions.

Therefore, mapping landslide susceptibility requires an understanding of the current land use/land cover (LULC) and how it is being utilized, along with precise monitoring over time (Caldwell, 2019). In this study, Landsat 8 (OLI) was processed and analyzed to derive the LULC classified layer. The study area was classified into four categories: vegetation (30.6%), built-up area (3.4%), roads (5%), and bare ground (61% of the total area) (back to Fig. 5d).

Additionally, the accuracy of the LULC resultant image map was assessed using 200 randomly generated points that fall under different LULC categories. The classified LULC raster data were compared with a validated reference dataset using high-resolution Google Earth images. The relationship between these two datasets was evaluated using an error matrix (confusion matrix) and the Kappa coefficient. An overall accuracy value above 85% and a Kappa coefficient between 0.7 and 0.8 are considered good for validation (Mas et al., 2022). In this study, the overall accuracy was 96.5% and the Kappa coefficient was 0.94, indicating a significantly high classification accuracy for LULC (Table 2).

2.4.2. Normalized Difference Vegetation Index (NDVI)

Natural vegetation is one of the commonly used factors in landslide susceptibility analysis, as it helps stabilize slopes and prevent landslides (Liu et al., 2018). NDVI is a scientific measurement index reflecting the density and vitality of vegetation cover, derived from remote sensing (RS) data. It is calculated using the difference between the red-spectrum reflection and the near-infrared (NIR) in the electromagnetic spectral radiation. Healthy vegetation has a high NDVI value, indicating its absorption of red light and reflection of a greater amount of NIR light (Gao, 1996). Conversely, areas with low NDVI values indicate bare soil or unhealthy vegetation, which may be more prone to landslides due to the lack of stabilizing roots (Liu et al., 2018).

In this study, the Normalized Difference Vegetation Index (NDVI) was calculated using 30-meter resolution satellite images from Landsat 8, utilizing Band 4 (Red) and Band 5 (Near Infrared), captured on April 10, 2023. These images are currently available for public download through NASA's Earth Explorer website. The selection of satellite images with minimal or no cloud cover (Cloud Cover = 0.07) or complete cloud removal was considered, as it is a crucial requirement for calculating the NDVI. This index highlights areas with potential landslide hazards (Alvarez-Mendoza et al., 2019). NDVI can also be combined and infused as a musk band layer

Table 2. Accuracy assessment of land use/ land cover (LULC) map using kappa coefficient (*k*).

LULC Classes		Built up area	Roads	Vegetation	Bare ground	Total (User)	User's Accuracy%	Producer's Accuracy %
		1	2	3	4			
Built up area	1	9	1	0	0	10	90.00	90.00
Roads	2	1	11	0	0	12	91.67	91.67
Vegetation	3	0	0	60	2	62	96.77	95.24
Bare ground	4	0	0	3	113	116	97.41	98.26
Total (Producer)		10	12	63	115	200		
Overall accuracy		96.50%						
kappa Coefficient(k)		0.94						

with other factors, such as topography and land use, to produce detailed landslide hazard maps (Liu et al., 2018). In this study, the NDVI map was derived using the Raster Calculator in ArcGIS 10.5 with the following equation (Eq. 5):

$$NDVI = \frac{NIR - RED}{(NIR + RED)} \quad \textbf{(Eq. 5)}$$

Where NIR is the reflectance of Near Infrared spectrum and Red is the reflectance of red spectrum. In this study, the NDVI values range from -0.0214 to 0.3229 (Fig. 5e), with negative values indicating areas devoid of vegetation and barren regions (such as rocky or sandy terrains and even landslide areas), while positive values indicate healthy green vegetation (Pradhan et al., 2017).

2.4.3. Distance from Roads

The construction of mountain roads is a significant anthropogenic factor affecting the stability of natural slopes (Xiao et al., 2019). This is due to the required engineering work, such as cutting or excavating slopes, which weaken and destabilize them, leading to landslides (Jennifer et al., 2021). Therefore, landslides are often distributed near constructed or under-construction roads (Pham et al., 2019). Thus, the distance from roads can be considered as a main factor contributing to landslide occurrence (Chen et al., 2019; Moragues et al., 2021; Ozioko & Igwe, 2020). In this study, a raster layer representing the distance from roads was created using the Euclidean distance tool in ArcGIS 10.5 and classified into seven categories. The maximum road distance in this study is 7770 meters (back to Fig. 5f).

3. Multicollinearity Assessment

Multicollinearity refers to the presence of a high correlation between two or more independent variables in a multiple regression model. The fundamental rule in selecting independent variables (landslide conditioning factors) is that they should exhibit a weak correlation among themselves but a strong correlation with the dependent variable (landslide locations). Factors that are highly correlated with each other have the same effect and respond in the same manner, which can impact the prediction model (Kalantar et al., 2020) and lead to incorrect systematic analysis (Dormann et al., 2013). Therefore, testing for the strength of any linear correlation

between two or more independent variables is essential to assess the robustness of the landslide model (Nwazelibe et al., 2022). The Variance Inflation Factor (VIF) and Tolerance Level (TOL) are commonly used to test for multicollinearity issues.

The theoretical critical value for the Variance Inflation Factor (VIF) is typically 5 or 10. Values greater than these indicate a very strong relationship between two or more independent variables, suggesting multicollinearity issues. Similarly, the theoretical critical value for the tolerance level (TOL) is either 0.1 or 0.2. TOL values below 0.2 indicate somewhat weak multicollinearity between the selected independent variables, while values below 0.1 indicate strong multicollinearity (Sujatha & Sridhar, 2021). Multicollinearity is measured using the following equation (Eq. 6):

$$\mathrm{VIF} = \frac{1}{1-R_i^2} = \frac{1}{\mathrm{TOL}} (i = 1, 2, 3 \cdots\cdots k) \qquad \textbf{(Eq. 6)}$$

Where *Ri* denotes the correlation coefficient when the independent variable *Xi* is a Regression coefficient on the remaining variables

4. Landslide Susceptibility Suggested Models

4.1. Artificial Neural Network (ANN) Model

The Artificial Neural Networks (ANN) is a machine learning algorithm that mimics the structure of neural networks in the human brain. It has been effectively used in the field of landslides and is widely employed to map landslide susceptibility (Selamat et al., 2023). The ANN model predicts future landslides based on the historical distribution of landslide occurrences, making it a valuable tool for assessing the likelihood and risks of landslides. Therefore, this model has been widely used in predictive studies of landslides (Hu et al., 2021). The ANN is essentially the development of an intelligent mathematical model that mimics human cognition and biological neural networks. It consists of interconnected units that form a specific structure. The ANN is characterized by its ability to recognize multiple sets of data within a wide range of datasets without the need for prior expertise, pre-existing knowledge, or a predefined framework for data training (Yao et al., 2008).

In this study, the back-propagation training algorithm was used as it is one of the most commonly employed ANN algorithms by researchers in the field of landslides (Zhao et al., 2022). The ANN utilized the multi-layer perceptron (MLP) architecture, which includes three components: the input layer, the hidden layer, and the output layer. The input layers are created based on the landslide conditioning factors selected for model development.

The input layer has a dedicated neuron for each landslide conditioning factor, which connects to the hidden layers. The hidden layers are intermediate components situated between the input and output layers. They receive data from the input neurons via interconnections, process this data, make predictions about the output neurons, and then transfer it to the output layer through these interconnections. The output layer represents the landslide predictions, which were used in this study to classify areas prone to landslides and areas that are not. The proposed designed structure of the model is shown in (Fig. 7)

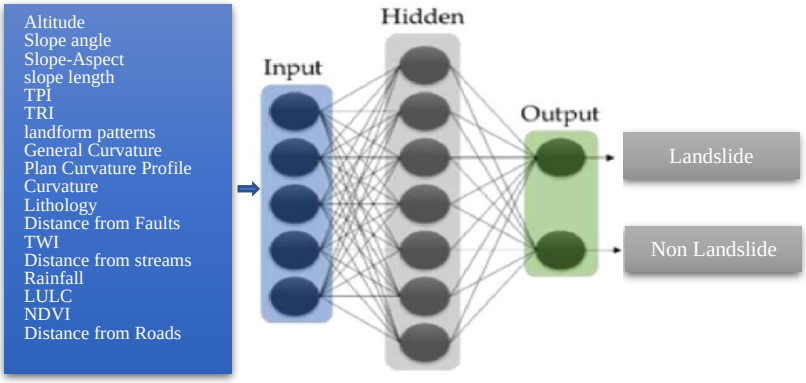

Fig. 7. The basic proposed designed structure of Machine Learning MLP-ANN of landslide susceptibility mapping of the study area.

4.2. Frequency Ratio (FR) Model

The frequency ratio (FR) is one of the most commonly used bivariate statistical analysis methods for assessing landslide susceptibility (Zhang et al. 2020). The FR method is characterized by its ease of implementation, compatibility with Geographic Information System (GIS) technology, and its ability to provide accurate results. The application of the FR method relies on the assumption of all future events can be predicted based on past information (Chimidi et al., 2017). The FR value is calculated by dividing the ratio of the area where landslides have occurred by the total study area. It also represents the ratio of the likelihood of landslide occurrence to non-occurrence within different classification categories for each landslide susceptibility assessment factor, thereby measuring the influence of each factor on landslides (Regmi et al. 2014). Generally, a ratio greater than one indicates a strong relationship between the conditioning factor and landslides, suggesting a high likelihood of landslide occurrence. Conversely, values less than one indicate a low relationship with the probability of landslides, while a value of one signifies a neutral relationship for landslide occurrence in the overall area. The FR value can be calculated using the following equation (Eq. 7):

$$FR_i = \frac{(LS_i/LS)}{(A_i/A)} \quad \textbf{(Eq. 7)}$$

Where *FRi* = frequency ratio of *i*th class, *LSi* = total landslide area (number of landslide pixels) in the *i*th class, *LS* = total landslide area (total number of landslide pixels) in the study area, *Ai* = area falling under *i*th class (total number of pixels of *i*th class) and *A* = total area (total number of pixels of the entire map).

These FR values of different classes of the conditioning factors (back to Table 1) are then used to obtain the prediction rate (PR) of each factor which depicts the weightage of the individual class, using Eq. (8-10) as follows:

$$RF_i = \left(FR_i \Big/ \sum FR\right) \quad \textbf{(Eq. 8)}$$

$$R_j = MAX(RF_{i,j}) - MIN(RF_{i,j}) \quad \textbf{(Eq. 9)}$$

$$PR_j = R_j / MIN(R) \quad \textbf{(Eq. 10)}$$

Where *RF* is the relative frequency, *MAX* (*RFi, j*) is the maximum value of *RF* of *j*th factor, *MIN(RFi, j)* is the minimum value of *RF* of *jth* factor, *PRj* is the prediction rate of *jth* factor. The FR values obtained by using Eq. (8) will act like the weight of each class (*wij, FR*), and *PRj* is converted to a percentage, which will be the weight of the *j*th factor, i.e. *Wj, FR*. To generate the landslide susceptibility index (LSI) using the FR method, the weightage to each class of every landslide conditioning factor is given as per the corresponding FR values obtained by Eq. (8) and then integrated with the corresponding weight of each element. Named after Claude Shannon, the Shannon Entropy (SE) Model is one of the most efficient bivariate statistical methods that used to measure the influence of conditioning factors on slope instability and landslide occurrence. It examines the relationship between the likelihood of landslides and their causative factors, as well as the categories of those factors. This technique has the advantage of allowing for the weighting of the factors and their categories (Constantin et al. 2011; Jaafari et al. 2014). The larger the SE index, the more significant the contributing factor to landslide occurrence (Sujatha, 2012). The SE value is calculated using the following equations (Eq. 11-13):

$$\left(P_{ij}\right) = \frac{\mathrm{FR}}{\sum_{i=1}^{m} \mathrm{FR}}, \quad \textbf{(Eq. 11)}$$

where *m* is the number of landslides that have occurred, and FR is the frequency ratio. As a result, the normalized decision matrix *Pij* can be defined as follows (Eq. 12) for each landslide criterion:

$$E_j = -K \sum_{i=0}^{m} P_{ij} \ln P_{ij}, \quad \textbf{(Eq. 12)}$$

where E_j is the entropy value, P_{ij} is the value of the *i*th landslide in the *j*th criterion, and *k* is a positive instant given as $(\ln m)^{-1}$. The weights (Wj*)* were assigned to the roles of the variables influencing the synthesis grade, with higher scores indicating greater significance of the variable's contribution within the rating scheme. Where V_j is defined $1 - E_j$. The weights were calculated by using (Eq. 13).

$$W_j = \frac{V_j}{\sum_{i=0}^{m} V_j}. \quad \textbf{(Eq. 13)}$$

Validating predictive models is an essential part of landslide susceptibility research (Hong et al. 2018; Pourghasemi & Rossi. 2017). A landslide susceptibility map becomes ineffective without model validation (Mersha and Meten, 2020; Pham et al. 2017). There are several statistical measures to evaluate the performance of

predictive landslide models, including: Receiver Operating Characteristic Curve (ROC) and Area Under the Curve (AUC) ,landslide Relative Density Index (R-index), a Model -Success and Predictive Rate Curve. the sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and kappa statistics (Azemeraw، 2021; Selamat, 2022).

This research study employs the ROC_AUC model to evaluate the performance of landslide susceptibility maps (LSM) derived from the FR, ANN, and SE models. AUC is one of the most widely used metrics for assessing the performance of predictive models (Zhao et al. 2022; Sengupta & Nath, 2022; Tingyu & Nath, 2022). Additionally, AUC represents a graph between the success rate curve, indicating the model's ability to classify areas into categories at risk of landslides using the training dataset, and the prediction rate curve, reflecting the model's ability to predict future landslide occurrences using validation data (Kalantar et al. 2018). The ROC graph was plotted using the "Calculate ROC Curves and AUC Values" tool from the ArcSDM5 toolbox (Mas et al. 2013). AUC values range from 0.5 to 1, and they are classified as follows: excellent (0.9-1.0); very good (0.8-0.9); good (0.7-0.8); average (0.6-0.7), and fair (0.5-0.6) (Yesilnacar& Topal. 2005). Eventually, AUC_ROC resultant values that are close to 1 indicate more accurate and reliable predictions in model performance, while the model is considered weak when the resultant values are less than or equal to 0.5 (Mfondoum et al., 2023). The AUC value can be calculated using the following equation (Eq 14):

$$AUC = \frac{(\sum TP + \sum TN)}{(P + N)} \quad \textbf{(Eq. 14)}$$

where TP (true) and TN (true negative) denote the correctly classified raster cells, P expresses the total number of landslide raster cells, and N represents the total number of non-landslide raster cells.

V. Results and Discussion

1. Multicollinearity Analysis

The Variance Inflation Factors (VIF) and Tolerance Level (TOL) indicators were used to detect and measure multicollinearity among the eighteen landslide conditioning factors. The VIF and TOL values were calculated using SPSS (Table 3). The results of the analysis showed that all VIF values were below the critical threshold, with the highest VIF value being 5.76 and the lowest being 1.29. Regarding TOL values, all were above the critical threshold, with the highest and lowest TOL being 0.77 and 0.17 for the lithology and landform patterns factors, respectively. This indicates that the multicollinearity assessment for the eighteen specified factors met the critical thresholds. Therefore, none of the identified conditioning factors for building the landslide prediction model exhibited multicollinearity.

Table 3. Variance Inflation Factors (VIF) and Tolerance (TOL) analysis.

Conditioning Factors	Multicollinearity analysis	
	Tolerance	VIF
Altitude	0.391	2.556
Slope angle	0.210	4.772
Slope-Aspect	0.762	1.312
slope length	0.698	1.433
TPI	0.193	5.176
TRI	0.673	1.486
landform patterns	0.174	5.760
General Curvature	0.702	1.424
Plan Curvature	0.654	1.529
Profile Curvature	0.564	1.772
Lithology	0.770	1.299
Distance from Fault	0.637	1.569
TWI	0.207	4.836
Distance from Stream	0.317	3.153
Rainfall	0.589	1.697
LULC	0.629	1.589
NDVI	0.646	1.547
distance from Road	0.581	1.721

2. Resultant Landslide Susceptibility Predictive Models

2.1. Integrating GIS with ML-Artificial Neural Networks (ML-ANN)

Understanding the factors causing landslides is crucial for efficiently managing their risks. Therefore, landslide studies are essential for improving landslide prevention and risk assessment. The landslide predictive model using the Machine Learning Artificial Neural Networks (ML-ANN) serves as a valuable intelligent modeling approach for identifying at-risk areas and predicting the likelihood of landslides. This model has been used by many researchers as a reliable predictive modeling algorithm (Jacinth Jennifer & Saravanan, 2022; Orhan et al., 2022; Mehrabi & Moayedi, 2021).

In this study, an approach of integrating Geographic Information Systems (GIS) with ML-ANN was employed to make predictions for the landslide model. This approach relied on pre-prepared spatial information as landslide causative factors, along with inventory data of past and present landslides, which are considered the best indicators for future predictions (Ma et al., 2021). Upon reviewing the literature on landslide studies, 18 factors influencing the occurrence of landslides were selected. These factors are: (Altitude, Slope angle, Slope-Aspect, slope length, TPI, TRI, landform patterns, “General, Plan and Profile Curvature”, Lithology, Distance from Fault, TWI, Distance from Stream, Rainfall, LULC, NDVI, distance from Road) (Back to Fig. 3, 4 and 5).

The final results indicated that the conditioning factors such as slope angle, slope aspect, landform patterns, distance from streams, distance from roads, TPI, lithology, slope length, and TWI were the most significant factors influencing the occurrence of landslides. Conversely, general curvature and rainfall were identified as the least important factors for landslide occurrence in the study area (Fig. 8a).

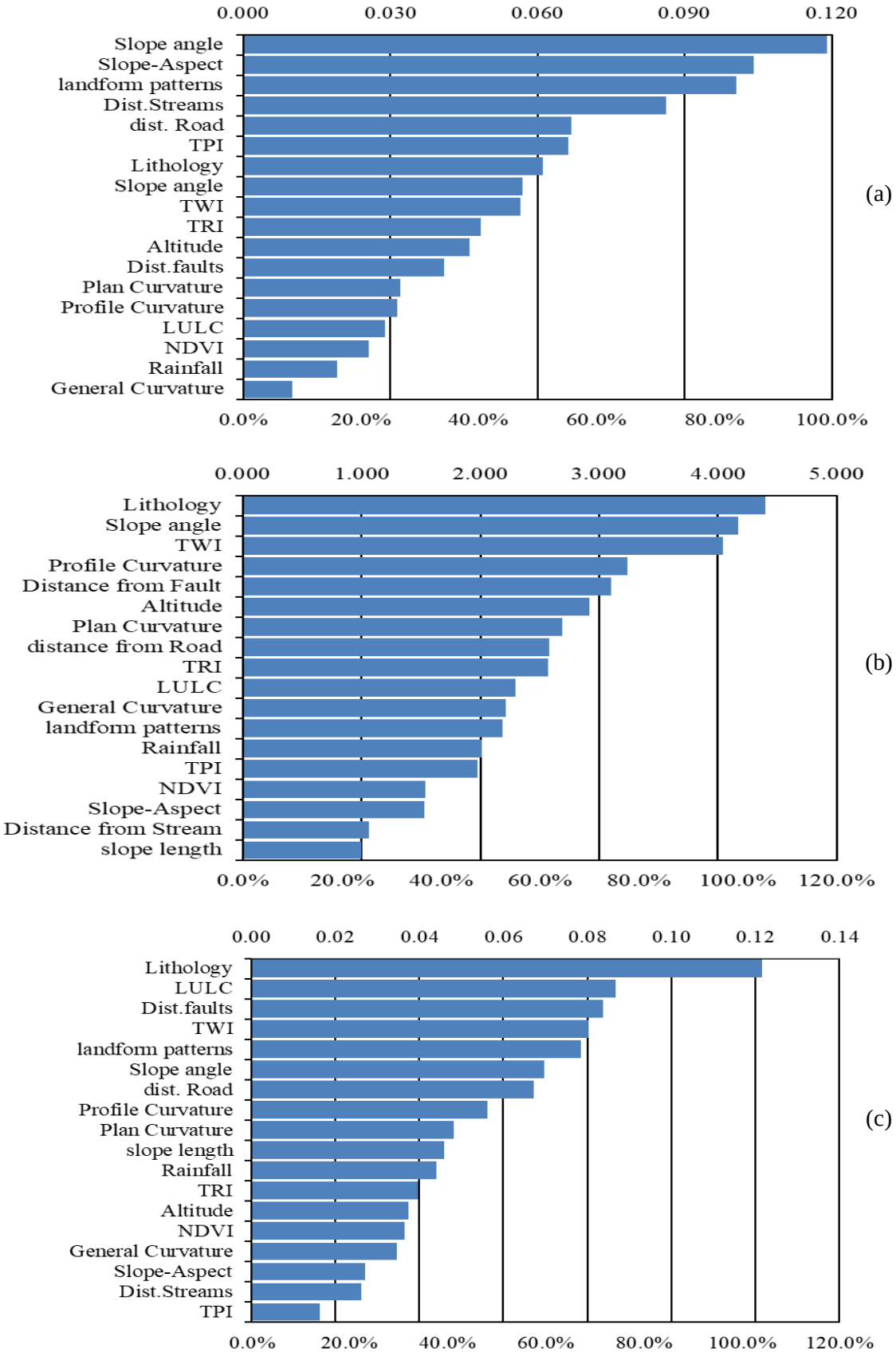

Fig. 8. Importance of influencing factors differentiation according to the examined three landslide utilized predictive modeling approaches (a) ANN, (b) FR and (c) SE.

According to ranking the resultant influence importance of the selected landslide conditioning factors that have been extracted and analyzed in an integrated GIS/ML-ANN method, the first ranking factor is the slope angle which is the most significant factor influencing landslides in Wadi Dil'ah basin.

An increase in slope angle leads to the instability of rock and soil masses (Jebur et al., 2014), making it a widely used factor in landslide studies. The results indicated that slope aspect is the second important ranking factor that effects landslide occurrence, as it affects various processes that have direct and indirect impacts on landslides, such as wind directions, rainfall, vegetation cover, sunlight exposure, evaporation, transpiration, and soil moisture concentration (Devkota et al., 2013).

The results also showed that landform patterns are the third most important ranking factor affecting landslides in the study area. The fourth most important factor is the distance from streams, as most landslide incidents in the study area occur in mountainous regions and near drainage networks.

Additionally, the distance from roads is a significant factor contributing to landslides. Most landslide incidents in the study area occur in mountainous regions and near roads. Road construction in mountainous areas negatively impacts slope stability because it always imposes an engineering load and damages the slope structure (Tien Bui et al., 2016). Therefore, any road construction activities involving the cutting of slopes steeper than 10 degrees cause soil and rock disruptions (Nohani et al., 2019), leading to landslides.

2.2. Frequency Ratio (FR) Model

As it is displayed in Table 1, the results demonstrated the spatial relationship between landslide locations and their causative factors using the Frequency Ratio (FR) model. The results indicated that lithology is the most influential factor in landslide occurrences (Fig. 8b), with a relative impact value of FR equal to 4.40. In terms of geological formations, the Bahah Group has an FR value of 1.334, while the other formations have zero values. This is attributed to the predominance of these formations in the study area, as they constitute 74.9% of the total area of Wadi Dil'ah basin. Additionally, many steep roads have been constructed through these units without any engineering specifications, and numerous landslides have been recorded along their lengths. Furthermore, these units are located in areas characterized by high and steep cliffs.

The slope angle factor ranked second in terms of relative impact, with a value of 4.17. The slope category of (45 – 67) degrees recorded an FR value of 34.73, followed by the slope category of (30 – 45) degrees with an FR value of 1.76, while the remaining categories recorded values below one. The TWI index ranked third with an FR value of 4.04, indicating that the category of TWI greater than 10 has the highest FR value of 4.41, while the other categories had values less than one. The profile curvature factor ranked fourth with a value of 3.24, where the concave category recorded the highest FR value of 1.59, while the Convex and Flat categories had the lowest FR values of 0.57 and zero, respectively. The distance from faults ranked fifth with a relative weight of 3.10, showing that the distance category of (0-750 m) recorded the highest FR value of 2.16. The probability of landslide occurrences decreases at distances greater than 750 meters, with the categories of 750-

1500 m recording a FR value of 0.544, followed by 1500-2250 m with 0.144, 2250-3000 m with 0.128, and 3000-3750 m with 0.088. While the other categories recorded zero values, the elevation category of 2000-2500 m had the highest FR value of 4.214, followed by the categories of 1000-1500 m (0.954), 1500-2000 m (0.938), and 500-1000 m (0.246). The elevation categories greater than 2500 m and less than 500 m recorded the lowest FR values (0.00). The results also indicate that the FR values for elevation categories do not show significant correlations with decreasing elevation in the study area. Pachauri and Pant (1992) noted that higher elevations are associated with a greater susceptibility to landslides. In the current study, the high FR values may be related to areas of steep slopes that are interspersed with many streams and gullies. Regarding the plan curvature factor, it had an impact ratio of 2.69.

The convex category recorded the highest FR value of 1.325, while the concave and flat categories recorded the lowest FR values of 0.842 and zero, respectively. Regarding the distance from roads, it can be observed that the general trend of FR values increases with decreasing distance. The results indicate that roads have a significant impact on landslides, primarily due to the rock-cutting processes associated with the construction of escarpment roads on slopes in the study area, which have weakened the stability of the rocks and consequently led to landslides along the roads.

As for the TRI index, the relative impact value for this factor was 2.57, with the category 0.12-0.16 showing the highest FR value of 18.065, followed by the categories 0.08-0.12 and 0.04-0.08, with values of 8.972 and 3.536, respectively, while the other categories recorded values less than one. For the LULC factor, bare ground had the highest FR value of 1.483, while the remaining LULC patterns recorded values less than one, indicating that barren and desolate areas are the most susceptible to landslides due to their exposure to erosion and soil moisture (Solaimani et al. 2013).

For the General Curvature factor, the convex category recorded the highest FR value of 1.733, followed by the concave category at 0.571, while the flat category had the lowest values. Regarding the Landform Patterns factor, the results showed that the Summit category had the highest FR value of 5.059, followed by the Ridge and Slope categories with values of 3.125 and 1.775, respectively. The Spur category had an FR value of 0.206, while the remaining landform patterns recorded zero values.

Concerning the rainfall factor, the results indicated that the rainfall category of 300-400 mm recorded the highest FR value of 1.215, followed by the 400-500 mm category at 0.979, and the 250-300 mm category at 0.468. The 200-250 mm category had a zero value. The results show that heavy rainfall, which frequently occurs along the escarpment area at higher elevations, increases the likelihood of landslides.

Regarding the TPI index, the results showed that the category (>70) had the highest FR value of 2.848, followed by the category (30-70) at 1.731, while the remaining categories recorded values below one, with the category (<-50) having the lowest value of 0.111. Concerning the NDVI index, the values between (0.0717: 0.0906) and (-0.0214: 0.0717) had the highest FR values of 1.432 and 1.226, respectively. This indicates that the categories with high FR values represent barren and sparsely vegetated areas, which are more prone to landslides. Abdi et al. (2010) pointed out that vegetation can protect slopes by reducing erosion effects,

strengthening the soil, enhancing slope stability, and limiting the occurrence of landslides.

Regarding the slope aspect, most landslides occurred in the eastern, southeast, and northeast directions. These directions recorded the highest FR values of 2.912, 1.427, and 1.381, respectively, indicating a higher likelihood of landslide occurrences. However, other slope aspect categories recorded values below one, suggesting a lower probability of landslides. The high frequency of landslides in specific slope aspect directions may be associated with local conditions, such as prevailing wind and storm directions, fault orientation, and rock structure.

Regarding the Distance from the stream factor, the distances between 200-300 meters, 100-200 meters, and 300-400 meters recorded the highest FR rates, with values of 1.338, 1.271, and 1.186, respectively. It is noteworthy that the general trend of FR values increases as the distance from water bodies decreases. Therefore, it can be said that proximity to streams is one of the most important factors contributing to slope instability and the frequency of landslides.

As for the slope length factor, it ranked last in terms of relative impact (1.0) on landslide occurrences. The relationship between slope length and landslide probability is evident; the category of slopes greater than 45 meters recorded the highest FR value (4.194), followed by the categories of 15-30 meters, 30-45 meters, and 5-15 meters, with FR values of 3.975, 3.330, and 2.968, respectively. Meanwhile, the category of slopes less than 5 meters recorded the lowest FR value (0.737). In general, the results indicate that the trend of FR values increases with increasing slope length.

2.3 Shannon Entropy (SE) Model

The Shannon Entropy (SE) method is one of the preferred modeling approaches for mapping landslides due to its flexibility and ease of computation. The higher the SE value corresponding to a particular factor, the greater the significance of that factor and its strong influence on the occurrence of landslides. Conversely, a lower SE value indicates a diminished discriminatory power of that factor in the decision-making process (Lotfi & Fallahnejad, 2010).

The weights in the SE Model were calculated based on the values of the Frequency Ratio (FR). The results of the SE analysis indicated that the lithology factor is the most supportive of landslide occurrence, with a value of (0.122), followed by LULC (0.087), distance from faults (0.084), TWI (0.080), landform pattern (0.078), slope angle (0.070), and distance from roads (0.067). On the other hand, there are factors that have less influence on landslide occurrence compared to the aforementioned factors, which are: plan curvature (0.048), slope length (0.046), rainfall (0.044), TRI (0.040), altitude (0.037), NDVI (0.037), curvature (0.035), slope aspect (0.027), distance from streams (0.026), and finally, the TPI index, which has a weight of (0.016), making it the least significant factor in terms of its impact on landslide occurrence in the study area (Fig. 8c).

The subcategories of the factors most influential in landslide occurrence have been identified (back to Table 1), where the Pij values for the lithology factor indicate that the Bahah Group formation is the most influential on landslide occurrence (Pij=1). The Pij values for the LULC factor show that bare ground is the area most prone to landslides, with the highest value of (0.598) Pij, followed by roads

(Pij=0.238), built-up areas (Pij=0.087), and then vegetation (Pij=0.077). The results for the distance from faults indicate that the category (0-750 m) recorded the highest Pij values (0.705), followed by the category from 750-1500 m (Pij=0.178), indicating a correlational relationship between proximity to faults and landslide occurrence.

The TWI index recorded the highest Pij values (0.919) for the category (> 10), indicating that this range is the most susceptible to landslides, followed by the category from 8-10 (Pij=0.081). The results show that the summit has the highest probability of landslide occurrence, with a value of (Pij=0.498), followed by the ridge (Pij=0.307), slope (Pij=0.175), and spur (Pij=0.020).

The Pij results revealed that the slope category (45–67) is the most susceptible to landslides, with a value of (Pij=0.949), followed by the slope category (30-45) with (Pij=0.048) and (18-30) with (Pij=0.003). The Pij values increase as the distance to roads decreases, indicating that road construction and vehicle movement contribute to the destabilization of rocks, leading to landslides. The highest Pij value (0.586) was recorded for the distance category from 0 to 1000 meters, followed by the category from 1000-2000 meters (Pij=0.282), the category from 2000-3000 meters (Pij=0.083), and the category from 3000-4000 meters (Pij=0.049).

Regarding profile curvature, it was found that concave (Pij=0.736) has the highest impact on the likelihood of landslides, followed by convex (Pij=0.264). The Pij values for plan curvature indicate that convex (Pij=0.612) is the most influential factor in landslide occurrence, followed by concave (Pij=0.338). The probability of landslides increases with the slope length; the category (>45 meters) recorded the highest Pij value (0.276), followed by the category between 15-30 meters (Pij=0.261), 30-45 meters (Pij=0.219), and 5-15 meters (Pij=0.195), while the category (<5 meters) recorded the lowest Pij value (Pij=0.048).

The results show a clear relationship between heavy rainfall and the increased likelihood of landslides, with the category of 300-400 mm recording the highest Pij value (0.456) for average annual rainfall, followed by the category of 400-500 mm (Pij=0.368) and 250-300 mm (Pij=0.176). The TRI results indicate that the category of 0.12-0.16 had the highest Pij values (0.585), followed by the categories of 0.08-0.12 (Pij=0.291) and 0.04-0.08 (Pij=0.115), suggesting that landslides are more likely to occur in these areas. The highest Pij value (0.663) was found in the altitude category of 2000-2500 m, followed by 1000-1500 m (Pij=0.150) and 1500-2000 m (Pij=0.148). The arid and sparsely vegetated areas in the NDVI layer represent the most susceptible regions for landslides, with the category between (0.0717: 0.0906) recording the highest Pij values (0.373), followed by the category (-0.0214: 0.0717) (Pij=0.320) and the category (0.0906 - 0.1109) (Pij=0.244).

Meanwhile, Pij values decrease in the remaining categories, indicating the role of dense vegetation in strengthening the soil, protecting slopes, stabilizing them, and reducing the likelihood of landslides. For the General Curvature factor, the highest Pij value (Pij=0.612) was for convex, followed by concave (Pij=0.213), while the lowest value (Pij=0.142) was for flat. The density of landslides increases on slopes facing east, southeast, and northeast, with Pij values for these directions being (0.347, 0.174, 0.164), respectively. Regarding Distance from Stream, the category between 200 and 300 meters recorded the highest Pij value (0.242), followed by the categories of 100-200 meters (0.230), 300-400 meters (0.214), 0-100 meters (0.119), 400-500 meters

(0.102), and 500-600 meters (0.093). TPI values indicate that the category (>70) has the highest Pij value (0.468), followed by the category of 30-70 (0.284), the category of -10 to 30 (0.164), the category of -10 to -50 (0.065), and the category (<-50) (0.018).

3. Creating Landslide Susceptibility Maps (LSMs)

The primary objective of this study is to create comprehensive landslide susceptibility maps and evaluate the effectiveness of the Machine Learning Artificial Neural Network (ML-ANN) model and Bivariate techniques (FR and SE) in predicting landslide susceptibility in Wadi Dil'ah basin study area. To achieve this goal, an approach of integrating Geographic Information Systems (GIS) and Remote Sensing (RS) with statistical models and artificial machine learning algorithms were adopted to produce distinct maps indicating the level of susceptibility to landslides.

The generated maps have provided valuable insights into the potential likelihood of future landslide occurrences. By comparing the maps obtained from the three models with the existing landslide data (actual points) and non-landslide data, we aimed to assess the performance of each model. To achieve this, a testing dataset (comprising 30% of the total landslide inventory and 30% of the total non-landslide data) was used.

To reveal the similarities and differences between the results of the models used in the study, the derived maps from the three models were converted to standardized values ranging from 0 to 1 using Fuzzy Membership. Following this, the derived maps were classified into five categories (very high, high, moderate, low, and very low) using the Natural Breaks (Jenks) classification tool in ArcGIS. The low to very low sensitivity areas on the maps indicate stable areas not prone to landslides (non-landslide areas), while the high to very high sensitivity areas indicate unstable regions.

The stable areas not prone to landslides are represented in the models as follows: 65.45% in the ANN model, 35.78% in the FR model, and 31.41% in the SE model. The moderate areas are represented as 11.93%, 30.11%, and 28.69%, respectively. Meanwhile, the unstable areas output by the ANN model constitutes the lowest percentage (22.62%), followed by the FR model (31.11%), with the SE model producing the highest percentage (39.90%) (Table 5, Fig. 9).

The percentage distributions revealed that the different LSM models handle the same conditioning factors differently to produce various landslide-prone areas. The Machine Learning algorithm (ANN Model) differs from the Bivariate Statistical models (FR, SE) in its outputs; it generates a higher number of stable areas and fewer unstable areas. In contrast, the statistical models allocate higher percentages to unstable areas. Highlighting the low, medium, and high areas, the ANN model reveals more consistent patterns in its results when compared to previous landslide incidents, while the FR and SE models appear to overestimate the occurrence of landslides.

4. Validation of Landslide Susceptibility Maps

Model validation is a crucial step in developing susceptibility maps and determining their predictive capabilities for natural hazard management (Alam et al.

2024). For this purpose, a dataset was used (70% training data, 30% testing data), validated to create an Area Under the Curve (AUC curve). The testing dataset is used

Table 5. Comparing percentages of pixels in each class of the selected different LSM models.

	ANN	FR	SE
Class of LS	% of Pixel in classes	% of Pixel in classes	% of Pixel in classes
Very low	64.30	11.04	10.22
Low	1.15	24.74	21.19
Moderate	11.93	30.11	28.69
High	3.97	22.55	27.71
Very high	18.65	11.56	12.19

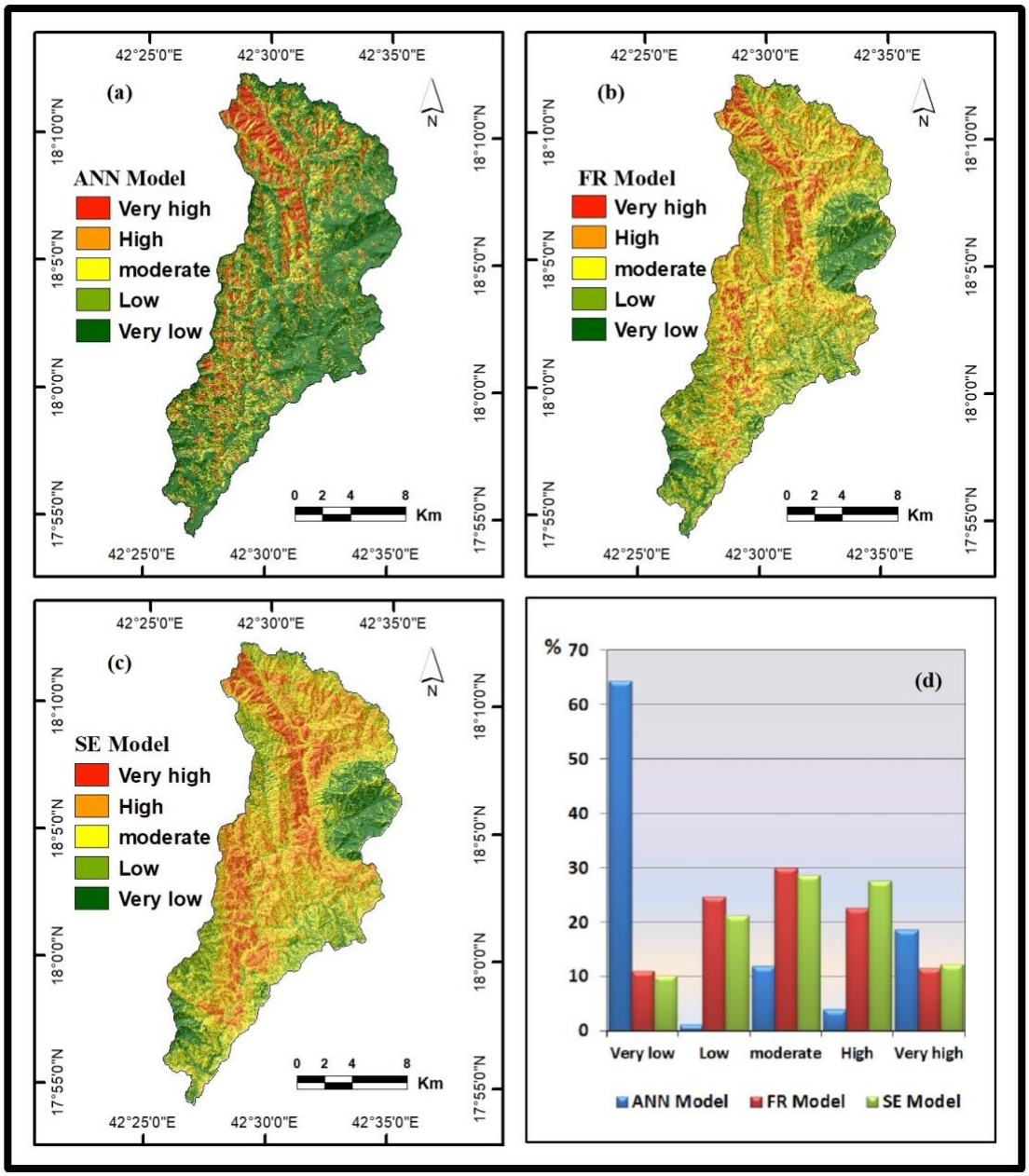

Fig. 9. Resultant Landslide susceptibility maps in Wadi Dil'ah Basin of the selected proposed models (a) Artificial Neural Network (ANN), (b) Frequency Ratio (FR), (c) Shannon Entropy (SE), (d) Landslide susceptibility class distribution bar chart of the three examined models.

for model validation purposes. The Receiver Operating Characteristic curves (ROC) demonstrated the relationship between sensitivity and specificity, and the results of the ROC-AUC model -statistical measures that evaluate the performance of all examined predictive landslide models- showed that the performance of the LSMs models is close to each other, with some minor differences. When comparing the validation of the training and testing inventory data, the performance accuracy was similar, with the AUC values for the testing inventory slightly outperforming the training inventory for all models. The ANN model achieved success in terms of training and testing accuracy (AUC=0.966, 0.983), followed by the FR model (AUC=0.955, 0.993), and the SE model (AUC=0.953, 0.971) (Fig. 10 and 11).

The training and testing performance for all examined models was found to be close, with the difference between the lowest and highest values (ANN, SE) not exceeding 1.3% and 2.2%, respectively. The smallest difference in training performance was observed between the FR and SE models (0.2%), indicating a high degree of similarity between these two models compared to the ANN model. On the other hand, the testing performance showed the smallest difference between the ANN and FR models (1%), suggesting a similarity between these two models. Although the SE model recorded the lowest performance values for both training and testing, it cannot be considered significantly different from the ANN and FR models, as the slight difference is not of substantial importance.

Overall, the performance of the ML-ANN model was superior to the bivariate statistical models (frequency ratio (FR) and Shannon Entropy (SE) models). It is noteworthy that all the models used had an AUC value higher than 0.9, indicating that they predict the probability of landslides with excellent reliability.

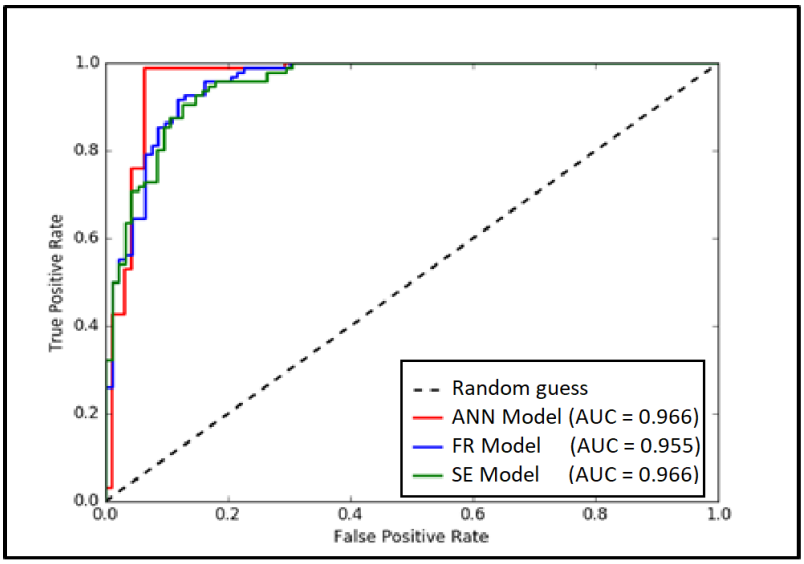

Fig. 10. ROC curve for all the three examined models using the ***Training Datasets***.

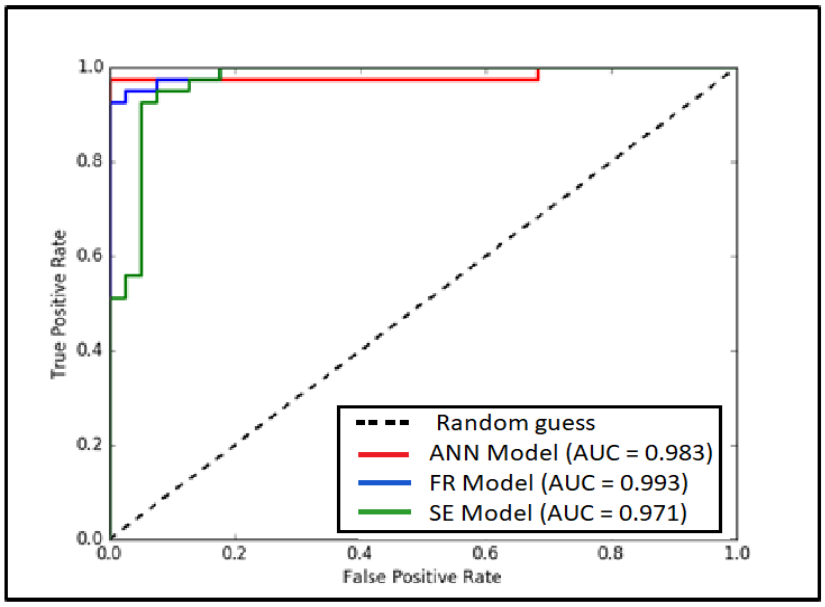

Fig. 11. ROC curve for all the three examined models using the ***Testing Datasets***.

VI. Conclusions

This study employs the Machine Learning Artificial Neural Network (ML-ANN) model and Bivariate Statistical Models to create a landslide susceptibility map for Wadi Dil'ah basin area in Asir region of Saudi Arabia. A total of 137 landslide locations were identified and distributed in a 70/30 ratio to form training and testing datasets. An equal number of locations that did not experience landslides (non-landslide areas with slope angles less than 2˚) were also identified. Eighteen influential factors contributing to the occurrence of landslides were selected, including Altitude, Slope angle, Slope-Aspect, slope length, TPI, TRI, landform patterns, "General, Plan and Profile Curvature", Lithology, Distance from Fault, TWI, Distance from Stream, Rainfall, LULC, NDVI, distance from Road.

Multicollinearity was assessed before building the predictive models. The results of the Variance Inflation Factor (VIF) and Tolerance Level (TOL) indicators demonstrated that the assessment of multicollinearity for the eighteen selected factors met the critical thresholds, indicating no multicollinearity issues among the factors.

The back-propagation training algorithm and the Multi-Layer Perceptron (MLP) architecture were used for the ANN. The results of the ANN indicated that the conditioning factors of slope angle, slope aspect, landform patterns, distance from streams, distance from roads, TPI, geological factors, slope length, and TWI are the most significant factors influencing the occurrence of landslides.

The results from the Frequency Ratio (FR) analysis indicated that lithology, slope angle, TWI, profile curvature, distance from faults, altitude, plan curvature, and distance from roads are significant factors. Meanwhile, the results from the Shannon Entropy (SE) analysis revealed that lithology, LULC, distance from faults, TWI, landform pattern, slope angle, distance from roads, and profile curvature are the most influential factors supporting the occurrence of landslides.

The landslide susceptibility map was categorized into five classes, where the low to very low sensitivity areas indicate stable regions that are not prone to landslides (non-landslide), while the high to very high sensitivity areas represent unstable

regions. The classification was performed using the Natural Breaks (Jenks) tool. The results showed that stable and non-landslide areas accounted for (ANN=83.65), (FR=43.84), and (SE=34.45). Moderate susceptibility areas were represented by (ANN=2.93), (FR=27.85), and (SE=28.51), while unstable areas were indicated by (ANN=13.42), (FR=428.57), and (SE=37.03).

The results of the ROC-AUC model showed that the ANN model outperformed both the FR and SE models, recording training and testing accuracies of (AUC=0.966, 0.983). This was followed by the FR model with (AUC=0.955, 0.993), and then the SE model with (AUC=0.953, 0.971), respectively.

This research utilized a broader range of conditioning factors, resulting in a more comprehensive approach to validate landslide susceptibility modeling (LSM). As a result, the study provides a framework that assists decision-makers, policy-makers and planners in effectively managing landslide risks in the region. This can be achieved by enhancing drainage infrastructure to manage excess rainfall and thus reduce soil saturation, modifying slope angles to mitigate landslide risks, and implementing structural reinforcements such as slope stabilization to enhance their stability.

In the future, landslide projects should include raising awareness about early warning signs and emergency evacuation procedures. Landslide susceptibility maps will support this effort as well. There is a necessity for research and development in new technologies especially artificial intelligence and strategies, along with collaboration with geological experts, engineers, and local authorities when planning and implementing landslide mitigation measures. This study has enhanced landslide susceptibility maps and identified high-risk areas in Wadi Dil'ah basin, necessitating urgent rehabilitation and management. This valuable information can significantly assist decision-makers in making informed decisions regarding infrastructure development and urban centers.

References:

1- Abdelrahman, K., Al-Otaibi, N., Ibrahim, E., & Binsadoon, A. (2021). Landslide susceptibility assessment and their disastrous impact on Makkah Al-Mukarramah urban Expansion, Saudi Arabia, using microtremor measurements. Journal of King Saud University-Science, 33, 101450, 1-12.
2- Abdelrahman, K., Al-Amri, A., Al-Kahtany, Kh., & Al-Otaibi, N. (2023). Landslide susceptibility mapping of Al Taif urban area, Saudi Arabia, using remote sensing data and microtremor measurements: integrated approach, Frontiers in Earth Science, 10.3389/feart.2023.1270061, 1-19.
3- Abdo, H. G. (2022). Assessment of landslide susceptibility zonation using frequency ratio and statistical index: a case study of Al-Fawar basin, Tartous, Syria. International Journal of Environmental Science and Technology, 19, 2599-2618.
4- Abdi, E., Majnounian, B., Genet, M., & Rahimi, H. (2010). Quantifying the effects of root reinforcement of Persian Ironwood (Parrotia persica) on slope stability; a case study: Hillslope of Hyrcanian forests, northern Iran. Ecological Engineering, 36(10), 1409-1416.
5- Abedi Gheshlaghi, H., & Feizizadeh, B. (2021). GIS-based ensemble modelling of fuzzy system and bivariate statistics as a tool to improve the accuracy of landslide susceptibility mapping. Natural Hazards, 107(2), 1981-2014.
6- Abedini, M., Ghasemian, B., Shirzadi, A., Shahabi, H., Chapi, K. & et al. (2019). A novel hybrid approach of bayesian logistic regression and its ensembles for landslide susceptibility assessment. Geocarto International, 34(13), 1427-1457.
7- Abu Abdullah MM, Youssef AM, Maerz NH, Abu-AlFadail E, & et al. (2020) A flood risk management program of Wadi Baysh dam on the downstream area: an integration of hydrologic and hydraulic models, Jizan Region. KSA Sustainability 12:1069.
8- Achu, A. L., Aju, C. D., Pham, Q. B., Reghunath, R., & et al. (2022a). Landslide susceptibility modelling using hybrid bivariate statistical-based machine-learning method in a highland segment of Southern Western Ghats, India. Environmental Earth Sciences, 81(13), 360.
9- Achu, A. L., Thomas, J., Aju, C. D., Remani, P. K., & et al. (2022b). Performance evaluation of machine learning and statistical techniques for modelling landslide susceptibility with limited field data. Earth Science Informatics, 16(1), 1025-1039.
10- Aghdam, I. N., Varzandeh, M. H. M., & Pradhan, B. (2016). Landslide susceptibility mapping using an ensemble statistical index (Wi) and adaptive neuro-fuzzy inference system (ANFIS) model at Alborz Mountains (Iran). Environmental Earth Sciences, 75, 1-20.
11- Alam, M., Siddiqui, A. N., Shamim, S. K., Ahmad, A., & et al. (2024). A Comparative Assessment of Landslide Prediction Capability of Machine Learning Methods using Frequency Ratio (FR), Shannon Entropy (SE), and Analytical Hierarchy (AHP) Techniques: A Case Study of Uttarakhand, India.
12- Alharbi, T., & El-Sorogy, A. (2023). Landslide Prediction in Mountainous Terrain Using Remote Sensing and GIS: A Case Study of Al-Hada Road, Makkah Province, Saudi Arabia. Water, 15, 3771, 1-13.
13- Ali, M. Z., Chu, H. J., Chen, Y. C., & Ullah, S. (2021). Machine learning in earthquake- and typhoon-triggered landslide susceptibility mapping and critical factor identification. Environmental Earth Sciences, 80(6), 233.
14- Al-Najjar, H. A., & Pradhan, B. (2021). Spatial landslide susceptibility assessment using machine learning techniques assisted by additional data created with generative adversarial networks. Geoscience Frontiers, 12(2), 625-637.

15- Al-Najjar, H. A., Kalantar, B., Pradhan, B., & Saeidi, V. (2019, October). Conditioning factor determination for mapping and prediction of landslide susceptibility using machine learning algorithms. In Earth resources and environmental remote sensing/GIS applications X (Vol. 11156, pp. 97-107). SPIE.
16- Alvarez-Mendoza, C. I., Teodoro, A., & Ramirez-Cando, L. (2019). Improving NDVI by removing cirrus clouds with optical remote sensing data from Landsat-8–a case study in Quito, Ecuador. Remote Sensing Applications: Society and Environment, 13, 257-274.
17- Anis, Z., Wissem, G., Vali, V., Smida, H., & et al. (2019). GIS-based landslide susceptibility mapping using bivariate statistical methods in North-western Tunisia. Open Geosciences, 11(1), 708-726.
18- Arabameri, A., Pradhan, B., Rezaei, Kh., & Lee, Ch-W. (2019). Assessment of Landslide Susceptibility Using Statistica- and Artificial Intelligence-Based FR-RF Integrated Model and Multiresolution DEMs. Remote Sensing, 11, 999,1-24.
19- Arabameri, A., Saha, S., Roy, J., Chen, W., & et al. (2020). Landslide susceptibility evaluation and management using different machine learning methods in the Gallicash River Watershed, Iran. Remote Sensing, 12(3), 475.
20- Azarafza, M., Azarafza, M., Akgün, H., Atkinson, P.M., & Derakhshani, R. Deep learning-based landslide susceptibility mapping. Sci. Rep. 2021, 11, 24112.
21- Bista, R. B. (2022, October). Does disaster change income and wealth distribution toward extremity of inequality and poverty? Analysis of flood and landslides in the vulnerable locations of nepal. In Forum for social economics (Vol. 51, No. 4, pp. 467-481). Routledge.
22- Cardinali, M., Reichenbach, P., Guzzetti, F., Ardizzone, F. & et al. (2002). A geomorphological approach to the estimation of landslide hazards and risks in Umbria, Central Italy. Natural Hazards and Earth System Sciences, 2(1/2), 57-72.
23- Chen, W., Pradhan, B., Li, S., Shahabi, H. & et al. (2019). Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. Natural Resources Research, 28(4), 1239-1258.
24- Chen, W., Pradhan, B., Li, S., Shahabi, H. & et al. (2019). Novel hybrid integration approach of bagging-based fisher's linear discriminant function for groundwater potential analysis. Natural Resources Research, 28, 1239-1258.
25- Chimidi, G., Raghuvanshi, T. K., & Suryabhagavan, K. V. (2017). Landslide hazard evaluation and zonation in and around Gimbi town, western Ethiopia—a GIS-based statistical approach. Applied Geomatics, 9, 219-236.
26- Chowdhury, S., Rahman, N., Sheikh, S., Sayeid, A., & et al. (2024). GIS-based landslide susceptibility mapping using Logistic Regression, Random Forest and Decision and Regression Tree models in Chattogram District, Bangladesh. Heliyon, 10:e23424, 1-19.
27- Cruden, D. (1991). A Simple Definition of a Landslide, Bulletin of the International Association, Engineering Geology,43.27-29. Full Text via CroossRef.
28- Cruden, D. M., & Varnes, D. J. (1996). landslide types and processes, transportation research board. US National Academy of Sciences, special report, 247: 36-75. Transp Res Board, 247, 36-57.
29- Cui, Y., Cheng, D., Choi, C. E., Jin, W., & et al. (2019). The cost of rapid and haphazard urbanization: lessons learned from the Freetown landslide disaster. Landslides, 16, 1167-1176.
30- Dahal, R. K., Hasegawa, S., Nonomura, A., Yamanaka, M., & et al. (2008). GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Environmental Geology, 54, 311-324.
31- Dam, N. D., Amiri, M., Al-Ansari, N., Prakash, & et al. (2022). Evaluation of shannon entropy and weights of evidence models in landslide susceptibility mapping for the

pithoragarh district of uttarakhand state, India. Advances in Civil Engineering, 2022(1), 6645007.

32- Dang, V. H., Dieu, T. B., Tran, X. L., & Hoang, N. D. (2019). Enhancing the accuracy of rainfall-induced landslide prediction along mountain roads with a GIS-based random forest classifier. Bulletin of Engineering Geology and the Environment, 78, 2835-2849.

33- Devkota, K. C., Regmi, A. D., Pourghasemi, H. R., Yoshida, K., & et al. (2013). Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. Natural hazards, 65, 135-165.

34- Dormann, C. F., Elith, J., Bacher, S., Buchmann, & et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography, 36(1), 27-46.

35- Farhan, Y. (2002). Slope Stability Problems in Central and Northern Jordan. The Arab Word Geographer, 5(4), 265-290.

36- Feizizadeh, B., Roodposhti, M. S., Jankowski, P., & Blaschke, T. (2014). A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping. Computers & geosciences, 73, 208-221.

37- Ganesh, B., Vincent, S., Pathan, S., & Benitez, S. R. G. (2023). Integration of GIS and machine learning techniques for mapping the landslide-prone areas in the state of Goa, India. Journal of the Indian Society of Remote Sensing, 51(7), 1479-1491.

38- He, W., Chen, G., Zhao, J., Lin, Y., & et al. (2023). Landslide susceptibility evaluation of machine learning based on information volume and frequency ratio: a case study of Weixin County, China. Sensors, 23(5), 2549.

39- Hong, H., Liu, J., Bui, D. T., Pradhan, B., & et al. (2018). Landslide susceptibility mapping using J48 Decision Tree with AdaBoost, Bagging and Rotation Forest ensembles in the Guangchang area (China). Catena, 163, 399-413.

40- Horton, R. E. (1945). Erosional development of streams and their drainage basins; hydrophysical approach to quantitative morphology. Geological Society of America bulletin, 56(3), 275-370.

41- Hu, X., Mei, H., Zhang, H., Li, Y., & Li, M. (2021). Performance evaluation of ensemble learning techniques for landslide susceptibility mapping at the Jinping county, Southwest China. Natural Hazards, 105, 1663-1689.

42- Jaafari, A., Najafi, A., Pourghasemi, H. R., Rezaeian, J., & Sattarian, A. (2014). GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. International Journal of Environmental Science and Technology, 11, 909-926.

43- Jacinth Jennifer, J., & Saravanan, S. (2022). Artificial neural network and sensitivity analysis in the landslide susceptibility mapping of Idukki district, India. Geocarto International, 37(19), 5693-5715.

44- Jakob, M. (2022). Landslides in a changing climate. In Landslide hazards, risks, and disasters (pp. 505-579). Elsevier.

45- Jasiewicz, J., & Stepinski, T. F. (2013). Geomorphons—a pattern recognition approach to classification and mapping of landforms. Geomorphology, 182, 147-156.

46- Jebur, M. N., Pradhan, B., & Tehrany, M. S. (2014). Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sensing of Environment, 152, 150-165.

47- Jenness, J. (2006). Topographic position index (tpi_jen. avx_extension for Arcview 3. x, v. 1.3 a, Jenness Enterprises [EB/OL].

48- Jenness, J. (2007). Some thoughts on analyzing topographic habitat characteristics. Jenness Enterprises, Flagstaff, AZ, USA, 26.

49- Jennifer, J. J., Saravanan, S., & Abijith, D. (2021). Application of frequency ratio and logistic regression model in the assessment of landslide susceptibility mapping for Nilgiris District, Tamilnadu, India. Indian Geotechnical Journal, 51(4), 773-787.
50- Kalantar, B., Pradhan, B., Naghibi, S. A., Motevalli, A., & Mansor, S. (2018). Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomatics, Natural Hazards and Risk, 9(1), 49-69.
51- Kalantar, B., Ueda, N., Saeidi, V., Ahmadi, K., Halin, A. A., & Shabani, F. (2020). Landslide susceptibility mapping: Machine and ensemble learning based on remote sensing big data. Remote Sensing, 12(11), 1737.
52- Khaliq, A.H., Basharat, M., Riaz, M.T., Riaz, M.T., & et al. (2022). Spatiotemporal landslide susceptibility mapping using machine learning models: A case study from district Hattian Bala, NW Himalaya, Pakistan. Ain Shams Eng. J., 101907.
53- Lee, D. H., Kim, Y. T., & Lee, S. R. (2020). Shallow landslide susceptibility models based on artificial neural networks considering the factor selection method and various non-linear activation functions. Remote Sensing, 12(7), 1194.
54- Li H, Xu Q, He Y, Fan X, Li S (2020) Modeling and predicting reservoir landslide displacement with deep belief network and EWMA control charts: a case study in three gorges reservoir. Landslides 17:693–707
55- Li Q, Song D, Yuan C, Nie W (2022a) An image recognition method for the deformation area of open-pit rock slopes under variable Content courtesy of Springer Nature, terms of use apply. Rights reserved.
56- Li R., Wu X., Tian H., Yu N., & Wang C. (2022b) Hybrid memetic pretrained factor analysis-based deep belief networks for transient electromagnetic inversion. IEEE Trans Geosci Remote Sens.
57- Li W., Zhu J., Fu L., Zhu Q., & et al. (2021b) An augmented representation method of debris flow scenes to improve public perception. Int J Geogr Inf Sci 35(8):1521–1544.
58- Li, W., Fang, Z., & Wang, Y. (2021a). Stacking ensemble of deep learning methods for landslide susceptibility mapping in the Three Gorges Reservoir area, China. Stochastic Environmental Research and Risk Assessment, 1-22.
59- Liu Q., Chen X., & Chen J. (2018) Landslide susceptibility analysis using the artificial neural network model and the normalized difference vegetation index. Environmental Earth Sciences, 77(17), 647. DOI: 10.1007/s12665-018-7187-4.
60- Lotfi, F., & Fallahnejad, R. (2010). Imprecise Shannon's entropy and multi attribute decision making. Entropy, 12(1), 53-62.
61- Ma, Z., Mei, G., & Piccialli, F. (2021). Machine learning for landslides prevention: a survey. Neural Computing and Applications, 33(17), 10881-10907.
62- Mas, J. F., Filho, B. S., Pontius Jr, R. G., Gutiérrez, M. F., & Rodrigues, H. (2013). A suite of tools for ROC analysis of spatial models. ISPRS International Journal of Geo-Information, 2(3), 869-887.
63- Mas, J. F., García-Álvarez, D., Paegelow, M., Domínguez-Vera, R., & Castillo-Santiago, M. Á. (2022). Metrics based on a cross-tabulation matrix to validate land use cover maps. In Land use cover datasets and validation tools: validation practices with QGIS (pp. 127-151). Cham: Springer International Publishing.
64- Masruroh, H., Leksono, A. S., & Kurniawan, S. (2023). Developing landslide susceptibility map using Artificial Neural Network (ANN) method for mitigation of land degradation. Journal of Degraded & Mining Lands Management, 10(3).
65- Mehrabi, M., & Moayedi, H. (2021). Landslide susceptibility mapping using artificial neural network tuned by metaheuristic algorithms. Environmental Earth Sciences, 80(24), 804.

66- Mekonnen, A. A., Raghuvanshi, T. K., Suryabhagavan, K. V., & Kassawmar, T. (2022). GIS-based landslide susceptibility zonation and risk assessment in complex landscape: A case of Beshilo watershed, northern Ethiopia. Environmental Challenges, 8, 100586.
67- Merghadi, A., Yunus, A. P., Dou, J., Whiteley, J., & et al. (2020). Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews, 207, 103225.
68- Mersha, T., & Meten, M. (2020). GIS-based landslide susceptibility mapping and assessment using bivariate statistical methods in Simada area, northwestern Ethiopia. Geoenvironmental disasters, 7, 1-22.
69- Mfondoum, A. H. N., Nguet, P. W., Seuwui, D. T., Mfondoum, J. V. M., & et al. (2023). Stepwise integration of analytical hierarchy process with machine learning algorithms for landslide, gully erosion and flash flood susceptibility mapping over the North-Moungo perimeter, Cameroon. Geoenvironmental Disasters, 10(1), 22.
70- Miller, S., Brewer, T., & Harris, N. (2009). Rainfall Thresholding and Susceptibility assessment of rainfall induced landslides: application to landslide management in St Thomas, Jamaica. Bull Int. Assoc. Eng. Geol. 68:539–550
71- Mitra, R., Saha, P., & Das, J. (2022). Assessment of the performance of GIS-based analytical hierarchical process (AHP) approach for flood modelling in Uttar Dinajpur district of West Bengal, India. Geomatics, Natural Hazards and Risk, 13(1), 2183-2226.
72- Mondini, A. C., Guzzetti, F., Chang, K. T., Monserrat, O., & et al. (2021). Landslide failures detection and mapping using Synthetic Aperture Radar: Past, present and future. Earth-Science Reviews, 216, 103574.
73- Moore, I. D., & Burch, G. J. (1986). Physical basis of the length-slope factor in the universal soil loss equation. Soil Science Society of America Journal, 50(5), 1294-1298.
74- Moragues, S., Lenzano, M. G., Lanfri, M., Moreiras, S., & et al. (2021). Analytic hierarchy process applied to landslide susceptibility mapping of the North Branch of Argentino Lake, Argentina. Natural Hazards, 105(1), 915-941.
75- Moragues, S., Lenzano, M. G., Lanfri, M., Moreiras, S., & et al. (2021). Analytic hierarchy process applied to landslide susceptibility mapping of the North Branch of Argentino Lake, Argentina. Natural Hazards, 105(1), 915-941.
76- Naceur, H. A., Abdo, H. G., Igmoullan, B., Namous, M., & et al. (2022). Performance assessment of the landslide susceptibility modelling using the support vector machine, radial basis function network, and weight of evidence models in the N'fis river basin, Morocco. Geoscience Letters, 9(1), 39.
77- Nguyen, V. V., Pham, B. T., Vu, B. T., Prakash, I., & et al. (2019). Hybrid machine learning approaches for landslide susceptibility modeling. Forests, 10(2), 157.
78- Nohani, E., Moharrami, M., Sharafi, S., Khosravi, K., & et al. (2019). Landslide susceptibility mapping using different GIS-based bivariate models. Water 11 (7): 1402.
79- 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.
80- Orhan, O., Bilgilioglu, S. S., Kaya, Z., Ozcan, A. K., & Bilgilioglu, H. (2022). Assessing and mapping landslide susceptibility using different machine learning methods. Geocarto International, 37(10), 2795-2820.
81- Ozioko, O. H., & Igwe, O. (2020). GIS-based landslide susceptibility mapping using heuristic and bivariate statistical methods for Iva Valley and environs Southeast Nigeria. Environmental monitoring and assessment, 192, 1-19.
82- Pachauri, A. K., & Pant, M. (1992). Landslide hazard mapping based on geological attributes. Engineering geology, 32(1-2), 81-100.

83- Panchal, S., & Shrivastava, A. K. (2020). Application of analytic hierarchy process in landslide susceptibility mapping at regional scale in GIS environment. Journal of Statistics and Management Systems, 23(2), 199-206.
84- Panchal, S., & Shrivastava, A. K. (2021). A comparative study of frequency ratio, Shannon's entropy and analytic hierarchy process (AHP) models for landslide susceptibility assessment. ISPRS International Journal of Geo-Information, 10(9), 603.
85- Panchal, S., & Shrivastava, A. K. (2022). Landslide hazard assessment using analytic hierarchy process (AHP): A case study of National Highway 5 in India. Ain Shams Engineering Journal, 13(3), 101626.
86- Paul S. K, & Bhowmik A (2016) A review of the different methods for landslide susceptibility mapping. Natural Hazards, 84(2), 581-605.
87- Pham, B. T., Bui, D. T., Prakash, I., & 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.
88- Pham, B. T., Prakash, I., & Bui, D. T. (2018). Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees. Geomorphology, 303, 256-270.
89- Pham, B. T., Prakash, I., Dou, J., Singh, S. K., & et al. (2020). A novel hybrid approach of landslide susceptibility modelling using rotation forest ensemble and different base classifiers. Geocarto International, 35(12), 1267-1292.
90- Pham, B. T., Prakash, I., Khosravi, K., Chapi, K., & et al. (2018). A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. Geocarto International, 34(13), 1385-1407.
91- Pham, B. T., Prakash, I., Khosravi, K., Chapi, K., & et al. (2019). A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. Geocarto International, 34(13), 1385-1407.
92- Pham, B. T., Tien Bui, D., Prakash, I., & Dholakia, M. B. (2016). Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. Natural Hazards, 83, 97-127.
93- Poudel, K., & Regmi, A. D. (2016). Landslide susceptibility mapping along Tulsipur-Kapurkot road section and its surrounding region using bivariate statistical model. Journal of Nepal Geological Society, 50(1), 83-93.
94- Pourghasemi, H. R., & Rahmati, O. (2018). Prediction of the landslide susceptibility: Which algorithm, which precision? Catena, 162, 177-192.
95- Pourghasemi, H. R., & Rossi, M. (2017). Landslide susceptibility modeling in a landslide prone area in Mazandarn Province, north of Iran: a comparison between GLM, GAM, MARS, and M-AHP methods. Theoretical and Applied Climatology, 130(1), 609-633.
96- Pradhan, B., Seeni, M. I., & Kalantar, B. (2017). Performance evaluation and sensitivity analysis of expert-based, statistical, machine learning, and hybrid models for producing landslide susceptibility maps. Laser scanning applications in landslide assessment, 193-232.
97- Rajakumar, P., Sanjeevi, S., Jayaseelan, S., Isakkipandian, G., & et al. (2007). Landslide susceptibility mapping in a hilly terrain using remote sensing and GIS. Journal of the Indian Society of Remote Sensing, 35, 31-42.
98- Regmi, A. D., Yoshida, K., Pourghasemi, H. R., DhitaL, M. R., & Pradhan, B. (2014). Landslide susceptibility mapping along Bhalubang—Shiwapur area of mid-Western Nepal using frequency ratio and conditional probability models. Journal of Mountain Science, 11, 1266-1285.
99- Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. intermountain Journal of sciences, 5(1-4), 23-27.

100- Roy, J., & Saha, S. (2019). Landslide susceptibility mapping using knowledge driven statistical models in Darjeeling District, West Bengal, India. Geoenvironmental Disasters, 6(1), 1-18.
101- Saha, S., Sarkar, R., Roy, J., Hembram, T. K., & et al. (2021). Measuring landslide vulnerability status of Chukha, Bhutan using deep learning algorithms. Scientific Reports, 11(1), 16374.
102- Selamat, S. N., Abd Majid, N., & Mohd Taib, A. (2023). A Comparative Assessment of Sampling Ratios Using Artificial Neural Network (ANN) for Landslide Predictive Model in Langat River Basin, Selangor, Malaysia. Sustainability, 15(1), 861.
103- Selamat, S. N., Majid, N. A., Taha, M. R., & Osman, A. (2022). Landslide susceptibility model using artificial neural network (ANN) approach in Langat river basin, Selangor, Malaysia. Land, 11(6), 833.
104- Sengupta, A., & Nath, S. K. (2022). GIS-based landslide susceptibility mapping in eastern boundary zone of northeast india in compliance with indo-burmese subduction tectonics. Geospatial Technology for Environmental Hazards: Modeling and Management in Asian Countries, 19-37.
105- Shahabi, H., Ahmadi, R., Alizadeh, M., Hashim, M., & et al. (2023). Landslide Susceptibility Mapping in a Mountainous Area Using Machine Learning Algorithms. Remote Sensing, 15, 3112,1-18.
106- Shano, L., Raghuvanshi, T. K., & Meten, M. (2021). Landslide susceptibility mapping using frequency ratio model: the case of Gamo highland, South Ethiopia. Arabian Journal of Geosciences, 14, 1-18.
107- Shu, H., Hürlimann, M., Molowny-Horas, R., González, M., & et al. (2019). Relation between land cover and landslide susceptibility in Val d'Aran, Pyrenees (Spain): Historical aspects, present situation and forward prediction. Science of the total environment, 693, 133557.
108- Sidle, R. C., Al-Shaibani, A. M., & Kaka, S. I. (2019). Geomorphic hazards in south-west Saudi Arabia: The human–environmental nexus. Area, 51(4), 670-680.
109- Solaimani, K., Mousavi, S. Z., & Kavian, A. (2013). Landslide susceptibility mapping based on frequency ratio and logistic regression models. Arabian Journal of Geosciences, 6, 2557-2569.
110- Sujatha, E. R. (2012). Geoinformatics based landslide susceptibility mapping using probabilistic analysis and entropy index of Tevankarai stream sub-watershed, India. DISASTER ADVANCES, 5(3), 26-33.
111- Sujatha, E. R., & Sridhar, V. (2021). Landslide susceptibility analysis: A logistic regression model case study in Coonoor, India. Hydrology, 8(1), 41.
112- Tian, Y., Xu, C., Hong, H., Zhou, Q., & Wang, D. (2019). Mapping earthquake-triggered landslide susceptibility by use of artificial neural network (ANN) models: an example of the 2013 Minxian (China) Mw 5.9 event. Geomatics, Natural Hazards and Risk, 10(1), 1-25.
113- Tien Bui, D., Tuan, T. A., Klempe, H., Pradhan, B., & Revhaug, I. (2016). Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13, 361-378.
114- Vayadande, K., Sadake, S., Sangwai, S., Patil, M., & et al. (2024). Landslide Susceptibility Prediction System. Research Square (https//doi.org/10.21203/rs.3.rs-3976209/v1)
115- Wang Q., Li W., Chen W., & et al. (2015). GIS-based assessment of landslide susceptibility using certainty factor and index of entropy models for the Qianyang County of Baoji city, China. Journal of Earth System Science 124: 1399-1415

116- Weiss, A. (2001, July). Topographic position and landforms analysis. In Poster presentation, ESRI user conference, San Diego, CA (Vol. 200).
117- Wubalem, A. (2021). Landslide susceptibility mapping using statistical methods in Uatzau catchment area, northwestern Ethiopia. Geoenvironmental Disasters, 8(1), 1.
118- Xiao, T., Yin, K., Yao, T., & Liu, S. (2019). Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, Three Gorges Reservoir, China. Acta Geochimica, 38, 654-669.
119- Yao, X., Tham, L. G., & Dai, F. C. (2008). Landslide susceptibility mapping based on support vector machine: a case study on natural slopes of Hong Kong, China. Geomorphology, 101(4), 572-582.
120- Yesilnacar, E., & Topal, T. A. M. E. R. (2005). Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). Engineering Geology, 79(3-4), 251-266.
121- Young, A. (1972) Slopes. Longman, London, UK, 268 p.
122- Youssef, A.M. (2015). Landslide susceptibility delineation in the Ar-Rayth area, Jizan, Kingdom of Saudi Arabia, using analytical process, frequency ratio, and logistic regression models. Environmental Earth Sciences, 73 (12), 8499-8518.
123- Youssef, A.M., Al-Kathery, M., & Pradhan, B. (2015a). Landslide susceptibility mapping at Al-Hasher Area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. Geosciences Journal, 19 (1), 113-134.
124- Youssef, A.M., Pradhan, B., Pourghasemi, H. R., & Abdullahi, S. (2015b). Landslide susceptibility assessment at Wadi Jawrah Basin, Jizan region, Saudi Arabia using two bivariate models in GIS. Geosciences Journal, 19 (3), 449-469.
125- Youssef, A. M., Pradhan, B., Dikshit, A., Al-Katheri, M. M., & et al. (2022). Landslide susceptibility mapping using CNN-1D and 2D deep learning algorithms: comparison of their performance at Asir Region, KSA. Bulletin of Engineering Geology and the Environment, 81(4), 165.
126- Youssef, K., Shao, K., Moon, S., & Bouchard, L. S. (2023). Landslide susceptibility modeling by interpretable neural network. Communications Earth & Environment 4: 162.
127- Zhang, Y. X., Lan, H. X., Li, L. P., Wu, Y. M., & et al. (2020). Optimizing the frequency ratio method for landslide susceptibility assessment: A case study of the Caiyuan Basin in the southeast mountainous area of China. Journal of Mountain Science, 17(2), 340-357.
128- Zhao, P., Masoumi, Z., Kalantari, M., Aflaki, M., & Mansourian, A. (2022). A GIS-based landslide susceptibility mapping and variable importance analysis using artificial intelligent training-based methods. Remote Sensing, 14(1), 211.
129- Zhao, X., & Chen, W. (2020). Optimization of computational intelligence models for landslide susceptibility evaluation. Remote Sensing, 12(14), 2180.
130- Zhou, C., Yin, K., Cao, Y., Ahmed, B., & et al. (2018). Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China. Computers & geosciences, 112, 23-37.
131- Zhou, X., Wen, H., Zhang, Y., Xu, J., & Zhang, W. (2021). Landslide susceptibility mapping using hybrid random forest with GeoDetector and RFE for factor optimization. Geoscience Frontiers, 12(5), 101211.

Journal homepage: https://mkgc.journals.ekb.eg/
ISSN: 2357-0091 (Print) 2735-5284 (Online)

ملخص:

تقييم قابلية حدوث الانهيارات الأرضية في حوض وادي ضلع (السعودية) من خلال دمج أساليب نظم المعلومات الجغرافية والاستشعار عن بعد والإحصاء ثنائي المتغيرات ونهج التعلم الآلي الاصطناعي

إعداد

د/ وليد شكري عبد الحميد يوسف

قسم الجغرافيا ونظم المعلومات الجغرافية، كلية الآداب، جامعة أسيوط، أسيوط، مصر

د/ نرمين أحمد شكري

قسم الجغرافيا، كلية الآداب، جامعة القاهرة، الجيزة، مصر

أ.م.د/ أحمد علي أحمد علي

قسم الجغرافيا ونظم المعلومات الجغرافية، كلية الآداب، جامعة أسيوط، أسيوط، مصر

تعد الانهيارات الأرضية أحد المخاطر الطبيعية التي تسبب الكثير من الضحايا وخسائر الممتلكات في العالم. تهدف هذه الدراسة إلى إنتاج خرائط قابلية حدوث الانهيارات الأرضية باستخدام دمج تقنيات نظم المعلومات الجغرافية والاستشعار عن بعد ونهج التعلم الآلي للشبكات العصبية الاصطناعية (ML-ANN) والأساليب الإحصائية ثنائية المتغيرات لنموذجي (نسبة التكرار FR، شانون انتروبي SE) في حوض وادي ضلع في الجزء الجنوبي الغربي من السعودية. تم إعداد ١٨ طبقة من العوامل المتعلقة بالانهيارات الأرضية (الارتفاع، زاوية الانحدار، اتجاه الانحدار، طول المنحدر، مؤشر الموقع الطبوغرافي، مؤشر وعورة التضاريس، أشكال سطح الأرض، "التقوس العام والأفقي والعمودي"، والتكوين الصخري، المسافة من الصدوع، مؤشر الرطوبة الطبوغرافي، المسافة من المجاري المائية، هطول الأمطار، الغطاء الأرضي/ استخدام الأرض، مؤشر الغطاء النباتي، والمسافة من الطريق)، ولكشف وقياس الارتباطات الخطية المشتركة بين العوامل المتعلقة بالانهيارات الأرضية، تم استخدام معامل تضخم التباين(VIF) ومؤشر التباين المسموح (TOL). تم تطبيق منحنى خصائص تشغيل المستقبل (ROC)، والمساحة تحت المنحنى (AUC). أظهرت نتائج مؤشر (VIF) و(TOL) عدم وجود علاقة خطية مشتركة بين العوامل المختارة. وأن نسبة AUC لمعدلات التدريب هي (٠.٩٦٦, ٠.٩٥٥, ٠.٩٥٣) على التوالي، في حين أن معدلات الاختبار هي (٠.٩٨٣, ٠.٩٩٣, ٠.٩٧١) لنماذج (ANN, FR, SE) على التوالي. تم تصنيف خريطة القابلية للانهيارات الأرضية إلى خمس فئات باستخدام toolK Natural Breaks (Jenks)، وتم حساب النسبة المئوية لكل فئة من فئات قابلية حدوث الانهيارات الأرضية. وكشفت النتائج أن أداء نموذج ML-ANN يعد الأفضل من حيث الدقة، ومن حيث المقارنة بالنسبة لكلا من النموذجين FR وSE، لذا يوصى بأن يكون نموذج مناسب لتطبيق خرائط قابلية الانهيارات الأرضية في المنطقة الجبلية. ومن المأمول أن تساعد نتائج هذه الدراسة صناع القرار والباحثين في التخفيف من الانهيارات الأرضية وديناميكياتها.

الكلمات المفتاحية: رسم خرائط قابلية حدوث الانهيارات الأرضية (LSM)، الشبكات العصبية الاصطناعية (ANN)، نسبة التكرار (FR)، شانون انتروبي (SE)، الانهيارات الأرضية، نظم المعلومات الجغرافية، المملكة العربية السعودية.