

**Earthquake-Induced Landslides Susceptibility Evaluation:
A Point Process Modelling Approach**



By

Zain Ahmed Bin Amir

(2019-NUST-MS-GIS-00000320185)


**A thesis submitted in partial fulfillment of the requirements for
the degree of Master of Science in Remote Sensing and GIS**

**Institute of Geographical Information Systems
School of Civil and Environmental Engineering
National University of Sciences & Technology
Islamabad, Pakistan**


August 2023


THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by Zain Ahmed Bin Amir (Registration No. MSRSGIS 00000320185), of Session 2019 (Institute of Geographical Information Systems) has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulation, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: 
Name of Supervisor: Dr. Salman Atif
Date: 23.8.23

Signature (HOD): 
Date: 23/08/2023
Dr. Javed Iqbal
Professor & HOD IGIS, SCEE (NUST)
H-12, Islamabad

Signature (Associate Dean): 
Date: 23.8.2023
Dr. Eliaz Hussain
Assoc. Dean IGIS, SCEE (NUST)
H-12, ISLAMABAD

Signature (Principal & Dean SCEE): 
Date: 24 AUG 2023
PROF DR MUHAMMAD IRFA
Principal & Dean
SCEE, NUST

ACADEMIC THESIS: DECLARATION OF AUTHORSHIP

I, **Zain Ahmed Bin Amir**, declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

Earthquake-Induced Landslides Susceptibility Evaluation: A Point Process Modelling Approach

I confirm that:

1. This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text;
2. Wherever any part of this thesis has previously been submitted for a degree or any other qualification at this or any other institution, it has been clearly stated;
3. I have acknowledged all main sources of help;
4. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
5. None of this work has been published before submission.
6. This work is not plagiarized under the HEC plagiarism policy.

Signed: 

Date: 25/08/2023

"In heartfelt dedication to my exceptional parents and beloved brother, whose constant support and unwavering cooperation have guided me to achieve this remarkable accomplishment!"

ACKNOWLEDGEMENTS

I extend my heartfelt gratitude to the **ALMIGHTY ALLAH** for guiding me at every step of this journey and for inspiring new thoughts that enhanced my work. I recognize that everyone who supported me, be it my parents or any other individual, did so by Your will, and thus, all praise belongs to You alone.

My sincere appreciation goes to my dear parents, who nurtured me from the time I took my first steps and have steadfastly supported me in every facet of my life.

A special acknowledgment is due to **Dr. Salman Atif**, my supervisor, for his unwavering assistance throughout my thesis and his exceptional guidance in the field of Disaster Management. His teaching has allowed me to delve into GIS subjects with unparalleled depth.

I am deeply grateful to Associate Dean, **Dr. Ejaz Hussain** and Head of Department, **Dr. Javed Iqbal** for their invaluable support and cooperation. Whenever I faced challenges, they provided solutions that were instrumental in completing my thesis. Their patience and guidance have been a constant source of motivation.

My thanks also extend to **Dr. Muhammad Azmat** and **Dr. Zeeshan Ali Khan** for their roles on my thesis guidance and evaluation committee. I express my special gratitude to **Dr. Fawad Najam** for his invaluable assistance.

I acknowledge the support and cooperation of **Mr. Kashif Ali** and **Mr. Rizwan**. Their contributions have been invaluable to me.

Lastly, I extend my heartfelt appreciation to all those who have lent their valuable assistance to my study.



Zain Ahmed Bin Amir

Table of Contents

THESIS ACCEPTANCE CERTIFICATE	i
ACADEMIC THESIS: DECLARATION OF AUTHORSHIP	ii
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	vii
LIST OF TABLES	x
LIST OF ABBREVIATIONS	xi
ABSTRACT	xii
Chapter 1	1
INTRODUCTION	1
1.1. SEISMICITY OF BADAKHSHAN REGION	2
1.2. LANDSLIDES AND EARTHQUAKES RELATION	2
1.3. LANDSLIDE SUSCEPTIBILITY MAPPING.....	3
1.4. OBJECTIVES	6
1.5. LITERATURE REVIEW	6
1.6. STATISTICAL ANALYSIS AND MODELLING	11
Chapter 2	13
MATERIALS AND METHODS.....	13
2.1. STUDY AREA	13
2.2. MATERIALS.....	17
2.3. ANALYTICAL FRAMEWORK	17
2.4. EARTHQUAKE DATA.....	21
2.5. LANDSLIDE INVENTORY.....	21
2.6. REMOTE SENSING/SATELLITE DATA.....	22
2.7. DIGITAL ELEVATION MODEL (DEM).....	23
2.8. HYDROLOGY DATA.....	24
2.9. GEOLOGICAL DATA	25
2.10. STATISTICAL ANALYSIS AND MODELLING	25
2.11. POINT PATTERN ANALYSIS IN R.....	32
2.12. LANDSLIDE SUSCEPTIBILITY MAPPING (LSM)	32

<i>Chapter 3</i>	34
RESULTS AND DISCUSSIONS	34
3.1. EARTHQUAKE MAPPING	34
3.2. LANDSLIDE INVENTORY	35
3.3. LANDSLIDE CAUSATIVE FACTORS	36
3.4. POINT PATTERN ANALYSIS AND MODELLING	52
3.5. LOGISTIC REGRESSION MODELLING.....	59
3.6. FREQUENCY RATIO MODELLING	61
3.7. LANDSLIDE SUSCEPTIBILITY MAPPING (LSMS)	66
3.8. DISCUSSIONS	70
<i>Chapter 4</i>	72
CONCLUSTION AND RECOMMENDATIONS	72
4.1. CONCLUSIONS.....	72
4.2. RECOMMENDATIONS.....	73
REFERENCES	75
APPENDICES	85
A. GOOGLE EARTH ENGINE CODE FOR LAND USE AND LAND COVER (LULC) CLASSIFICATION OF THE STUDY AREA FOR THE YEAR 2016 FROM LANDSAT 8 IMAGERY.....	86
B. GENERALIZED CODE FOR POINT PATTERN ANALYSIS IN "R" STUDIO	88

LIST OF FIGURES

Figure 2.1. Map of the area under study showing the district Chitral of the Khyber Pakhtunkhwa province of Pakistan. It is situated towards the northwestern border of Pakistan with Afghanistan and is mostly run by the Hindu Kush mountain ranges of the country.....	14
Figure 2.2. Detailed methodology flowchart showing step by step collection of data and processing with results	20
Figure 3.1. Earthquake map showing the exact location of the epicenter and the affected region of Chitral, Pakistan	34
Figure 3.2. Landslide Inventory map of Chitral showing locations of the landslides (Modelling samples that were used to train different models Validation Samples that were used to validate the models after training them)	35
Figure 3.3. Map of Elevation Profile Chitral divided into 5 classes with equal intervals	37
Figure 3.4. Figure showing relationship between the Elevation classes and the landslide pixel % in corresponding class	37
Figure 3.5. Slope map of the study area showing different slope angles in different classes	38
Figure 3.6. Relationship between slope angle and landslide pixel % in each class	39
Figure 3.7. Aspect map of the study area divided into classes from 0 to 360 degrees	40
Figure 3.8. Relationship between landslide pixel % in each class of Aspect	40
Figure 3.9. Geological Map of Chitral shows different lithologies and the main fault lines in Chitral	41
Figure 3.10. Relationship between the landslide pixel % and the lithologies in the study area	42
Figure 3.11. The above Map shows different soil types in the region	43
Figure 3.12. The relationship between the soil types and percentage landslide occurrence in each class.....	43
Figure 3.13. LULC map of Chitral showing different land covers in the study area ..	44

Contd.

Figure 3.14. Relationship between different land covers and the landslide occurrence in each class 45

Figure 3.15. Map of the study area showing Vegetation cover in Chitral 46

Figure 3.16. Relationship between landslide pixel % in vegetated class and non-vegetated class 46

Figure 3.17. Map of the study area showing different rivers and streams flowing in the Chitral district 47

Figure 3.18. Map showing proximity to river with multiple buffer zones 48

Figure 3.19. Distance to rivers and the landslide pixels percentage in each class 48

Figure 3.20. Map showing proximity to fault lines with multiple buffer zones 49

Figure 3.21. Relationship between the fault lines and landslide pixel percentage in each buffer class 50

Figure 3.22. Road Network Map of Chitral showing main highways in the study area 51

Figure 3.23. Proximity to Road map showing multiple buffer zones 51

Figure 3.24. Percentage landslide occurrence with respect to distance from roads ... 52

Figure 3.25. Map showing number of landslides in each quadrat 53

Figure 3.26. Map showing landslides density in each quadrat 54

Figure 3.27. Map showing Kernel Density Estimate (KDE) with the bandwidth of 50km disc. function 54

Figure 3.28. Quadrat density estimation for elevation ranges 55

Figure 3.29. Kernel Density Estimate for elevation ranges 55

Figure 3.30. Average Nearest Neighbor Analysis (ANN), K, L and G functions on a randomly chosen 20 landslide points to establish a relationship between elevation, slope and aspect parameters. 56

Figure 3.31. The landslide point distribution in each causative factor 57

Figure 3.32. Distribution of landslide points in Elevation, Slope and Aspect according to the values of each class 58

Contd.

Figure 3.33. Cluster map of the region under study showing clusters of points at different locations indicating high density of landslides in most of the regions.....	58
Figure 3.34. Landslide Susceptibility Map generated by LR Method showing areas from Low to Very High Susceptible	66
Figure 3.35. Landslide Susceptibility Map generated by FR Method showing areas from Low to Very High Susceptible	67
Figure 3.36. ROC Curve of LR and FR Model.....	69
Figure 3.37. Sensitivity vs Specificity Analysis of LR and FR Models.....	69

LIST OF TABLES

Table 2.1. List of datasets with its purposes	18
Table 2.2. Tools used for Analysis and Processing.....	19
Table 3.1. Landslide causative factors and the calculated Regression coefficients from the LR method.....	60
Table 3.2. Calculated weights for each landslide causative factor and its classes for FR Method	62

LIST OF ABBREVIATIONS

Abbreviation	Explanation
GIS	Geographic Information System
DTM	Digital Terrain Model
GEE	Google Earth Engine
LULC	Land Use Land Cover
DEM	Digital Elevation Model
NDVI	Normalized Difference Vegetation Index
ML	Machine Learning
PPA	Point Pattern Analysis
CSR	Complete Spatial Randomness
KDE	Kernel Density Estimation
ANN	Average Nearest Neighbor
USGS	United States Geological Survey
FAO	Food and Agriculture Organization
MKT	Main Karakoram Thrust
NHA	National Highways Authority
LR	Logistic Regression
FR	Frequency Ratio
LSM	Landslide Susceptibility Mapping
AUC	Area Under Curve
ROC	Receiver Operating Characteristics

ABSTRACT

Landslides pose a significant threat to human lives and infrastructure in mountainous regions such as the Chitral district of Northern KPK, Pakistan. This study aims to assess the earthquake induced landslides in the region by integrating the creation of a landslide inventory with the application of Point Pattern Analysis (PPA), Logistic Regression (LR), and Frequency Ratio (FR) models. The research began by compiling a comprehensive landslide inventory through the analysis of high-resolution satellite imagery and temporal study in Google Earth Pro. The inventory includes about 210 landslides that occurred due to the high seismicity in the region. To analyze the spatial distribution and clustering of landslides, Point Pattern Analysis (PPA) was applied. Furthermore, Logistic Regression (LR) and Frequency Ratio (FR) models were employed to develop Landslide Susceptibility Maps (LSMs). These models utilized the landslide inventory data along with relevant terrain, geological, and environmental variables as causative factors or predictors. The results of the study revealed significant spatial patterns and hotspots of landslide occurrence in the Chitral district. Logistic Regression model shows that 18.1 % of the region is highly susceptible towards landslide whereas the Frequency Ratio model reveals 26.1 % of the study area falls in Very high susceptible class in Landslide Susceptibility Mapping. Accuracy assessment through Receiver Operating Characteristics (ROC) Analysis reveals that the overall accuracy of LR model (85.34 %) was better than FR model (78.56 %). The findings of this research provide valuable insights into the spatial distribution of landslides in the Chitral district and contribute to improved understanding of the underlying factors influencing landslide susceptibility. The generated landslide susceptibility maps can be utilized for land-use planning, infrastructure development, and disaster risk management in the region.

Keywords: *Landslide Susceptibility Map (LSM), Point Pattern Analysis (PPA), Logistic Regression (LR) Model, Frequency Ratio (FR) Model, causative factors, landslide inventory, Receiver Operating Characteristics (ROC)*

INTRODUCTION

Pakistan has a diverse and complex geology, shaped by millions of years of tectonic activity and geological processes. The country is situated at the intersection of several major tectonic plates, including the Eurasian, Indian, Arabian, and African plates. This has led to the formation of a variety of geological features, including mountain ranges, valleys, plateaus, and plains. The geology of Pakistan is closely related to earthquakes, as the country is situated a region that is known to experience seismic activity most frequently. Indian and Eurasian tectonic plates collision is the main reason, Pakistan is situated within a zone that experiences heightened seismic activity. The collision of these two plates has resulted in the formation of the Himalayas and other mountain ranges in the region. This process has also led to the formation of several active fault zones throughout the country, where tectonic stresses build up and are eventually released in the form of earthquakes.

Pakistan has experienced several devastating earthquakes in recent history, including the 2005 Kashmir earthquake and the 2013 Baluchistan earthquake. These earthquakes were caused by the movement of tectonic plates and fault zones in the region and resulted in huge loss to life buildings. Pakistan Geological composition also contributes to the occurrence and spreading of earthquakes in the country. The northern area of Pakistan, situated in the Himalayan area, encounters frequent seismic activity caused due to convergence of Indian and Eurasian plates. Whereas, southern part of the country, which is located on the Arabian Sea coast, experiences fewer earthquakes, but those that do occur can be more intense due to the type of rocks and sediments in the region.

The relationship between geology and earthquakes in Pakistan is complex and highlights the importance of understanding the geological

processes and tectonic activity in the region to better predict and prepare for future seismic events.

1.1 Seismicity of Badakhshan region

The Badakhshan province of Afghanistan, close to the Pakistan border, was hit by a strong earthquake on October 26, 2015, which had a magnitude of 7.5. The earthquake occurred at a depth of approximately 212 km, making it a deep-focus earthquake. The quake was felt across a wide area, including Pakistan, India, and Tajikistan. The earthquake initiated major loss in the region, with reports of hundreds of fatalities and thousands of injuries. In Pakistan, the hardest-hit area was the Chitral region, where dozens of people were killed, and hundreds were injured. The earthquake triggered landslides in the mountainous region, which blocked roads and caused further damage to infrastructure.

The earthquake induced landslides in Chitral were particularly significant, as they blocked several key roads and access routes in the region. In addition to this, the landslides resulted in harm to bridges and other buildings as well making it difficult for relief and rescue teams to access affected areas. In addition to the landslides, the earthquake also caused significant damage to buildings and other infrastructure in the region, including homes, schools, and hospitals. The impact of the earthquake and its aftermath was felt for weeks and months afterward, as communities worked to rebuild and recover from the disaster.

1.2 Landslides and Earthquakes relation

Earthquakes and landslides are often related in Pakistan. Earthquakes can trigger landslides by shaking loose rocks and soil from mountain slopes, which can then cause landslides and rockfalls. This is particularly common in the mountainous regions of northern Pakistan, where earthquakes are relatively frequent and the terrain is steep and prone to landslides.

The 2005 Kashmir earthquake in Pakistan led to the occurrence of several landslides in the affected region, resulting in significant damage and

contributing to the high death toll. Similarly, landslides were set off in the area affected by the 2013 Balochistan earthquake, which exacerbated the damage and fatalities caused by the quake.

Pakistan frequently experiences landslides as a natural hazard, especially in its northern mountainous regions. The topography of these areas, combined with heavy rainfall and seismic activity, increases the likelihood of landslides occurring. Pakistan has witnessed several significant landslides in recent years, including:

- 1. Attabad landslide:** In January 2010, a landslide took place in the Hunza Valley of Gilgit-Baltistan. The landslide blocked the Hunza River and created a natural dam, which flooded nearby areas and displaced thousands of people.
- 2. Kohistan landslide:** This landslide occurred in Kohistan district in Khyber Pakhtunkhwa in February 2010. The landslide resulted in unfortunate loss of over 20 lives and inflicted significant damage to infrastructures in the area.
- 3. Mansehra landslide:** This landslide occurred in Mansehra district in Khyber Pakhtunkhwa in June 2019. This led to the unfortunate demise of over 20 individuals and damage to infrastructure in the affected region.
- 4. Golen Gol landslide:** This landslide occurred in Chitral district in Khyber Pakhtunkhwa in July 2015. The landslide caused by the natural disaster led to the loss of over 15 lives and inflicted substantial damage to the infrastructures in the affected area.

1.3 Landslide Susceptibility Mapping (LSM)

LSM involves the identification of regions that are more susceptible to landslides. The process typically involves an analysis of causes such as slope, geology, and land cover, soil type and precipitation which are known to contribute to landslides. By mapping these factors and their potential interactions, it is possible to identify high risk areas prone to landslides.

Landslide susceptibility mapping has been an important field of study for several decades, with early efforts focused on identifying and characterizing

landslide-prone areas using geological and geomorphological criteria. From the 1970s to the 1980s, there was a transition towards the use of quantitative techniques in the creation of maps that predict the likelihood of landslides occurring in a particular area. This approach involved using statistical models to analyze different environmental factors and their relationships with the occurrence of landslides.

Geographic Information System (GIS) is crucial in landslide susceptibility mapping. GIS is a software platform that enables the collection, analysis, and visualization of spatial data. This makes it possible to combine and analyze various types of data like soil, landcover, geology and topography to look for areas that are more susceptible to landslides. The ability of GIS to integrate different data sources and analyze them spatially makes it a valuable tool in landslide susceptibility mapping. By using GIS tools, analysts can input and analyze data on the various factors that contribute to landslides. GIS software can also be used to generate maps that depict areas that are at higher risk of landslides. These maps can be used by planners and policy makers to identify areas that may need additional mitigation efforts for risk reductions analysis and assessments.

Mapping involves assessing geological, topographical, and environmental factors to determine the areas at higher risk of landslides. This is achieved by analyzing data on these factors and producing a map that highlights the areas that are more susceptible to landslides. The map can be a useful resource for land-use planning, risk management, and hazard evaluation.

To generate a landslide susceptibility map, various data sources can be utilized, such as geological maps, soil surveys, topographical maps, rainfall data, and land-use maps. These sources are integrated and assessed using statistical analysis, remote sensing, and GIS technology. Subsequently, the map displays the regions that are more prone to landslides based on the characteristics of the contributing factors. This map is then useful for land-use planning, hazard assessment, and risk management purposes.

In a research conducted in the Zerqa River Basin in Jordan, the susceptibility of landslides was mapped by analyzing contributing causes such as geology, elevation, slope angle, soil type, rainfall, and land use. GIS technology was employed to integrate and analyze these factors, leading to the development of a map that illustrates areas of high and low susceptibility to landslides. The map was subsequently utilized to identify high-risk areas and to formulate suitable risk management and land-use planning strategies.

The integration of multiple data sources and the landslide spatial distribution can be achieved through GIS technology, which is vital for susceptibility mapping. GIS software allows for the combination of diverse data sources, statistical analysis, and the creation of susceptibility maps displaying areas of high and low susceptibility. These maps can be used to prioritize areas for risk mitigation, identify locations that require further investigation, and support land-use planning and emergency management policymaking.

Landslide susceptibility mapping is a great instrument for identifying high-risk areas and informing decision-making for land-use planning and risk management. GIS technology is essential for analyzing and visualizing the contributing factors and creating susceptibility maps that display areas of high and low risk.

Landslide susceptibility mapping plays a crucial role in mitigating the hazards and risks associated with landslides. Landslide susceptibility mapping is a great instrument for identifying high-risk areas and informing decision-making for land-use planning and risk management. GIS technology is essential for analyzing and visualizing the contributing factors and creating susceptibility maps that display areas of high and low risk. By identifying the region that are in danger of landslides occurrences, decision-makers can prioritize risk management and mitigation, such as developing early warning systems, improving land-use planning practices, and implementing measures to stabilize slopes. This can help reduce the impact of landslides on communities and infrastructure, ultimately saving lives and minimizing property damage. By mapping these factors and their potential interactions, it is possible to identify

high risk areas prone to landslides. This is tool for proactive hazard management and can contribute to building more resilient communities.

1.4 Objectives

Main aim of this study is:

Identification of potential areas at high risk of landslides and develop a susceptibility map that can be used on future for disaster risk reduction and management.

- a. To create a current landslide inventory in the study region for future investigations.
- b. To analyze the factors contributing to landslides and develop a better understanding of their interactions and relationships.
- c. To establish and check new approaches and technologies for susceptibility mapping, with a focus on improving accuracy and efficiency.

1.5 Literature Review

The analysis of earthquake-induced landslide hazards, specifically examining the forms and magnitudes of bulk movements in regions with tectonic activity, was conducted in one of the early studies by Keefer in 1984 (Garcia et al., 2008). The majority of research on landslides triggered by earthquakes has concentrated on identifying, describing, and recording the landslides that occur during and immediately after the earthquake, especially those resulting from major seismic events (Khattak et al., 2014).

Garcia et al., (2008) used a Geographic Information System (GIS) to create a database that included multiple layers or coverages to investigate seven factors that influence landslides: elevation, lithology (bedrock and soil), slope, mean annual precipitation, slope aspect, land use and terrain roughness lines (faults, escarpments, dikes, paleo-riverbeds, and mineral seams), Polygons (pedogenic, geologic and volcanic classes), and points (volcanic, fumaroles, and fossils classes) were used to organize the lithology data. 20 meter contour lines on 1:25000 topographic map being used for model with a 100m cell size, which was perfect for regional work. The slope gradient and aspect of the terrain were

generated using the Digital Terrain Model (DTM). The landslide inventory was created using data from the SNET on slope displacement caused by the 2001 El Salvador earthquakes. Landslide density study was done to figure out how slope and elevation impacted landslides. According to the findings, places with soft terrain at lower elevations had a lower density of landslides than those with higher elevations. Receiver Operating Characteristics (ROC) curves was created for identifying likelihood of landslides, and logistic regression modeling was used to create a model of landslide susceptibility.

In the early post-seismic phase, new semi-automatic methods for detecting landslides using satellite images with high resolution which have the potential to help in landslides risk assessment (Harp et al., 2011, Martha et al., 2010;). In contrast to all-weather, day-night operational satellite radar sensors, optical satellites rely on solar illumination and favorable weather conditions to capture imagery (Wasowski, Keefer, & Lee, 2011). Developing precise landslide inventories is crucial in assessing the temporal patterns of co-seismic landslides and identifying the dominant hazard and risk. These inventories play a crucial role in analyzing the geographical, geological, seismic, hydrological, anthropogenic, and climatic factors that contribute to their occurrence and activation. Furthermore, landslide inventories are vital in devising effective long-term strategies for mitigating and restoring the impacted areas (Shafique, 2020). Logistic regression analysis was employed to assess the landslide risk in Izmir, Turkey, utilizing various factors, including slope aspect, lithology, slope gradient and proximity to fault lines, drainage and roadways. Based on the coefficients presented, lithology emerged as the most significant contributor to the incidence and distribution of landslides (Kıncal et al., 2017). Petley et al., (2006) reported that tension cracks have created intricate, interlocking patterns with both bedrock and colluvium near faults, resulting in high-density arrays.

In the recent decades, numerous studies have contributed to creating landslide susceptibility maps utilizing qualitative and quantitative methodologies. The mechanism of landslides is investigated using various parameters such as DEM, geology, and others. In landslide research, the quality

of data plays a crucial role, and achieving more reliable results depends on the adequacy and precision of the data across a broad range of parameters (Li et al., 2021). Comprehensive areal model of earthquake-induced landslides (CAMEL) has been established for decision-making in disaster risk reduction planning. CAMEL uses GIS and fuzzy logic systems to create a comprehensive framework capable of simulating various types of earthquakes induced landslides. The primary goal of CAMEL is to simplify the process of representing terrain conditions in both quantitative and qualitative terms and to aid in understanding the influence of these characteristics on the expected concentration area of each landslide type (Miles & Keefer, 2009).

The assessment of geo-hazards, debris flow after earthquake, and the impact of landslides on geomorphological advancement rely on the determination of the volume of landslides generated by significant earthquakes, making it a critical parameter. This value is often challenging to detect due to the hidden sliding surfaces underground and the three-dimensional nature of the landslides. Keefer fitted a power law relationship between the total volume of quake-triggered landslides and earthquake magnitude based on past 15 events (Xu et al., 2016).

Hubbard and Shaw (2009) conducted a study on a city in China, which is situated 30km away from the epicenter of earthquake. The Sichuan earthquake that occurred on 12 May 2008 was caused by the movement of the Longmen Shan fault, primarily on Yingxiu-Beichuan fracture (Yang & Chen, 2010). Many rocks and soils slid down onto communities, highways, and rivers across the broad territory along the 300 km northeast of the LFZ as a result of the Wenchuan Earthquake (W. Zhang, Lin, Peng, & Lu, 2010). More than 93% of the land of Sila Greca, is covered by forest, was selected as a testing ground for the GIS-based model, 4SLIDE. The model demonstrated strong predictive potential in identifying areas susceptible to shallow landslides, as confirmed by the sensitive Receiver Operating Characteristic (ROC) analysis (Moresi et al., 2020).

Shahabi & Hashim (2015) used the Analytical Hierarchy Process (AHP) which aids in decision making by assigning weights to parameters based on their importance. To use AHP, a pair-wise comparison matrix is created. The weighted linear combination (WLC), which assigns weight to each parameter based on the significance determined by the user, strikes a balance between qualitative and quantitative methodologies. Spatial multi-criteria evaluation (SMCE) enables operators to execute multi-criteria analysis using a location based approach. While (Marcus Nossner, 2002) used the Newmark's approach which calculates the increasing movement of a mass as a function of the acceleration time history of an earthquake. If the computed sum of the static and dynamic driving forces surpasses the block's shear resistance, the mass begins to move. Additionally, when the induced accelerations exceed a critical acceleration, permanent deformation can occur.

According to (Chacon et al., 2006; Spiker and Gori, 2000) landslide management and damage assessment is a critical step in formation of the landslides inventory map, which displays the contours & positions of landslides and their classification on larger scale maps. The following stage involves generating a landslide susceptibility map (LSM), which shows the spatial distribution of event-controlling factors including slope gradient and lithology which lead to slope breakdown. It helps to define locations that are prone to landslides without consideration to time, and identifies places where landslides are expected to occur in the future (Chacon et al., 2006). Last step is to produce a map of landslides hazard that shows temporal framework, indicating the likelihood of a landslides occurring in a given time frame, which differs from landslide susceptibility mapping.

An approach known as PARSIFAL approach is a probabilistic approach used to generate earthquake-induced landslide scenarios for seismic microzonation studies. The approach considers various causes like as topography, geology, seismicity, soil type and hydrology to generate expected scenarios that depend on seismic input and saturation conditions. The scenarios generated by the PARSIFAL approach can be used to map instability areas and

assess the potential impact of landslides on infrastructure and human settlements (Martino et al., 2019). MCE (Multi-Criteria Evaluation) is a widely used method that employs a variety of methodologies. Other methods, like multivariate statistical approach (MSA) and bivariate statistical analyses (BSA) developed by Ayalew (2005), are solely statistical in nature and use statistical correlations between landslides and environmental factors to identify the most important factors that influence landslide occurrence (Growley). OBIA approaches require high-resolution satellite images, which may not be available immediately after a landslide event. However, they can be used for landslide inventory and susceptibility evaluation in pre-disaster or post-disaster situations, if remote sensing data is available. Prediction rate curve is a useful tool for assessing the accuracy of statistical models. Weight of Evidence modeling is a widely used method for landslide susceptibility mapping and can provide valuable information for landslide disaster mitigation planning (Bacha, 2019).

Identifying and categorizing landslides through the application of convolutional neural network (CNN) algorithms on satellite imagery is a commonly used and highly effective approach in geological hazard assessments (Qin et al., 2021). SAR is an active microwave instrument that can collect electromagnetic signals reflected from the earth's surface and generate 2D map based on the sensor target distance and direction of flight (Colesanti and Wasowski 2006). PALSAR sensor utilizes L-band microwaves (wavelength = 23.6 cm) that can penetrate through leaves and tree branches. This allows for the detection of ground deformation that may be concealed by vegetation, making it more visible than with shorter wavelength microwaves. SAR is widely used in geological hazard assessments (Sato & Harp, 2009).

Most researches on landslides triggered by earthquakes has focused on identifying and describing coseismal landslides, specifically those resulting from catastrophic earthquakes, with little understanding of the relationship between post-seismic and coseismal landslide activities in most mountainous regions. In northern Pakistan's western Himalaya, the 2005 Kashmir earthquake resulted in landslides which occurred because of snowmelt and rainfall on the

slopes. At 68 different sites in AJK and NWFP, earthquake-triggered landslides were investigated. Kamp et al. (2008) used pre-earthquake 2001 ASTER data to map landslides and compared it to 2005 ASTER data. Total landslides increased from 369 to 2252 from 2001 (covering 8.2 km²) to October 2005 (covering 61.1 km²), according to (Kamp et al. 2008).

In domain of GIS, Multi-Criteria Evaluation (MCE) is an essential decision-making tool used for choosing alternatives or determining priorities such as landslide susceptibility. MCE integrates multiple factors to create a single composite to aid decision-making based on a specific goal. In this study, the declared objective of MCE is to analyze the selected research region to estimate landslide susceptibility. MCE uses different methodologies, some qualitative like AHP (Saaty, 1980) and WLC, and some exclusively employing statistical techniques, such as bivariate statistical analyses (BSA) and the approach of multivariate statistics (MSA) (Ayalew, 2005). Their study utilizes the AHP approach due to its precision and integration with the software used to conduct the analysis.

Pakistan faces a significant challenge from landslides that are triggered by earthquakes. The mountainous regions of northern Pakistan are particularly prone to natural disasters, and the Kashmir earthquake of October 8th, 2005 resulted in thousands of landslides across India-Pakistan, concentrated in 6 locations. A spatial database of 2,252 landslides was created and studied using satellite images of ASTER and GIS technologies. The significance of parameters that triggered the landslides was determined using a multi-criterion analysis. To manage and assess the damage caused by landslides. Multi-criteria evaluation (MCE) combines these criteria to create an only complex that can be used to make decisions for a precise goal (Malczewski, 1999) (Kamp et al., 2008). Based on the massive fracturing of dolomites near Muzaffarabad, it is highly probable that future heavy rainfall or low-intensity earthquakes could trigger slope failures in the area. Therefore, this location is prone to experience slope failures (van der Meijde et al., 2010). A study has been conducted on the Karakoram Highway (KKH), which passes through most seismically active

regions. The presence of active thrusts and strike-slip faults in the area leads to earthquakes, which in turn trigger landslides. Since 1904, there have been 317 documented occurrences of earthquakes with magnitudes greater than 5, and 10 of those have had magnitudes greater than 7, along the KKH. Notable events include the Muzaffarabad earthquake of October 2005 with a magnitude of 7.6 and the Afghanistan earthquake of October 2015 with a magnitude of 7.5 (Ali et al., 2019).

1.6 Statistical Analysis and Modelling

Statistical analysis involves the collection, analysis, interpretation, presentation, and organization of data. It involves applying mathematical formulas and techniques to describe and understand data. Different categories of statistical analysis include inferential statistics, descriptive statistics, time series analysis, regression analysis and cluster analysis.

In GIS-based landslide susceptibility modelling, statistical analysis and modeling have an important role in establishing the correlation between landslides and environmental factors such as geology, land use, soil, vegetation and topography. By analyzing these relationships, predictive models can be developed to estimate the likelihood of landslides occurring in a specific area.

Landslide susceptibility modeling in GIS involves statistical analysis and modeling to determine the correlation between landslide incidents and various environmental factors like topography, geology, soil, vegetation, and land-use. Predictive models are developed from these correlations to estimate the possibility of landslides happening in a particular location. Support vector machines, logistic regression (LR), artificial neural networks (ANN), random forests (RF) stand out as commonly employed statistical frameworks for the purpose of landslide susceptibility modeling. To evaluate the performance of these models, statistical analysis is carried out by comparing predicted landslide probabilities to actual landslide occurrences and computing measures of model accuracy, including sensitivity, specificity, and AUC-ROC.

MATERIALS AND METHODS

2.1 Study Area

The region under investigation consists of Chitral district and is situated in the northwestern part of Pakistan, in the Khyber Pakhtunkhwa province (Figure 2.1). To the north and west lies Afghanistan, while the eastern border of this place is shared with the Gilgit-Baltistan region of Pakistan. The region is mountainous, with several major mountain ranges intersecting in the area. The Hindu Kush range runs along the northern and western borders of the district, while the Pamir range extends along its eastern border. The Chitral River flows through the region, originating in the glaciers of the Hindu Kush and emptying into the Kabul River in Afghanistan. The district has a total area of around 14,850 square kilometers and is divided into several sub-districts or tehsils.. The district headquarters is in the town of Chitral. The Chitral region's altitude ranges from 1065 meters to 7701 meters above the sea level and is known for its rugged terrain, with steep mountains, deep valleys, and narrow gorges. The highest peak in the region is Tirich Mir, which is located in the Hindu Kush range.

2.1.1 Climate and Topography

The climate of the Chitral district is characterized by cold and temperate conditions with significant variations in temperature and precipitation. The region is situated in the mountainous northwestern part of Pakistan and is subject to a wide range of weather patterns for whole year. With temperatures ranging from 10 to 30 degrees Celsius (50 to 86 degrees Fahrenheit), the summers are warm and pleasant while, winters are chilly and snowy, with lows well below freezing. Snowfall is common in the winter months, especially at higher elevations. The district receives most of its precipitation during the summer monsoon season, which runs from June to September. During this time,

the region experiences heavy rainfall and occasional flooding, particularly in low-lying areas. The monsoon season is critical for agriculture in the district, providing much-needed water for crops and livestock. The rugged topography of the region is defined by steep peaks, narrow gorges, and valleys. Situated at the convergence of multiple prominent mountain ranges such as the Pamir, Karakoram, and Hindu Kush, the region bears this distinctive feature. These mountains form a natural barrier that separates the region from the rest of Pakistan and contributes to its isolation and unique identity.

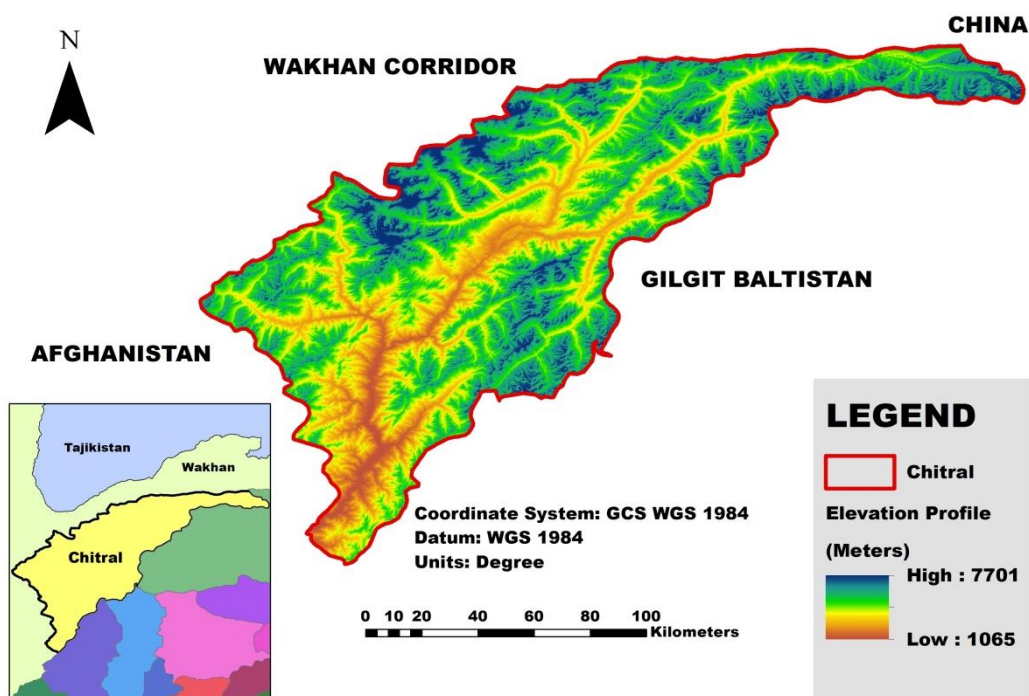


Figure 2.1. Map of the area under study showing the district Chitral of the KPK province of Pakistan. It is situated towards the northwestern border of Pakistan with Afghanistan and is mostly run by the Hindu Kush mountain ranges of the country.

The region's topography has a significant impact on its natural resources and land use patterns. The mountains provide habitat for a wide range of wildlife, including the snow leopard, Himalayan Ibex, and Markhor. The valleys and river basins are home to forests, agricultural lands, and pastures for grazing

livestock. The Chitral district's climate and geography create unique challenges and opportunities for the region's residents and natural resources. Harsh winter conditions, combined with rugged mountain terrain, can make travel and access to remote areas difficult. However, the region's natural resources, including forests, rivers, and wildlife, provide important economic and ecological benefits to the local communities.

2.1.2 Geology and Tectonics

The geology of Chitral is characterized by complex and a diverse range of rock formations, reflecting the region's geological history. It is situated within the Himalayan orogenic belt and consists of sedimentary, igneous, and metamorphic rocks that have undergone significant heat and pressure during the mountain-building process. The oldest rocks in the Chitral region are Precambrian in age and include granite, gneiss, and schist. Above these rocks are younger sedimentary formations, including sandstone, shale, and limestone. The Paleozoic and Mesozoic eras are represented by the Hazara Formation, which consists of limestone, shale, and sandstone. The region's Cenozoic rocks include the Murree and Nagri formations, which are composed of sandstone and shale. These formations are important for their fossil content, which includes plant and animal remains that provide insight into the region's past environments and ecosystems. The region is located at the interface of the Indian and Eurasian plates, which have been converging for millions of years. This collision has led to the creation of the Himalayan Mountain range which also includes the Mount Everest. The Chitral region is characterized by complex tectonic structures, including thrust faults, folds, and uplifted blocks, which have resulted from the compression and deformation of the Earth's crust due to the plate collision. The region's geology is also marked by the presence of ophiolites, which are fragments of oceanic crust and upper mantle that were thrust onto the continent during the collision. The region is still tectonically active, with ongoing deformation and seismic activity. As a result of the convergence between the Indian and Eurasian plates, the Himalayan Mountain range formed along with

the development of multiple faults throughout the region. One of the major faults in the Chitral region is the Reshun Fault. It is a thrust fault that runs along the Chitral valley and is capable of producing large earthquakes. The fault has been responsible for several earthquakes in the past, including the earthquake of 2015, on the Richter scale it had a magnitude of 7.5. The Reshun Fault is a part of larger Hindu Kush-Himalayan (HKH) seismic belt, which extends from Afghanistan to Myanmar. This region is known for its high seismicity, and a lot of major earthquakes have happened in this region in the past. The region's complex tectonic history and ongoing activity make it an important area for studying the processes that shape the Earth's crust and the hazards associated with seismic activity.

2.1.3 Landuse and Landcover (LULC)

The classification of the Chitral region's land cover reveals a varied landscape, which is dominated by glaciers, natural vegetation, agriculture, and bare ground. Total area of the district is almost 14, 866 km². The region's LULC can be broadly classified into forests, agriculture, settlements, and water bodies (glaciers and snow). Almost 40% of the area is covered with glaciers and snow. Forests and rangelands are also in abundance in the district and covering almost 35% of the area with more than 5175 km². Agriculture is practiced in the fertile river valleys and terraced hillsides which covers almost 3.2% of the district and is mainly focused on growing crops such as wheat, maize, and potatoes. Barren land comprises of 20% of the district area. Settlements are scattered in about 1% of the total land cover and are concentrated around the larger towns and villages, characterized by a mix of residential, commercial, and industrial land use. The region also has several water bodies, including rivers, lakes, and wetlands, which are important for irrigation, hydropower generation, and biodiversity.

2.2 Materials

2.2.1 Satellite Images

To overcome field data collection in the inaccessible study area, Landsat 8 data was used as it is considered a powerful tool for collecting information on landslides and understanding the processes that drive them. Landslides were identified by analyzing changes in the terrain and vegetation cover. Areas of bare ground and areas where vegetation has been disrupted or removed were initially considered as landslides. Satellite imagery provided the base for Landslide Susceptibility Mapping and establishing of the Landslide Catalog of the region. Table 2.1 shows the detailed information.

2.2.2 Tools and Software

Table 2.2 lists the tools and software utilized for data analysis and processing.

2.3 Analytical Framework

2.3.1 Methodology

This research is conducted in four major sections, which comprise a literature review and data collection using Landsat satellite imagery (landslide inventory and exploratory data) to generate various maps. Google Earth Engine (GEE) was then used for change detection and Classification of LULC was performed from 2001 to 2021. Statistical analysis was performed, and a Landslide Susceptibility Model was developed for the study. Detailed methodology is shown in (Figure 2.2), and a step-by-step explanation is given below:

Step 1: A comprehensive literature review was conducted on earthquake-triggered landslides and landslide susceptibility mapping to gain a thorough understanding of the current state of research, methodologies, and best practices. All necessary data was gathered for the study area, including geological and topographical maps, satellite imagery, digital elevation models (DEMs), and earthquake data. Other exploratory data collection included hydrology, slope angle, aspect, lithology, and vegetation cover.

Step 2: Preprocessing of satellite imagery was done. Landslide inventory was created through time series in Google Earth Pro and DEM was used to extract the required information of elevation, slope angle and aspect. Hydrology data was used to digitize all the rivers in study area and geological information was extracted from the geology datasets. Different maps were generated from these datasets to further incorporate it in the statistical analysis.

Step 3: In this step, LULC Classification was performed to observe the changes in land cover. Then Point Pattern Analysis was performed for landslide susceptibility mapping that allowed to identify spatial patterns and clustering of landslide occurrences in the Chitral district.

Step 4: Finally, Landslide Susceptibility Model was generated in “R” by incorporating all important factors that contributed to landslide susceptibility. Ultimately, the validation of the landslide susceptibility map involved the utilization of receiver operating characteristic (ROC) analysis and the assessment of the area under the curve (AUC). These methodologies were instrumental in appraising the precision and trustworthiness of the generated map.

Table 2.1. List of datasets with its purposes.

Materials/Data	Purpose
Landsat 5 and 8 Imagery	Landslide Inventory Landuse/Landcover (LULC) Vegetation Index (NDVI)
DEM (SRTM)	Elevation Slope Angle Aspect Stream Network

Table 2.2. *Tools used for Analysis and Processing.*

Software/Tools	Purpose
ERDAS Imagine/QGIS	Preprocessing of Landsat Satellite Imagery
ArcMap/QGIS	Landslide Inventory, Processing of DEM, Landslide Inventory Map, Elevation Map, Slope Angle, Aspect Map, Geological Map, Hydrology Map, Landslide Susceptibility Maps
Google Earth Pro	Time series study of landslides, Digitization of rivers and streams
Google Earth Engine	Landuse/Landcover Classification (LULC), Vegetation Index (NDVI)
R-Studios, Python and Microsoft Excel	Statistical Analysis and Modeling
Microsoft Word	Thesis Report Writing
Microsoft Power Point	Presentations

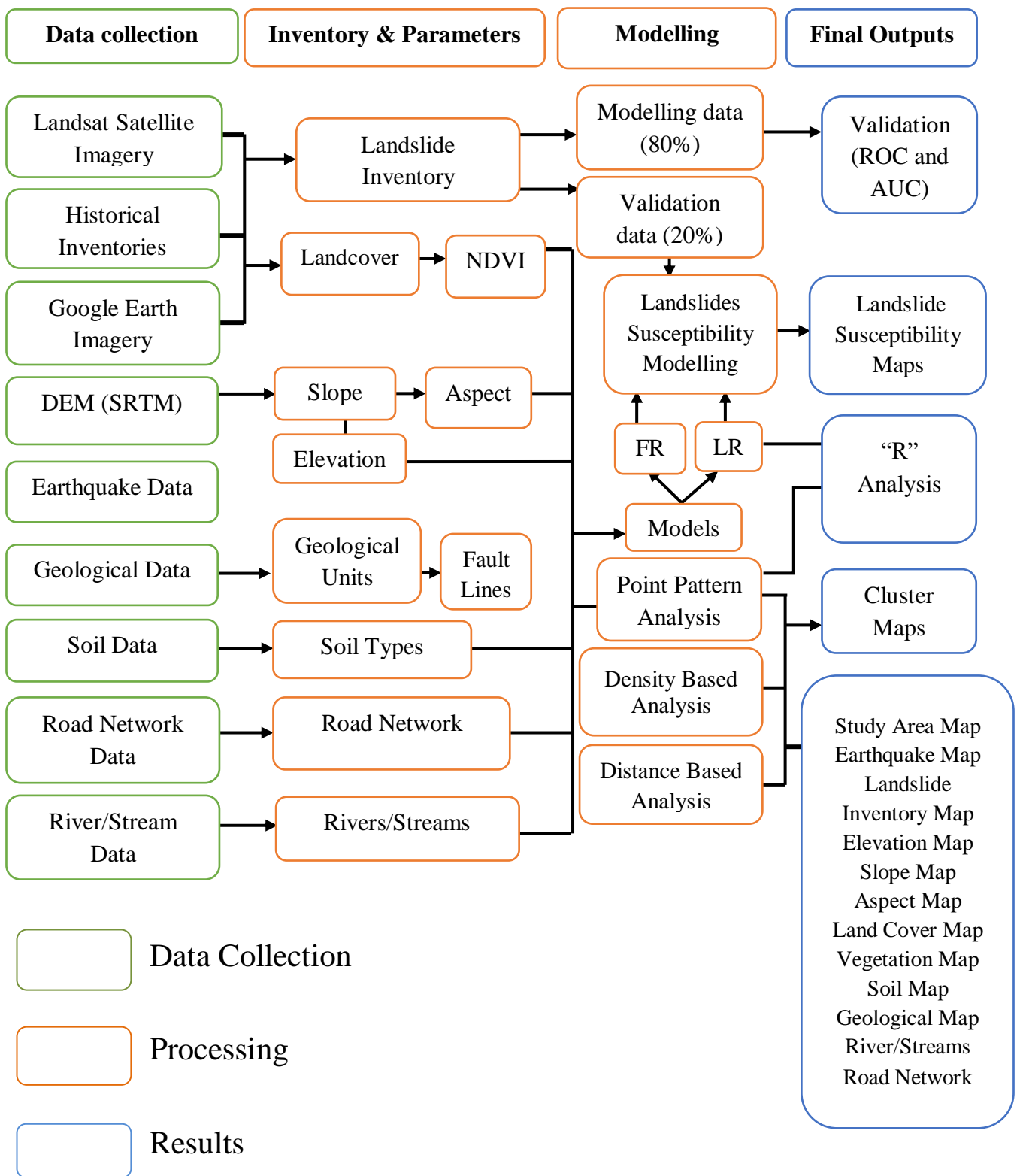


Figure 2.2. Detailed methodology flowchart showing step by step collection of data and processing with results.

2.4 Earthquake Data

Earthquake data was collected from USGS Earthquake Catalogue (<https://earthquake.usgs.gov/earthquakes/map/>) for mapping the events in the study area that provided valuable insights into the seismic activity and earthquake risk in the region. Earthquake event of 26th October 2015 having magnitude of 7.5 with 115 other earthquakes above 4.5 magnitudes on Richter scale were mapped in such a way that the landslides triggered by these earthquakes in a radius of 400 kilometers could be identified. This data was further analyzed to identify patterns and trends in the seismic activity in and helped in developing the Landslide inventory for study area.

2.5 Landslide Inventory

The utilization of satellite imagery can be highly effective in recognizing potential landslide zones and creating a comprehensive inventory of landslides for a given study area. However, different other sources of data and information can be combined to create a comprehensive landslide catalog.

The best alternative for developing a landslide catalog of an inaccessible region is Satellite imagery. Making a landslide inventory requires various steps. The USGS Earth Explorer website (earthexplorer.usgs.gov) provides free access to Landsat 8 imagery. First of all, Landsat 8 imageries were obtained for the study area of Chitral district in such a way that Earthquake event of October 26 2015, was set as a reference event. Before and after event imageries with cloud cover less than 10% were downloaded for the year 2015. Preprocessing of the imagery was done in ERDAS Imagine software in which atmospheric corrections and radiometric calibrations were performed to ensure that the image data is accurate. Different bands were combined to create a composite image to highlight different features and identify potential landslide areas. Initially the visual interpretation helps in interpretation of the areas with distinctive features such as scarps, hummocky terrain, or steep slopes. Identified landslides can be verified by comparing it with the past landslide inventories if ground validation

is not possible. Once identified and verified the landslides, one can generate a landslide inventory for this study.

Another way or procedure for development of a landslide inventory is the Google Earth Pro. It can be used for the identification and analysis of landslides using time series data. Time series data involves the analysis of changes over time, which can be helpful in identifying the characteristics of landslides and their potential impact. Time series (temporal) study was performed for the identification of landslides in Google Earth Pro in number of steps. First step included the visual interpretation of the historical and present high resolution satellite imagery provided by the Google Earth Pro. “Time Slider” tool was used to view the images over time and compare them to identify any changes. This tool allowed viewing multiple images of the same locality at dissimilar points in time. Diverse areas in the study area were identified where there were noticeable changes in topography, such as new or larger cracks or fissures, changes in slope or vegetation cover, or other indications of ground movement. Some other tools for measurement were used to measure the dimensions and locations of the identified landslide features, as well as any potential areas at risk of future landslides. A total of 210 landslides were identified and confirmed through comparison with historical landslide records of the area.

2.6 Remote Sensing/Satellite Data

Google Earth Engine (GEE) has a wide range of satellite imagery data available for free, such as Landsat, Sentinel, and MODIS. Data can be accessed through the GEE Code Editor. Different satellite imagery was selected as required for analysis. For Landuse/Landcover (LULC) classification, Landsat satellite data was used and for Normalized Difference Vegetation Index (NDVI) calculations, Sentinel 2 imagery was used.

2.6.1 LULC Classification in GEE

For Landuse/Landcover classification (LULC) of the study area, three Landsat images of two different years were selected. Landsat 5 satellite image of year 2001 was selected and Landsat 8 image was selected for the year 2021. Both the satellite images have a spectral resolution of 30 m. Images were pre-processed by removing cloud cover, shadows, and atmospheric corrections by applying different tools and functions in Google Earth Engine. Supervised Classification was performed on each image by using traditional Machine Learning (ML) algorithms and five different classes were extracted for these years namely Built-up area, snow cover, soil, vegetation, and water. GEE code is attached in **Appendix A**.

2.6.2 NDVI Calculations in GEE

For NDVI calculations and vegetation cover of the study area, satellite data of Sentinel-2 for the year 2021 was imported in Google Earth Engine. The image was pre-processed by removing the cloud cover and a function to clip the image according to the study area was applied. After that, the Red and Near-Infrared (NIR) bands were selected from the satellite data. The red band is denoted by B4 and the NIR band is represented by B8. NDVI was calculated by using following formula:

$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (2.1)$$

2.7 Digital Elevation Model (DEM)

Digital Elevation Model (DEM) which is a comprehensive digital dataset of the Earth's terrain, covering the entire globe was generated.

Specifications:

- Data Collection Date: February 11, 2000 – February 22, 2000
- Data Source: Space Shuttle Endeavour using the SIR-C/X-SAR radar instrument.
- Spatial Resolution: 30 meters
- Vertical Accuracy: +/- 16 meters (relative) and +/- 20 meters (absolute)

First of all, pre-processing was done on the downloaded SRTM DEM. SRTM DEM data is provided in tiles that cover different parts of the Earth's surface. It was necessary to mosaic the tiles together to create a seamless elevation dataset that covers the entire study area of Chitral district. Once DEM was mosaicked, it was further used for extraction of river networks in the study area plane curvature, aspect, slope angle and elevation,

2.8 Hydrology Data

To map the drainage networks, river data was needed for the study area. DEM was used to extract rivers. First the DEM was pre-processed by filling in some sinks, which are areas of the DEM where the elevation is lower than the surrounding area, so that water can flow out of them. This step is important because sinks can interfere with the hydrologic analysis. Then the flow direction and flow accumulation grids were derived from the preprocessed DEM. These grids were used to identify the flow path of water and the accumulation of flow at each cell. Using the flow direction and flow accumulation grids, the stream network was extracted by thresholding the flow accumulation grid at a suitable value to isolate streams and rivers. Further the extracted stream network was refined by smoothing and simplifying the lines and by using tools such as edge detection and image segmentation. This helped to remove any noise or inaccuracies in the original data and create a more visually pleasing result. Missing rivers were digitized as accurately as possible, using high-resolution imagery and taking care to follow the exact course of the river to ensure the resulting data to be accurate as possible. By incorporating river data into

landslide susceptibility mapping, as it has the potential to provide a broader insight of the potential triggers and locations of landslides. This information can be used to develop effective mitigation strategies and land-use planning.

2.9 Geological Data

Geological data is crucial for mapping landslides because the geology of an area has a substantial impact in the stability of slopes and the likelihood of landslides. The Survey of Pakistan has produced a geological map of Pakistan that provides information on the geological structures and formations found in the country. The geological map depicts the different geological formations and rock units found in the country. The map also provides information on the age and composition of these rock formations. The map is divided into different regions, each with its unique geological features. Main Fault lines and geological formations were digitized manually by the help of this map. Other geological datasets were downloaded and incorporated into the digitized data to develop a geological map of the region. The USGS Geological datasets provided a lot of information about the geological formations of the Chitral district. It also provided information about the composition and structure of rock and soil layers, which helped in identifying the potential triggers and locations of landslides. Geological data can be used to map areas that are prone to landslides based on their geological characteristics. For example, areas with steep slopes, loose soils, or weak rock layers are more likely to experience landslides. By incorporating geological data into landslide mapping, it is possible to grow a more complete understanding of the potential triggers and locations of landslides in future.

2.10 Statistical Analysis and Modeling

Statistical analysis and modeling are important components of geographic information systems (GIS) that help to extract meaningful information from spatial data. Statistical analysis in GIS involves the use of various statistical

tools to analyze data and detect patterns or trends. Modeling involves the use of mathematical models to simulate various spatial phenomena.

2.10.1 Point Pattern Analysis

Point pattern analysis is used to identify patterns and clustering of events or phenomena in a particular geographic region. Point pattern analysis can be applied in the field of landslides to determine the spatial arrangement of landslides and pinpoint regions with a higher risk of landslide activity. The approach involves assessing the concentration of landslide occurrences and examining the spatial association between individual landslides.

Descriptive Statistics

In point pattern analysis (PPA); descriptive statistics are used to measures of dispersion and central tendency. According to O'Sullivan and Unwin (2010), central tendency refers to the determination of the central location of a point pattern and provides information of how points are distributed. The commonly measures of central tendency in PPA are mean and median center (Gimond, 2019). (x, y) coordinates of all points are used as an average by mean center in the study area, as shown below.

$$(\mu_x, \mu_y) = \left(\frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n} \right) \quad (2.2)$$

The coordinates of the mean center are denoted by (μ_x, μ_y) , while (x_i, y_i) are coordinates of point i . while n is total number of points.

Median center is basically location that minimizes the total distances traveled to all points. This center is calculated by the help of the procedure developed by Kuenne & Kulin in 1962. Algorithm begins with an initial location, such as the median center, and then updates its coordinates (x', y') through the following process, as explained by Rogerson (2019):

$$x' = \frac{\sum_{i=1}^n \frac{w_i x_i}{d_i}}{\sum_{i=1}^n \frac{w_i}{d_i}}, y' = \frac{\sum_{i=1}^n \frac{w_i y_i}{d_i}}{\sum_{i=1}^n \frac{w_i}{d_i}} \quad (2.3)$$

Where as " d_i " is distance, between the point (x_i, y_i) and " w_i " is weight of point.

For computing an unweighted mean center, " W_i " remains constant across all locations.

Iterative process is carried out till the difference between the previously computed median centers and newly computed is no longer significant, i.e., the distance between the two centers is less than a pre-defined threshold.

a) Standard distance:

Standard deviations and the standard distance have comparable definitions. Both these indicators are used to determine the extent of spread or dispersion of a set of points around their mean center (Gimond, 2019). The calculation of standard distances involves utilizing Equation 2.4.

$$d = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu_x)^2 + \sum_{i=1}^n (y_i - \mu_y)^2}{n}} \quad (2.4)$$

The equation takes into account the coordinates of the mean center, represented by (μ_x, μ_y) , along with the coordinates of a specific point " i ", represented by (x_i, y_i) , and the total number of points denoted by " n ".

b) Standard deviational ellipse:

Although it can reflect the level of dispersion of a point pattern, standard distance only computes an isotropic metric and fails to demonstrate any directional influence. Researchers use standard deviational ellipses to compute independent standard distances on two axes that are perpendicular for this purpose. The mean centre of the ellipse serves as the major axis, and the total length of orthogonal axis is calculated by the matching standard distance along the direction with the highest dispersion. The standard deviational ellipse is very useful for displaying point patterns with a directed orientation (ESRI, 2018).

According to Gatrell et al. (1996), the standard deviational ellipse can be calculated using point locations or by giving various points weights (w) based on their characteristics. That is known as weighted standard deviational ellipse. A weighted directional distribution's rotational semi-minor (y) and semi-major (x) axes may be determined as follows (Wang, Shi, & Miao, 2015):

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = w \cdot \begin{pmatrix} x_i \\ y_i \end{pmatrix} - \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad (2.5)$$

$$\sigma_x = \sqrt{\frac{1}{n} \sum_1^n (\tilde{y}_i \sin \theta + \tilde{x}_i \cos \theta)^2} \quad (2.6)$$

$$\sigma_y = \sqrt{\frac{1}{n} \sum_1^n (\tilde{y}_i \cos \theta + \tilde{x}_i \sin \theta)^2} \quad (2.7)$$

The equation considers the weight matrix denoted by " w ", along with the coordinates of point " i " represented by (x_i, y_i) , the (weighted) mean center indicated by (μ_x, μ_y) , and " θ " is rotation angle determined by:

$$\tan \theta = \frac{(\sum_{i=1}^n \tilde{x}_i^2 - \sum_{i=1}^n \tilde{y}_i^2) + \sqrt{\sum_{i=1}^n \tilde{x}_i^2 - \sum_{i=1}^n \tilde{y}_i^2 + 4(\sum_{i=1}^n \tilde{x}_i \tilde{y}_i)^2}}{2 \sum_{i=1}^n \tilde{x}_i \tilde{y}_i} \quad (2.8)$$

Distance-based Measures

By computing the separations between pairs of points, distance-based measures are used to analyse the spatial arrangement of points. These measurements are frequently considered to be the most accurate predictors of the second-order attribute. There are several distance-based measures that use distances other than Euclidean distances, even though the majority of these measures use Euclidean distances (Lamb et al., 2016; Okabe & Yamada, 2001).

a) Nearest-Neighbor Distance:

A point's distance from its closest neighbour is referred to as the nearest-neighbor distance (NND). The first order nearest neighbour is another name for NND. In addition to NND, the kth nearest neighbour, commonly known as the kth-order NN or KNN, can also have its distance determined. The mean NND among all point sets is frequently employed as a global indicator to assess a point set's overall pattern (Clark & Evans, 1954). The mean NND of a certain point collection can be compared to the anticipated NND produced from points that follow complete spatial randomness (CSR), which can help establish the significance of the pattern. A point pattern known as a CSR has points that all appear to be randomly dispersed within the research region. This comparison helps in assessing the level of clustering or dispersion present in the point pattern. Equation 2.9 can be used to calculate the mean NND (\bar{D}) for a point set.

$$\bar{D} = \frac{\sum_{i=1}^n d_i}{n} \quad (2.9)$$

Where " d_i " is Distance of point " i ", as " n ", is total points. Moreover, the mean NND in a scenario of complete spatial randomness (CSR) can be determined as:

$$\bar{D}_E = \frac{0.5}{\sqrt{\frac{n}{A}}} \quad (2.10)$$

The area of a point set is used in the following equation to calculate the z-score of mean NND (represented by A):

$$z = \frac{\bar{D} - \bar{D}_E}{SE} \quad (2.11)$$

Where

$$SE = \frac{0.261356}{\sqrt{n^2/A}}$$

b) Distance Functions:

i.G Function:

The mean NND offers a single-value metric for determining how clustered a set of points is, but it provides only a limited understanding of the complex nature of point pattern at various spatial scales. In order to account for more complex variations in a point pattern, multiple distance functions have been developed. Among them, the G function is the simplest approach used to measure cumulative frequency. And it is represented as:

$$G(d) = \frac{\text{sum}(D_{ij} < d)}{n} \quad (2.12)$$

In the G function equation, " $\text{sum}(D_{ij} < d)$ " denotes the count of points (i, j) that have smaller distance as compared to d .

ii.F Function:

The F function involves generating a small set of random points (referred to as "P"), where as computing the min. distance between random points & any of the original points (referred to as "O"). It expressed as follows:

$$F(d) = \frac{\text{sum}[d_{\min}(p_i, s) < d]}{n} \quad (2.13)$$

$F(d)$ = value of the function at distance d .

$\text{sum}[d_{\min}(p_i, s) < d]$ = count of points in P .

iii.K Function:

Since the F & G functions focus solely on the NN of every point and do not take into account distances to other points, they are not suitable for analyzing point patterns at multiple scales or for reflecting local variations.

While Ripley's K function can be used to identify multi-scale point patterns. The K function is expressed as follows:

$$K(d) = \frac{R}{n^2} \sum \sum_{i \neq j} \frac{I_d(d_{ij})}{w_{ij}} \quad (2.14)$$

The extent of the study region is denoted by R , and the indicator “ w_{ij} ” accounts for edge correction.

Density-based Analysis

Density-based measures are employed by researchers to investigate how point densities vary across space. There are two types of density measurements: local density and global density. Ratio of the observed points to area of the research region is called global density and can be easily calculated by given equ.:

$$\lambda = \frac{n}{a} \quad (2.15)$$

The equation calculates the global density using the number of points (n) and the area of the study district (a).

Local density, on the other hand, considers variations in point density across the study region. Two density based measures used to study local density are quadrat density and kernel density.

a) Quadrat Density:

Divided into smaller subregions called quadrats, the research area is subjected to a quadrat density analysis, which involves calculating the point density for each quadrat. According to (Gimond 2019) quadrats can be hexagons, squares, triangles, or Thiessen polygons among other shapes. However, the shape, size, and number of quadrats used have a major effect on the results of quadrat density. Small quadrats may result in multiple quadrats with no or few points,

but exceptionally big quadrats may not be able to record fine-scale changes (Anderson & Marcus, 1993).

b) Kernel Density:

Kernel Density Estimation (KDE) differs from quadrat density analysis in that it assumes that every location has a density, and density estimation is done by kernel. By counting the number of events in a search window centred at the density calculation site, KDE determines the local density of points. Only selected point window are counted, and nearby points often carry more weight than far-off places. A number of kernel functions, such as gaussian kernel, polynomial kernel, exponential kernel, uniform kernel, and the linear kernel are available to assign weights to the points. KDE is useful for transforming discrete data into continuous variables, as it generates a continuous surface of point densities.

2.11 Point Pattern Analysis in R

R-Studio is a comprehensive IDE designed specifically for the R programming language, providing a user-friendly interface for writing and executing R code, as well as a suite of tools for data analysis, visualization, and package development. R language is being widely used and designed for statistical analysis, and it is frequently employed in point pattern analysis along with various other data analyses. “R” code is attached in the *Appendix B*.

2.12 Landslide Susceptibility Mapping (LSM)

Landslide susceptibility mapping aims to identify areas with a higher possibility of landslide occurrences based on the analysis of various contributing factors. These factors may include environmental, geological, and topographic factors that affect landslide occurrences. The outcome of this analysis is a map that demonstrates the susceptibility level of an area to landslides. The initial step in landslide susceptibility mapping involves collecting data on various factors that can influence landslides, including topographical maps, soil surveys, geological

maps, land use maps, and rainfall data. Following that, the collected data is analyzed using different methods such as remote sensing, statistical analysis, and GIS technology, which can aid in the identification of areas with greater risks of landslides based on the analyzed factors. The results of the analysis shows susceptibility map that indicates areas that are at higher risk of landslides. This map can assist in the development of land use planning, engineering design, and emergency management. Nonetheless, it can provide useful information that can help to mitigate the risks related to landslides, decrease potential damage, and prevent loss of life.

RESULTS AND DISCUSSIONS

3.1 Earthquake Mapping

Google Earth Pro was utilized to map the earthquake event of October 26, 2015 so that seismic activity within 400 kilometers of the epicenter could be observed. The epicenter of the earthquake was seen 45 kilometers south of Feyzabad, in the vicinity of Jarm, Badakhshan, Afghanistan (Figure 3.1). It was a deep-focus earthquake with a magnitude of 7.5 that happened at a depth of around 212 kilometers. A large area, including Pakistan, India, and Tajikistan, felt the quake. There were hundreds of reported fatalities and thousands of reported injuries from the earthquake that left the area significantly damaged. The Chitral region of Pakistan was worst affected, with results of fatalities and hundreds of injuries.

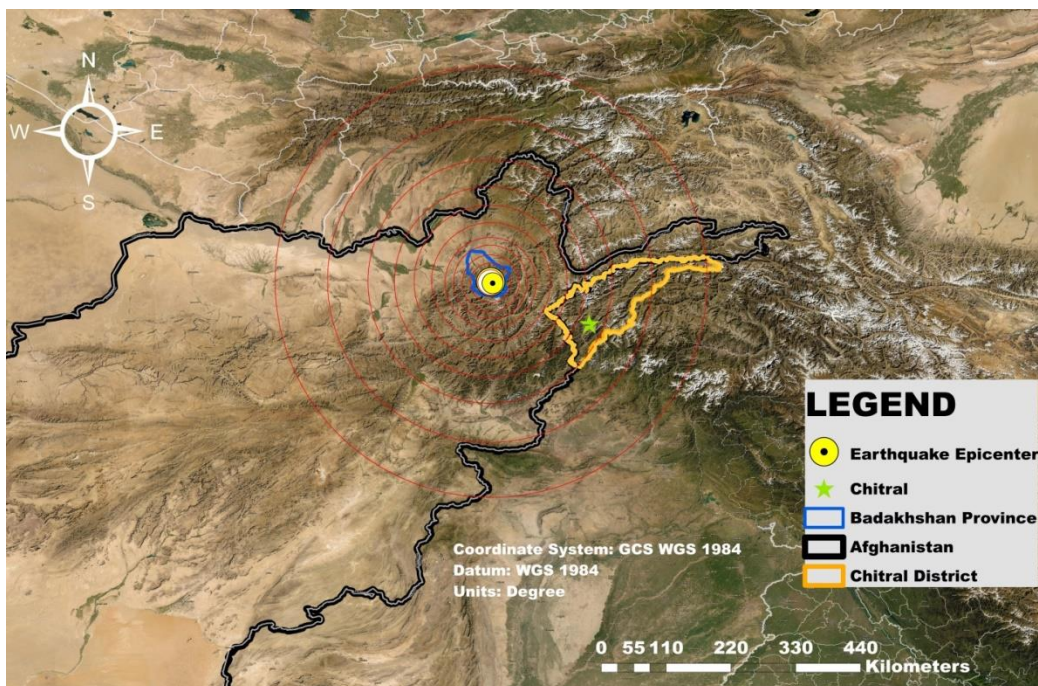


Figure 3.1. Earthquake map of 26th October, 2015 mapped showing the exact location of the epicenter and the affected region of Chitral, Pakistan.

3.2 Landslide Inventory

Landslide catalog was created by plotting a total of 210 landslides that are scattered almost all over the study area. Landslides were identified using satellite imagery and Google Earth Pro by using the time series study. Temporal images of 20 years from the year 2001 to 2021 were visualized to identify the landslide prone regions. Most of the landslides identified were along the major riverbanks. The study area is surrounded by deep valleys and the mountainous terrain which makes it difficult to identify and mark the landslides locations. Some landslides were identified along the roads and fault lines due to different events like anthropogenic and seismic activities taking place in the region. Identified landslides were validated with the NASA landslides catalog and Global landslides viewer to check the accuracy with the historical landslide inventories. Furthermore, the identified landslides were categorized into modelling datasets (80%) and validation datasets (20%) for further processing and modelling of the Landslide Susceptibility Mapping.

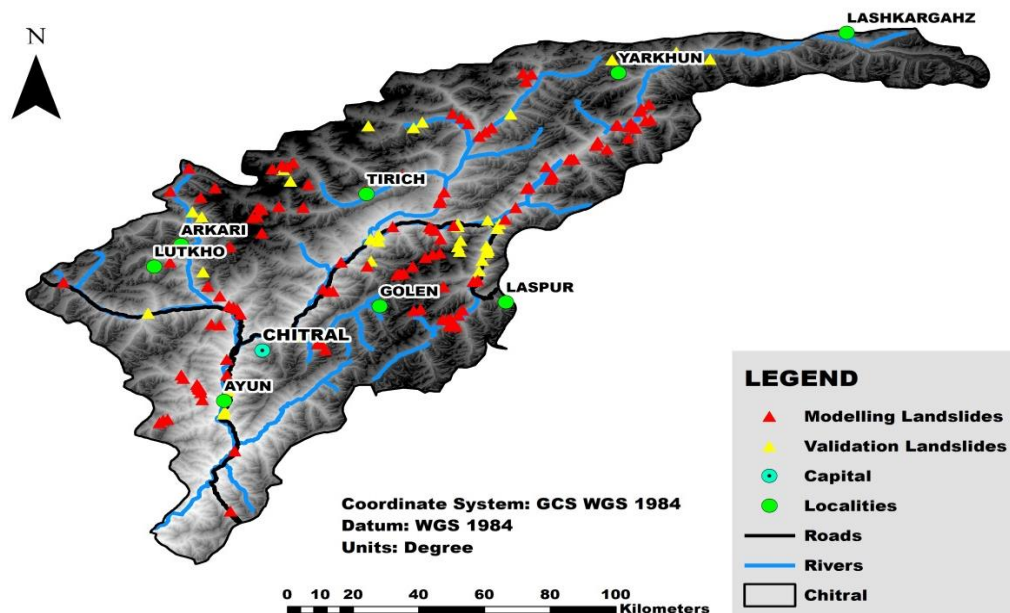


Figure 3.2. Landslide Inventory map of Chitral showing locations of the landslides (Modelling samples that were used to train different models Validation Samples that were used to validate the models after training them).

3.3 Landslide Causative Factors

Landslides can occur in many different forms, ranging from small debris flows to catastrophic rockslides. However, all landslides share a common feature: they involve the movement of soil, rock, or other materials down a slope. To better understand and manage the risk of landslide, it is very important to identify and understand the factors that contribute to their occurrence. A total of 10 causative factors for landslides were identified that includes elevation, slope angle, aspect, land cover, geological/lithological units, soil type, vegetation cover, distance to faults, rivers/streams and roads. Given factors plays crucial role in identification of the potential landslide occurrences and landslide susceptible zones when combined and studied under different statistical models.

Elevation Profile

Elevation plays an important role as a landslide causative factor. Elevation was mapped from the SRTM Digital Elevation Model which is freely available from the USGS website, and it has a resolution of 30 m. The elevation profile of Chitral district ranges from 1065m -7701m above the sea level and it was further classified into 5 classes of equal intervals. These 5 classes (1065 – 2392.2, 2392.2 – 3718.4, 3719.4 – 5046.6, 5046.6 – 6373.8 and 6373.8 – 7701) were mapped to show the landslide occurrence in each class with respect to the area of each class (Figure 3.3, 3.4).

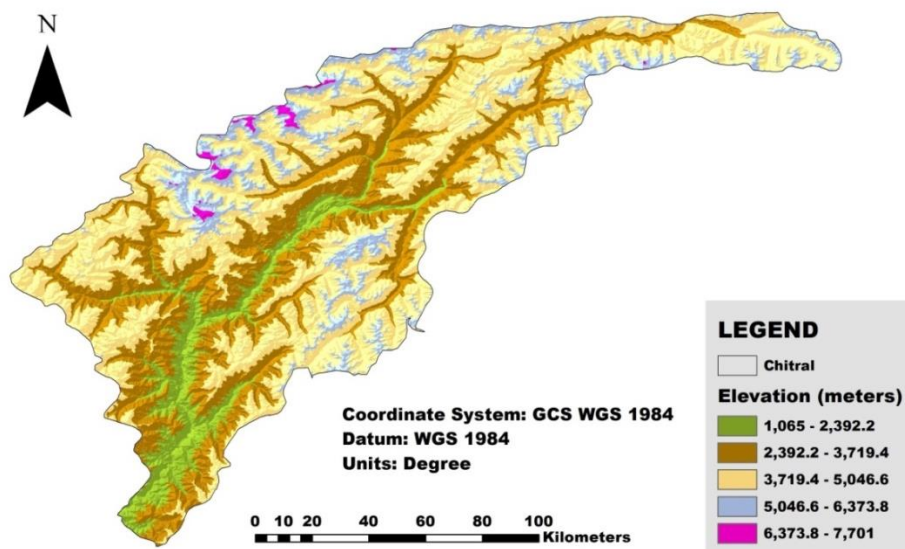


Figure 3.3. Map of Elevation Profile Chitral divided into 5 classes with equal intervals.

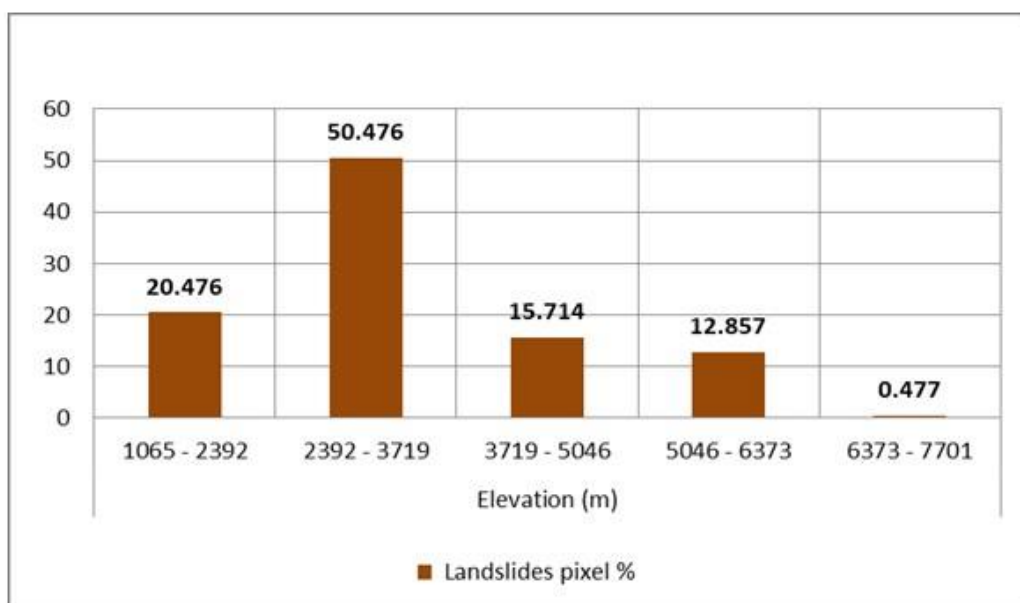


Figure 3.4. Relationship between the Elevation classes and the landslide pixel % in corresponding class.

Slope Angle

The greater the slope angle, the more likely landslides are to occur. The equilibrium between shear pressures and resistance to shear determines how stable a slope is. Landslides may occur when the slope angle is steep due to an increase in movement of mass (Guillard and Zezere, 2012). According to studies, the shifting of the landslip mass and the lateral forces exerted on the slope of the hill are both proportional to how steep a slope is (Tien Bui et al. 2017). Additionally, slope gradient is important for subsurface movement and influences the amount of soil moisture, both of which have a direct bearing on the likelihood of landslides (Magliulo et al. 2008). A map of slope angle of Chitral was created using a Digital Elevation Model (DEM) with 30 meters spatial resolution (Figure 3.5). Map covered a range of slope angles from 0 to 80.6 degrees and was further classified into 5 classes ranging from (< 15 degrees, 15 – 30, 30 – 45, 45 – 60 and > 60 degrees).

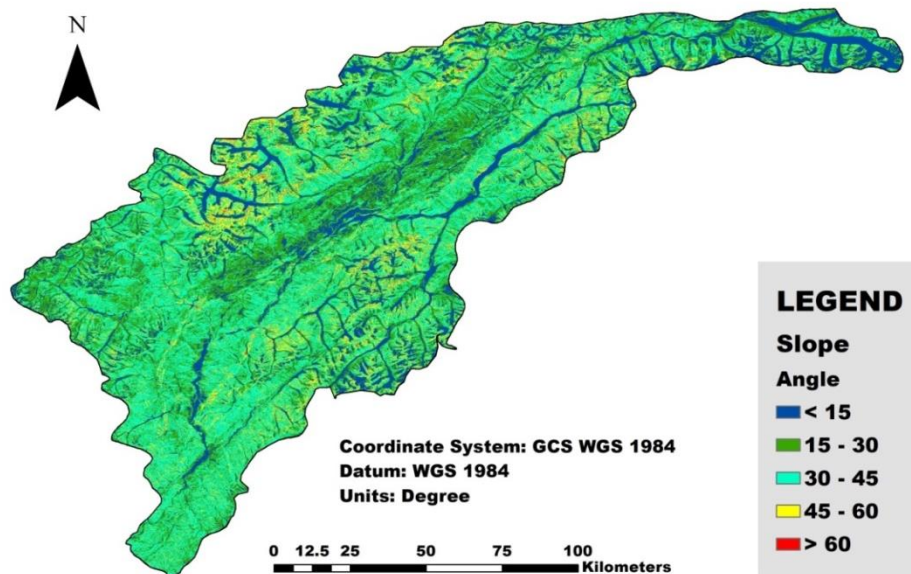


Figure 3.5. Slope map of the study area showing different slope angles in different classes.

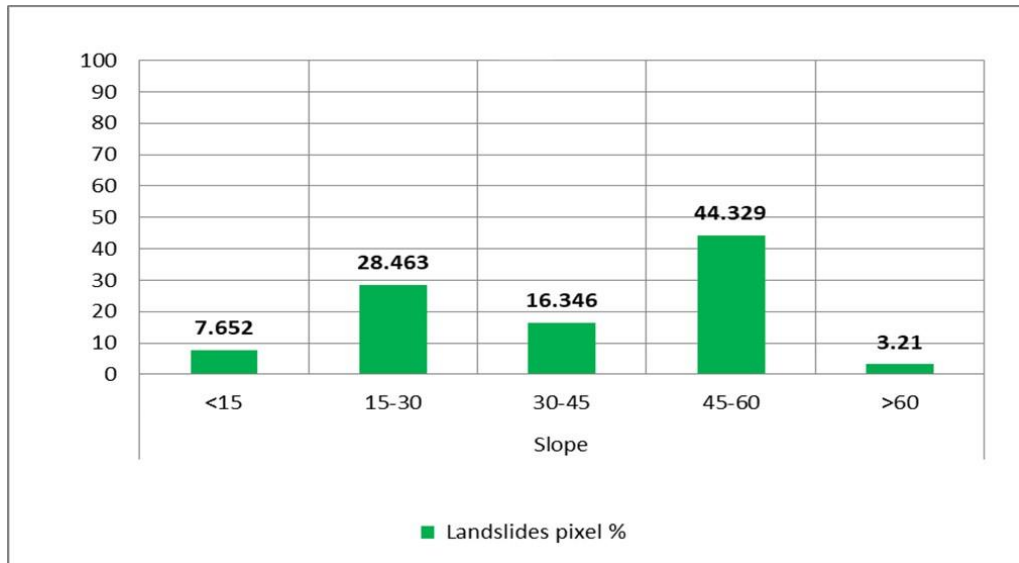


Figure 3.6. Relationship between slope angle and landslide pixel % in each class.

Aspect

The aspect can have both direct and indirect impacts on landslides through various processes. These processes include sunlight exposure, precipitation patterns, vegetation, orientation of discontinuities, evapotranspiration, and concentration of soil moisture, hydrological processes and wind directions (Devkota et al., 2013; Quan and Lee 2012; Neuhcauser et al., 2012). To examine the slope aspect in the region, a Digital Elevation Model (DEM) was utilized to create a classification system with nine categories. The categories included flat (1) West (247.5 – 292.5), Southeast (112.5 – 157.5), Northeast (292.5 – 337.5), South (157.5 – 202.5), Southeast (202.5 – 247.5), Northeast (22.5 – 67.5), North (0 – 22.5 and 337.5 – 360), East (67.5 – 112.5).

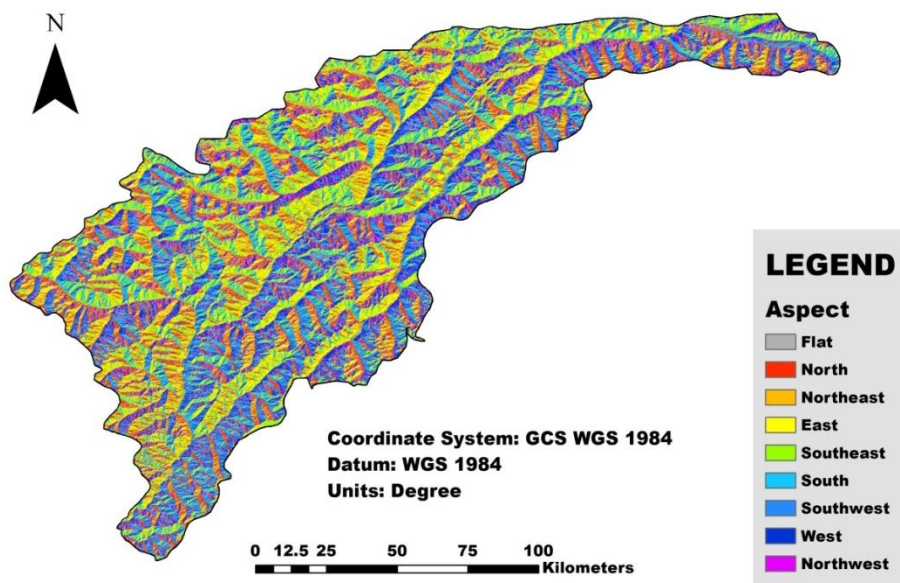


Figure 3.7. Aspect map of the study area divided into classes from 0 to 360 degrees.

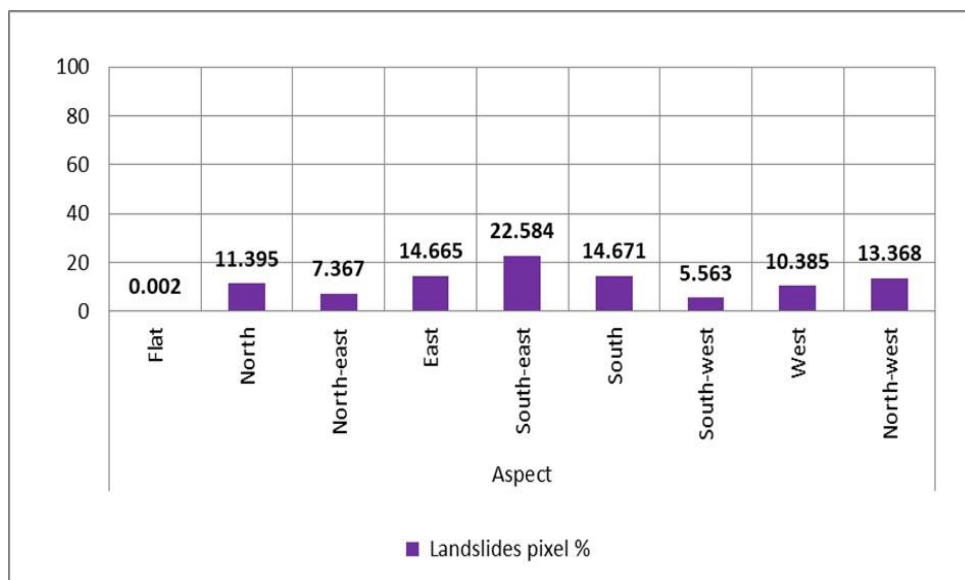


Figure 3.8. Relationship between landslide pixel % in each class of Aspect.

Geological/Lithological Units

Rahmati et al. (2016) found that changes in lithology significantly affect a variety of geo-hazards, including land subsidence and landslide. Different physical and mechanical properties, such as degree of weathering, permeability, durability, strength, type and density are to blame for these variances in lithology as highlighted by (Henriques et al. 2015). Geological map was taken from the Geological Survey of Pakistan at a 1:20,00,000 scale. Furthermore, the map was digitized and lithological units were marked according to the standards of USGS. The results identified 12 lithological units, including Glaciers, Chitral Slate, Karakoram Metamorphic Complex, Kohistan Batholiths, Wakhan Formation, Permian Massive Limestone, Shamran Volcanic Group, Pre-collision Intrusive rocks, Post-collision Granitic rocks, Karakoram Batholiths, Mirkhani Batholiths and Chalt Group.

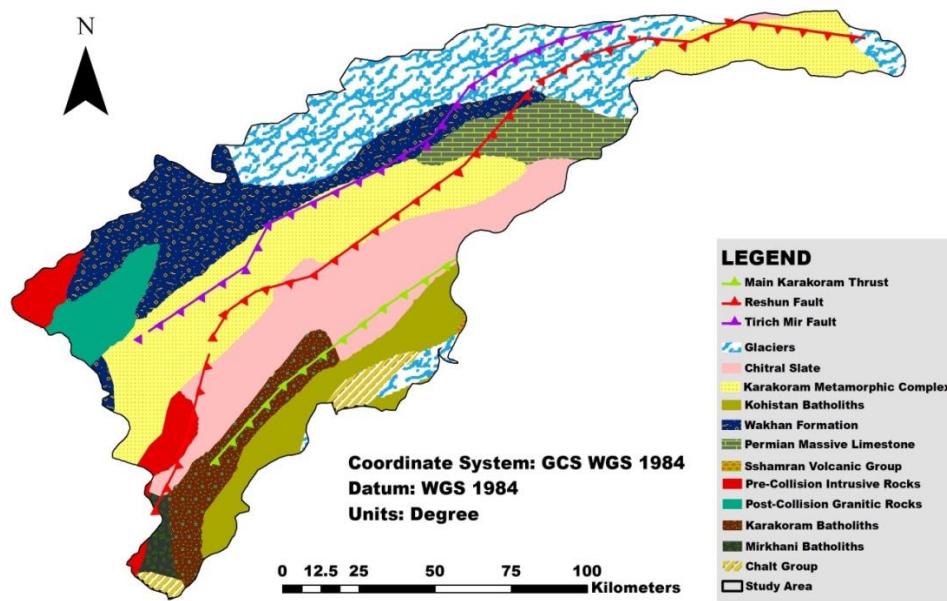


Figure 3.9. Geological Map of Chitral shows different lithologies and the main fault lines in Chitral district.

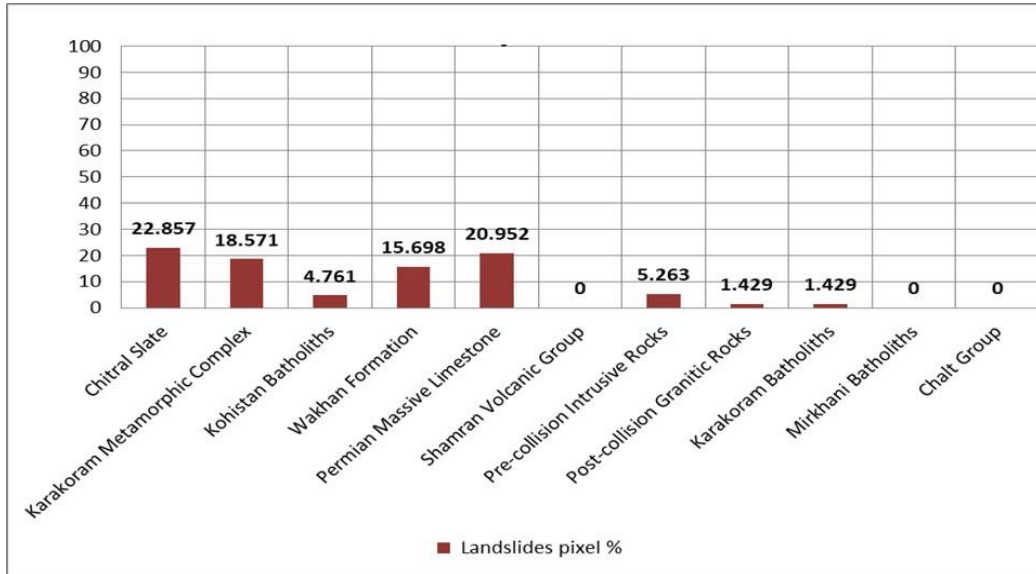


Figure 3.10. Graph showing relationship between the landslide pixel % and the lithologies in the study area.

Soil Type

Soil is considered as a significant factor in landslide susceptibility and occurrence. Different types of soils have varying physical and mechanical properties that can affect the slope's stability. For example, highly weathered and poorly consolidated soils such as clay and silt are more prone to landslides compared to more compacted soils like sand and gravel. Soil data was obtained from website of the Food and Agriculture Organization (FAO) of United Nations. Soil was classified on the basis of type and class. The data was classified into 3 sub-classes namely Sand and Silts (Lithosols), Sands (Haplic Xerosols) and Glaciers.

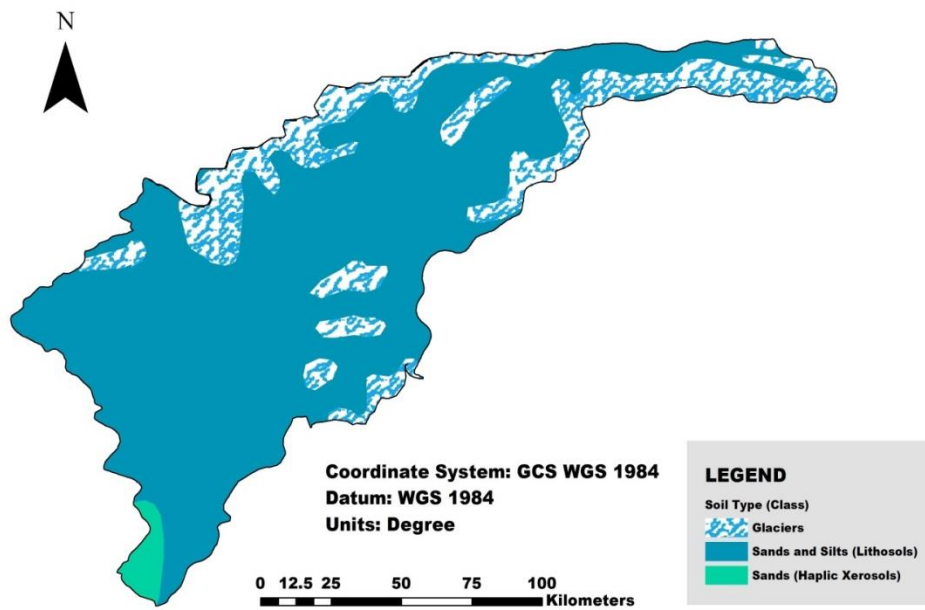


Figure 3.11. The above Map shows different soil types in the study region that includes Sands and silts also known as Lithosols and Sands also known as Haplic Xerosols along with the glacier debris.

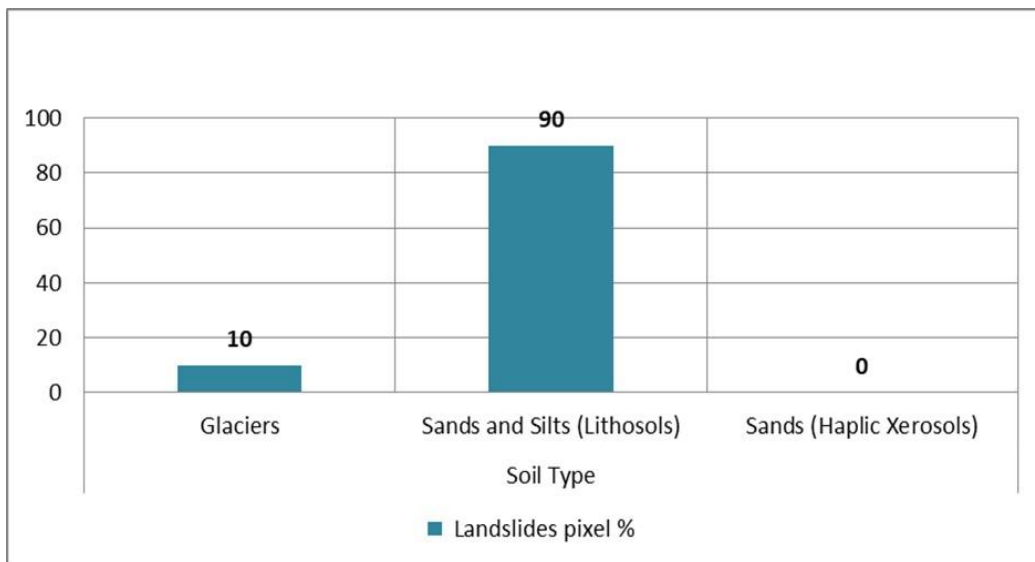


Figure 3.12. The relationship between the soil types and percentage landslide occurrence in each class.

Land Cover

Land cover information is an important input for landslide susceptibility mapping and modeling. Different types of land cover feature like deforestation and man-made structures like roads, buildings, and other infrastructures can also alter the land cover and slope stability, making slopes more vulnerable to landslides. Land cover data was obtained from the Satellite Imagery of the LANDSAT 8 through Google Earth Engine. It was classified into 5 classes on the basis of statistical differences in the spectral characteristics of pixels, otherwise known as Unsupervised Classification Technique. The classes identified were: Built up areas, Soil, Vegetation, Water and Snow/Ice.

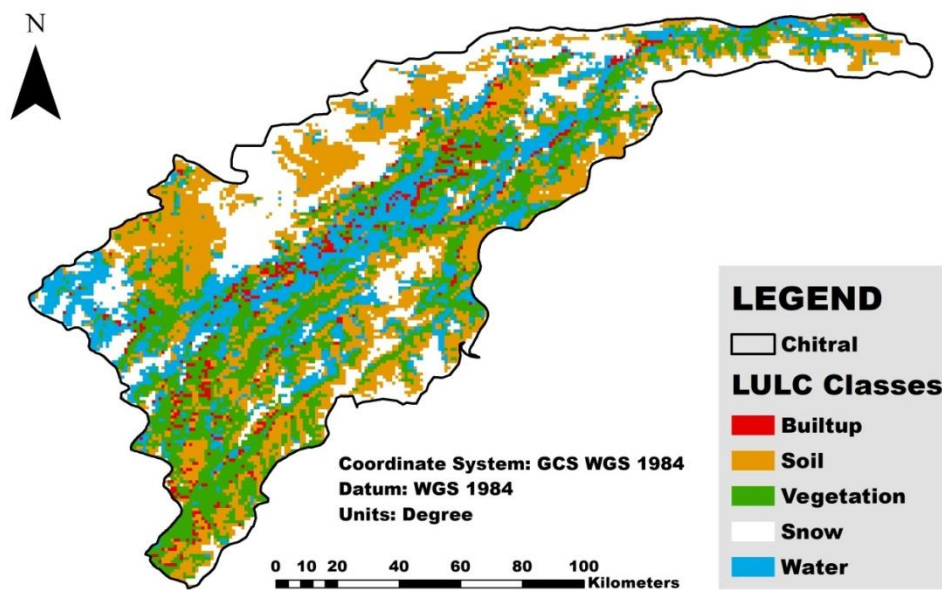


Figure 3.13. LULC map of Chitral showing different land covers in the study area.

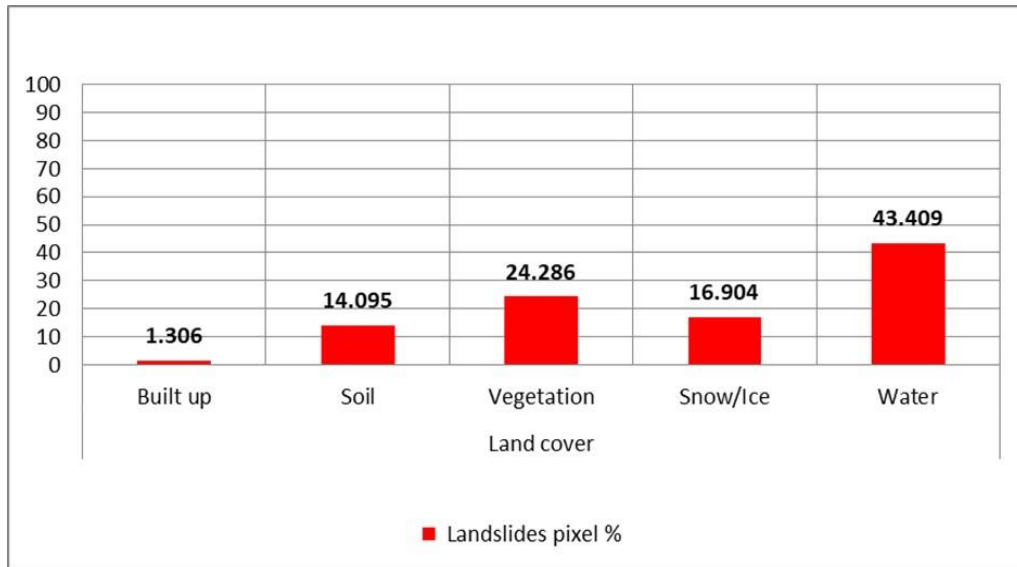


Figure 3.14. Relationship between different land covers and the landslide occurrence in each class.

NDVI/Vegetation Cover

According to Althuwaynee et al. (2012), the NDVI is a significant factor in predicting the likelihood of landslides. To assess vegetation density, the NDVI was utilized in this research. Typically, the NDVI values are in between -1 to +1, with higher values means +ive values indicating dense vegetation coverage. For calculating NDVI values, Landsat satellite imagery was utilized and processed through Google Earth Engine. The values for Normalized Difference Vegetation Index lies between -0.642141 to 0.667333 where negative values representing barren land and positive values indicating vegetated regions. Vegetation cover map was extracted from the NDVI map and 2 classes were assigned as Vegetation and Non Vegetation.

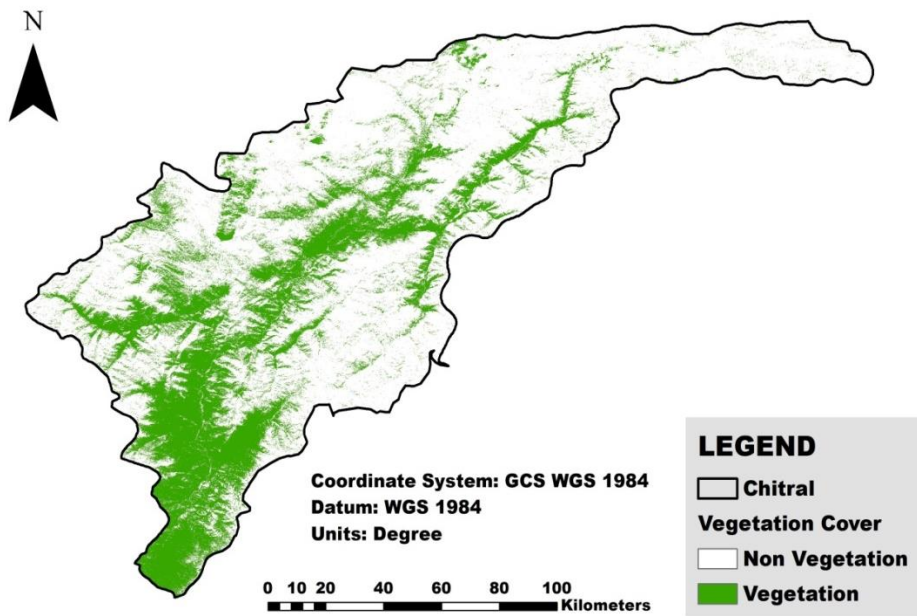


Figure 3.15. Map of the study area showing Vegetation cover calculated through the NDVI values in Chitral district.

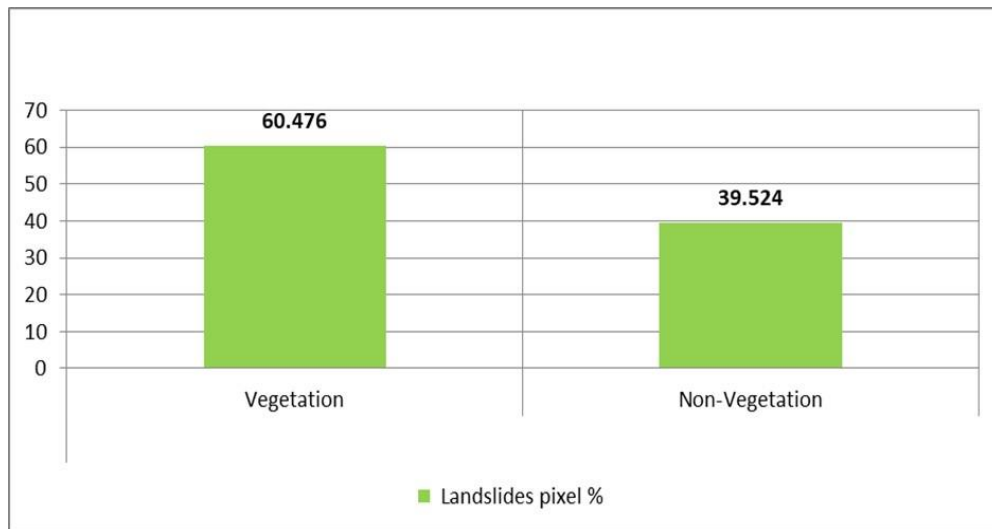


Figure 3.16. Relationship between landslide pixel % in vegetated class and non-vegetated class.

Distance to Rivers/Streams

The present investigation involves utilization of a DEM and Google earth images to identify primary streams in the area. It is worth noting that runoff through these streams is critical in triggering undercutting phenomena, raising pore water pressure in regions near these streams, and ultimately leading to landslide events (Hadji et al., 2013). The presence of streams and rivers serves as a crucial contributing factor in the susceptibility of landslides (Pradhan et al., 2010a, b). The digitized streams resulted into 27 rivers flowing all over the region with Yarkhun and Mastuj rivers being the longest. Buffer Analysis was performed on all the rivers and 5 classes were identified on the basis of their Euclidean distances ranging from (less than 1000 m, 1000 to 2000 m, 2000 to 3000 m, 3000 to 4000 m and greater than 4000 m).

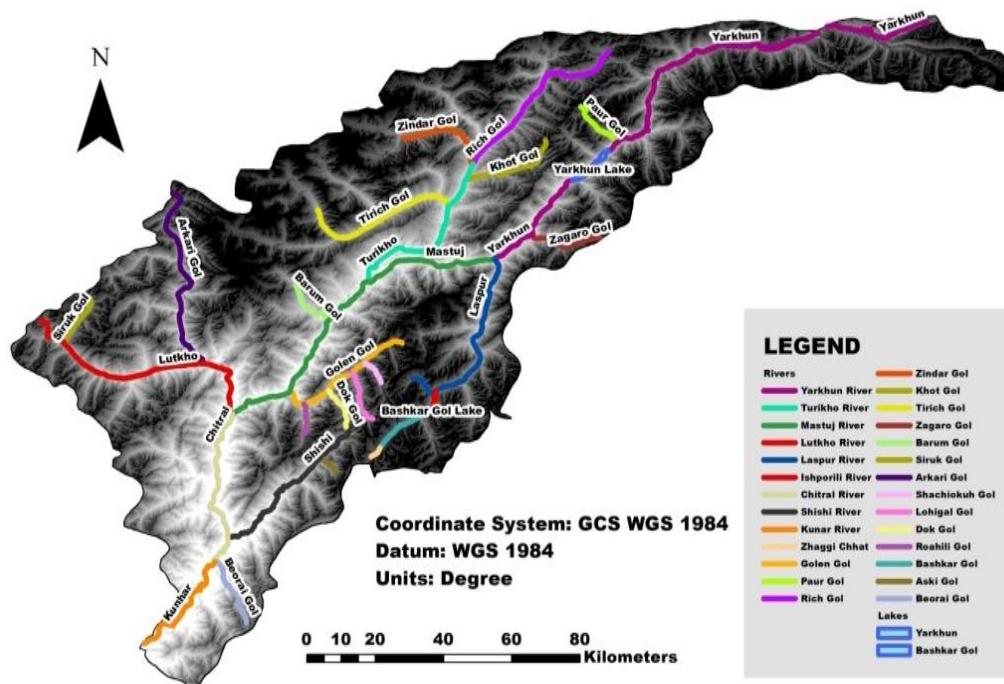


Figure 3.17. Map of the study area showing different rivers and streams flowing in the Chitral district.

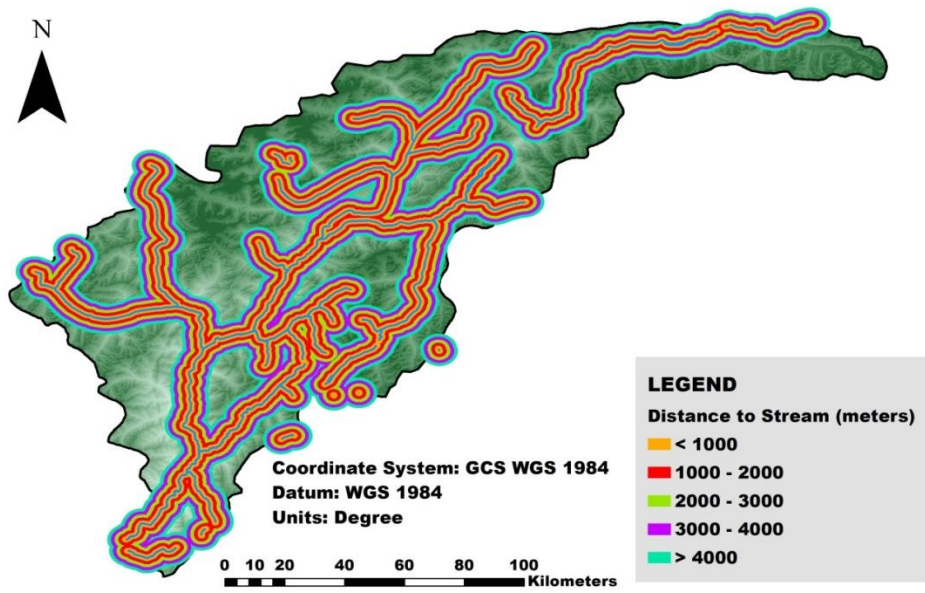


Figure 3.18. Map showing proximity to river with multiple buffer zones.

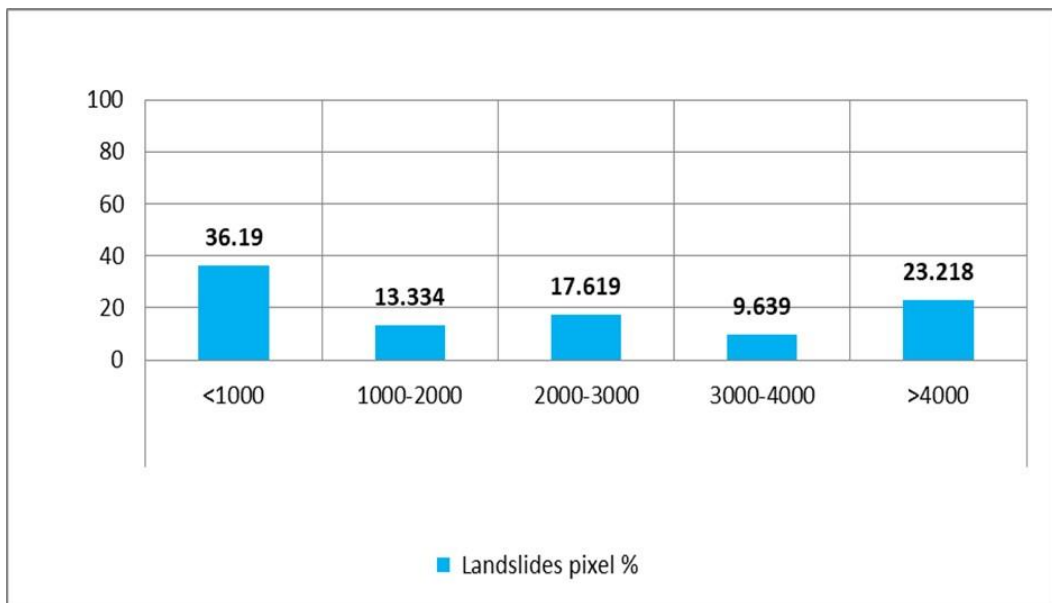


Figure 3.19. Figure showing distance to rivers in meters and the landslide pixels percentage in each class.

Distance to Faults

The existence of structural discontinuities, such as joints, shear zones, fractures, folds, and faults is crucial aspect in the occurrence of landslides and destabilization of rock masses (Bucci et al., 2016; Kanungo et al., 2006; Lee et al., 2002). Thus, the proximity to faults may serve as a useful predictor of landslide susceptibility. Apart from the extraction of geological/lithological units, extraction of fault lines from the geological map dataset. The Chitral district comprises of three main fault lines namely the Main Karakoram Thrust Fault (MKT), Reshun Fault and the Tirich Mir Fault. Buffer Analysis was performed on the three main faults and 5 classes were identified on the basis of their Euclidean distances ranging from (less than 1000 m, 1000 to 2000 m, 2000 to 3000 m, 3000 to 4000 m and greater than 4000 m).

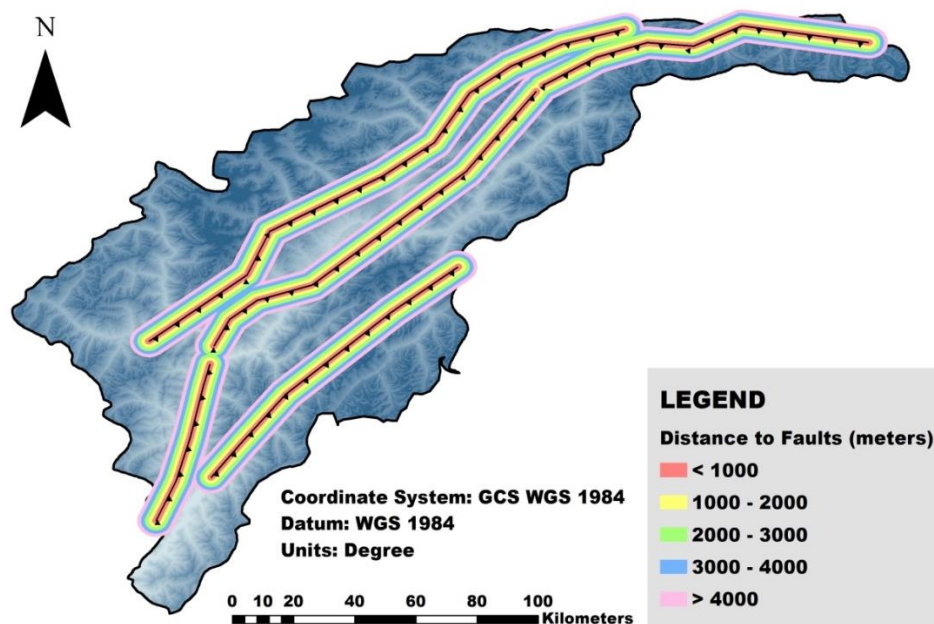


Figure 3.20. Map showing proximity to fault lines with multiple buffer zones.

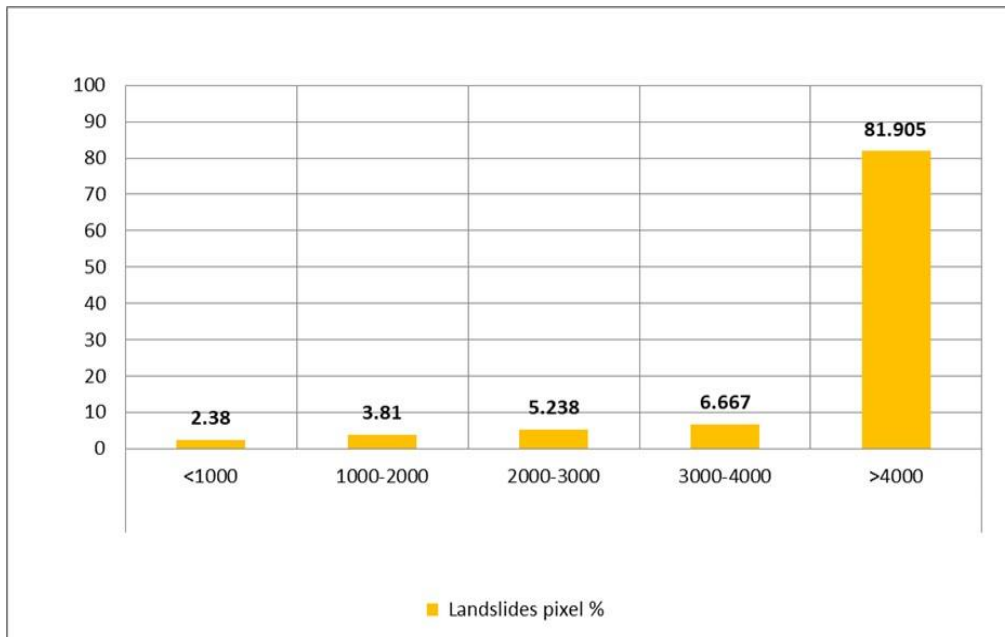


Figure 3.21. Graph showing the relationship between the fault lines and landslide pixel percentage in each buffer class.

Distance to Roads

Engineering tasks excavating slopes during the construction of mountain roads, also known as escarpment roads, have the potential to change the initial geological setup and decrease the natural stability of rock slopes. Therefore, these doings may significantly reduce the likelihood of landslides (Xiao et al., 2019; Wang et al., 2016). Main roads data was obtained from the National Highways Authority (NHA) official website and then the Buffer Analysis was performed on the data and 5 classes were identified on the basis of their Euclidean distances ranging from (less than 1000 m, 1000 to 2000 m, 2000 to 3000 m, 3000 to 4000 m and greater than 4000 m).

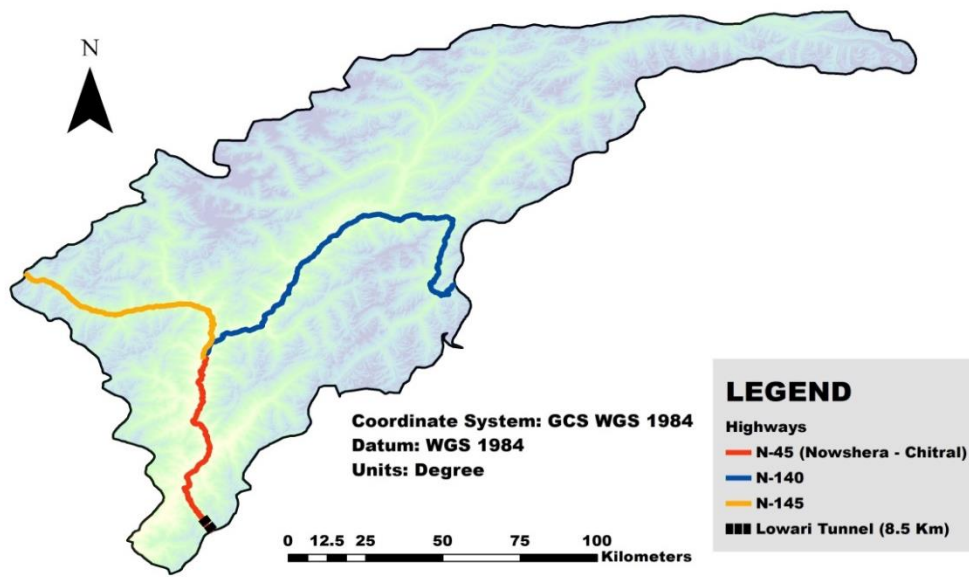


Figure 3.22. Road Network Map of Chitral showing main highways in the study area.

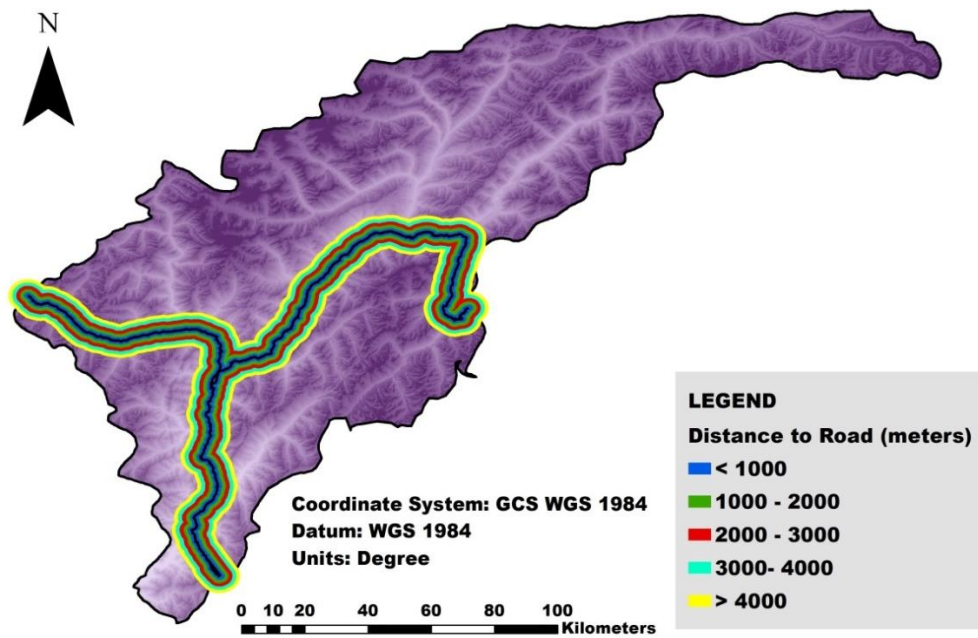


Figure 3.23. Proximity to Road map showing multiple buffer zones.

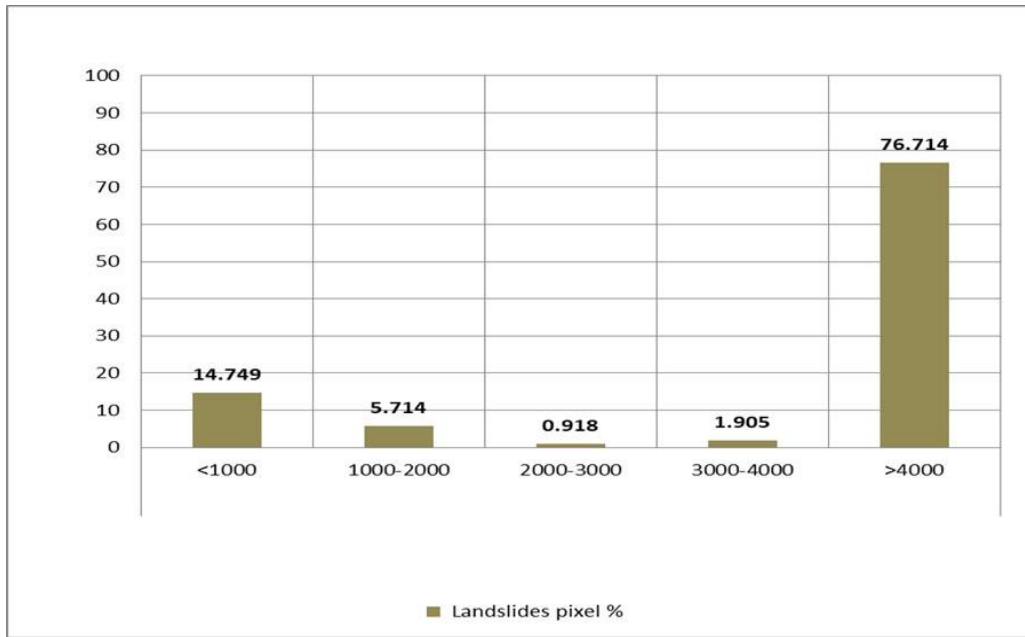


Figure 3.24. Percentage landslide occurrence with respect to distance from roads.

3.4 Point Pattern Analysis and Modelling

By analyzing the spatial distribution of past landslide event, it was possible to identify areas with a high concentration of landslides or areas that are spatially clustered. These areas of high landslide density were considered as high-risk zones. This analysis was also used to identify possible spatial dependence or clustering between different factors that contribute to landslide susceptibility, such as included factors contributing to landslides like elevation profile, slope angle, slope aspect, geology, soil, land cover, vegetation, and proximity to rivers, faults and roads. By analyzing the spatial relationships between these factors and landslide events, different patterns were identified, and cluster map was generated.

Density Based Analysis

Density based analysis was performed in R-Studios where two types of analysis were done. Study region was divided into smaller regions for Quadrat density analysis, are known as Quadrats and point density was calculated for each quadrant. Some quadrats showed zero landslide occurrences whereas the highest number of events occurred in a single quadrat recorded was 66. The resulting map showed the density of historical landslides that occurred in each region. Kernel Density Estimate (KDE) is another method to perform Density based analysis. In this method isotropic kernel intensity estimate of a point pattern was computed using the spatstat package in R-Studios. Kernel window extent was set by assigning the bandwidth of 50km. This kernel defaults to a smoothing function and then it is converted into any of three; quartic, disc or epanechnikov function. By including different covariates in KDE, landslides point pattern intensity was estimated against each covariate or landslide causative factors.

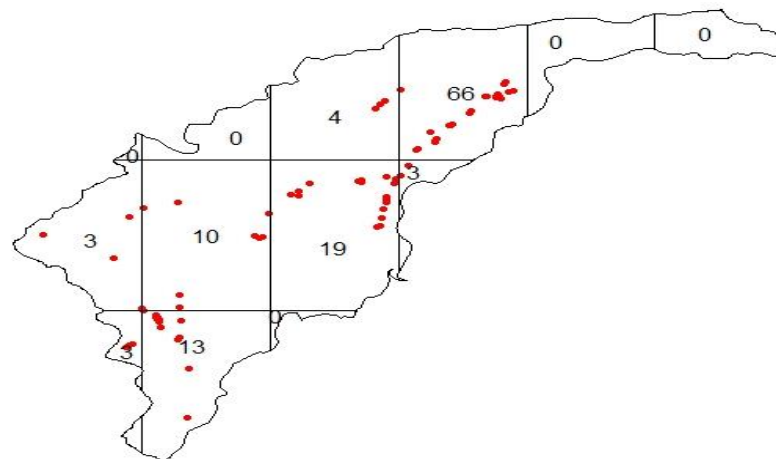


Figure 3.25. Map showing number of landslides in each quadrat.

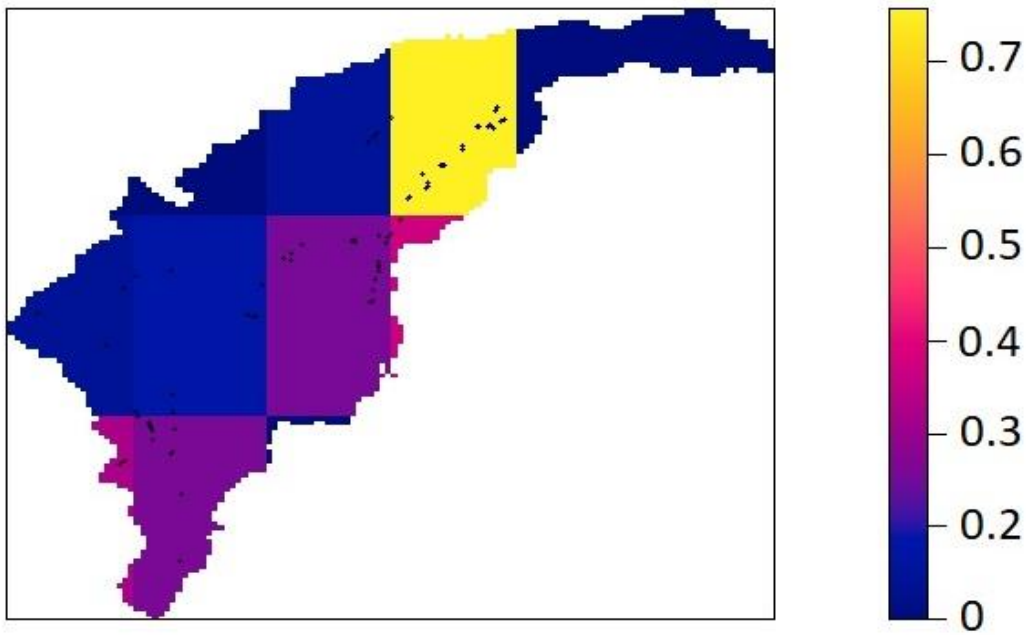


Figure 3.26. Map showing landslides density in each quadrat.

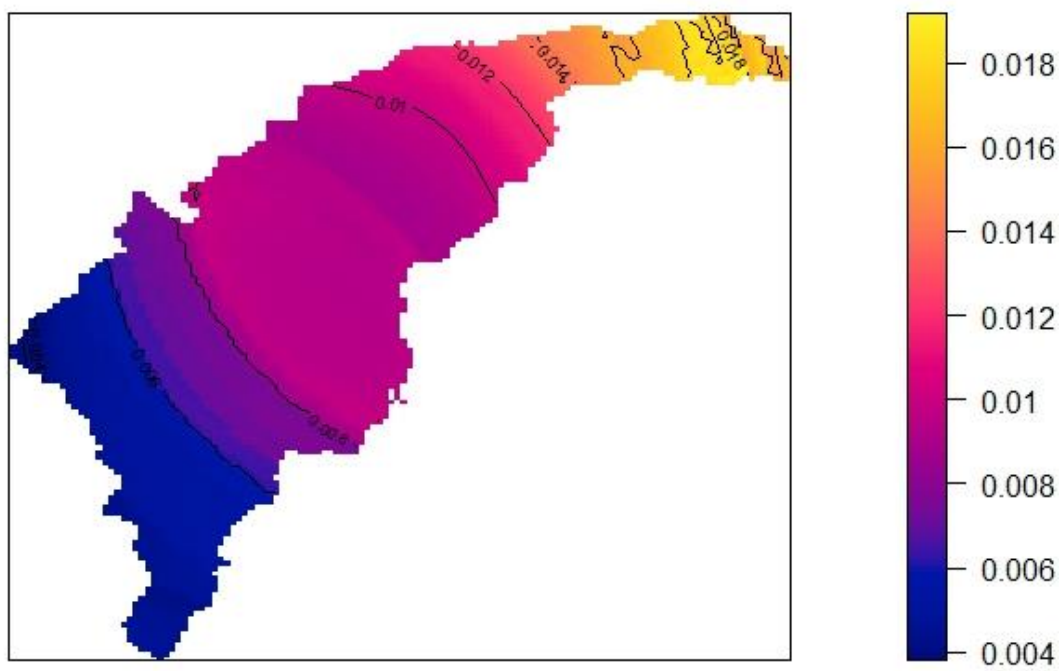


Figure 3.27. Map showing Kernel Density Estimate (KDE) with the bandwidth of 50km disc. function.

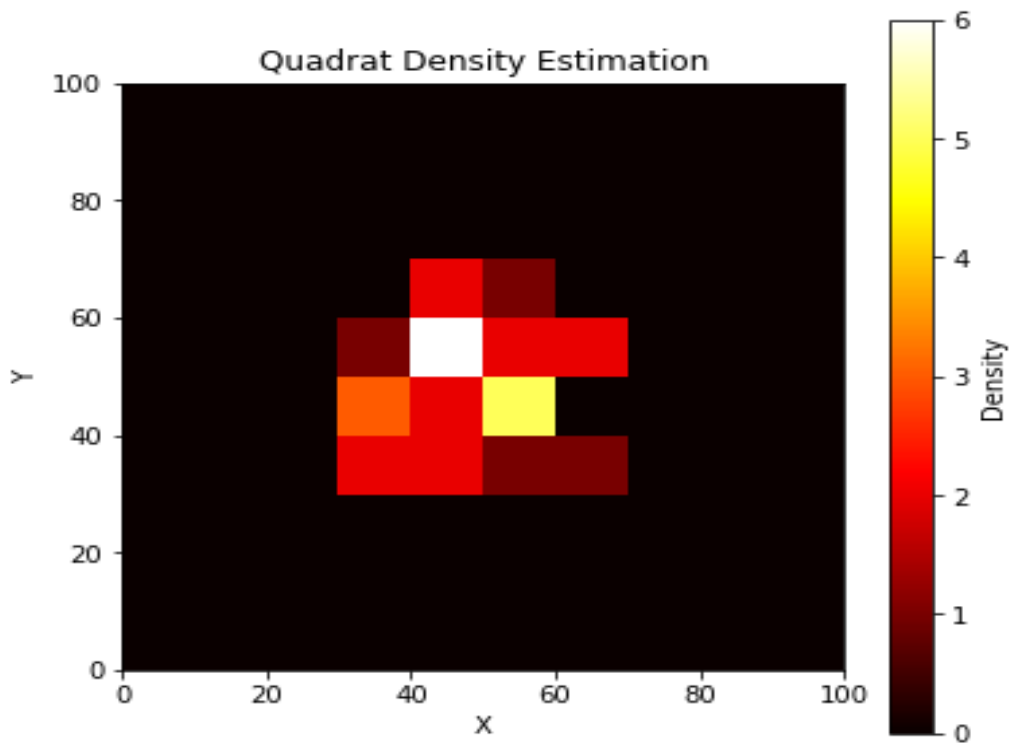


Figure 3.28. Quadrat density estimation for elevation ranges.

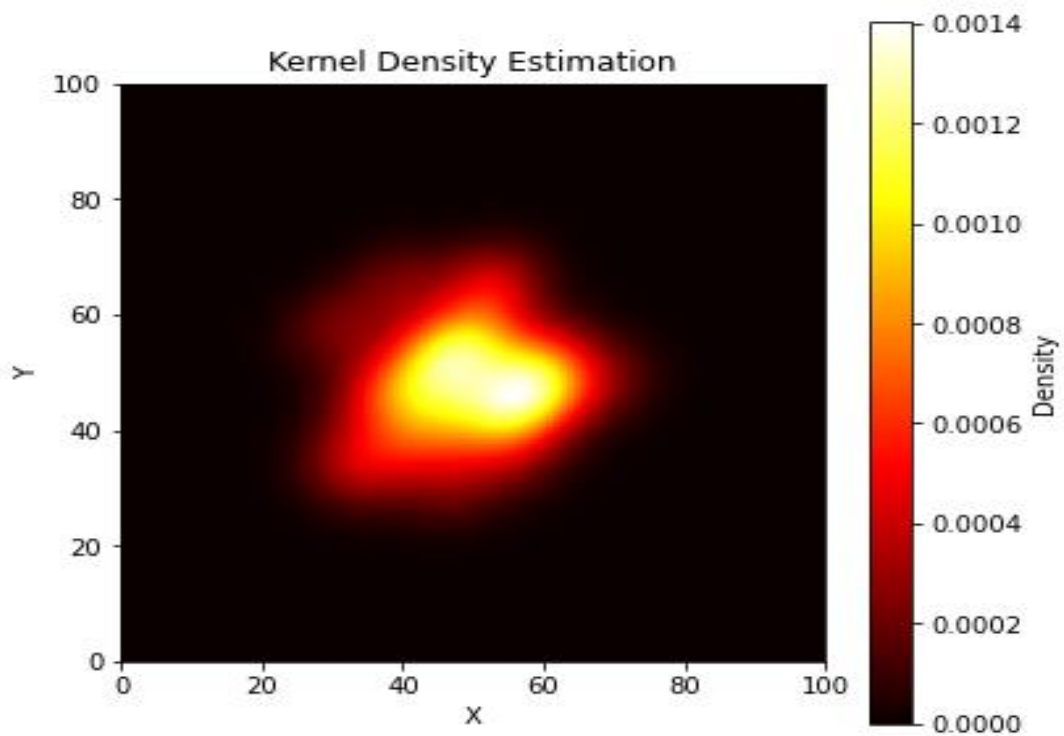


Figure 3.29. Kernel Density Estimate for elevation ranges.

Distance based Analysis

Distance based analysis was performed on the spatial data on the basis of the Euclidean distances of respective covariates and their relationship with the landslides occurrence. Average Nearest Neighbor Analysis (ANN) function was applied to observe the relation between ANN and Neighbor order where first hundred closest neighbors were assessed to find the k^{th} order for prediction of the landslide's occurrences. To determine the significance of a point pattern, such as its level of clustering or dispersion, mean nearest-neighbor distance (NND) of the point collection were compared to the expected NND of points that conform to complete spatial randomness (CSR). A Poisson distribution was formed by the mean nearest-neighbor distance (NND) of points in a complete spatial randomness (CSR) arrangement. The rapid increase of the G-Function at the short distances shows the clustering of the point pattern. K-Function and L-Function were computed by applying the default isotropic, translate and border edge correction techniques. Values of K and K_{expected} were compared at lower distances where more values were greater than 0 which indicates clustering of the point pattern. A cluster map was generated after all the density based and distance based measurements were performed for the data. The map shows the high clustering of landslide points in different parts of the study region. Apart from that, some landslides points are randomly distributed.

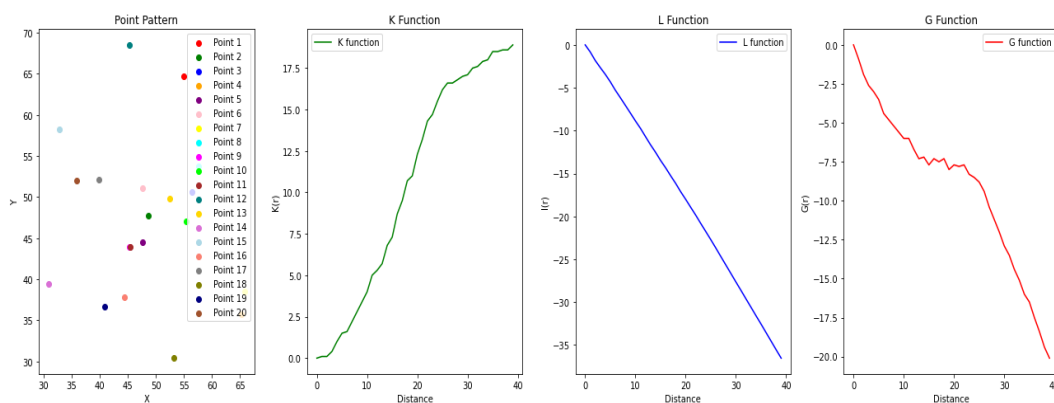


Figure 3.30. Average Nearest Neighbor Analysis (ANN), K, L and G functions on a randomly chosen 20 landslide points to establish a relationship between elevation, slope and aspect parameters.

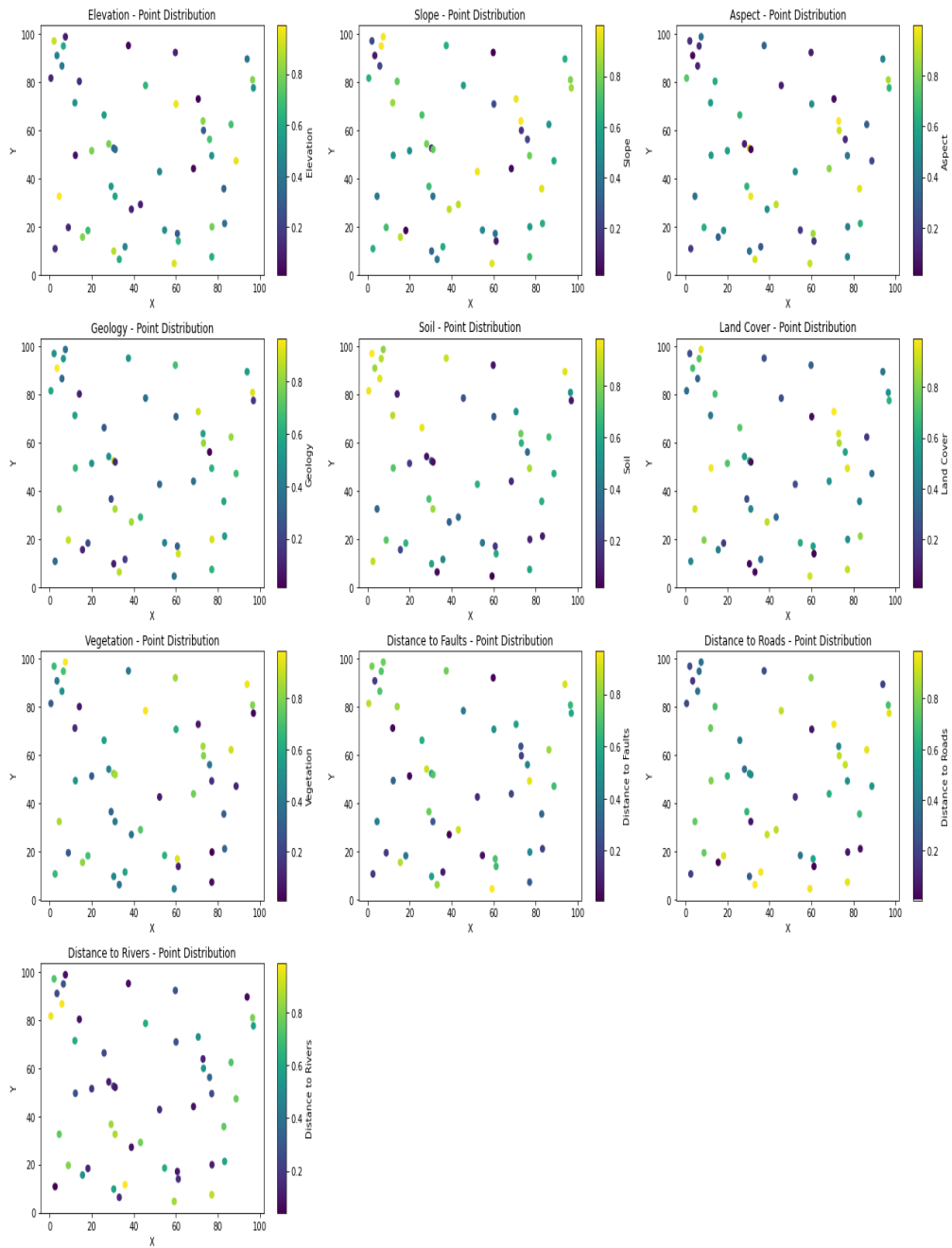


Figure 3.31. Figure shows the landslide point distribution in each causative factor.

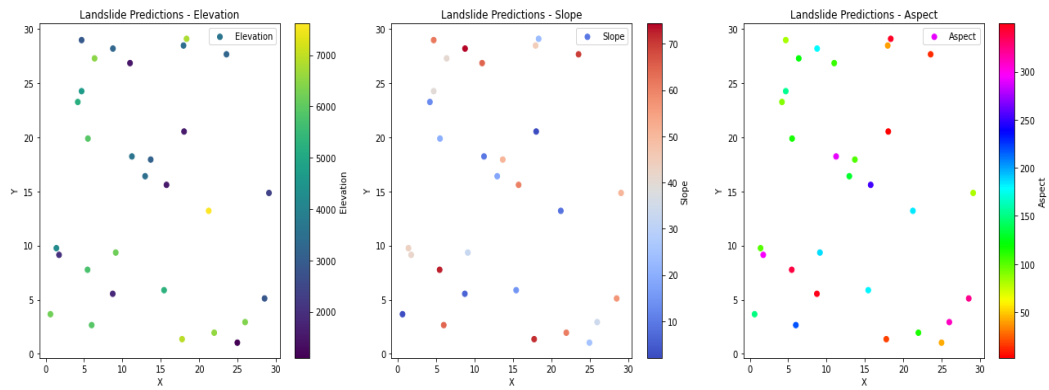


Figure 3.32. Distribution of landslide points in Elevation, Slope and Aspect according to the values of each class.

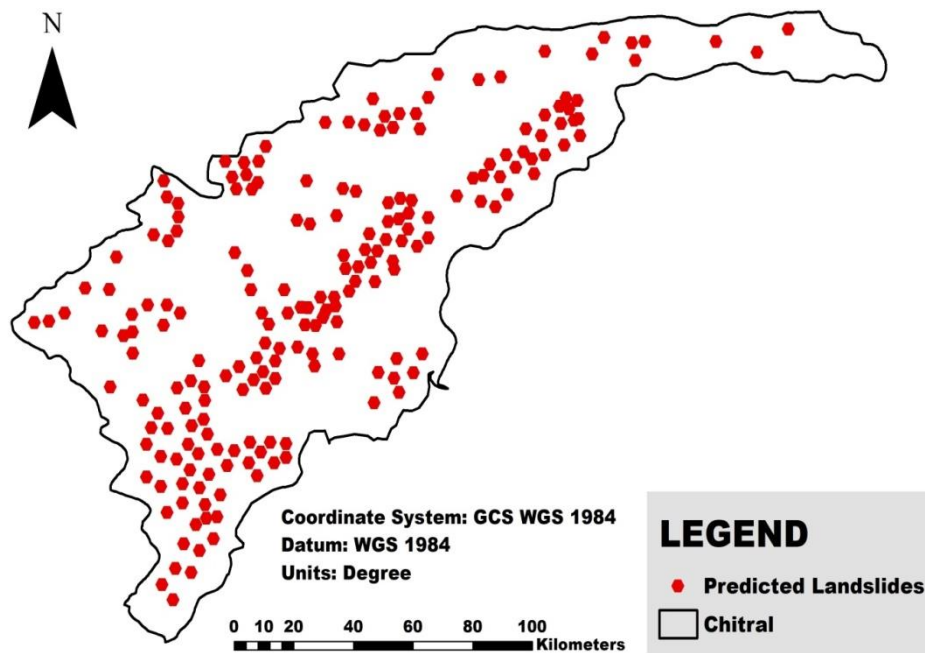


Figure 3.33. Cluster map of the region under study showing clusters of points at different locations indicating high density of landslides in most of the regions.

3.5 Logistic Regression Modelling

Based on the logistic regression analysis conducted on the collected data, several key findings were observed. The coefficients of the predictor variables provided insights into their relationship with the likelihood of landslide occurrences. The slope variable had a positive coefficient of 1.320 in the class greater than 60 degrees, indicating that an increase in slope was associated with higher odds of landslides. The aspect variable, after converting it into useable variables, showed that the “southeast” aspect had a positive coefficient of 0.9824, suggesting an increased likelihood of landslides compared to the reference category (“north”). In terms of elevation, a higher elevation was associated with a decrease in the odds of landslides, as indicated by the negative coefficient. The geology/lithology variable, specifically the “Permian Massive Limestone” class, had a positive coefficient with value greater than 1, indicating a higher likelihood of landslides compared to the reference category (“Chitral Slate”). The distance to rivers/streams variable showed a negative coefficient, indicating that greater distance from streams was associated with lower odds of landslides. Distance to faults, on the other hand, had a positive coefficient, suggesting that proximity to faults increased the likelihood of landslides. The inclusion of vegetation cover as a predictor revealed a negative coefficient with value of -1.7960, indicating that denser vegetation cover was associated with reduced odds of landslides. Land cover variables were also significant, with certain classes showing positive coefficients, suggesting an increased likelihood of landslides compared to the reference category. Soil type had an impact on landslide prediction, with certain soil types exhibiting positive coefficients, indicating a higher likelihood of landslides. Distance to road exhibited a negative coefficient, suggesting that greater distance from roads was associated with lower odds of landslides. These results were statistically significant, as evidenced by the p-values associated with the coefficients. Overall, this model provided insights into the relationships between the predictor variables and the prediction of landslide occurrences, highlighting the importance of land cover,

elevation, slope, lithology, soil, distance to rivers, faults, and road, aspect, and vegetation cover, in landslide susceptibility mapping. Table 3.1 shows the detailed relationship between each group of elements leading to landslides and the regression coefficients derived from respective classes.

Table 3.1. *Landslide causative factors and the calculated Regression coefficients from the LR method.*

Landslide Causative Factors	Classes	Regression Coefficients
Elevation	1065 - 2392	-0.6337
	2392 - 3719	0.9859
	3719 - 5046	-0.1599
	5046 - 6373	0.5476
	6373 - 7701	1.1547
Slope	<15	-0.455
	15-30	-0.513
	30-45	0.6018
	45-60	0.6495
	>60	1.320
Aspect	Flat	-0.1599
	North	0.2487
	North-east	0.5476
	East	0.4859
	South-east	0.9824
	South	-0.585
	South-west	-0.7916
	West	0.6082
Geological/Lithological units	North-west	0.6510
	Glaciers	1.018
	Chitral Slate	2.003
	Karakoram Metamorphic Complex	-0.829
	Kohistan Batholiths	-0.250
	Wakhan Formation	0.141
	Permian Massive Limestone	1.356
	Shamran Volcanic Group	0
	Pre-collision Intrusive Rocks	-0.403
	Post-collision Granitic Rocks	0.761
Karakoram Batholiths	0.410	
Mirkhani Batholiths	0	

Landslide Causative Factors	Classes	Regression Coefficients
	Chalt Group	0
Land cover	Built up	1.550
	Soil	-0.014
	Vegetation	-0.5850
	Snow/Ice	-0.0273
	Water	0.649
Soil Type	Glaciers	0.320
	Sands and Silts (Lithosols)	1.4397
	Sands (Haplic Xerosols)	0
Vegetation Cover	Vegetation	-1.7960
	Non-Vegetation	0.1714
Distance to Fault	<1000	0.6082
	1000-2000	0.4522
	2000-3000	-0.2487
	3000-4000	-0.8023
	>4000	1.3569
Distance to Rivers	<1000	0.6871
	1000-2000	0.8783
	2000-3000	0.5623
	3000-4000	-0.1567
	>4000	-1.6223
Distance to road	<1000	0.5963
	1000-2000	-0.2342
	2000-3000	-0.6435
	3000-4000	-0.7534
	>4000	1.2657

3.6 Frequency Ratio Modelling

To assess the link between factors causing landslides and their occurrence, the Frequency Ratio model were employed. The analysis reveals the impact of slope on landslide distribution, with increasing slope gradient corresponding to an escalation in landslide activities within the area. Specifically, slope angle classes of 15 – 30, 30 – 45, and > 60 demonstrate a positive correlation with landslide occurrences, as indicated by their FR values of 1.17, 1.343, and 2.149, respectively. The aspect direction model reveals that the Southeast, East, South and Northwest classes exhibit Frequency Ratio values >1, indicating a higher

correlation with landslide events, while the Flat class shows the lowest association. Analysis of elevation reveals a normal correlation, which shows neither landslides increase with higher elevation nor decreases with lower elevations. Notably, the elevation class of 2392 – 3719 meters exhibits the strongest correlation, with a value of 2.259. Geologically, the Permian Massive Limestone formation displays the highest number of 2.831, while some other formations also exhibit a positive correlation with landslide inventory. Landslide occurrences increase significantly with increasing distance to faults. Streams have a significant influence on landslide susceptibility, with buffer classes of < 1000 m and 1000 – 2000 positively correlated with landslide occurrences, having values of 2.898 and 1.034, respectively. In terms of land cover, the water class shows the highest landslide occurrence (43%) with an FR ratio of 2.097, indicating a strong correlation. High vegetation cover class, account for approximately 60% of landslides. Along roads, a significant number of landslides were detected, with the FR model demonstrating a decreasing frequency ratio as the distance to the road increases. The buffer class of < 1000 m exhibits the highest FR value of 1.921, while the highest distance category (> 4000) has a value of 1.052. Table 3.2 shows the detailed information about the landslide pixels and percentages with respect to each class of the landslide causative factors.

Table 3.2. *Intended weights for each landslide causative factor and its classes for FR method.*

Landslide Causative Factors	Classes	Pixels in each class	% Pixel in each class	Landslides pixels	% Landslides pixel	Frequency Ratio (FR)
Elevation	1065 - 2392	309834 9	18.88 2	73423	20.476	1.084
	2392 - 3719	366724 9	22.34 9	180998	50.476	2.259
	3719 - 5046	630532 5	38.42 6	56348	15.714	0.409

Landslide Causative Factors	Classes	Pixels in each class	% Pixel in each class	Landslides pixels	% Landslides pixel	Frequency Ratio (FR)
	5046 - 6373	2498599	15.227	46103	12.857	0.844
	6373 - 7701	1167337	7.114	1710	0.477	0.067
Slope	<15	2883391	17.572	27439	7.652	0.435
	15-30	3992148	24.329	102064	28.463	1.17
	30-45	1997797	12.175	58614	16.346	1.343
	45-60	7290523	44.430	158957	44.329	0.998
	>60	245151	1.494	11511	3.21	2.149
Aspect	Flat	4266	0.026	8	0.002	0.077
	North	2147447	13.087	40861	11.395	0.871
	North-east	1904758	11.608	26417	7.367	0.635
	East	1835676	11.187	52586	14.665	1.311
	South-east	2252793	13.729	80983	22.584	1.645
	South	2163035	13.182	52608	14.671	1.113
	South-west	2069668	12.613	19948	5.563	0.441
	West	1981552	12.076	37239	10.385	0.86
	North-west	2049813	12.492	47936	13.368	1.07
Geological/Lithological units	Glaciers	2663674	16.233	32416	9.04	0.557
	Chitral Slate	1698660	10.352	81962	22.857	2.208
	Karakoram Metamorphic Complex	3179081	19.374	66593	18.571	0.959

Landslide Causative Factors	Classes	Pixels in each class	% Pixel in each class	Landslides pixels	% Landslides pixel	Frequency Ratio (FR)
	Kohistan Batholiths	1446618	8.816	17072	4.761	0.54
	Wakhan Formation	2358467	14.373	56291	15.698	1.092
	Permian Massive Limestone	1214267	7.40	75131	20.952	2.831
	Shamran Volcanic Group	46109	0.281	0	0	0
	Pre-collision Intrusive Rocks	635684	3.874	18872	5.263	1.359
	Post-collision Granitic Rocks	689835	4.204	5125	1.429	0.34
	Karakoram Batholiths	1142395	6.962	5125	1.429	0.205
	Mirkhani Batholiths	466508	2.843	0	0	0
	Chalt Group	867798	5.288	0	0	0
Land cover	Built up	776310	4.731	4683	1.306	0.276
	Soil	4018073	24.487	50542	14.095	0.576
	Vegetation	3565021	21.726	87086	24.286	1.118
	Snow/Ice	4653430	28.359	60615	16.904	0.596
	Water	3396172	20.697	155658	43.409	2.097
Soil Type	Glaciers	3056998	18.630	35858	10.00	0.537
	Sands and Silts (Lithosols)	12812809	78.084	322726	90.00	1.153

Landslide Causative Factors	Classes	Pixels in each class	% Pixel in each class	Landslides pixels	% Landslides pixel	Frequency Ratio (FR)
	Sands (Haplic Xerosols)	627809	3.826	0	0	0
Vegetation Cover	Vegetation	7014030	42.745	216857	60.476	1.415
	Non-Vegetation	9394977	57.255	141727	39.524	0.69
Distance to Fault	<1000	1235675	9.422	8534	2.380	0.253
	1000-2000	1207888	9.571	13662	3.810	0.398
	2000-3000	1134555	8.720	18783	5.238	0.601
	3000-4000	1095788	8.395	23907	6.667	0.794
	>4000	1875665	63.892	293698	81.905	1.282
Distance to Rivers	<1000	2048993	12.487	129772	36.190	2.898
	1000-2000	2115449	12.892	47814	13.334	1.034
	2000-3000	3243240	19.765	63179	17.619	0.891
	3000-4000	3055521	18.621	34564	9.639	0.518
	>4000	5945804	36.235	83256	23.218	0.641
Distance to road	<1000	1259884	7.678	52888	14.749	1.921
	1000-2000	1070852	6.526	20489	5.714	0.876
	2000-3000	1024414	6.243	3292	0.918	0.147
	3000-4000	1092348	6.657	6831	1.905	0.286
	>4000	11961509	72.896	275084	76.714	1.052

3.7 Landslide Susceptibility Mapping (LSMs)

3.7.1 LSM by Logistic Regression Model (LSM – LR)

Assessment of the terrain's potential for landslides is crucial for assuring sustainable development and reducing the likelihood of landslide-related disasters. The analysis's causative factors included elevation, slope, aspect, lithological units, land cover, NDVI, proximity to faults, streams and rivers, and road network. The study found that 18.1% of the district's land area is particularly vulnerable to landslides. Landslides are highly vulnerable in 21.3% of the region, moderately susceptible in 25.5% of the territory, and not at all susceptible in 35.1% of the area. The observed landslide susceptibility map was compared in order to verify the accuracy of the susceptibility map. The verification findings revealed a satisfactory level of agreement between the landslides location data currently available and the susceptibility map.

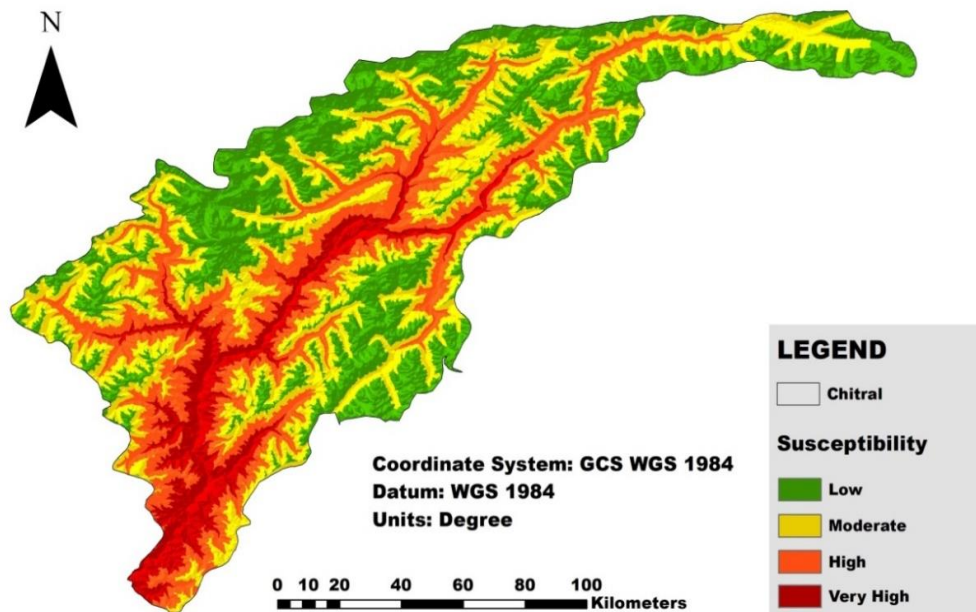


Figure 3.34. Landslide Susceptibility Map generated by LR Method showing areas from Low to Very High Susceptible.

3.7.2 LSM by Frequency Ratio Model (LSM – FR)

The map shows the results derived from frequency ratio model that 26.1% of the district's land area is particularly vulnerable to landslides. Landslides are highly vulnerable in 23.5% of the region, moderately susceptible in 31.5% of the territory, and not at all or low susceptible in 18.9% of the area. A robust chronological association between landslides and the validation set for the Frequency ratio (FR) model was produced. The overlay analysis of the training and validation set of Frequency ratio derived maps, which depicts a gradual rise from the lowest to extremely high susceptibility classes, reveals a satisfactory degree of match. Similar results are seen when comparing the training and validation sets of the FR-derived susceptibility map for all susceptibility categories.

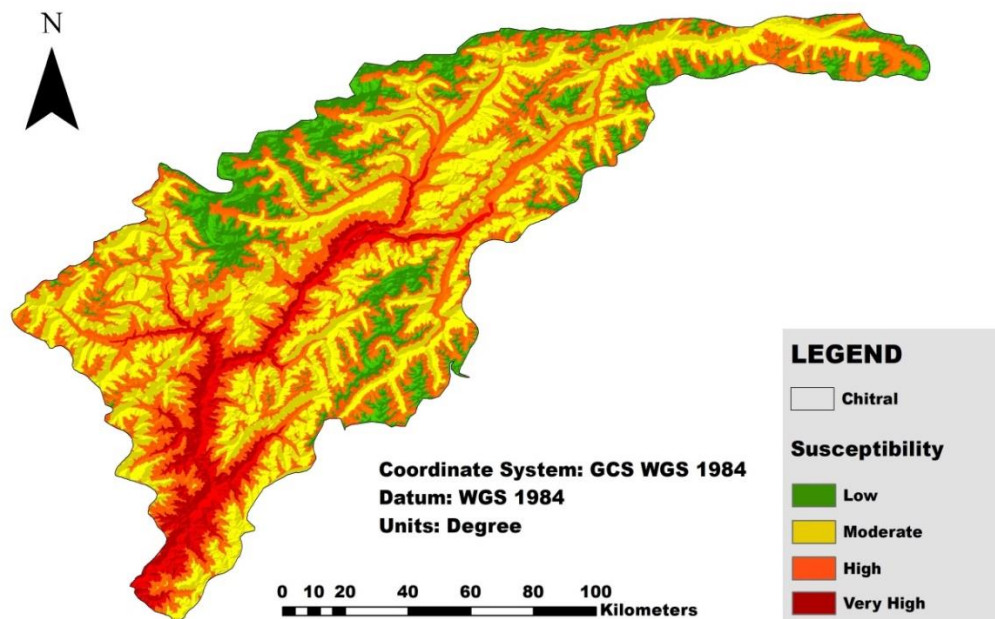


Figure 3.35. Landslide Susceptibility Map generated by FR Method showing areas from Low to Very High Susceptible.

3.7.3 AUC – ROC Analysis

To evaluate the performance of the models, Receiver Operating Characteristic (ROC) curve and Area under Curve (AUC) analysis was applied. The AUC analysis was performed on the data by using Python codes. Data was generated for the Logistic Regression (LR) model coupled with for the Frequency Ratio (FR) model. The LR model showed a higher AUC of 0.8534, suggesting better discrimination compared to the FR model with an AUC of 0.7856. This implies that the Logistic Regression model, on average, has a greater probability of assigning higher scores to positive instances and lower scores to negative instances, making it more effective in distinguishing between the two classes. The specificity vs sensitivity curve showcases the relationship between sensitivity (successful positive recognition) and specificity (successful negative recognition) for different classification thresholds. Y-axis & x-axis represents specificity and sensitivity respectively. By displaying the values in percentage on the specificity vs sensitivity graph, we can interpret the curve more intuitively. The sensitivity values range from 0% to 100%, representing the correct identification percentage for true positives. Similarly, the specificity values also range from 0% to 100%, representing the correct identification percentage for true negative values. By evaluating the AUC values and analyzing the ROC curve, the effectiveness of the LR and FR models in classifying positive and negative instances was assessed. The higher AUC value of LR (85.34 %) suggests that it has better discriminatory power, while the lower AUC value of FR (78.56 %) indicates comparatively lower performance. These values provide valuable insights into the models' abilities to distinguish between classes and aid in the validation and selection of the most suitable model for the research.

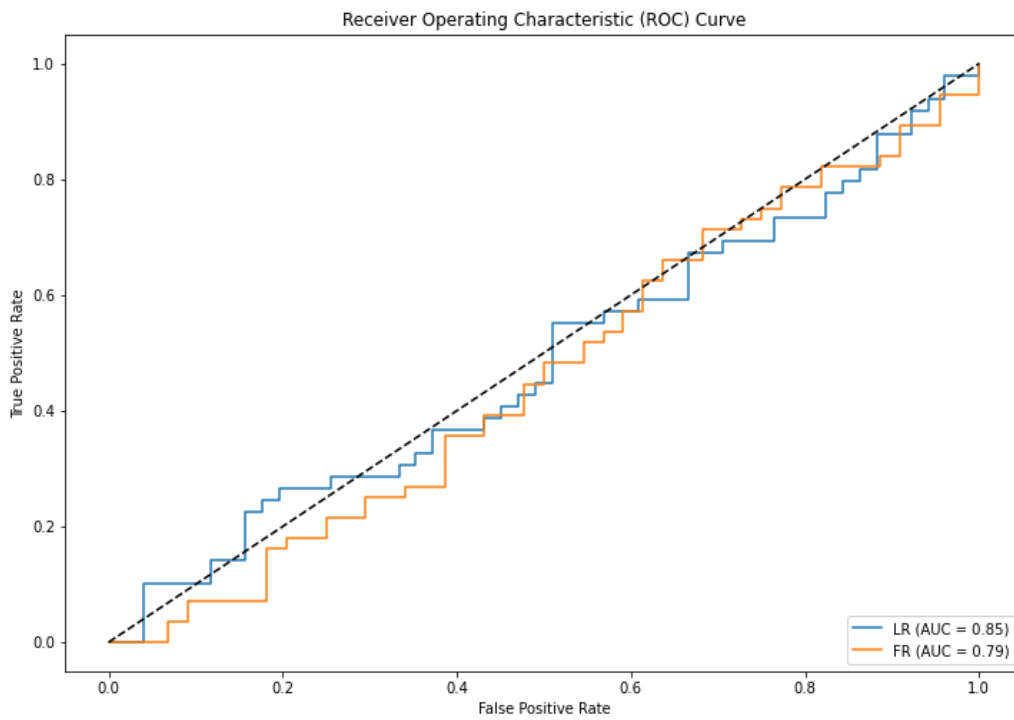


Figure 3.36. ROC Curve of LR and FR Model showing that LR model has a value of 0.85 and FR model shows value of 0.79 respectively.

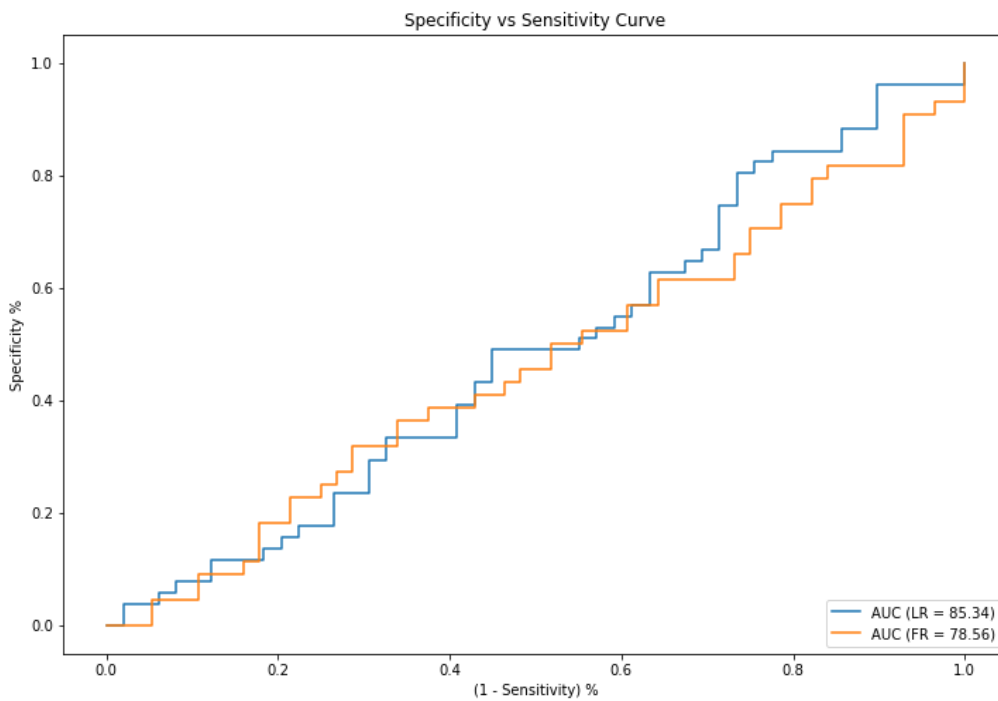


Figure 3.37. Sensitivity vs Specificity Analysis of LR and FR Models.

3.8 Discussions

This research aimed to assess and compared the effectiveness of three widely used methods, namely Logistic Regression (LR), Frequency Ratio (FR), and Point Pattern Analysis (PPA), for landslide susceptibility mapping considering various causative factors. These factors included elevation, slope, aspect, lithology/geology, land cover, soil type, vegetation cover, and distance to faults, rivers, and roads.

The application of Logistic Regression revealed significant relationships between landslide occurrences and the considered causative factors. The model highlighted the importance of terrain characteristics, such as slope and elevation, as key contributors to landslide susceptibility. In addition, lithology/geology, soil type, and land cover were found to be influential factors. The analysis further indicated that proximity to faults, rivers, and roads had a major impact on landslide occurrence. The resulting landslide susceptibility map generated by Logistic Regression exhibited distinct zones of susceptibility, with high-risk areas characterized by steep slopes, certain lithologies, and proximity to roads and faults.

The Frequency Ratio method focused on analyzing the frequency of landslide events within different terrain units and their association with the causative factors. The resulting landslide susceptibility map revealed areas with high occurrence rates, emphasizing the significance of terrain characteristics in landslide susceptibility. Steep slopes, specific lithologies, and proximity to rivers and faults were found to be closely related to increased landslide susceptibility. The map generated by Frequency Ratio exhibited different levels of susceptibility, ranging from low to high, enabling the identification of areas at the risk of landslides according to past event occurrences and their associated causative factors.

Furthermore, the Point Pattern Analysis method examined the spatial clustering of landslide events and its relation to the considered causative factors. The analysis identified significant spatial clusters with high landslide densities,

indicating areas of increased susceptibility. Steep slopes, specific lithologies and proximity to rivers and faults were found to be strongly associated with the observed clustering patterns.

The combined results of Logistic Regression, Frequency Ratio, and Point Pattern Analysis provide a comprehensive understanding of susceptibility of landslides in the region. These methods consider various causative factors, including terrain characteristics, lithology/geology, soil type, land cover, and proximity to natural features and infrastructure. The generated landslide susceptibility maps integrate these factors, enabling the identification of different susceptibility zones. The maps provide valuable information for land management, spatial planning, and decision-making processes aimed at mitigating landslide hazards. By considering multiple methods and causative factors, a more accurate and reliable assessment of landslide susceptibility can be achieved, supporting effective risk reduction strategies and promoting safer land use practices.

CONCLUSIONS AND RECOMMENDATIONS

4.1 Conclusion

Each model offers advantages of its own, but the frequency ratio (FR) model is particularly simple to use and makes it simple to interpret the findings of earlier research (Lee and Talib et, al 2005 and Lee and Pradhan et al, 2007). Frequency ratio (FR) model is capable of being used as a general method to evaluate a landslide's susceptibility. A popular statistical technique for mapping landslip vulnerability is the linear regression (LR) model. Inferring the binary output from a set of input variables using the model is helpful. The benefit of the linear regression (LR) model is that by supplementing the standard linear regression model with an appropriate link function, the variables may take either a discrete or continuous shape and are not required to follow normal distributions.

It was determined that the logistic regression model effectively illustrates the connection between the likelihood of landslide & instability components. The regression coefficient can be used to express the relative weights of independent variables. It was discovered that, compared to other parameters, land use had the strongest correlation with slope collapse occurrences. Particularly, characteristics like "distance from river" and "distance to roads" exhibit zero weight; this is possible given the study's minimal impact on the frequency of landslides. The majority of the observed landslide locations were found to be in the two susceptibility classes of high and very high, according to the validation findings from the landslides susceptibility study. As a result, it may be concluded that the map and actual field conditions (collected data) coincide well. To the chosen areas, the model appears to be trustworthy.

The Point Pattern Analysis approach identified spatial clusters of landslide events highlighting areas with concentrated landslide densities. The analysis further confirmed the influence of terrain characteristics, lithology/geology, and proximity to natural features on the occurrence of

landslides. The identification of these clusters enables targeted interventions in high-risk zones.

The research area's north-western and south-eastern regions tend to have the highest landslides occurrence, which is related to elevations between 1065m and 3719m. The majority of the landslides in the region are connected to slope angles between 30° and 60°. The south-east slope aspect exhibits the greatest density of landslides.

In conclusion, the integration of Logistic Regression, Frequency Ratio, and Point Pattern Analysis provides a comprehensive understanding of landslide susceptibility and its underlying causative factors. The study underscores the importance of considering multiple methods and variables to accurately assess and map landslide susceptibility. The findings of this study contribute to improved landslide risk management, facilitating informed decision-making, and promoting safer land use practices in areas prone to landslides. Further research and continuous monitoring are recommended to improve the precision and efficiency of landslide susceptibility mapping and to support proactive measures for landslide risk reduction.

4.2 Recommendations

The landslide susceptibility map's precision and predictability could provide us with vital knowledge for future city planning, infrastructure building, and agricultural activities in our area or in other areas with comparable characteristics. The landslides susceptibility map can be used for a variety of things, including regional planning, hazard mitigation strategies, and decision-making on slope repair measures. On the other hand, by adding more variables, the susceptibility map's quality can be raised even further. Additionally, any alteration of the natural environment brought about by human intervention, such as the deployment of development projects or deforestation, may alter the region's current susceptibility to landslides. Therefore, such maps need to be updated on a regular basis.

Risk analysis can be done if information about the susceptibility of people, buildings, villages, and other property is available. By combining the results with information about time, a landslides hazard zonation map can then be created. Continuous monitoring and updating of the susceptibility maps is essential to ensure accuracy and relevance over time. Incorporating new data and monitoring changes in the landscape will enable the identification of evolving susceptibility patterns and facilitate adaptive management approaches. To effectively manage landslide risks, an integrated risk management framework is recommended. This framework should incorporate the findings from landslide susceptibility mapping and integrate early warning systems, land use planning regulations, and targeted mitigation measures.

Collaboration among stakeholders, including government agencies, researchers, and local communities, is essential for the successful implementation of risk reduction strategies. By implementing these recommendations, the field of landslide susceptibility mapping can advance in its ability to accurately assess and manage landslide risks. Continued research, data improvement, and collaboration among various stakeholders will contribute to more effective landslide risk reduction strategies and the protection of vulnerable communities and infrastructure.

REFERENCES

1. Ali, S., Biermanns, P., Haider, R., & Reicherter, K. (2019). Landslide susceptibility mapping by using a geographic information system (GIS) along the China–Pakistan Economic Corridor (Karakoram Highway), Pakistan. *Natural Hazards and Earth System Sciences*, *19*(5), 999-1022. doi:10.5194/nhess-19-999-2019
2. Anusha, N., & Bharathi, B. (2020). Flood detection and flood mapping using multi-temporal synthetic aperture radar and optical data. *The Egyptian Journal of Remote Sensing and Space Science*, *23*(2), 207-219. doi:10.1016/j.ejrs.2019.01.001
3. Anderson, S., & Marcus, L. F. (1993). Effect of Quadrat Size on Measurements of Species Density. *Journal of Biogeography*, *20*(4), 421-428.
4. Azarafza, M., Azarafza, M., Akgun, H., Atkinson, P. M., & Derakhshani, R. (2021). Deep learning-based landslide susceptibility mapping. *Sci Rep*, *11*(1), 24112. doi:10.1038/s41598-021-03585-1
5. Bacha, A. S. (2019). Semi-Automated Landslide Detection and Susceptibility Modelling Using Geospatial Tools in Northern Pakistan.
6. Borchers, A., & Pieler, T. (2010). Programming pluripotent precursor cells derived from *Xenopus* embryos to generate specific tissues and organs. *Genes (Basel)*, *1*(3), 413-426. doi:10.3390/genes1030413
7. Carabella, C., Cinosi, J., Piattelli, V., Burrato, P., & Miccadei, E. (2022). Earthquake-induced landslides susceptibility evaluation: A case study from the Abruzzo region (Central Italy). *Catena*, *208*, 105729. doi:10.1016/j.catena.2021.105729

8. Chew, Victor. "Confidence, prediction, and tolerance regions for the multivariate normal distribution." *Journal of the American Statistical Association* 61.315 (1966): 605-617.
9. Das, I., & Stein, A. (2016). Application of the Multitype Strauss Point Model for Characterizing the Spatial Distribution of Landslides. *Mathematical Problems in Engineering*, 2016, 1-12. doi:10.1155/2016/1612901
10. Dhakal, G., Karim, H., Haque, A. K. E., Nepal, M., Rahman, S. M. M., & Pervin, I. A. (2020). Adapting to urban flooding: a case of two cities in South Asia. *Water Policy*, 22(S1), 162-188. doi:10.2166/wp.2019.174
11. Dubeau, P., King, D., Unbushe, D., & Rebelo, L.-M. (2017). Mapping the Dabus Wetlands, Ethiopia, Using Random Forest Classification of Landsat, PALSAR and Topographic Data. *Remote Sensing*, 9(10), 1056. doi:10.3390/rs9101056
12. Fisher, N. I., T. Lewis, and B. J. J. Embleton. *Statistical Analysis of Spherical Data*. 1st ed. Cambridge: Cambridge University Press, 1987. Cambridge Books Online. Web. 26 April 2016.
13. García-Rodríguez, M. J., Malpica, J. A., Benito, B., & Díaz, M. (2008). Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression. *Geomorphology*, 95(3-4), 172-191. doi:10.1016/j.geomorph.2007.06.001
14. Gómez-Rubio, V. (2016). Spatial Point Patterns: Methodology and Applications with R. *Journal of Statistical Software*, 75(Book Review 2). doi:10.18637/jss.v075.b02
15. Haq, M., Akhtar, M., Muhammad, S., Paras, S., & Rahmatullah, J. (2012).

Techniques of Remote Sensing and GIS for flood monitoring and damage assessment: A case study of Sindh province, Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, 15(2), 135-141. doi:10.1016/j.ejrs.2012.07.002

16. Hermas, E., Gaber, A., & El Bastawesy, M. (2021). Application of remote sensing and GIS for assessing and proposing mitigation measures in flood-affected urban areas, Egypt. *The Egyptian Journal of Remote Sensing and Space Science*, 24(1), 119-130. doi:10.1016/j.ejrs.2020.03.002
17. Hussain, M., Tayyab, M., Zhang, J., Shah, A. A., Ullah, K., Mehmood, U., & Al-Shaibah, B. (2021). GIS-Based Multi-Criteria Approach for Flood Vulnerability Assessment and Mapping in District Shangla: Khyber Pakhtunkhwa, Pakistan. *Sustainability*, 13(6), 3126. doi:10.3390/su13063126
18. Hussain, M. A., Chen, Z., Wang, R., & Shoaib, M. (2021). PS-InSAR-Based Validated Landslide Susceptibility Mapping along Karakorum Highway, Pakistan. *Remote Sensing*, 13(20), 4129. doi:10.3390/rs13204129
19. Hussain, S., Hongxing, S., Ali, M., Sajjad, M. M., Ali, M., Afzal, Z., & Ali, S. (2021). Optimized landslide susceptibility mapping and modelling using PS-InSAR technique: a case study of Chitral valley, Northern Pakistan. *Geocarto International*, 37(18), 5227-5248. doi:10.1080/10106049.2021.1914750
20. Ili, I., Loupasakis, C., & Tsangaratos, P. (2018). Land subsidence phenomena investigated by spatiotemporal analysis of groundwater resources, remote sensing techniques, and random forest method: the

case of Western Thessaly, Greece. *Environ Monit Assess*, 190(11), 623. doi:10.1007/s10661-018-6992-9

21. Kamp, U., Growley, B. J., Khattak, G. A., & Owen, L. A. (2008). GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. *Geomorphology*, 101(4), 631-642. doi:10.1016/j.geomorph.2008.03.003
22. Khan, H., Shafique, M., Khan, M. A., Bacha, M. A., Shah, S. U., & Calligaris, C. (2019). Landslide susceptibility assessment using Frequency Ratio, a case study of northern Pakistan. *The Egyptian Journal of Remote Sensing and Space Science*, 22(1), 11-24. doi:10.1016/j.ejrs.2018.03.004
23. Khan, S. D., & Glenn, N. F. (2007). New strike-slip faults and litho-units mapped in Chitral (N. Pakistan) using field and ASTER data yield regionally significant results. *International Journal of Remote Sensing*, 27(20), 4495-4512. doi:10.1080/01431160600721830
24. Khattak, G. A., Owen, L. A., Kamp, U., & Harp, E. L. (2010). Evolution of earthquake-triggered landslides in the Kashmir Himalaya, northern Pakistan. *Geomorphology*, 115(1-2), 102-108. doi:10.1016/j.geomorph.2009.09.035
25. Kincal, C., Li, Z., Drummond, J., Liu, P., Hoey, T., & Muller, J.-P. (2017). Landslide Susceptibility Mapping Using GIS-based Vector Grid File (VGF) Validating with InSAR Techniques: Three Gorges, Yangtze River (China). *AIMS Geosciences*, 3(1), 116-141. doi:10.3934/geosci.2017.1.116
26. Kulin, H. W., & Kuenne, R. E. (1962). An Efficient Algorithm for the Numerical Solution of the Generalized Weber Problem in Spatial

Economics. *Journal of Regional Science*, 4(2), 21-33.
doi:10.1111/j.1467-9787.1962.tb00902.x

27. Lamb, D. S., Downs, J. A., & Lee, C. (2016). The network K-function in context: examining the effects of network structure on the network K-function. *Transactions in GIS*, 20(3), 448-460. doi:10.1111/tgis.12157
28. Li, B., Wang, N., Chen, J., & Ji, J. (2021). GIS-Based Landslide Susceptibility Mapping Using Information, Frequency Ratio, and Artificial Neural Network Methods in Qinghai Province, Northwestern China. *Advances in Civil Engineering*, 2021, 1-14. doi:10.1155/2021/4758062
29. Liu, W., Yamazaki, F., & Maruyama, Y. (2019). Detection of Earthquake-Induced Landslides during the 2018 Kumamoto Earthquake Using Multitemporal Airborne Lidar Data. *Remote Sensing*, 11(19), 2292. doi:10.3390/rs11192292
30. Marcus Nasser, w. B. D. (2002). Incipient Landslides in the Jhelum Valley, Pakistan Following the 8th October 2005 Earthquake.
31. Martino, S., Battaglia, S., D'Alessandro, F., Della Seta, M., Esposito, C., Martini, G., . . . Troiani, F. (2019). Earthquake-induced landslide scenarios for seismic microzonation: application to the Accumoli area (Rieti, Italy). *Bulletin of Earthquake Engineering*, 18(12), 5655-5673. doi:10.1007/s10518-019-00589-1
32. Melancon, A. M., Molthan, A. L., Griffin, R. E., Mecikalski, J. R., Schultz, L. A., & Bell, J. R. (2021). Random Forest Classification of Inundation Following Hurricane Florence (2018) via L-Band Synthetic Aperture Radar and Ancillary Datasets. *Remote Sensing*, 13(24), 5098. doi:10.3390/rs13245098

33. Meunier, P., Hovius, N., & Haines, A. J. (2007). Regional patterns of earthquake-triggered landslides and their relation to ground motion. *Geophysical Research Letters*, *34*(20). doi:10.1029/2007gl031337
34. Miguez, M., Veról, A., de Sousa, M., & Rezende, O. (2015). Urban Floods in Lowlands—Levee Systems, Unplanned Urban Growth and River Restoration Alternative: A Case Study in Brazil. *Sustainability*, *7*(8), 11068-11097. doi:10.3390/su70811068
35. Miles, S. B., & Keefer, D. K. (2009). Evaluation of CAMEL — comprehensive areal model of earthquake-induced landslides. *Engineering Geology*, *104*(1-2), 1-15. doi:10.1016/j.enggeo.2008.08.004
36. Moresi, F. V., Maesano, M., Collalti, A., Sidle, R. C., Matteucci, G., & Scarascia Mugnozza, G. (2020). Mapping Landslide Prediction through a GIS-Based Model: A Case Study in a Catchment in Southern Italy. *Geosciences*, *10*(8), 309. doi:10.3390/geosciences10080309
37. Mitchell, Andy. *The ESRI Guide to GIS Analysis*, Volume 2. ESRI Press, 2005.
38. O'Driscoll, M., Clinton, S., Jefferson, A., Manda, A., & McMillan, S. (2010). Urbanization Effects on Watershed Hydrology and In-Stream Processes in the Southern United States. *Water*, *2*(3), 605-648. doi:10.3390/w2030605
39. Okabe, A., & Yamada, I. (2010). The K-Function Method on a Network and Its Computational Implementation. *Geographical Analysis*, *33*(3), 271-290. doi:10.1111/j.1538-4632.2001.tb00448.x

40. Oyana, T. J., & Margai, F. M. (2010). Spatial Patterns and Health Disparities in Pediatric Lead Exposure in Chicago: Characteristics and Profiles of High-Risk Neighborhoods. *The Professional Geographer*, *62*(1), 46-65. doi:10.1080/00330120903375894
41. Pourghasemi, H. R., Kariminejad, N., Gayen, A., & Komac, M. (2020). Statistical functions used for spatial modelling due to assessment of landslide distribution and landscape-interaction factors in Iran. *Geoscience Frontiers*, *11*(4), 1257-1269. doi:10.1016/j.gsf.2019.11.005
42. Qin, S., Guo, X., Sun, J., Qiao, S., Zhang, L., Yao, J., . . . Zhang, Y. (2021). Landslide Detection from Open Satellite Imagery Using Distant Domain Transfer Learning. *Remote Sensing*, *13*(17), 3383. doi:10.3390/rs13173383
43. Rehman, M. U., Zhang, Y., Meng, X., Su, X., Catani, F., Rehman, G., . . . Ahmad, I. (2020). Analysis of Landslide Movements Using Interferometric Synthetic Aperture Radar: A Case Study in Hunza-Nagar Valley, Pakistan. *Remote Sensing*, *12*(12), 2054. doi:10.3390/rs12122054
44. Rong, G., Alu, S., Li, K., Su, Y., Zhang, J., Zhang, Y., & Li, T. (2020). Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China. *Water*, *12*(11), 3066. doi:10.3390/w12113066
45. Rozina, Ahmad, M., Khan, A. M., Abbas, Q., Arfan, M., Mahmood, T., . . . Ameen, M. (2022). Implication of scanning electron microscopy as a tool for identification of novel, nonedible oil seeds for biodiesel

production. *Microsc Res Tech*, 85(5), 1671-1684.
doi:10.1002/jemt.24027

46. Rybarczyk, G., & Wu, C. (2010). Bicycle facility planning using GIS and multi-criteria decision analysis. *Applied Geography*, 30(2), 282-293.
doi:10.1016/j.apgeog.2009.08.005
47. Saba, S. B., van der Meijde, M., & van der Werff, H. (2010). Spatiotemporal landslide detection for the 2005 Kashmir earthquake region. *Geomorphology*, 124(1-2), 17-25.
doi:10.1016/j.geomorph.2010.07.026
48. Saber, M., Abdrabo, K. I., Habiba, O. M., Kantosh, S. A., & Sumi, T. (2020). Impacts of Triple Factors on Flash Flood Vulnerability in Egypt: Urban Growth, Extreme Climate, and Mismanagement. *Geosciences*, 10(1), 24. doi:10.3390/geosciences10010024
49. Sato, H. P., & Harp, E. L. (2009). Interpretation of earthquake-induced landslides triggered by the 12 May 2008, M7.9 Wenchuan earthquake in the Beichuan area, Sichuan Province, China using satellite imagery and Google Earth. *Landslides*, 6(2), 153-159. doi:10.1007/s10346-009-0147-6
50. Shafique, M. (2020). Spatial and temporal evolution of co-seismic landslides after the 2005 Kashmir earthquake. *Geomorphology*, 362, 107228. doi:10.1016/j.geomorph.2020.107228
51. Shahabi, H., & Hashim, M. (2015). Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. *Sci Rep*, 5, 9899. doi:10.1038/srep09899
52. Shano, L., Raghuvanshi, T. K., & Meten, M. (2020). Landslide

susceptibility evaluation and hazard zonation techniques – a review. *Geoenvironmental Disasters*, 7(1). doi:10.1186/s40677-020-00152-0

53. Usma, A., Ahmad, M., Ramadan, M. F., Khan, A. M., Zafar, M., Hamza, M., . . . Yaseen, G. (2022). Micro-morphological diversity of pollen among Asteraceous taxa from Potohar Plateau-Pakistan. *Microsc Res Tech*, 85(7), 2467-2485. doi:10.1002/jemt.24102
54. Wang, B., Shi, W., & Miao, Z. (2015). Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. *PLoS One*, 10(3), e0118537. doi:10.1371/journal.pone.0118537
55. Wasowski, J., Keefer, D. K., & Lee, C.-T. (2011). Toward the next generation of research on earthquake-induced landslides: Current issues and future challenges. *Engineering Geology*, 122(1-2), 1-8. doi:10.1016/j.enggeo.2011.06.001
56. Xu, C., Xu, X., Shen, L., Yao, Q., Tan, X., Kang, W., . . . Li, K. (2016). Optimized volume models of earthquake-triggered landslides. *Sci Rep*, 6, 29797. doi:10.1038/srep29797
57. Yang, X., & Chen, L. (2010). Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides. *International Journal of Applied Earth Observation and Geoinformation*, 12(6), 487-495. doi:10.1016/j.jag.2010.05.006
58. Youssef, A. M., & Pourghasemi, H. R. (2021). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geoscience Frontiers*, 12(2), 639-655. doi:10.1016/j.gsf.2020.05.010
59. Zhang, K., Wu, X., Niu, R., Yang, K., & Zhao, L. (2017). The assessment

of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. *Environmental Earth Sciences*, 76(11). doi:10.1007/s12665-017-6731-5

60. Zhang, W., Lin, J., Peng, J., & Lu, Q. (2010). Estimating Wenchuan Earthquake induced landslides based on remote sensing. *International Journal of Remote Sensing*, 31(13), 3495-3508. doi:10.1080/01431161003727630

APPENDICES

Appendix A. Google Earth Engine code for Land use and Land cover (LULC) classification of the study area for the year 2016 from Landsat 8 imagery.

```
#GEE code for LULC classification
```

```
Map.addLayer(table, {}, "Study Area");
```

```
Map.centerObject(table,6);
```

```
var Images=ee.ImageCollection("LANDSAT/LC08/C02/T1_TOA");
```

```
var s2_composite = Images.filterBounds(table)
```

```
  filterDate('2016-05-01', '2016-08-30')
```

```
  filter(ee.Filter.lte('CLOUD_COVER', 10))
```

```
  median());
```

```
var comp=s2_composite.clip(table)
```

```
Map.addLayer(comp, {Bands:["B1","B2","B3"]}, 'Study Area Imagery');
```

```
print(comp);
```

```
var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7'];
```

```
var samples = snow.merge(water.merge(veg.merge(builtup.merge(soil))));
```

```
print ('Sample',samples);
```

```
var points = comp.select(bands).sampleRegions({
```

```
  collection: samples,
```

```
  properties: ['Landcover'],
```

```
  scale: 30
```

```
}).randomColumn();
```

Contd.

```
//-----  
var training = points.filter(ee.Filter.lt('random', 0.7));  
var validation = points.filter(ee.Filter.gte('random', 0.7));  
var classifier = ee.Classifier.smileRandomForest(100).train({  
  features: training,  
  classProperty: 'Landcover',  
  inputProperties: bands  
});  
//-----  
// CLASSIFY THE IMAGE.  
var Classified_imagery = comp.select(bands).classify(classifier);  
Map.addLayer(Classified_imagery.clip(table), {min: 0, max:4, palette: [  
'BLUE', 'Green', 'LightGreen', 'Brown', 'YELLOW']}, 'Classified');  
  
print('error matrix: ', classifier.confusionMatrix());  
print('accuracy: ', classifier.confusionMatrix().accuracy());  
  
Export.image.toDrive({  
  image: Classified_imagery ,  
  description: 'LULC2016',  
  region: table  
});
```

Appendix B. Generalized code for Point Pattern Analysis in "R" Studio.

Here are some steps to calculate measures of spatial clustering and patterns in R using point pattern analysis:

- 1. Loading the packages:** To perform point pattern analysis in R, we need to load the necessary packages, such as spatstat, rgdal, and raster.

```
install.packages("spatstat")
```

```
install.packages("rgdal")
```

```
install.packages("raster")
```

- 2. Loading the data:** Load the point data into R and preprocess it by checking for completeness and accuracy, removing duplicates, and transforming it if necessary.

```
library(spatstat)
```

```
library(rgdal)
```

```
library(raster)
```

```
data <- read.shp("path/to/data.shp")
```

```
coordinates(data) <- c("longitude", "latitude")
```

- 3. Creating a point pattern object:** Create a point pattern object using the 'ppp' function in spatstat. This function creates a point pattern object that can be used for further analysis.

```
Pp <- ppp(data$longitude, data$latitude, marks=data$attribute)
```

Contd.

- 4. Calculating measures of spatial clustering:** Calculate measures of spatial clustering and patterns using functions such as `Kest`, `pcf`, and `envelope` in `spatstat`.

```
K <- Kest(pp)
plot(K, main="Ripley's K-function")
```

- 5. Developing a statistical model:** Use a statistical model, such as a Poisson or negative binomial regression, to identify the environmental factors that contribute to landslide occurrence.

```
Model <- glm(attribute ~ variable1 + variable2, data=data, family=poisson)
summary(model)
```

- 6. Validating the model:** To check the how much the model is accurate and reliable, it is necessary to compare its predictions with actual landslide events in the study area that were not used in developing the model. This process, known as model validation, ensures that the model can accurately predict events beyond the ones it was trained on.

```
Validation <- data[validation_set, ]
prediction <- predict(model, newdata=validation, type="response")
```