Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia

Leilt Wossen Bekele

A Thesis Submitted to the Department of Applied Geology

School of Applied Natural Science

Office of Graduate Studies

Adama Science and Technology University

November, 2024

Adama, Ethiopia

Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia

Leilt wossen

Advisor: Dr. Shankar Karuppannan

Co-Advisor: Dr. Leulalem Shano

A Thesis Submitted to the Department of Applied Geology,

School of Applied Natural Science

Office of Graduate Studies

Adama Science and Technology University

November, 2024

Adama, Ethiopia

DECLARATION

I declare that this thesis entitled **"Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia**" is my own work and has not been submitted to any university for a similar purpose. The references used in this proposal are duly recognized by proper citations.

Leilt Wossen	____________________	____________________
Name of student	Signature	Date

RECOMMENDATION OF ADVISORS

We, the advisors of this thesis, at this moment, certify that we have closely advised the student while developing this thesis and read the draft thesis entitled "**Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia**" under our guidance by Leilt Wossen. Therefore, we recommend the submission of the proposal to the department for further review and evaluation.

Dr. Shankar Karuppannan __________________ __________________

Major advisor's name Signature Date

Dr. Leulalem Shano __________________ __________________

Co-advisor's name Signature Date

APPROVAL SHEET

We, the advisors of the thesis entitled "**Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia**" developed by Leilt Wossen Bekele hereby certify that the recommendations and suggestions made by the board of examiners are appropriately incorporated into the final version of the thesis.

Dr. Shankar Karuppannan	__________________	__________________
Major Advisor	Signature	Date
Dr. Leulalem Shano	__________________	__________________
Co-advisor	Signature	Date

We, the undersigned, members of the Board of Examiners of the thesis by Leilt Wossen Bekele have read and evaluated the thesis entitled "**Landslide Susceptibility Modeling using Machine Learning Algorithm: The Case of Sile and Elgo Catchments, Gamo Highland, Southern Ethiopia**", and examined the candidate during open defense. This is, therefore, to certify that the thesis is accepted for partial fulfilment of the requirement of the degree of Master of Science in Engineering Geology.

__________________	__________________	__________________
Chairperson	Signature	Date
__________________	__________________	__________________
Internal Examiner	Signature	Date
__________________	__________________	__________________
External Examiner	Signature	Date

Final approval and acceptance of the thesis is contingent upon submission of its final copy to the Office of Postgraduate Studies (OPGS) through the Department Graduate Council (DGC) and School Graduate Committee (SGC).

________________________	__________________	__________________
Department Head	Signature	Date
________________________	__________________	__________________
School Dean	Signature	Date
________________________	__________________	__________________
Office of Postgraduate Studies, Dean	Signature	Date

ACKNOWLEDGMENT

First and foremost, I express my deepest gratitude to Almighty God, whose grace, guidance, and strength have been my foundation throughout this journey. I also extend my heartfelt thanks to His Blessed Mother, whose intercessions have provided me with peace and support. I am deeply indebted to my advisor, Dr. Shankar Karuppannan, whose invaluable guidance, wisdom, and support have greatly contributed to my research. His insightful advice and encouragement have been vital in shaping the direction of this work. I would also like to sincerely thank my co-advisor, Dr. Leulalem Shano, for his constructive feedback, constant support, and expert advice, which have enriched the quality of this thesis. I am profoundly thankful to my family, whose unwavering love, encouragement, and patience have been my pillar of strength throughout my academic pursuits. Their sacrifices and belief in me have been instrumental in the completion of this thesis.

I also appreciate the Geological Survey of Ethiopia for its invaluable resources and data that were crucial for my research. The National Meteorological Agency of Ethiopia has provided essential meteorological data, and I am immensely grateful for their assistance and cooperation. I also extend my gratitude to the Geospatial Institute of Ethiopia for their support and access to geospatial data that was vital for the analysis in this study. I am equally grateful to my friends, whose companionship, encouragement, and understanding have made this journey more enjoyable and manageable. Teir academic and, has been a constant source of motivation.

Lastly, I would like to thank Adama Science and Technology University (ASTU) for allowing me to pursue my MSc studies and for their financial support for my thesis work.

TABLE OF CONTENTS

Contents **Pages**

LIST OF FIGURES

LIST OF TABLES

LIST OF PLATES

LIST OF ABBREVIATIONS ACRONYMS

ANN	Artificial Neural Network
ArcGIS	Arc Geographic Information System
AUC	Area Under Curve
DEM	Digital Elevation Model
DTL	Distance To Lineament
DTS	Distance To Stream
GIS	Geographic Information System
GIE	Geological Institute of Ethiopia
IDW	Inverse Distance Weighting
LSM	Landslide Susceptibility Mapping
LULC	Land Use Land cover
ML	Machine learning
MER	Main Ethiopian rift
NASA	National Aeronautics and Space Administration
QGIS	Quantum Geographic Information System
RS	Remote Sensing
ROC	Receiver Operating Characteristic
SMER	Southern Main Ethiopian Rift
KNN	K -Nearest Neighbors
SVM	Support Vector Machine
TOL	Tolerance
VIF	Variance Inflation Factor

ABSTRACT

Landslides are one of the most prevalent global geologic hazards, causing significant damage to people and property. They are prevalent in the highlands of Ethiopia. The primary purpose of this study was to develop a comprehensive landslide susceptibility modeling utilizing machine learning algorithms in the Sile and Elgo catchment areas. This study selected ten factors influencing landslides for landslide susceptibility mapping. These factors included elevation, slope, aspect, curvature, distance to streams, distance to lineaments, soil type, lithology, land use and land cover, and rainfall. Data for these parameters were primarily derived from Google Earth imagery, Digital Elevation Models (DEM), and Sentinel-2 imagery. Through QGIS and Google Earth imagery interpretation, 1068 landslides were identified within the study area. The analysis revealed that soil type, followed by slope are the most significant factors influencing landslide occurrence, as indicated by the frequency ratio value. The Support Vector Machine (SVM) and K-nearest neighbors (KNN) models were employed to assess landslide susceptibility. Model accuracy was evaluated using the receiver operating characteristic (ROC) curve, and the area under the curve (AUC) values for a success rate of 0.799 for SVM and 0.81 for KNN. The results indicate that the Support Vector Machine (SVM) model is the most effective for assessing landslide susceptibility in the study area, and highlights the efficacy of machine learning. The study demonstrates the efficacy of machine learning in landslide susceptibility mapping while suggesting potential advancements through deep learning techniques.

Keywords: *Landslide, Machine Learning, SVM, KNN, Sile and Elgo catchment*.

CHAPTER ONE

1. INTRODUCTION

1.1 Background

A landslide is a geological phenomenon that depends on the type of movement (topple, spread, flow, and deformation of a slope), and the kind of material (debris, rock, or earth). Variables, such as geology, hydrogeology, earthquakes, and human contact, can cause landslides (Mebrahtu et al., 2021). It is commonly thought in landslide research that these elements may cause landslides in a given location. As a result, assessing these factors and their relationship to previous landslides in a specific area may serve as the foundation for future landslide prediction (Shano et al., 2021).

According to World Disaster Statistics (2020), landslide hazards are responsible for 13.9% of all catastrophe fatalities worldwide caused by natural disasters in the last several decades. They cause significant damage to the global economy, environment, and human lives (Mahmoud and Gan, 2018). They are caused by natural causal variables, including slope geometry, slope material, structural discontinuities, land use land cover, and groundwater conditions. In addition, external causal elements like precipitation, volcanism, seismicity, and human activity are also causes of landslide disasters (Tadesse et al., 2024).

Landslides are a common cause of property loss in the steep regions of Eastern Africa (Thi Ngo et al., 2021). Because of the topography, rainfall patterns, and changes in land use, the area is quite prone to landslides (Ibrahim et al., 2022). In Ethiopia, landslides are also a frequent issue in hilly areas, posing a risk to people's lives, property, and infrastructure (Shano et al., 2020). Over 600 sites are known to be at risk of landslides, which has grown to be a major concern for government planners and decision-makers at all levels (Tadesse et al., 2024). The study area falls in the southern Ethiopian highlands and is part of the Ethiopian highlands, severely affected by landslide incidences in recent years. In the region, landslides are primarily caused by a range of slope instabilities resulting from geological, hydrological, geomorphological, climatic, and human variables. Besides, problems such as misusing landslide susceptibility and hazard, other problems related to landslide studies include a random selection of independent factors and inappropriate mapping scales (Shano et al., 2022).

There are two types of landslide prediction models: data-driven models and numerical models. Several computer models with a physical basis have been put forth in recent years to simulate the beginning of landslides (Montgomery and Dietrich, 1994; Montrasio and Valentino, 2007). Geographic information systems (GIS) and remote sensing (RS) have advanced technologically in recent years. The use of remote sensing data and GIS spatial analytic tools has improved the accuracy of landslide susceptibility mapping. Here, thorough landslide inventory data and understanding the elements that condition landslides are essential for knowledge-based and data-driven spatial modeling techniques (Ghorbanzadeh et al., 2020). When compared to numerical models, data-driven models are frequently more widely used due to their perceived ease of use, higher prediction accuracy, demonstrated strong performance, and cheaper costs (Corominas et al., 2005). Many people have used data-driven models, particularly machine learning techniques, to "learn" the relationship between landslides' occurrence and the associated factors (Liu et al., 2021).

Machine learning (ML) techniques offer key advantages, such as an objective statistical foundation, consistency, accuracy in analyzing factors influencing landslide development, and adaptability for regular updates. Accordingly, this thesis aims to identify more reliable algorithms that can be applied across various spatial scales while addressing gaps in previous research. Furthermore, the chosen algorithms, including the widely recognized KNN and SVM approaches, introduce a new perspective for Ethiopia, as these methods have not yet been applied in the country's landslide studies.

1.2 Statement of problem

Most of the Gamo Highlands' districts have experienced landslides, which pose a serious hazard to life (Tadesse et al., 2024). As part of the Ethiopian highland, landslides threaten lives, infrastructure, and ecosystems in mountainous regions such as the Sile and Elgo Catchments.

Seismicity and seasonal precipitation are causes of landslides in this area, primarily characterized by active faulting creating steep slopes. The study area also has weak geological formations that are affected highly by geological structures, making the area sensitive to landslides (Martínek et al., 2021). Even though landslide hazards are the most prevalent issue in South Ethiopia, not many noticeable studies have yet been carried out in this field (Shano et al., 2022).

Despite being critical, landslide susceptibility mapping remains challenging due to the complex interactions of geological, topographical, and environmental factors contributing to landslide occurrences. Therefore, this study is necessary to fill the gaps in the research area by using the machine learning algorithm and generating information that helps reform the current spatial plan and gives insight to local government in managing present and future land use practices.

1.3 Objectives

1.3.1 General Objective

The main objective of this study is to develop a comprehensive landslide susceptibility model utilizing machine Learning algorithms in Sile and Elgo catchment areas.

1.3.2 Specific Objectives

To meet the general objective of this study, the following specific objectives are designed:

- To understand processes and factors leading to slope failure in the study area.
- To produce a landslide inventory map of the area.
- To check the validity of the prepared landslide susceptibility map.
- To drive possible suitable recommendations based on the results of the present study.

1.4 Significance of the study

This study helps to:

- Upgrade the information to create a better awareness of landslide impacts, particularly for rural people who live near the toe of the hill slopes;
- Acquire tangible information that demonstrates how human activities can facilitate and increase landslide hazard;
- To obtain data and develop possible recommendations to reduce the impact of landslides.

The study is significant for the local community around the study area. It is located in the landslide-prone area of the southern highland of the country, where a community dislocated from that area, removed fertile soils, and lost families and animals' lives. Moreover, results obtained from this study will also be used as a reference to other related hazards, which are intended to be conducted in similar areas.

1.5 Limitations of the Study

The investigation faced significant limitations that impacted its completeness and accuracy, primarily due to restricted access to remote locations caused by inadequate infrastructure, rugged terrain, and seasonal weather conditions such as heavy rainfall and flooding. These challenges hindered field surveys and resulted in data collection gaps. Additionally, the sparse distribution of meteorological data in the area limited the accuracy and resolution of critical data. Remote sensing technique, particularly Google Earth imagery, were employed to compile the landslide inventory to mitigate these challenges. While this method provided valuable insights and helped identify landslide-prone areas, it lacked the precision of direct field observations, leading to potential inaccuracies in detecting smaller-scale landslides and validation challenges in the absence of ground truthing.

1.6 Organization of the Thesis

The current study comprises five chapters. The first chapter introduces the background of landslides, outlines the problem statement, objectives, and significance of the study, and addresses its limitations. The second chapter conducts a literature review focusing on landslide studies globally and in southern Ethiopia, providing an overview of landslide conditions and comparing machine learning with other methodologies. The third chapter offers an overview of the study area and discusses the methodology, including types of software, data collection procedures, and assigning of machine learning methods. The fourth chapter presents results and discussions, including assigning weights to causative factors and producing the final landslide susceptibility map. Finally, chapter five presents conclusions and provides recommendations based on the output.

CHAPTER TWO

2. LITERATURE REVIEW

2.1 Landslide definition and classification

Landslide definitions vary among academics, with geologists and engineers often using distinct terminologies. These differences reflect the complexity of studying landslide phenomena (Mersha and Meten, 2022). Landslide is generally defined as nearly any type of mass movement on a slope, including ones that entail little to no actual sliding, like rockfalls, topples, and debris flows (Shano et al., 2021).

According to the United States Geological Survey (USGS), mass movement and material involved in the slide are important factors in landslide classification. Landslides can be classified based on the type of material involved, such as rock, earth, soil, mud, or debris, and the type of movement that occurs. The movement can take different forms, including falls, topples, slides, spreads, or flows. Additionally, there is a "complex" category, where more than one type of movement happens simultaneously or sequentially, contributing to the downward motion of the landslide mass (Varnes, 1978*)*.

Table 1: Classification of slope movement (Varnes, 1978).

Type of Movement			**Type of Material**		
			Bedrock	Soils	
				Predominantly course	Predominantly fine
,Falls			Rock fall	Debris fall	Earth Fall
Topples			Rock Topple	Debris topple	Earth Topple
Slides	Rotational	Few units	Rock Slump	Debris slump	Earth slump
	Translational	Many units	Rock block slide Rock slide	Debris block slide Debris slide	
Lateral spreads			Rock spread	Debris spread	Earth spread
Flows			Rock flow	Flows	Rock flow
Complex			Combination of two or more principal type of movements.		

2.2 Landslide Studies Worldwide

According to World Disaster Statistics, in recent decades, landslide hazards have been responsible for 13.9% of all catastrophe fatalities worldwide caused by natural disasters (Mahmoud and Gan, 2018). From 2004–2016, more than 55,000 people lost their lives to landslides, and this does not include deaths caused by seismically triggered landslides. Overall losses were estimated at USD 20 billion annually (Sim et al., 2022).

"Landslide inventory," "landslide susceptibility," "landslide hazard," and "landslide risk" are commonly used in landslide investigations to reduce the landslide risk. Each of these terminologies has its uses and objectives (Shano et al., 2021). Thus, assessing these variables and their relationship to previous landslides in a region could serve as the foundation for forecasting future landslides (Chimidi et al., 2017).

Numerous methodologies, including deterministic, statistical, machine learning, and heuristic approaches, can be used to describe or model different landslide types (Cao et al., 2016; Reichenbach et al., 2018; Shano et al., 2021).

2.3 Landslide Studies in Ethiopia

One of Ethiopia's main geohazards is landslides, which typically take in the form of rock falls, earth slides, debris flows, and mudflows. These occur most frequently in the steep and rocky highlands at elevations above 1500 meters (Tesfa and Woldearegay, 2021). Numerous studies attempted to evaluate the circumstances behind these incidents and their effects in various regions of Ethiopia, including the country's northern, southern, and western highlands. These regions are divided into four blocks, the first of which is block A, which encompasses the northeastern and northern regions of the nation. Gojam and Gondar, both east and west, are included in Block B. Block C represents south and southwest Ethiopia, while Block D represents the eastern portion of the Main Ethiopian Rift (Mewa and Mengistu, 2022).

Shano et al. (2022), the Ethiopian Highland region has seen several landslide incidents in recent decades, usually during the rainy season. Rainfall-induced slope failures of various kinds and sizes often result from the hilly and mountainous terrains of Ethiopia's highlands, which are distinguished by varying topographical, geological, hydrological (surface and groundwater), and land-use conditions. Ethiopia's highlands are seeing more landslides as a result of heavy rainfall, high relative relief, complex, unstable geology, and increased human activity like building highways at top slopes (Woldearegay, 2013; Raghuvanshi et al., 2014). According to an analysis of earlier research, landslide dangers have been linked to fatalities, structural failures, damage to agricultural areas, and environmental harm. Due to the complicated geomorphological, hydrological, and geological setting of the Ethiopian landmass, reactivated historical landslides have regularly devastated the country's steep terrains (Woldearegay, 2013). There is no set method for choosing which component to employ in landslide susceptibility mapping; instead, it relies on the region's characteristics and the data availability (Ayalew and Yamagishi, 2005).

2.4 Landslide studies in the southern highland of Ethiopia

Southern Ethiopian Highland region has seen several landslide incidents in the past few decades, during nearly every wet season. Landslide hazard zonation must be implemented in the study region to lessen the issues these landslides bring.

According to statistical analysis, the parameters that have the biggest impact on the occurrence of landslides are the distance to the fault, the distance to the stream, the groundwater zones, the lithological units, and the aspect (Shano et al., 2022). The Gamo Zone is challenged by diverse geomorphological features such as altitude variations, rock weathering, active tectonics, and heavy rainfall. These, along with deforestation, lead to soil erosion, land degradation, landslides, and floods, threatening infrastructure, settlements, agriculture, food security, and even human lives. Surface erosion degrades farmlands while flooding along Lake Abaya and Lake Chamo.

Slope instabilities and mass movements are exacerbated by the region's topography, climate, and weathering, often triggered by extreme weather events, earthquakes, and human interference. Lakes Abaya and Chamo's catchments are found in the Southern Main Ethiopian Rift. Sediment buildup in the lakes and significant land degradation in the highlands are the area's defining features (Teffera, 2016). Additionally, landslides play a significant role in erosion by supplying rivers in hilly catchments with debris (Whipp and Ehlers, 2019).

According to Shano et al. (2024), the landslide catastrophe that occurred on July 22, 2024, in the villages of Kencho, Shacha, and Gozdi in Gofa Zone, Ethiopia, had widespread impacts. It caused significant geological and environmental disturbances, damaged infrastructure, and displaced communities. The socioeconomic repercussions included loss of life, destruction of property, and long-term economic instability for the affected populations. The investigation delves into the geological, geotechnical, geophysical, and geospatial factors contributing to the disaster.

2.5 Landslide controlling factors in Gamo highlands of Ethiopia

According to Zheng et al. (2021), landslides usually happen because of different things like geological factors, hydrological conditions, topography, land cover, human activities, and so on. Different studies of landslides around the world found that things like how high the land is, which direction it faces, the shape of the land, what the ground is made of, how it's shaped, what's growing on it, and how far it is from rivers and roads can all affect whether landslides happen.

Intrinsic and external factors are two categories of factors that lead to land instability (Carrera et al., 2023). Intrinsic parameters are like the basic features that naturally make a slope unstable, such as its shape, the type of soil it's made of, and any cracks or gaps in the structure. On the other hand, external factors are things like rain, earthquakes, and human activities (Chimidi et al., 2017). Precipitation and human activities are typically the main triggers in the Gamo Highlands, whereas slope, elevation, aspect, distance from the stream, drainage density, soil type, distance from the road, and land use/cover are the predictors (Shano et al., 2021). Table (2.) shows all the intrinsic and external factors that were used by different literatures in Gamo Highlands.

According to Kebeba et al. (2024), geospatial analysis, frequency ratio (FR), and analytical hierarchy process (AHP) were effectively integrated to assess landslide susceptibility in the Maze catchment, Omo Valley, southern Ethiopia. The study used GIS and remote sensing to analyze key factors like Slope, Aspect, Curvature, Distance to stream, Distance to lineament, Distance to the road, LULC, NDVI, Lithology, and Distance to fault. FR identified statistical relationships between landslides and these factors, while AHP weighed them based on expert judgment. Slope and rainfall were found to be the most significant contributors. The resulting susceptibility map offers valuable insights for land-use planning and disaster mitigation.

Table 2: Intrinsic and external factors that were used by different literatures in Gamo Highlands.

NO	Paper name	Authors name	Influencing factors
1	Landslide susceptibility mapping using frequency ratio model: the case of Gamo highland, South Ethiopia	Shano et al.,2021	Aspect, Slope, Curvature, Lithology, Elevation, Proximity to lineaments, Proximity to stream erosion, and Groundwater.
2	Landslide Hazard Zonation using Logistic Regression Model: The Case of Shafe and Baso Catchments, Gamo Highland, Southern Ethiopia	Shano et al.,2021	Lithology, Groundwater conditions, Distance to faults, Slope, Aspect, Curvature, and Land use/land cover.
3	Landslide Assessment and Hazard Zonation in the Birbir Mariam District, Gamo Highlands, Rift Valley Escarpment, Ethiopia	Oyda and Regasa, 2023	Slope geometry, Slope material (rock/soil), Structural discontinuities, Land use/land cover, and Groundwater conditions.
4	Landslide vulnerability mapping using multi-criteria decision-making approaches: in Gacho Babba District, Gamo Highlands Southern Ethiopia	Tadesse et al., 2024	Slope, Elevation, Aspect, Stream distance, Drainage density, Soil type, road length, and land use/cover.
5	Integration of geospatial analysis, frequency ratio, and analytical hierarchy process for landslide susceptibility assessment in the maze catchment, Omo Valley, southern Ethiopia	Kebeba et al., 2024	Slope, Aspect, Curvature, Distance to stream, Distance to lineament, Distance to road, LULC, NDVI, Lithology, Distance to fault.
6	Assessing landslide susceptibility in Lake Abya catchment, Rift Valley, Ethiopia: A GIS-based frequency ratio analysis	Oyda et al., 2024	Slope, Aspect, Distance from the stream, Curvature, LULC, Lithology, Elevation.
7	Landslide susceptibility modeling in the Kulfo River catchment, Rift Valley, Ethiopia: An integrated geospatial and statistical analysis	Mulugeta et al., 2024	LULC, slope, aspect, elevation, curvature, lithology, proximity to lineament, and NDVI.

2.5.1 Geological (lithological and structural) settings

Geological factors encompass lithology, soil type, and geological structures/lineaments. The influence of lithology on landslide occurrence greatly relies on factors such as composition, stratigraphy, degree of weathering, and the existence of geological discontinuities (Wubalem et al., 2022).

The geological pattern in southern Ethiopia exhibits a basement which is composed of metamorphic rocks from the Neoproterozoic era, which are overlain by extensive volcanic sequences. These volcanic sequences encompass various phases of volcanic activity, including pre-rift Cenozoic volcanism and the volcanism associated with the Main Ethiopian Rift (MER). This geological setup is typical of areas influenced by tectonic activity, such as rift zones, where the magma from the Earth's mantle rises to the surface, leading to volcanic eruptions and the formation of new crust (Martínek et al., 2021).

As stated by different researchers e.g. (Abebe et al., 2010), the occurrence of landslides in the rift margins of the Ethiopian highlands is linked to complex, and deep-rooted structural deformations. The geological and structural contexts are essential to accurately identify the causes of these failures and give effective mitigation strategies.

Geostatistical study indicates that the slope instabilities observed here, which primarily consist of landslides and rockfalls, are located near major faults and in places with significantly steeper slopes and a larger density of faults and lineaments. Large downthrown blocks, parallel sets of erosional valleys, asymmetrical ridges with SSW–NNE trending, and straight fault scarps with triangular facets are some of the morphological features of active tectonics that are strongly linked to the majority of the large-scale slope instabilities in this area. These structures are linked to massive displacement active normal faults (Martínek et al., 2021).

2.5.2 Topographic factors

Topographic characteristics are influenced by various slope-related factors such as gradient, aspect, curvature, and elevation, which are derived from digital elevation models (DEMs). Steeper slopes are more susceptible to instability compared to gradual ones. The terrain in the study region ranges from sharp, steep hills and tall cliffs to narrow valleys and flat plains (Hussain et al., 2022).

2.5.3 Human-induced factors

In addition to natural triggering factors, man-made activities increase the risk of slope instability (Oyda and Regasa, 2023). Human-induced landslides happen when people's actions directly cause landslides. These actions include things like changing the shape of the land, altering how water flows, changing how land is used, and letting infrastructure get old (Jaboyedoff et al., 2018). This is causing a lot of damage to people's lives and the environment. Both physical damage (like buildings collapsing) and social and economic damage (like people losing their homes or jobs) are happening in places where a lot of people live (Ahmed et al., 2022). It's really important to control things that make slopes erode. When there are lots of trees and plants, their roots hold the soil in place, so erosion happens less. However, when slopes don't have any plants, erosion can easily occur, making the slopes unstable (Jennifer, 2022).

2.5.4 Hydrological factors

Landslides can be triggered by seasonal rainfall and how reservoirs are managed. When there is a heavy rain, the soil around the reservoir gets weaker thereby increasing the risk of landslides. But it's not just rain that matters; how the reservoir is filled and emptied also plays a big role. For example, if the reservoir is filled or emptied too quickly, it can put extra pressure on the slopes, making landslides more likely. The type of soil, how steep the slopes are, and how fast the reservoir's water level changes affect how likely landslides are to happen (Zhang et al., 2021). On the other hand, the lowering of a lake or reservoir due to rapid withdrawal by water supply wells or the lowering of the groundwater table can cause slope failure when the buoyancy offered by the water declines and seepage gradients occur. The main effect of water is a reduction in the shear strength of failure surfaces as the effective normal stress (σ_n) decreases (Wubalem, 2021).

The recent landslide in Gerese is a pertinent example of these dynamics. Heavy rainfall likely played a significant role in saturating the soil, weakening its structure. Additionally, if there were rapid changes in the water levels of nearby water bodies or reservoirs, these could have further destabilized the slopes, contributing to the landslide event. Understanding and managing these factors is essential for preventing future landslides, particularly in regions with significant rainfall and vulnerable landscapes. Monitoring soil moisture levels, managing reservoir water levels carefully, and assessing the stability of slopes can help mitigate the risk of such disasters.

Plate 1: Recent landslide triggered by heavy rainfall around Gerese, Gofa Zone.

2.6 Landslide Susceptibility Mapping Methods

Over the past few decades, several landslide susceptibility mapping (LSM) techniques have been implemented (Reichenbach et al., 2018). These methods can be broadly divided into two categories: qualitative and quantitative. Heuristic methods and geomorphological analysis are examples of qualitative approaches , while statistical, artificial intelligence and deterministic methods are examples of quantitative approaches (Shano et al., 2020).

Machine learning is an effective class of data-driven tools that have become increasingly popular among statistical modeling techniques in recent years. Researchers have applied machine learning because their ability to adjust their internal structure to the existing landslide data, their capabilities to extract knowledge from huge databases in an automatic way, and they can build classification and regression, to provide an accurate landslide model, their models are more cost efficient and rapid than conventional models and can be extended to a large area, validation, and the prediction of non-linear problems by using mathematically verifiable and standardized techniques with relaxed assumptions (Youssef and Pourghasemi, 2021). They also focus on prediction by employing general-purpose learning algorithms to identify patterns in frequently rich and unmanageable data, while statistical techniques concentrate on inference, which is accomplished through constructing and fitting a problem-specific probability model (Tehrani et al., 2022). In this sense, machine learning techniques may be more effective at predicting landslide patterns. The majority of the problems indicated to clarify the models' performance are frequently addressed whenever geospatial data-driven investigations are carried out outside of the specialized ML literature.

Abraham et al. (2021) compared five machine learning algorithms for landslide susceptibility mapping (LSM). Their study found that KNN, Random Forest (RF), and

SVM performed better than Naïve Bayes (NB) and Logistic Regression (LR), especially with landslide inventory data. Machine learning methods, particularly KNN and SVM, captured complex patterns more effectively than traditional statistical approaches. These methods provided more accurate and reliable LSM results. Machine learning's ability to handle large datasets offers a clear advantage for such applications.

Yuan et al. (2022) conducted a study using certainty factor-based machine learning methods for landslide susceptibility mapping in Wenchuan County, China. They employed several algorithms, including logistic regression (LR), support vector machine (SVM), and random forest (RF). The study utilized causative factors like: slope, elevation, land use, and rainfall. The results demonstrated that the random forest model provided superior accuracy and predictive capability compared to traditional statistical methods like logistic regression, which had lower generalization ability.

Agrawal and Dixit (2023) used machine learning methods to assess landslide susceptibility in the Meghalaya-Shillong Plateau region. They employed several algorithms, including artificial neural networks (ANN), extreme gradient boosting (XGBoost), and k-nearest neighbors (KNN).

In, conclusion the literature review highlights the urgent need for better landslide studies in Ethiopia, focusing on using machine learning techniques. Traditional methods have provided a basic understanding of using machine learning techniques. However, Ethiopia's complex and varied nature of landslides requires more advanced methods. Machine learning, which can handle large amounts of data and find hidden patterns, offers a promising way to improve landslide prediction and risk assessment.

Table 3: Methods used for landslide analyses (advantages and disadvantages) including the working scales.

Method	**Type**	**Advantages**	**Disadvantages**	**Working Scale**
Logistic Regression	Statistical	-Simple and interpretable - Well-suited for binary classification - Low computation cost	- Assumes linearity - Limited in handling complex non-linear relationships	Local to regional
Frequency Ratio	Statistical	- Simple and easy to implement - Requires minimal data	- Ignores interaction between variables - Limited prediction accuracy	Local to regional
Weight of Evidence (WoE)	Statistical	- Good for qualitative data - Handles multi-variable relationships well	- Assumes conditional independence - Performance decreases with dependent variables	Local to regional
Random Forest	Machine Learning	- High accuracy - Robust to overfitting - Can handle large, noisy data	Computationally expensive - Less interpretable than linear models	Local to regional
Support Vector Machine	Machine Learning	- Effective for high-dimensional data - Works well for both linear and non-linear data	- Difficult to tune parameters - Slow with large datasets	Local to regional
K-Nearest Neighbors	Machine Learning	- Simple and effective	-Computational complexity - Sensitive to outliers	Local to regional

2.6.1 Machine Learning Methods

Machine learning methods rely on certain statistical ideas. These methods follow preset rules and assumptions to produce results. They work well when there isn't a clear, direct mathematical connection between what causes something and its effects (Chowdhury and Sadek, 2012). There are a number of machine learning methods that can be used for landslide studies. These can be categorized as Support Vector Machine (SVM), Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and Random Forest (RF) (Đurić et al., 2019; Kavzoglu et al., 2019).

Tehrani et al. (2022), there has been a significant increase in the number of scientific studies published in recent years focusing on landslide susceptibility assessment through the use of machine learning (ML) algorithms. A basic search conducted in the Scopus database on November 16, 2020, using the keywords "landslide" and "machine learning," restricted to journal articles, yielded 286 results. Among these, approximately 64% (186 articles) were dedicated to applying ML algorithms in landslide susceptibility modelling.

Rainfall-induced landslides occur when heavy rains reach a certain level, making landslides likely. Over the past few decades, researchers have mostly figured out these thresholds by observing them happen or using machine learning to analyze rainfall patterns (Segoni et al., 2018). Recently, Machine learning methods have been investigated for this purpose. For example, many reserchers have utilized the Support vector machine (SVM) to establish rainfall thresholds (Tehrani et al., 2022).

While there has been notable progress in machine learning methods, there remains a gap in their application within the specific study area, hindering advancements in this field. This thesis aims to provide machine learning techniques to fill the gaps in the study area.

2.6.1.1 Support Vector Machine (SVM)

Cortes and Vapnik (1995) were the ones who brought up support vector networks in the beginning. Based on statistical analysis, a support vector machine (SVM) determines which hyper-plane is optimal for splitting two groups. We refer to the optimal decision boundary as a hyperplane. The optimal hyper-plane to split the dataset into two classes, one with landslides and one without, is found using the SVM training approach. Support vectors are the observations that are near the border planes, and the margin is the space that separates them. Another name for it is the support vector network. Although it is a non-probabilistic binary linear classifier, implicit mapping allows it to be used as a non-linear classifier. It uses hyperplanes to show linear separability in high dimensions. It functions similarly to a maximal marginal classifier in this sense. The SVM classifier does not penalize small when used for regression; instead, it conducts linear regression in the high-dimensional vector space. The goal of SVM is to maximize the margin between several classes. As a result, it functions as a discriminative classifier. SVM is a supervised learning technique that uses labeled training data to generate projected plotting functions from a dataset (P. K. Singh et al., 2019).

Support Vector Machine (SVM) is the most popular state-of-the-art machine learning technology. Support-vector machines in machine learning models that examine data for regression and classification along with related learning techniques. Using a technique known as the "kernel trick," SVMs can effectively conduct non-linear classification in addition to linear classification by implicitly translating their inputs into high-dimensional feature spaces. In essence, it draws borders between the various classes. The margins are created to minimize the classification error by creating the greatest possible space between the margin and the classes (Mahesh, 2018). The SVM model functions effectively when dealing with unstructured data we are unaware of. The SVM algorithm's undeniable strength is its kernel approach, which helps solve any challenging problem (Jennifer, 2022).

2.6.1.2 Artificial Neural Network (ANN) Method

ANN method represents a mapping of landslide susceptibility from one multivariate space of information into another by providing a set of data or information relating to representative mapping (Garrett, 1994; Nefeslioglu et al., 2008; Pradhan and Lee, 2010). The backpropagation learning algorithm of the artificial neural network (ANN) approach is utilized to solve such non-linear complicated interactions between the components and the landslides. The ANN method may create rules for weight assignment for each factor.

Backpropagation is employed in artificial neural networks (ANNs) to ascertain a gradient necessary to compute the network's weights. Backpropagation is a shortcut for backward propagation of landslide susceptibility and hazard zonation mistakes. To do this, an error is calculated at the output and propagated backwards via the network's layers. The artificial neural network (ANN) simplifies the management of databases containing qualitative or quantitative data. Data from the entry nodes may be direct, reclassified, qualitative, or quantitative. Landslide densities of various kinds or data on determinant causal factors might be assigned to the entrance nodes for landslide susceptibility studies.

2.6.1.3 K-Nearest Neighbors (KNN)

K-Nearest Neighbors has been recently used in landslide susceptibility mapping and is the easiest of all ML algorithms (Miner et al., 2010). The k most similar training illustrations within the given period make up the majority of the input in a KNN. The outcomes are the potential to become members; in categorized snags, the potential to become a member denotes the ambiguity of an item assigned to a given class (Hussain et al., 2022). In KNN (K-Nearest Neighbors), an entity is classified based on the majority class of its closest neighbors. The entity is placed into the class most of its nearest neighbors belong to, "k" representing the number of neighbours considered. Typically, k is a small, positive number. If k = 1, the entity is simply assigned to the same class as its single closest neighbor (Kumar et al., 2023).

KNN is often described as an instance-based, nonparametric, "lazy" algorithm. The term "lazy learning" means that the model doesn't generalize or learn from the data until it needs to make a prediction, unlike "eager learning," where the model actively learns from the data during the training phase (Hussain et al., 2022).

2.6.1.4 Random Forest (RF)

RF is a data mining approach that uses an ensemble of decision trees to reliably classify vast volumes of data (Breiman, 2001). Decision trees are prediction models that identify a target variable by applying a set of binary rules. The target variable, which is being predicted, plus a collection of predictor variables; make up the training data for the model. Using the predictor variables, the decision tree's goal is to divide the data into homogeneous datasets that are related to the objective variable. In the simplest scenario, the target variable will be a binary classification (e.g., landslides present or absent). The root node contains both target and predictor data. Next, the algorithm evaluates each predictor variable's capacity to split the target variable into two classes. The procedure is repeated for each additional node, starting with the predictor variable that produces the most accurate classification (Taalab et al., 2018).

CHAPTER THREE

3. MATERIALS AND METHODS

3.1 Description of the study area

3.1.1 Location and Accessibility

Location of the study area for this proposed research was located in South Ethiopia Regional State (Figure 1). It is founded in the southern parts of the Ethiopian highlands, around 437 km from Addis Ababa, the capital city of Ethiopia. Geographically, it lies between 5°47’0” to 6°4’30” N and 37°17’0” to 37°32’0” E. The study area covers 556 square kilometres.

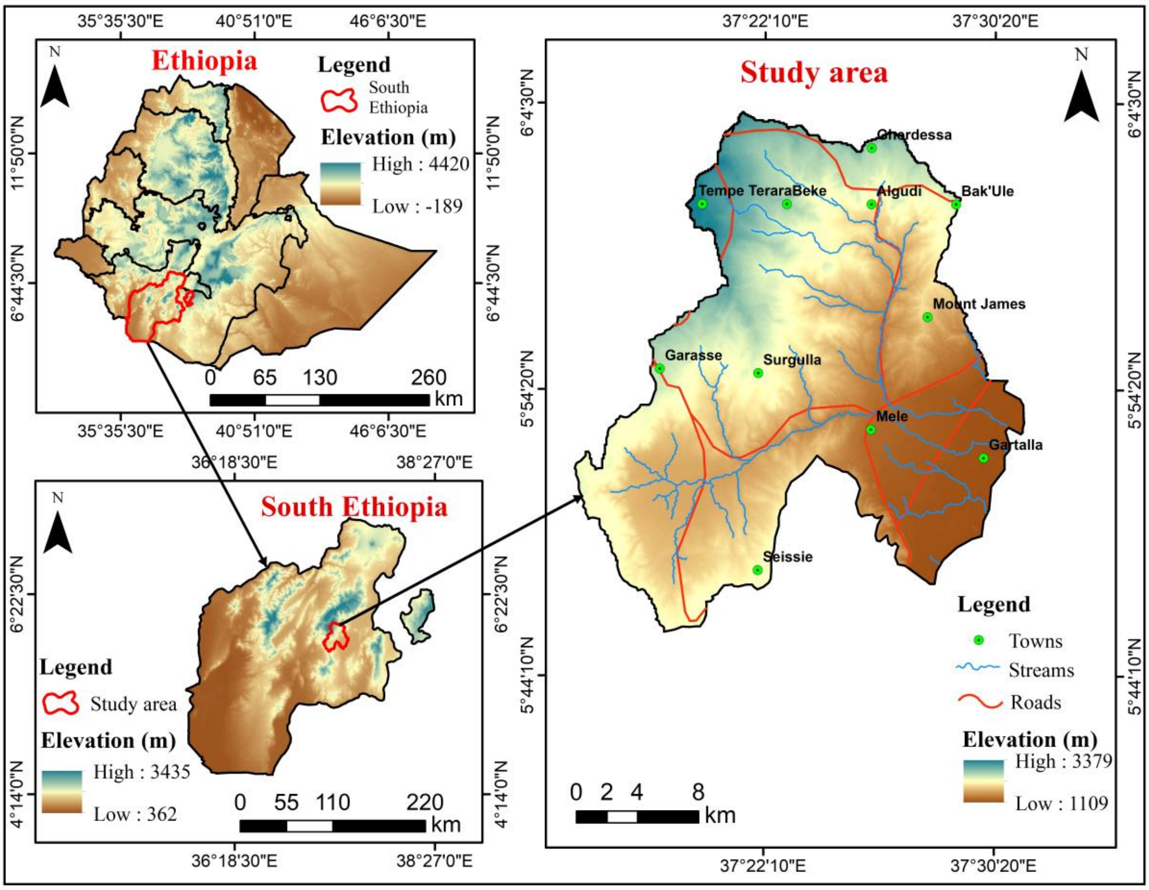


Figure 1: Location and accessibility map of the study area.

3.1.2 Physiography

The Southern Ethiopian highlands, concentrated on Gamo Highland, which has a peak elevation of roughly 3379 meters and a lowest point of 1109 meters above mean sea level, making up the majority of the study area's physiography. The highest elevation is near Guge Mountain, while the lowest point is near Chamo Lake.

Steep scarps, cliffs, jagged ridges, rough slope faces, deep gorges, and valleys are common in the area (Figure 2), indicating that surface dynamics and landslide events are likely to occur here.

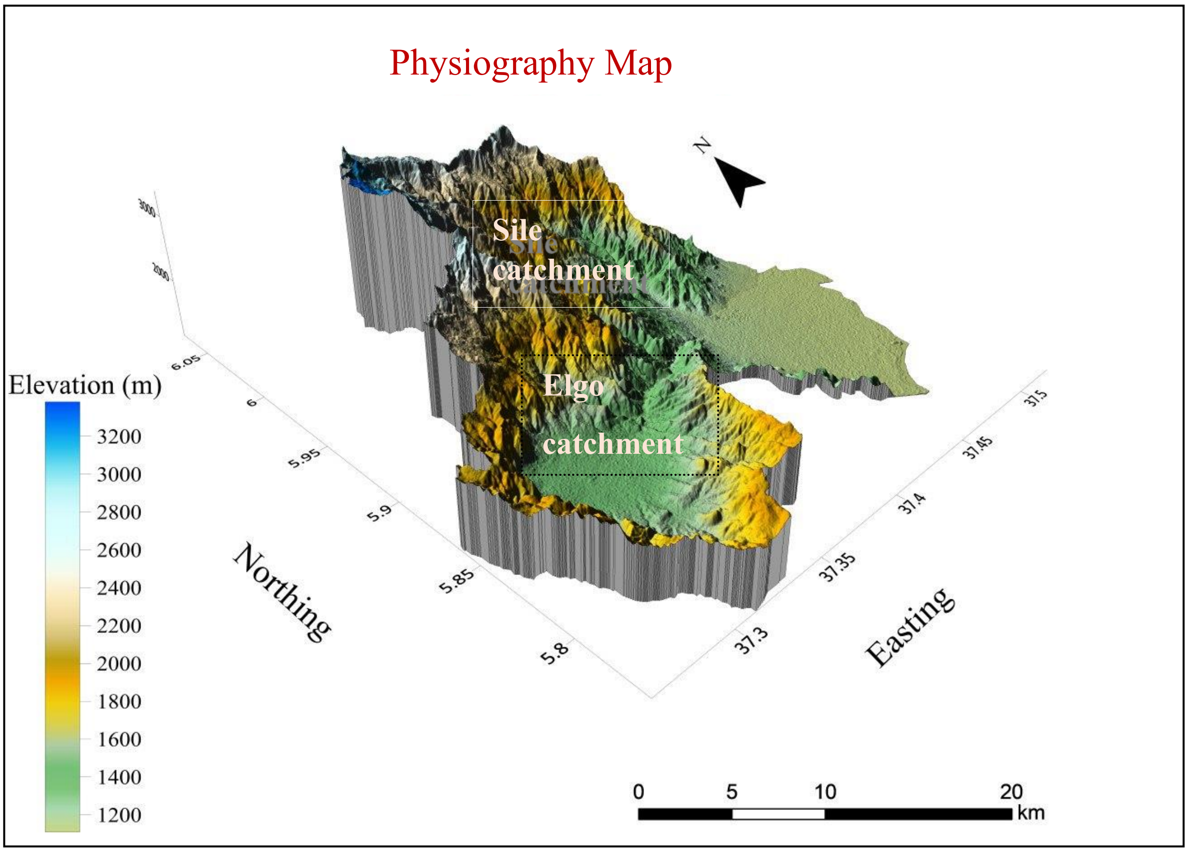


Figure 2: Physiography map of the study area.

3.1.3 Drainage network

The drainage network within the study area forms a vital component of the local hydrological system. The study area is a critical source for collecting and distributing surface water within the region. The network comprises an arrangement of streams and rivers, dynamically interconnected to facilitate water flow towards Lake Chamo (Figure 3); with its diverse topography characterized by steep scarps, rugged slopes, and deep valleys, the drainage network navigates through various terrain types, shaping the landscape and influencing the distribution of water resources. The convergence of runoff from elevated regions to lower-lying areas, contributes to the formation of streams and rivers, enhancing the drainage network's efficacy in channeling water toward the lake Chamo. Moreover, natural features like cliffs and gorges further influence the flow patterns, potentially facilitating erosion and sediment transport processes along the drainage pathways. Overall, the drainage network of Sile and Elgo catchments plays a pivotal role in regulating water movement and sustaining the hydrological balance within Lake Chamo.

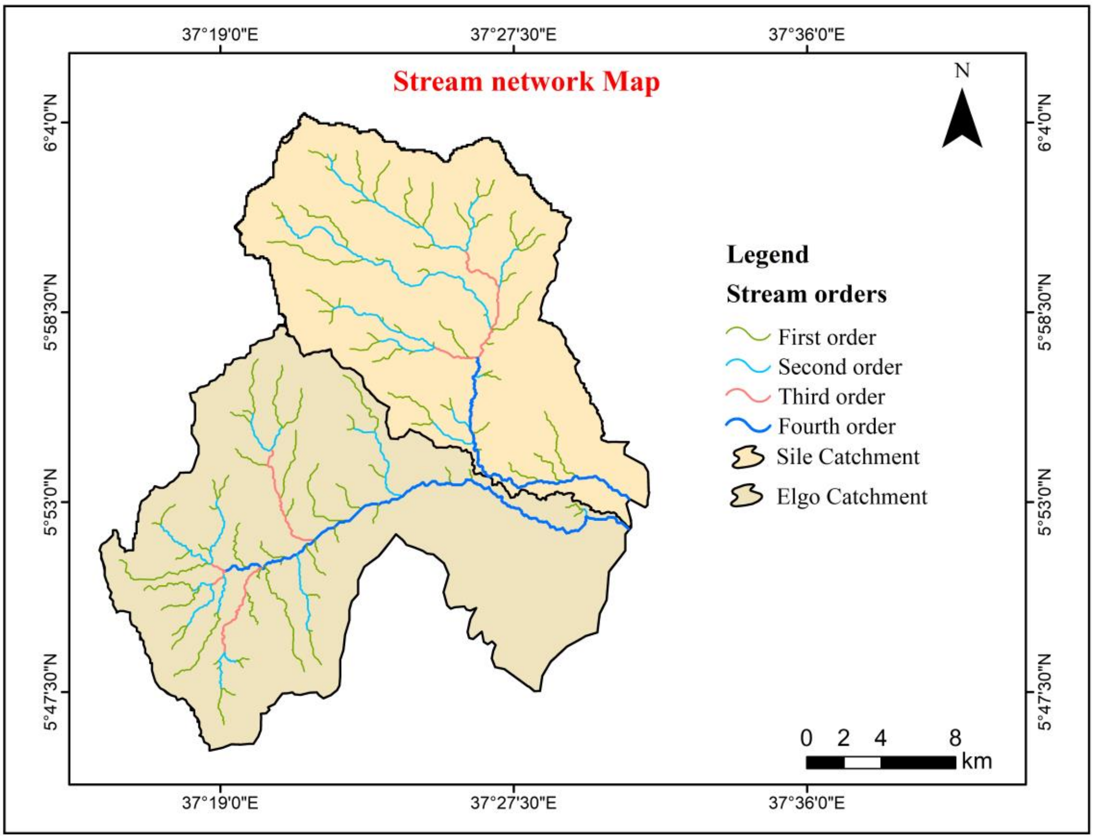


Figure 3: Stream network map of the study area.

3.1.4 Rainfall distribution

The study, used NASA satellite-based rainfall data to enhance and verify the rainfall records from local meteorological stations in the Gamo Zone. This data was especially useful in regions where ground-based measurements were limited or incomplete. By incorporating NASA's satellite data, was able to cross-reference and validate the annual and seasonal rainfall trends observed at stations in towns such as Arba Minch, Arfaide, Chencha, Kemba, Malanaha, Gerese, and Arguba. The maximum annual rainfall recorded at these stations is 1866.9 mm/10yrs, 1391.7 mm/10yrs, 1254.9 mm/10yrs, 1806.6 mm/10yrs, 1484.4 mm/10yrs, 2030.26 mm/10yrs, 1722.7 mm/10yrs respectively.

The Gamo Zone is mostly found in zone C, which has distinct seasons with bimodal precipitation patterns interspersed with dry periods. April and May saw the first rainfall period, followed by September and October (GIE, 2023). As indicated in (Figure 4a) in this particular region, rainfall predominantly occurs in two distinct periods throughout the year. The first rainy season typically falls between April and May, while the second occurs during September and October (Shano et al., 2021).

In the rainfall assessment across various stations, April emerged as the month with the highest precipitation for several locations. Arba Minch recorded its peak rainfall in April at approximately 169.64 mm, while Arfaide saw a maximum of 183.86 mm during the same month. Chencha reached its highest rainfall in May, measuring about 167.4 mm. Kemba observed the highest overall rainfall with 286.73 mm in April, indicating a significant peak for the region. Similarly, Malanaha recorded its maximum in April with 201.85 mm, Gerese experienced a peak of 282.88 mm, and Arguba (as shown in Figure 4b) reached 176.68 mm. Among all these stations, Kemba registered the greatest rainfall, marking April as the primary month for the rainy season in the area.

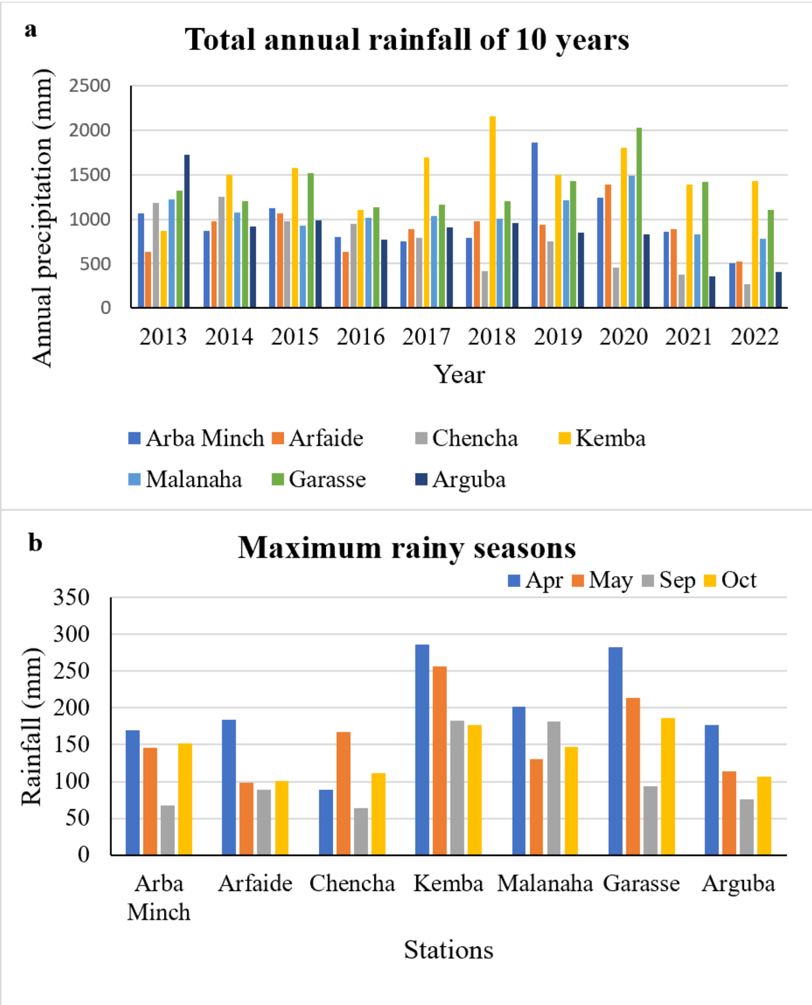


Figure 4: (a) Total yearly Rainfall, (b) Average maximum monthly rainfall (Source: NASA).

3.1.5 Temperature

A slope's stability can be impacted by temperature fluctuations in several ways. Temperature variations can affect groundwater levels and the number of plants. The variation is reflected in the study area's climate characteristics, which are based on data from the National Meteorological Agency's meteorological stations in the towns of Arba Minch, situated on low land, and Chencha, located on highland area. For instance, the typical yearly temperature of Arba Minch is about 21 ºC, with March and April being the warmest months with mean daily maximum temperatures that sometimes reach above 30 ºC. July, on the other side, has the lowest mean daily minimum temperatures, dropping to 18 ºC, making it the coldest month. Generally, the temperature difference between day and night is between 16 and 32 ºC (GIE, 2023).

3.1.6 Seismicity of the study area

Seismicity is a common geohazard risk phenomenon in the Gamo Zone environment, in most cases associated with normal faults with few exceptions of strike-slip faults based on focal mechanism-based inversions in southwestern Ethiopia, including the southern MER. Four earthquakes of magnitudes ranging from 5.3 to 6.3 were recorded between 1985 and 1987 in the vicinity of the Gamo Zone (Ayele and Arvidsson, 1997). Most earthquakes in Ethiopia, the southern Main Ethiopian Rift, are usually linked to normal faults, with a few occurring on strike-slip faults. According to Figure (5) the study area has been located in the zones of moderate to considerable damage.

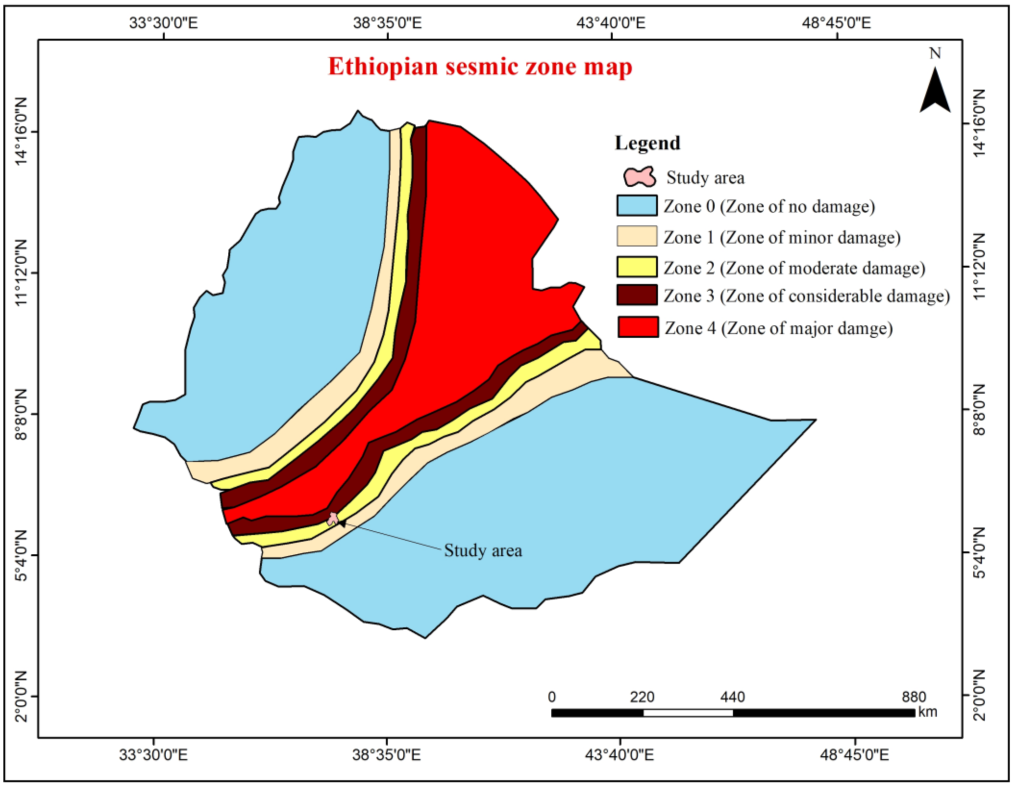


Figure 5: Seismic hazard map of Ethiopia for a 100-year return period (EBCS, 1995).

3.1.7 Geology

3.1.7.1 Regional Geology

The Gamo Zone's geological story shows how the land changed over time. The oldest rocks are in the southwest, and the newest ones are in the east, near the Main Ethiopian Rift and Abaya and Chamo lakes. Long ago, when the Gondwana supercontinent formed, a process called the East African Orogeny made different kinds of rocks under high heat and pressure.

We can still see signs of this process in the Proterozoic gneissic complex and plutonic rock formations, especially in the Hammar Domain in the southwest of the Ethiopian Shield (Verner et al., 2021).

In the southern main Ethiopian rift, volcanic activity is categorized into three distinct episodes: pre-rift volcanic activity during the Eocene to Oligocene period (approximately 45 to 27 million years ago), syn-rift volcanic activity during the Miocene period (around 24 to 11 million years ago), and post-rift volcanic activity during the Pleistocene epoch.

The pre-rift volcanic sequence, associated with basic volcanoes, has a minimum thickness of about 600 meters and is primarily composed of tholeiitic and transitional basalts. Additionally, the Shole Ignimbrites, or Amaro Tuffs, which have a rhyolitic composition, overlay this sequence with a thickness of approximately 200 meters. The pre-rift volcanic sequence also consists of amygdaloidal basalt lavas and pyroclastic deposits, known as the Amaro-Gamo Basalt, which date back to the Eocene period. These formations are primarily located around Lake Abaya and are followed by Eocene to Lower Oligocene rhyolitic ignimbrites, termed the Shole Ignimbrite, which are distributed across the sutudy area. Additionally, minor rhyolite lava flows and breccias, known as the Ugayo rhyolite, are present in the region. The Shole Ignimbrites are subsequently covered by several extensive basalt lava flows from the Lower Oligocene age, predominantly found in the northeastern part of the Arba Minch map sheet (George and Rogers, 2002; Rooney, 2010).

The origin of the syn-rift volcanic sequence is represented by minor deposition of the Dorze Ignimbrite (~30 meters in thickness) and the Mimo Trachyte lava flow (~10 meters in thickness), both revealing the Miocene age associated with the initial stage of the rift subsidence (GIE, 2023).

The Pleistocene post-rift volcanic deposits, known as the Dino Formation, mark the culmination of the region's volcanic activity. These deposits, situated in the lowest part of the rift structure, primarily consist of basaltic lava flows and remnants of cinder cones. A notable feature is the presence of unconsolidated pyroclastic deposits on the rift floor, formed during the early Pleistocene and subsequently disturbed by tectonic shifts and erosion. This phenomenon exemplifies the post-rift volcanic processes characteristic of the southern main Ethiopian rift (Martínek et al., 2021). The area's geology is mostly composed of intrusive and extrusive igneous rocks, such as ignimbrite with a trachytic to rhyolitic composition (ITR), basalts, and trachyte lava flows (BTL) (Martínek et al., 2021; Belayneh et al., 2022).

3.7.1.2 Geological structures

The geological map of the Gamo Zone is situated in the southern part of the NNE–SSW trending Southern Main Ethiopian Rift belonging to the East African Rift System. Due to the extension between the Nubian and Somalian lithospheric plates, the rift structure reflects a characteristic evolution of continental rifting from early plateau basalt lava flows (pre-rift stage), followed by the forming of fault-dominated rift morphology in the early stages of the continental extension (Martínek et al., 2021).

Across the study area, several sets of brittle structures (e.g., normal faults to faults covered by younger deposits, lineaments, beddings, and magmatic foliations) were observed. Regional faults and fault zones related to the Main Ethiopian Rift significantly affect the geological framework of the area and are mostly parallel with the NNE (N) to SSW (S) oriented axis of the rift and morphological escarpments (Figure 6).

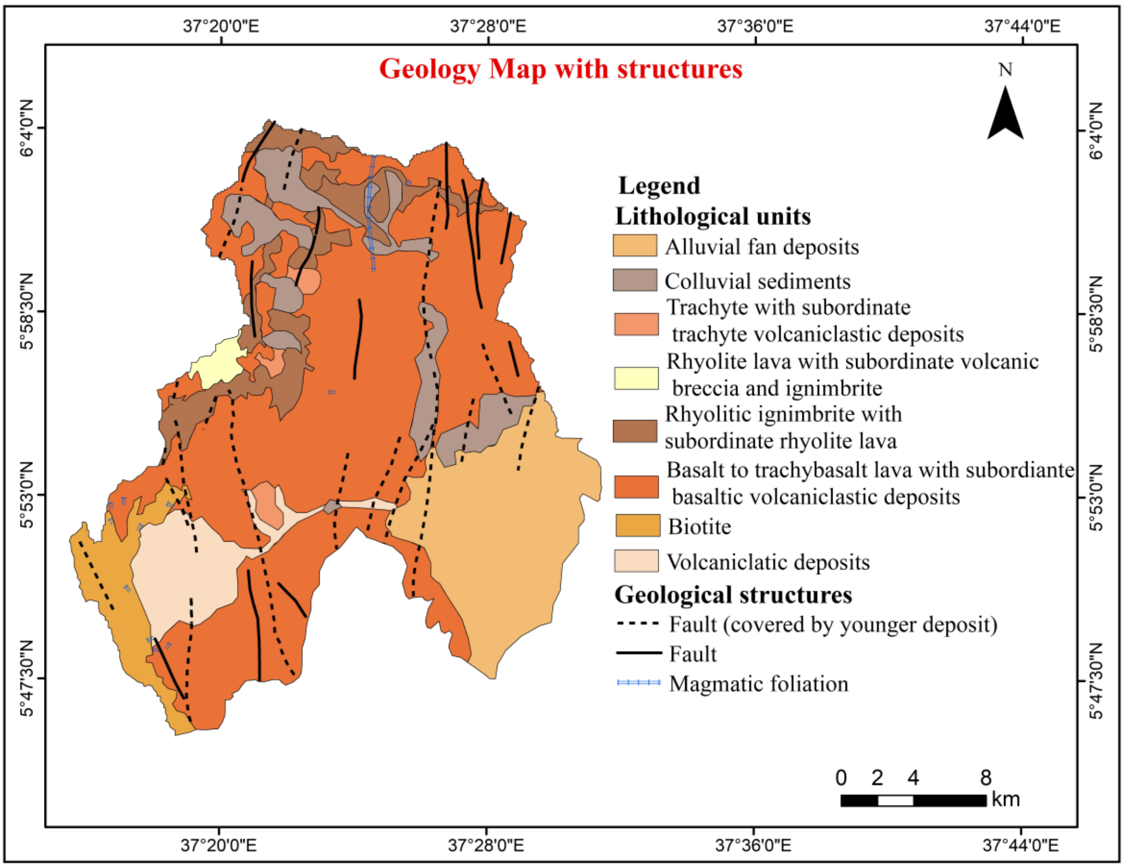


Figure 6: Lithology with Geological Structures Map of the Study area.

3.2 Materials

3.2.1 Software's

As indicated in (Table 4) in this thesis, various software tools were employed to address different aspects of landslide susceptibility analysis. Google Earth Pro is used to digitize landslide points and lineaments, utilizing its high-resolution satellite imagery and user-friendly interface to mark and analyze geographical features related to landslides accurately. Sulfur assists in preparing a physiography map, which is essential for understanding the physical landscape supporting the assessment of landslide-prone areas. QGIS (version 3.14), an open-source geographic information system, is instrumental in preparing landslide points and assigning weights to causative factors with inventory points. This tool is crucial for plotting landslide occurrences, analyzing spatial relationships, and integrating various spatial data. ArcGIS (version 10.8) used to create detailed causative factor maps, highlighting the different factors contributing to landslides. RStudio was utilized to generate landslide susceptibility maps because of its robust statistical and graphical capabilities to implement complex models and algorithms for predicting landslide susceptibility. By analyzing spatial and non-spatial data, RStudio helps produce detailed maps highlighting areas at risk, aiding in targeted mitigation efforts. Microsoft Excel organizes rainfall and landslide data, making use of its powerful data management and analysis features to handle large datasets efficiently. IBM SPSS performs multicollinearity analysis for causative factors and calculates the Area Under the Curve (AUC) for susceptibility maps. SPSS's advanced statistical analysis capabilities help identify and address multicollinearity issues among causative factors, ensuring the reliability of predictive models. Additionally, By integrating these tools, this thesis comprehensively addresses the various aspects of landslide susceptibility, from data collection and mapping to advanced statistical analysis and model validation.

Table 4: Different software's and their purpose.

Software's	**purpose**
Google Earth pro	To digitize Land slide pointsand lineament.
Surfur	To prepare a Physiography map.
QGIS (3.14)	To prepare landslide points and to weight causative factors with inventory points.

ArcGIS (10.8)	To prepare causative factor maps.
RStudio (4.2)	To generate landslide susceptibility maps.
Microsoft Excel	To organize rainfall and landslide data.
IBM Spss 22	To prepare multicollinearity analysis for causative factors, and to create AUC for both susceptibility maps.

3.3 Methods

3.3.1 Data collection

This part involved the collection of spatial and non-spatial data from primary and secondary sources. The initial steps for the landslide study. Sources and types of data are presented in Table 5.

Table 5: Primary and secondary data sources.

Types of data	Sources	Purpose
Published and unpublished papers	Google Scholar	To have a general background and a conceptual framework for the study.
Geological map	Geological Institute of Ethiopia (GIE)	To digitize Lithology.
DEM Images	USGS	To digitize Slope, Aspect, Curvature.
Metrological data	Power larc NASA	To prepare an Annual rainfall Map.
Google Earth images	Google Earth pro	To digitize landslide inventory, Lineament.
Soil map	GIE	To digitize soil.
LULC map	Sentinel satellite image	To prepare LULC.

This research data collection involved gathering both primary and secondary data from various sources. Secondary data included a comprehensive literature review of published and unpublished papers. Regional geological maps, topographic maps, and meteorological data were obtained from the Geological Institute of Ethiopia, the Ethiopian Space Science and Geospatial Information Agency, and the National Metrology Agency of Ethiopia, respectively. A Digital Elevation Model (DEM) with a spatial resolution of 30 × 30 m was used to generate topographic parameters such as slope, aspect, and curvature.

Primary data, including landslide inventory maps, land use/land cover (LULC), and lineaments, were prepared using Google Earth Pro and Sentinel satellite images, respectively. The relationship between factor classes in the study area was analyzed using frequency ratio. Weights were generated using both causative factors and landslide inventory data within QGIS. The mode value was used for categorical data, while the mean value was used for continuous data. These values were then transferred to an Excel sheet to produce the final landslide susceptibility map using RStudio.

Validation was conducted using the susceptibility index obtained from the two models. The accuracy of the results from the two models was evaluated through Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC).

3.2.2 Data preparation

This phase is crucial for analyzing the thematic data in the case study area, such as records of landslides and factors that contribute to them, and for building a dataset that is appropriate for building a machine learning model. The preparation of the data step is presented in (Figure 7).

The procedures utilized in this phase were tried to provide the most effective set of input variables (features) that correspond with the output variable (goal). The primary step in initiating the analysis involved gathering landslide inventory data by utilizing Google Earth images and preparing causative factor maps.

3.2.2.1 Landslide Inventory Mapping

An extensive database recording previous landslides' frequency, features, and characteristics is called a landslide inventory. It is used as the main dataset in landslide susceptibility mapping. The creation of landslide inventories is a crucial step in the training and testing of the model.

A unique identification code is usually given to each entry in the inventory, which also contains important information, including the type of landslide, its activity level, the date it occurred, and the material involved. How these inventories are organized depends on the type of event being recorded; they might be polygonal or point-based (Loche et al., 2022).These maps display the land's apparent landslides (Yalcin, 2005). Making the landslide susceptibility maps requires accurate determination of the landslide's position and area (Yalcin and Bulut, 2007).

Using Google Earth imagery is a straightforward way to identify landslide boundaries and types and to gather a time-series inventory of landslides across different sizes by utilizing time-series slides (Hamedi et al., 2022).

Google Earth Pro image interpretation was used to get landslide data for the study area with respect to the landslide distribution area and type of landslides.

In the process of creating a landslide susceptibility map, landslide and non-landslide points were divided into two sets: training and validation datasets. The training dataset was utilized to construct the model, while the validation dataset was employed for prediction.

The study area predominantly experiences landslides in its western, northern, and central regions, characterized by rugged, hilly terrain. This terrain features steep slopes, deep gorges, and weathered rocks that are fractured and exhibit a high relative relief.

3.1.2.2 Landslide Influencing Factors

In this analysis, ArcGIS 10.8 was utilized to create a map of landslide causative factors. Detailed remote sensing analysis was conducted to identify the relevant factors, resulting in the selection of ten key causative factors: slope, aspect, curvature, distance to stream, distance to lineament, soil type, lithology, rainfall, elevation, and land use/land cover (LULC). These factors were classified using Natural Breaks (Jenks).

All maps were projected using the Adindan UTM Zone 37N coordinate system and converted into raster format with a uniform cell size of 30 by 30 meters. Subsequently, a layer-overlay procedure was executed to consolidate and combine the data layers. Values for the chosen ten conditioning factors were extracted in a GIS environment using landslide inventories. For continuous data, the mean values were computed, while the mode values were computed for categorical data variables such as LULC, geology, and soil type, a data pre-processing technique was applied to convert these variables into numerical format to ensure compatibility with the RStudio environment, which requires all variables to be represented numerically. The weight of each causative factor was then transferred into an Excel sheet. Finally transferred into RStudio to map the final susceptibly map.

3.2.3 Multicollinearity Analysis

Multicollinearity was assessed using the variance inflation factor (VIF) in the least squares regression analysis. A variable is considered perfectly multicollinear if it can be completely predicted by other independent variables, indicating no margin of error. Conversely, a variable is entirely uncorrelated with different independent variables if its tolerance value is 1.

The study identified multicollinearity based on VIF and Tolerance values, which were calculated as follows:

$$\text{Tolerance} = 1 - R\,j^2 \quad \ldots\ldots (1)$$

$$\text{VIF} = 1 / \text{Tolerance} \quad \ldots\ldots (2)$$

Where, Rj^2 = coefficient of determination (R-squared) of the model of the descriptive variable j as the response variable and other descriptive variables as independent variables.

3.2.4 Correlation between landslide and causative factor using frequency ratio

The Frequency Ratio (FR) method explored and quantified the non-linear correlations between landslide occurrences and various environmental factors. According to Shu et al. (2021), this method involves dividing causative factors into intervals using different division methods. The frequency ratio value for each interval is calculated by determining the probability of landslide occurrence within those intervals. This statistical analysis enhances the accuracy of interval classification, making it a robust tool for understanding the relationship between landslides and their contributing factors (Z. Huang et al., 2023). When FR > 1, it indicates that the grading interval is conducive to landslide hazards. If FR = 1, the likelihood of landslide hazards in the grading interval matches the average probability across the study area. Conversely, if FR < 1, the grading interval is less favourable for landslide hazards(Mandal and Mandal, 2018). The formula is represented as follows:

$$\text{Frequency Ratio (FR)} = \frac{\text{Slide Ratio}}{\text{Class Ratio}} \quad \ldots\ldots (3)$$

$$\text{Slide Ratio} = \frac{\text{Number of landslide pixel in class}}{\text{Total number of landslide pixel}} \quad \ldots\ldots (4)$$

$$\text{Class Ratio} = \frac{\text{Number of pixel in individual class}}{\text{Total number of pixels in the whole class}} \quad \ldots\ldots (5)$$

3.2.5 Assigning machine learning techniques

Optimizing hyperparameters can enhance the accuracy of machine learning models by selecting the best parameter values based on evaluation metrics. Models built with machine learning algorithms can become more accurate with the use of hyperparameter optimization. The procedure seeks to determine the best hyperparameter values based on the assessment index (Adnan et al., 2020).

3.2.5.1 Support Vector Machine (SVM)

The SVM algorithm's undeniable strength is its kernel approach, which helps to solve any challenging problem. The kernel function and ideal parameters must be determined to assess landslide susceptibility using SVM (Huang and Zhao, 2018; Jennifer, 2022).

The basic concepts of support vector machine (SVM) were followed by Lewis et al. (2006).

- The separating hyper-plane: The idea of SVM is to determine a line or hyper-plane in higher-dimensional space that can separate the input data into classes;
- Maximum-margin hyper-plane: We can have many hyper-planes to separate different classes. Training the SVM aims to determine the hyper-plane with maximum margin where the margin represents the distance from the hyper-pane to the nearby data points (also known as support vectors). The hyper-plane with maximum margin from the data points is the best class-separator selected for the classification;
- The soft margin: If the data points are linearly separable, the maximum-margin hyperplane can separate the classes, but data in real-life applications might not be linearly separable. Soft-margin tolerates few miss-classification or anomalies, allowing the model to generalize better in the case of linearly inseparable data;
- The kernel function: The kernel in SVM is a mathematical function that accepts the input data and transforms the data into the desired form. Generally, the data are transformed to a higher dimension space, where non-linear data becomes separable.

The primary SVM classification function is defined in Equation (6) where a number of coordinates are given by n, parameters in vector x in original space are given by xi, class label by y, parameters of the hyper-plane are weight w and bias b, and the sign function by sgn.

$$y = \mathrm{sgn}(f(x)) = \mathrm{sgn} \sum_{i=1}^{n} (w_i x_i + b) = \mathrm{sgn}(w \cdot x + b) \quad \text{……………………………………(6)}$$

To deal with real-time noisy and non-linearly separable classification datasets, a soft margin is introduced to accommodate empirical errors in the classification model. Equation (6) Now becomes Equation (7) The weight is given as $w = \sum_{i=1}^{m} \alpha_i y_i x_i$, $i = 1, ..., m$ where α_i denotes the weight of the ith training example as a support vector.

$f(x) = \text{sgn } m \sum i=1\ \alpha iyi(xi \cdot x)+b$ ………………………………………………..................(7)

Again, to deal with non-linear decision boundaries, SVM can map data points (xi) into a higher-dimensional space where the data can be linearly separated, and xi is replaced by φ(xi) where φ produces the higher-dimensional mapping. Furthermore, by using the kernel function, which works on lower-dimension vectors (xi, xj) it is possible to make a value identical to the dot product of higher-dimensional vectors (φ(xi) · φ(xj)). The decision function using the kernel trick is given as Equation (8):

$f(x) = \text{sgn } m \sum i=1\ \alpha iyik(xi , x) + b$………………………………………………………. (8)

3.2.5.2 K-Nearest Neighbors (KNN)

The majority vote of its neighbors determines an entity in KNN; it is assigned to the broadest class of its nearest neighbors.

As stated by (Hussain et al., 2022), the KNN model requires the following steps:

- Receiving an unsystematic data set;
- Calculating the distance using a variety of methods;
- Determining the k-closest neighbors; in order to vote for labels.

Finding the training dataset for the K neighbors that are closest to the place of concern is the only task that matters. Determining the K nearest neighbors for a particular data point is hence crucial in the KNN technique. This is based on the idea that related features are closer together. Close describes a measurement of distance that could be as simple as the Euclidean distance between two places (Rebala et al., 2019).

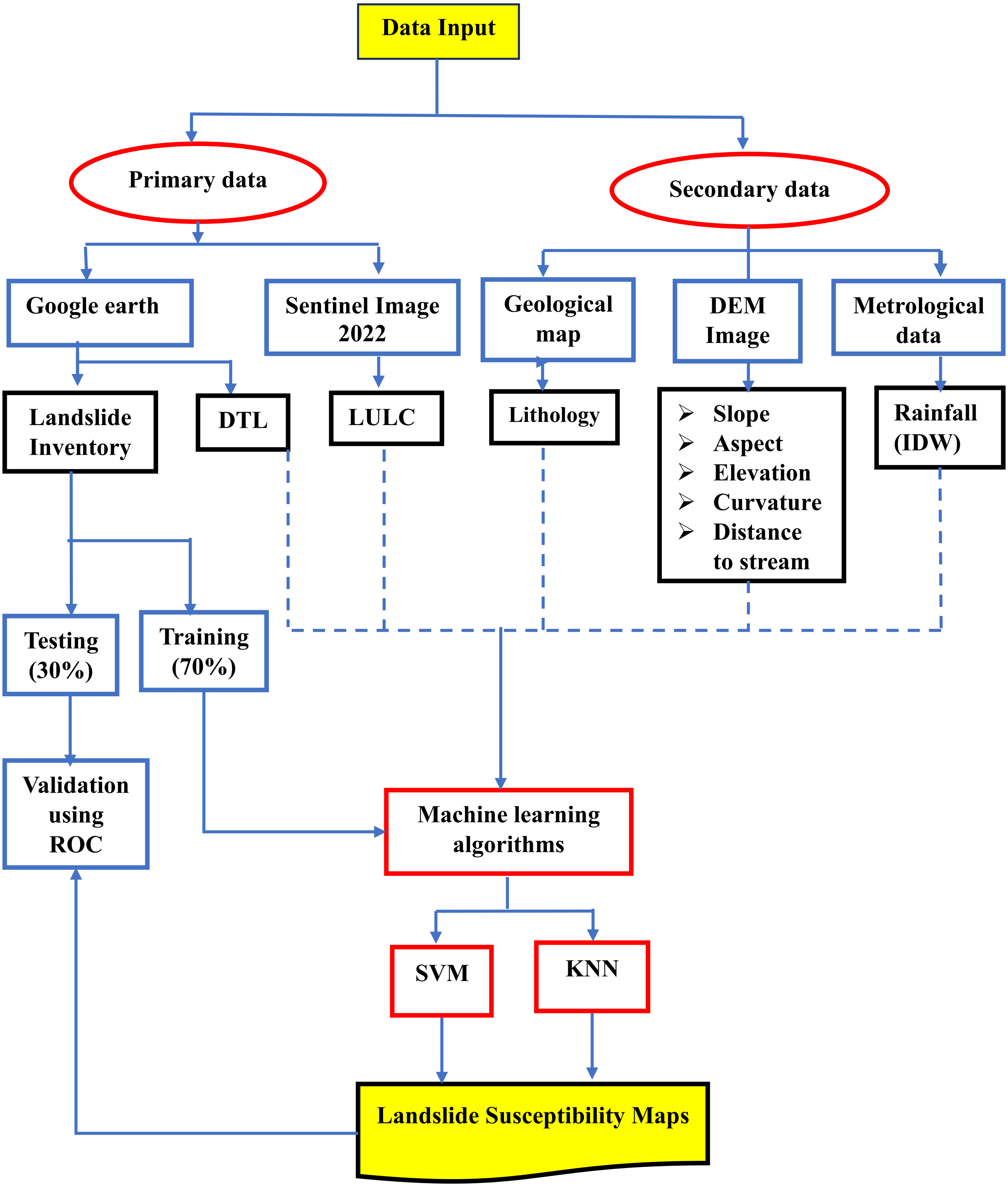


Figure 7: Methodological flow chart.

CHAPTER FOUR

4. RESULTS AND DISCUSSION

4.1 Landslide Inventory Map

This study employs classification-dependent landslide susceptibility mapping by integrating binary data; 534 landslide posints were digitally recorded using Google Earth images. To ensure unbiased and accurate results, an equal proportion of 534 non-landslide locations, identified through QGIS by using the K-means clustering algorithm, were selected to match the positive landslide samples.

ArcGIS software was used to divide the data into training and validation sets. The dataset, was split into training (70%) and testing (30%) sets as indicated in (Figure 8). For binary classification, non-landslide locations were assigned a value of 0 and landslide locations a value of 1.

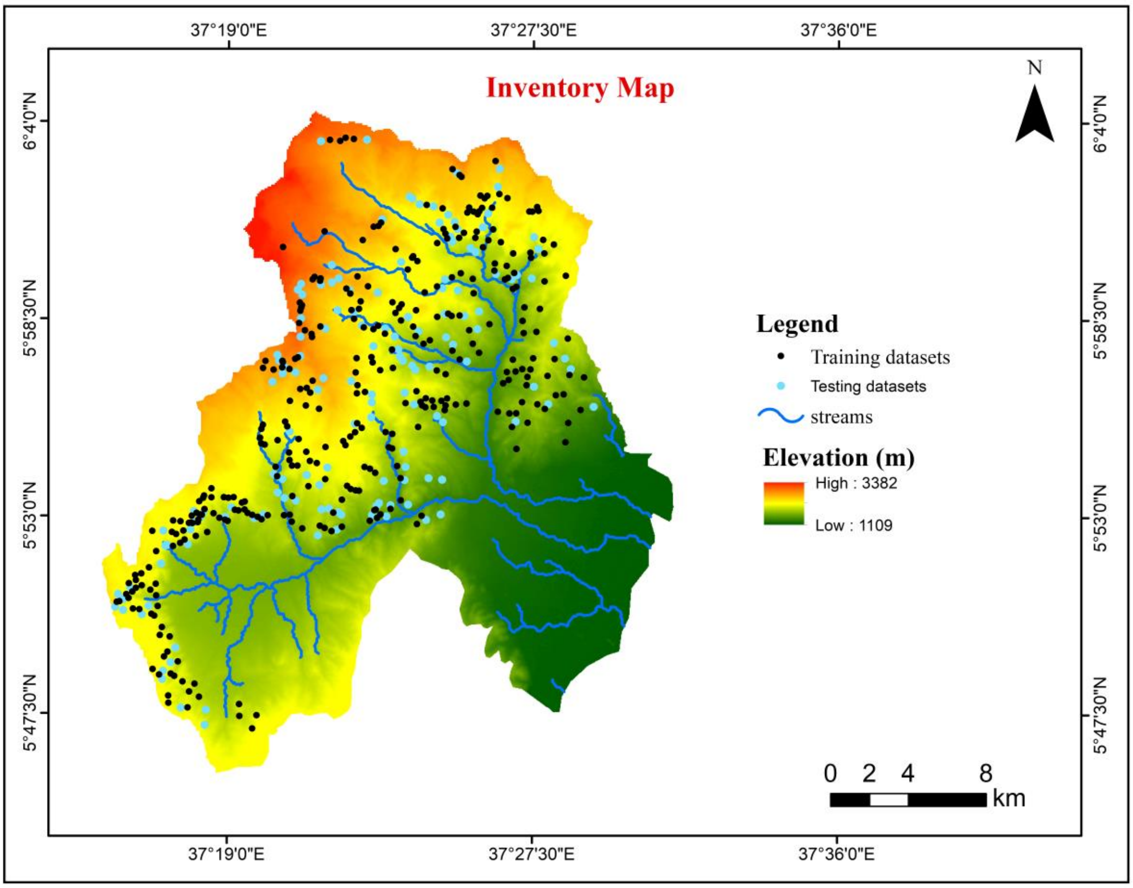


Figure 8: Landslide inventory map of the study area.

4.2 Landslide influencing factors

4.2.1 Slope

The degree of slope inclination significantly impacts various aspects of the terrain. Steeper slopes are associated with increased shear stress, particularly concerning soil and fragile rock materials, leading to a high risk of instability and potential landslides (Shano et al., 2021).

The slope map of the Sile and Elgo catchments is derived from the DEM image using ArcGIS software by changing its pixel size (30,30) m resolution in the form of raster format.

In this study (Figure 9), slope angle was classified into six classes (0-5$^{o)}$, (5-11 $^{o)}$, (11-17$^{o)}$, (17-24^{o}), (24-32^{o}), and >32 ^{o}. This classification was reclassified by using natural breaks (Jenks). The classification that had a lower slope angle (0-5^{o}) was found on alluvial fan deposits and volcaniclastic deposits, which were found in less steep slopes.

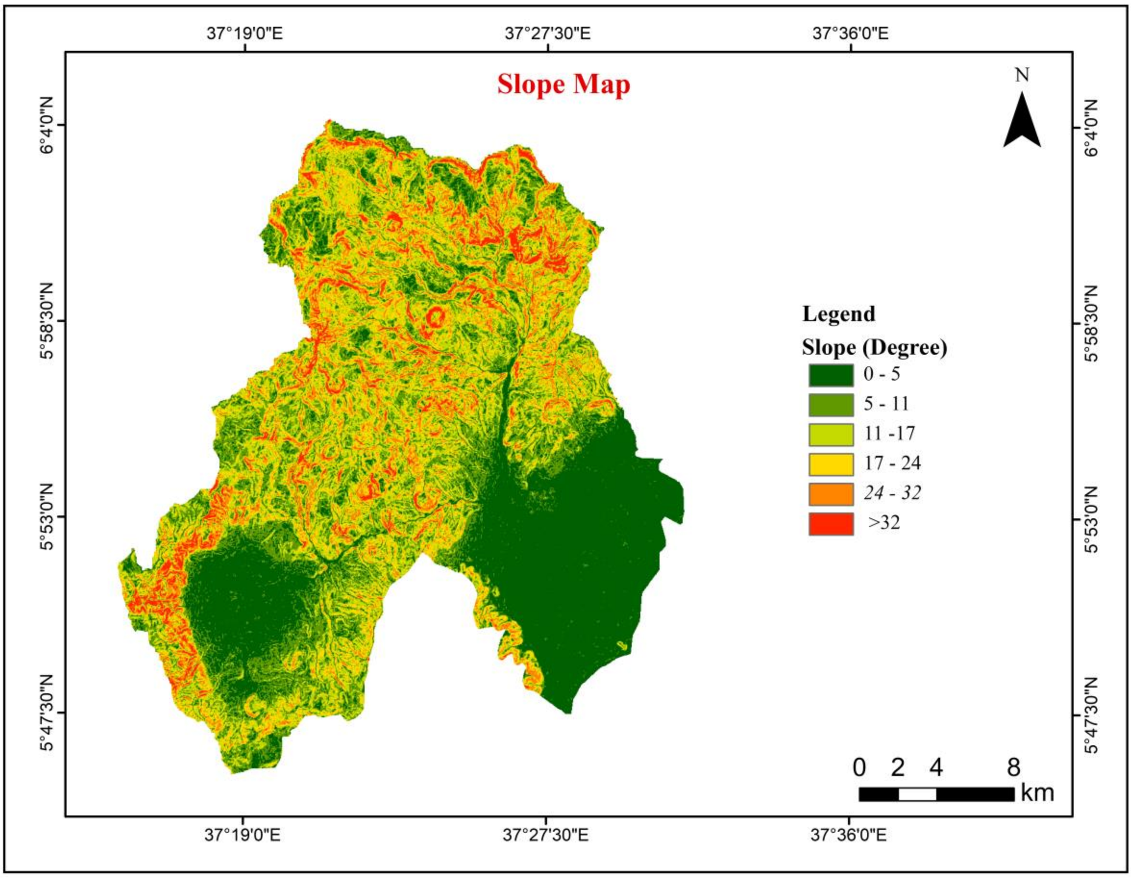


Figure 9: Slope map of the study area.

4.2.2 Aspect

Aspect signifies the orientation of the slope surface, ranging from 0 to 360 degrees (Shano et al., 2021).The orientation of slopes indirectly impacts the distribution of landslide occurrences by influencing the overall physiographic layout of the region and the primary direction of precipitation (Adnan et al., 2020).

The aspect map of the study area was created using a Digital Elevation Model (DEM) with a resolution of 30 meters by 30 meters. For this thesis, an aspect map (Figure 10) with nine classes was used to categorize the aspect as, Flat (-1), North (0 - 22.5), Northeast (22.5 - 67.5), East (67.5 - 112.5), Southeast (112.5 - 157.5), South (157.5 - 202.5), Southwest (202.5 - 247.5), West (247.5 - 292.5), and Northwest (292.5 - 337.5) using the “Natural breaks (Jenks)” in the spatial analyst’s tool using ArcGIS 10.8.

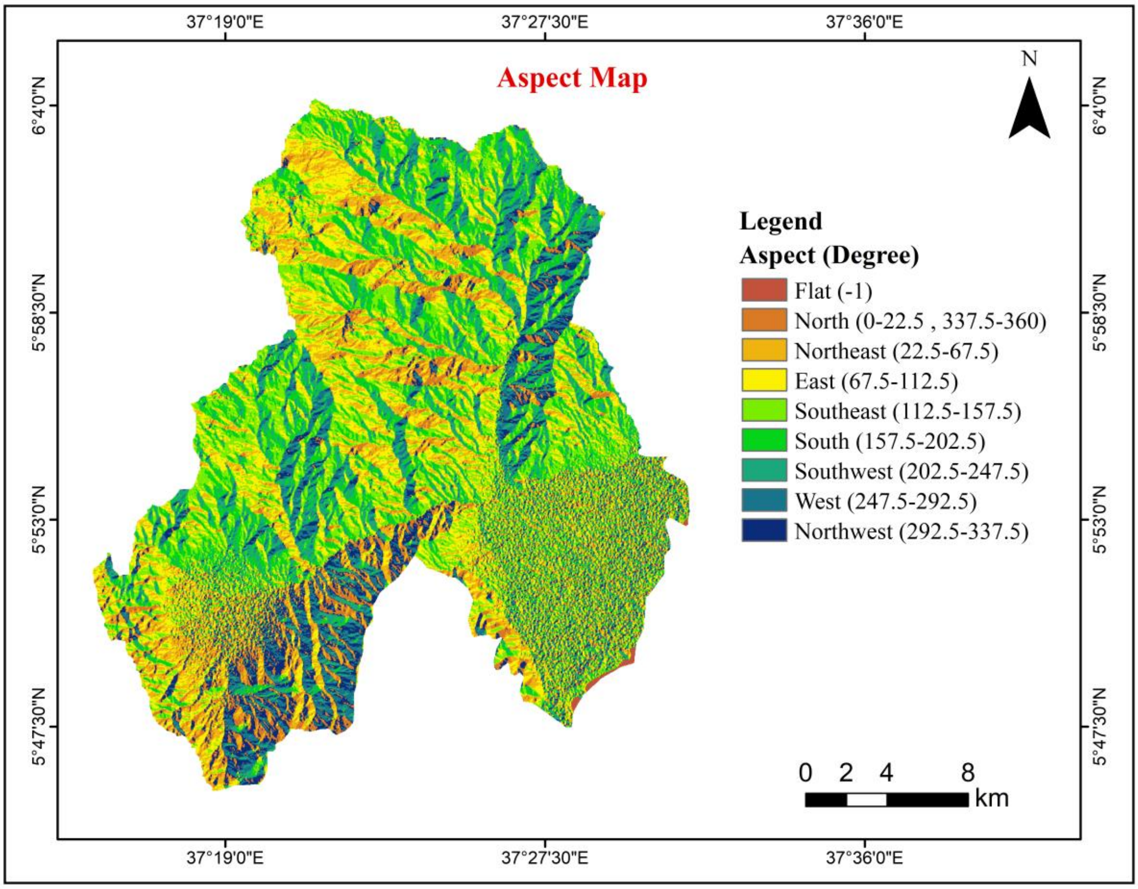


Figure 10: Aspect map of the study area.

4.2.3 Curvature

The slope's shape has a significant impact on the likelihood of landslides (Getachew and Meten, 2021). This contributes to landslides as it directly regulates the speed of water flow, thereby defining erosion boundaries (Adnan et al., 2020).

If the curvature of the slope is concave, there's a greater likelihood of groundwater presence, affecting slope stability. Conversely, if the slope's curvature is convex, it creates conditions conducive to erosion, making it the primary factor influencing slope instability in the area (Shano et al., 2021).

The study area's curvature map was generated using digital elevation model (DEM) data with a spatial resolution of 30 meters by 30 meters. It was classified into three classes (Figure 11a): concave or negative curvature (-11.888889 - -0.668846), flat (-0.668846 - 0.138998), and convex or positive curvature (0.138998 - 11).

4.2.4 Distance to stream

Understanding how slope stability and drainage features interact is crucial for studying landslides. Active stream incision can worsen slope instability in several ways. For instance, the erosion weakens the bases of slopes along riverbanks, increasing the risk of collapse. Additionally, water seeping into slope materials reduces their strength, making them more unstable. Water also erodes slopes, making them steeper and potentially causing more instability in already unstable areas (Getachew and Meten, 2021).

Numerous tributaries flowing into the Chamo River have led to multiple landslides in the vicinity of the river. Consequently, this factor was considered a key contributor in the investigation of landslide susceptibility. Regions characterized by steep slopes within the districts were identified as the primary areas prone to landslides. The majority of landslides in the present research site were observed near active springs, leading to the formation of a failed surface (Shano et al., 2021).

The stream map was generated from a Digital Elevation Model (DEM), using Euclidean distance buffering with a 408m buffer distance. The stream map was then reclassified into six classes in (Figure 11b), using the Natural Breaks (Jenks) (0 – 408m), (408 – 849m), (849 – 1321m), (1321 – 1848m), (1848 – 2461m), (2461 – 3893m).

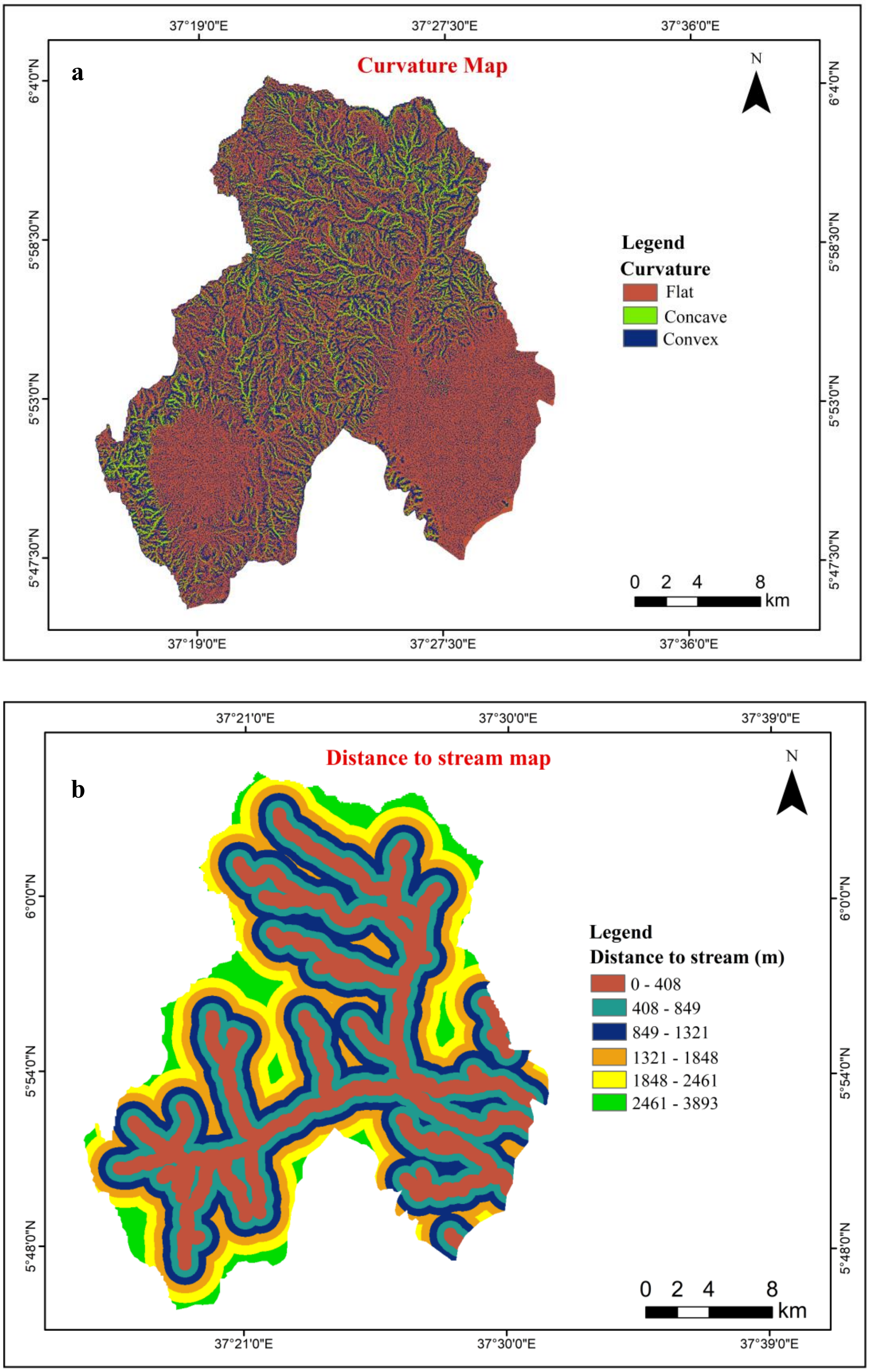


Figure 11: (a) Distance to stream map of the study area. (b) Curvature map of the study area.

4.2.5 Distance to lineament

The occurrence of landslides in the Ethiopian highlands is closely tied to the presence of lineaments, which are linear geological features like faults and fractures. These lineaments, often found without obvious signs of displacement, increase the susceptibility to slope failure. In the part of Ethiopia, where steep slopes and deep valleys mark the terrain, the proximity to such lineaments significantly influences landslide risk (Getachew and Meten, 2021).

Lineament analysis is essential for assessing its impact on landslide occurrences and comprehending the broader tectonic features of the terrain. These linear features often pinpoint potential weak planes in the landscape where the integrity of slope materials is compromised, thereby increasing the likelihood of slope failure (K. Singh et al., 2020).

Lineaments are manually digitized for this study using Google Earth Pro software after being checked against the Regional Geological Map received from GIE. The whole lineament map was subjected to Euclidian distance analysis using ArcGIS 10.8 at intervals of 633 meters. The findings were categorized in (Figure 12a) into six classes using the Natural Breaks (Jenks) method: (0 -633) m, (633 - 1512) m, (1512 - 2638) m, (2638 - 4151) m, (4151 - 6016) m and (6016 – 8972) m.

4.2.6 Soil type

Landslides are more common in highland and hilly locations with dense soil and volcanic terrains. Landslides in volcanic landscapes are caused by the breakdown of highland terrain under the effect of volcanic parent rocks, which aid in the formation of hydrothermal alterations. In regions experiencing hydrothermal changes, landslides are especially common due to the presence of clay minerals, which can make slope angles and rock formations more susceptible to landslides (Noviyanto et al., 2020).

The soil type of the study area was modified after the Geological Institute of Ethiopia (GIE,2023). Seven types of soils (Figure 12b) that were including the study are fluvisols, leptsols, vertisols, cambisols, nitisols, regosols, and luvisols.

Fluvisols (Alluvial soils)

Fluvisols, in an engineering geological context, are classified as alluvial soils, which are composed of unconsolidated sediments deposited by rivers or other running water. These soils typically exhibit poor grading with varying grain sizes and may lack cohesion, depending on the fine content. Alluvial soils are prone to settlement and pose challenges for construction due to high compressibility and potential for liquefaction in seismic prone areas. Flood hazards increase the risk of soil instability, especially in regions with seasonal heavy rainfall. Severe gully erosion exacerbates the vulnerability of structures in these areas. Flood hazards are a common constraint on fluvisols, especially in areas with seasonal rainfall patterns. Severe erosion and land degradation have occurred in the mapped area: the gullies are deeply incised into the soils and underlying layers.

Leptosols (Colluvial Soil)

Occur throughout the whole area where slopes are too steep to make any significant degree of soil development possible. An additional factor is severe recent erosion on cultivated slopes. They can be found on rocks that are resistant to weathering or where erosion has kept. They have a resource potential for grazing and as forests. Leptosols have poor load-bearing capacity and high erosion potential.

Vertisols (Expansive Clay Soils)

Are heavy clay soils, and a high proportion of swelling clays. The alternate swelling and shrinking of expansive clay soils result in deep cracks in the dry season and the formation of slickensides and wedge-shaped structural elements in the subsoil. Vertic horizons are clayey and, when dry, often have a hard to very hard consistency. Vertisols are recognized by their clayey textures, dark colors, and special physical attributes. Usually, Vertisols occupy the flatter landforms, which include depressions and lower sections of the wide parallel valleys, the lateral valleys of the undulating plateaus, and the volcanic plateaus predominantly in the low-lying positions of the mapped area. Vertisols' agricultural uses range from extensive (grazing) through smallholder crop production to small-scale and large-scale agriculture. They are often known as “black cotton soils”. Cotton is known to perform well on Vertisols, apparently because cotton has a vertical root system that is not damaged severely by the racking of the soil. On the slopes, they are affected by colluvial processes.

Cambisols (Weakly Developed Soils)

The horizons are generally richer in organic matter and therefore, have a darker color. In this case, some well-developed soil structure is needed to prove pedogenetic alteration. Cambisols in the humid tropics can form in a period of a few years. They have a faster rate of weathering and soil formation. However, in the areas with active erosion, they may occur in association with mature tropical soils and with very young and weakly developed soils. Cambisols occur where conditions are not favorable for soil-forming processes other than weathering to take place, and they often occur on steep slopes. They are generally unsuitable for heavy structures without proper soil reinforcement. Large occurrences of the Cambisols in the study area concentrated around Mount James.

Nitisols (Well-Drained, Stable Clay Soils)

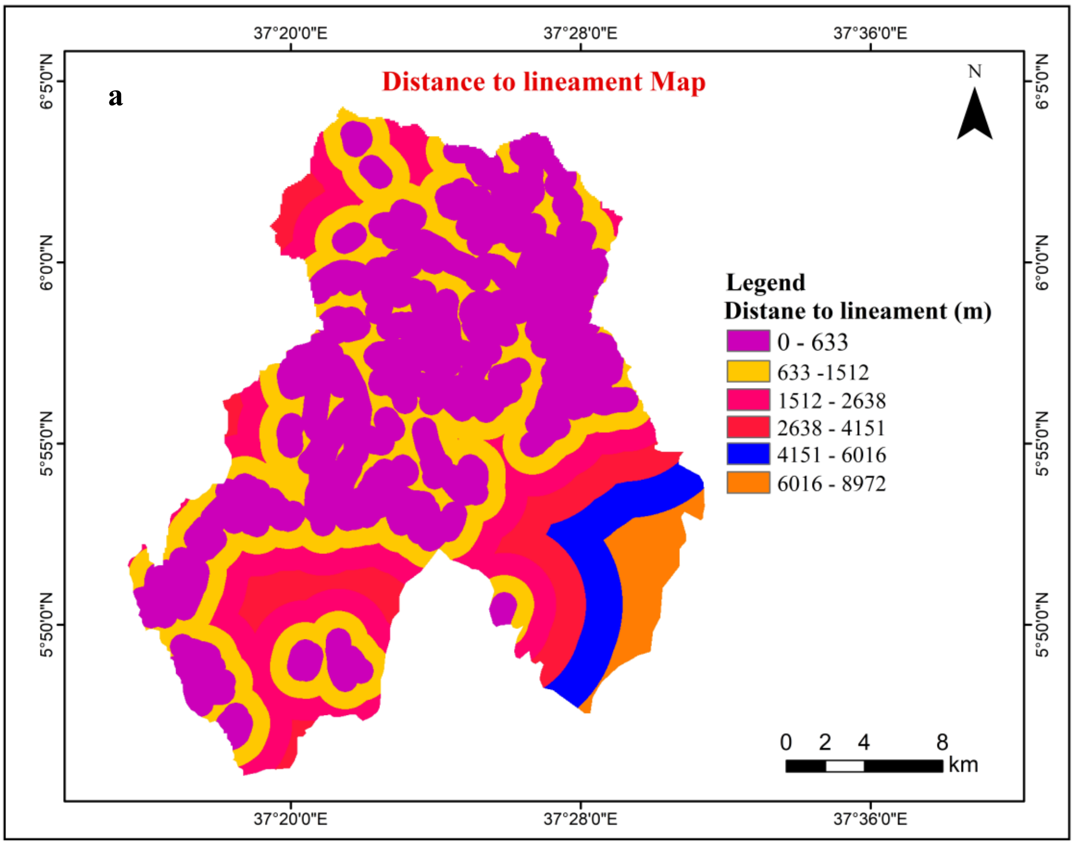
Nitisols are deep, well-drained, red tropical soils. The nitic horizon is a clay-rich subsurface horizon with at least 30 percent clay. The difference in clay content with the overlying and underlying horizons is gradual. Nitic horizons have a typically angular blocky structure breaking into polyhedral flat-edged or nut-shaped elements. The development of these typical nitic properties is assumed to be related to alternating micro-swelling and shrinking (Driessen et al., 2000). The soil moisture regime plays an important role because it influences the degree of swell-shrink behavior, which will be strongest in a climate with contrasting wet and dry seasons.

Regosols (Loose, Unconsolidated Soils)

They cover heterogeneous formations of colluvial sediments at the foot of steep eroded slopes throughout the mapped area. Regosols are weakly developed soils in unconsolidated material without diagnostic horizons. They show only slight signs of soil development of organic matter producing a darker topsoil is often the only evidence of soil formation. Regosols are generally associated with degrading and eroding surfaces. It has been transported as a result of erosion, wash, or soil creep, and the transport may have been accelerated by particular land use practices (deforestation). Regosols suffer severe erosion and are often covered with tree and shrub vegetation.

Luvisols (Moderately Weathered Soils)

Luvisols are moderately weathered, and occur on relatively young surfaces in the mapped area. Luvisols are fertile soils and suitable for a wide range of agricultural uses because of their moderate degree of weathering and high base saturation. Luvisols on steep slopes require erosion control measures.



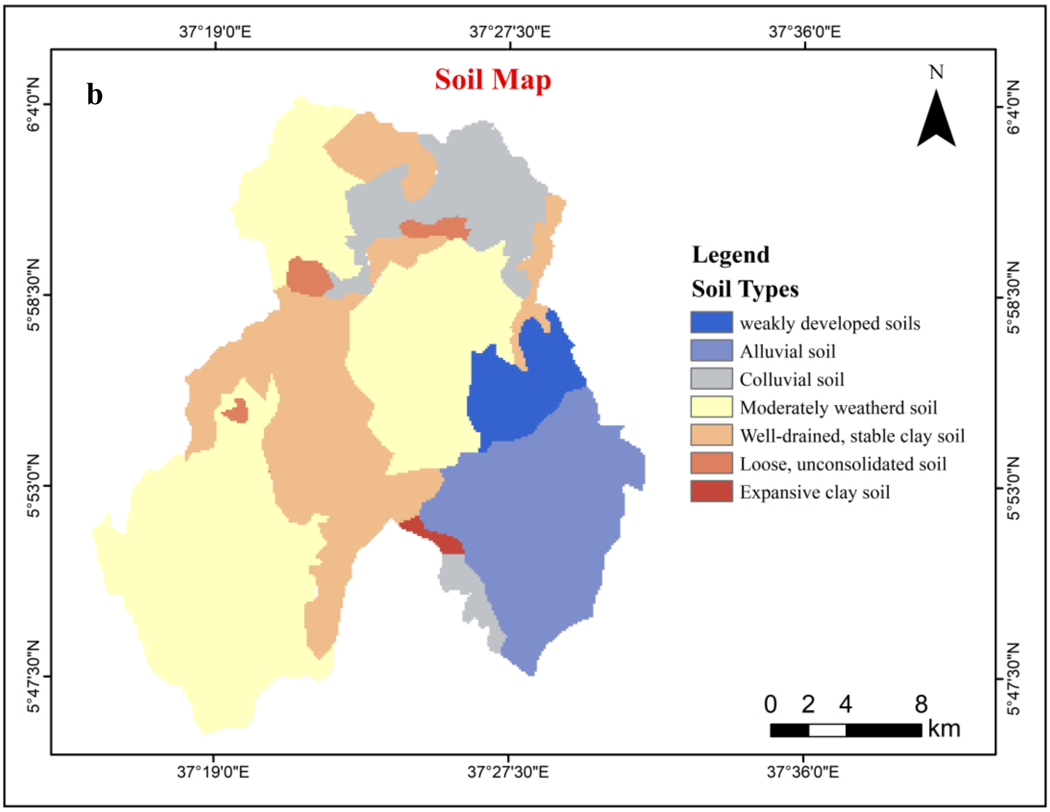


Figure 12: (a) Distance to lineament map of the study area. (b) Soil map of the study area.

4.2.7 Lithology

The geological composition of an area plays a crucial role in determining the likelihood of landslides (Getachew and Meten, 2021). This acknowledgement highlights how the type of rock formations present profoundly influences the landscape and the slopes' dynamics. Rocks vary in strength and stability, with softer or more weathered ones being more prone to landslides. Because different types of rock formations exhibit varying degrees of susceptibility, they contribute differently to slope instability and the occurrence of landslides (GIE,2023).

The geological map utilized in this study was modified after from the original version compiled by the Geological Institute of Ethiopia (GIE,2023). Lithological units were classified into eight units (Figure 13a). Alluvial fan deposits, Colluvial sediments Trachyte with subordinate trachyte volcaniclastic deposits, Rhyolite lava with subordinate volcanic breccia and ignimbrite, Rhyolitic ignimbrite with subordinate rhyolite lava, Basalt to trachybasalt lava with subordinate basaltic volcaniclastic deposits, Biotite, and Volcaniclastic deposits. From all the above lithological units, ignimbrite's study area is dominantly occupied.

Plate 2: (a) River channel with gravel bars incised in the large alluvial fan, north of Shele town. (b)Typical outcrop of basalt lava flow (GIE, 2023).

Alluvial fan deposits

Are preserved as accumulations of older phases of alluvial fans currently exposed to erosion processes or as preserved gravel accumulations on structural terraces. Alluvial fan deposits are represented by 17.65% silt to sandy silts or very fine sands, with indistinct fine- to medium-grained sand intercalations.

Colluvial sediments

Colluvial sediments are composed of a wide range of non-cohesive sediments of various grain sizes which is exposed to 8.1% of the total study area. The most common colluvial sediments found in the study area are associated with large landslides. These sediments are mostly matrix-supported, with angular to subangular-shaped cobbles and boulders up to several meters in diameter. The boulders are formed by competent rocks such as weathered ignimbrites, trachytes, and basalts. The matrix has the character of fine-grained sandy to loam soils. Rockfall colluvial accumulations (Plate 4.1A) associated with outcrops of weathered or only slightly weathered competent rocks are also numerous. These accumulations have a character of angular cobbles and boulders up to several meters in size without a fine-grained matrix. The most important accumulations of colluvial sediments are found below the fault scarps located west of Chamo Lake and north of Gerese town.

Trachyte with subordinate trachyte volcaniclastic deposits

It covers an area of 6.2 km^2 and 1.1% of the total study area and forms volcanic plugs and dykes, as well as small lava flows, northeast of Mele and Surgulla town. They consist of hydrothermally altered blocks, volcanic bombs, and angular trachyte clasts in an ash matrix.

Rhyolite lava with subordinate volcanic breccia and ignimbrite

They form small lava flows, lava domes, and ignimbrite layers. They are porphyritic or aphanitic rocks with fluidal textures, which are exposed to a smaller part of the study around 0.1%.

Rhyolitic ignimbrite with subordinate rhyolite lava

They are exposed on steep slopes of an escarpment, mainly along the northwestern edge of the study area, around Gerese, and the northern edge of the study area, which covers 8.1 % of the area. Rhyolite lava flows, and small domes are spatially related to a strongly welded crystal-rich ignimbrite, locally having a Fluid texture. Brecciated lava prevails in outcrops. Individual strata rhyolitic ignimbrite is densely welded to non-welded and locally alternated with up to 0.7 m thick strata nonwelded ash fall deposits.

Basalt to trachybasalt lava with subordinate basaltic volcaniclastic deposits.

The study area predominantly comprises rocks of this type, spanning the central, northeast, and southwest regions, constituting approximately 53.6% of the total coverage.

Characterized by subangular to angular basalt fragments embedded in a matrix ranging from coarse to fine-grained tuff, these rocks have undergone weathering processes and are interspersed with red paleosols. These paleosols are notably observable between individual lava flows, forming a distinctive feature throughout much of the study area.

Biotite

It covers 5.1% of the study and forms a body at the contact of the gneisses and the amphibolite southwest of Gerese.

Volcaniclastic deposits

This soil is deposited in the western parts of Mele town. It covers 6.3% of the study area.

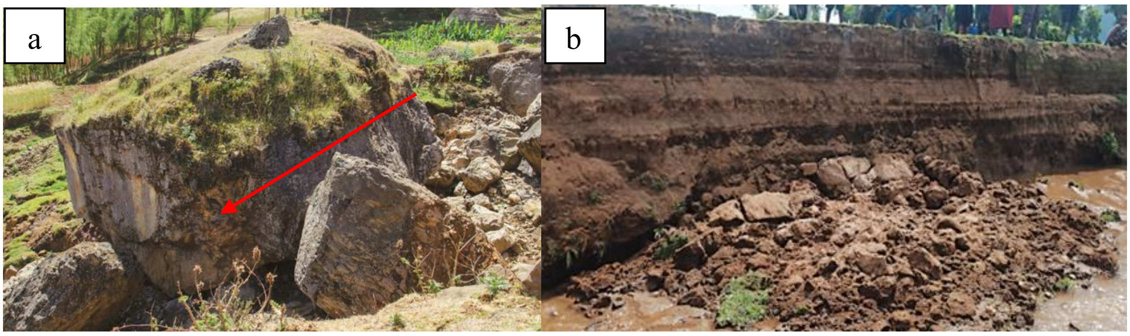


Plate 3: (a) Colluvial boulders accumulate as a result of rock fall, south to Gerese town. The red arrow indicates the direction of rock fall (b), Collapse of the river bank due to lateral erosion in the alluvial fan south of Gerese town (GIE, 2023).

4.2.8 Land Use Land Cover (LULC)

Landslides are mostly caused by changes in land use, and they are especially likely to occur on bare terrain. Because it helps to adhere and connect the slope material, vegetation is essential to preserving slope stability (Getachew and Meten, 2021).

In catchment areas, population growth has soared over the last thirty years, establishing new towns in locations that are extremely vulnerable to landslides. This growth increases the risk of landslides, which endangers people, property, and infrastructure (Shano et al.,2021). The land use/land cover map (Figure 13b) was prepared from the sentinel satellite image by using the ArcGIS 10.8 tool. The current study area has five land-use/land cover types: water body, trees, grassland, settlement, and agricultural land.

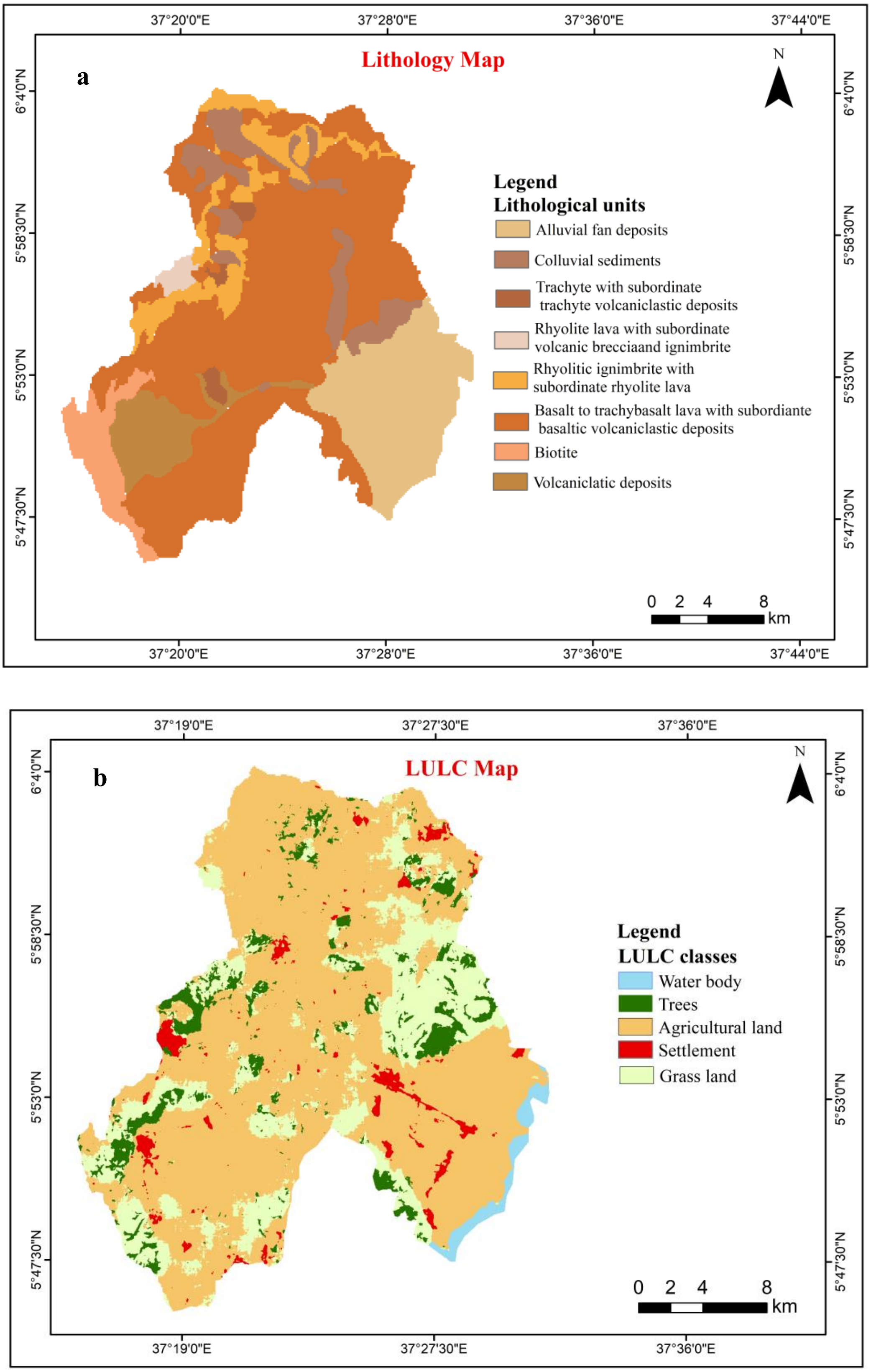


Figure 13: (a) Lithology map of the study area. (b) Land use Land cover map of the study area.

4.2.9 Rainfall

Rainfall stands out as one of the most crucial factors that trigger landslides (Wubalem et al., 2022).In mountainous areas, heavy rainfall often causes landslides by increasing moisture content in the soil, which weakens its stability. The intensity and distribution of rainfall directly impact the nature of landslides (Noviyanto et al., 2020).

The rainfall map was created by utilizing the GIS interpolation technique (IDW). IDW interpolation technique was chosen over Kriging due to its simplicity and straightforwardness compared to the more complex and intricate statistical modeling required by Kriging, making it more accessible and easier to comprehend.

The data was collected from seven different rainfall stations managed by NASA. The map was saved in a raster format, all possessing identical coordinate systems (Adindan UTM Zone 37 N) and pixel dimensions (30x30 meters).

The rainfall map of the study area was classified into seven classes (Figure 14a), (1104 - 1163), (1163 - 1221), (1221 -1280), (1280 - 1338), (1338 - 1397), (1397 - 1455), and (1455 - 1513) rainfall in mm (using the Natural Breaks (Jenks).

4.2.10 Elevation

Elevation plays a significant role in the occurrence of landslides due to its influence on various natural and human-induced factors. Geological tectonics, which are influenced by elevation, contribute to the spatial variability of landslides. Additionally, elevation affects factors like slope and curvature, which further influence landslide occurrence (Adnan et al., 2020).

Elevation indirectly impacts slope instability or landslides by influencing both rainfall patterns and the extent of vegetation coverage (Shano et al., 2021).

The elevation data was acquired directly from the Digital Elevation Model (DEM). By using ArcGIS software, it reclassified into six classes in (Figure 14b), (1109 -1308), (1308 - 1584), (1584 – 1859), (1859 – 2195), (2195 - 2653), and (2653 - 3380) in m.

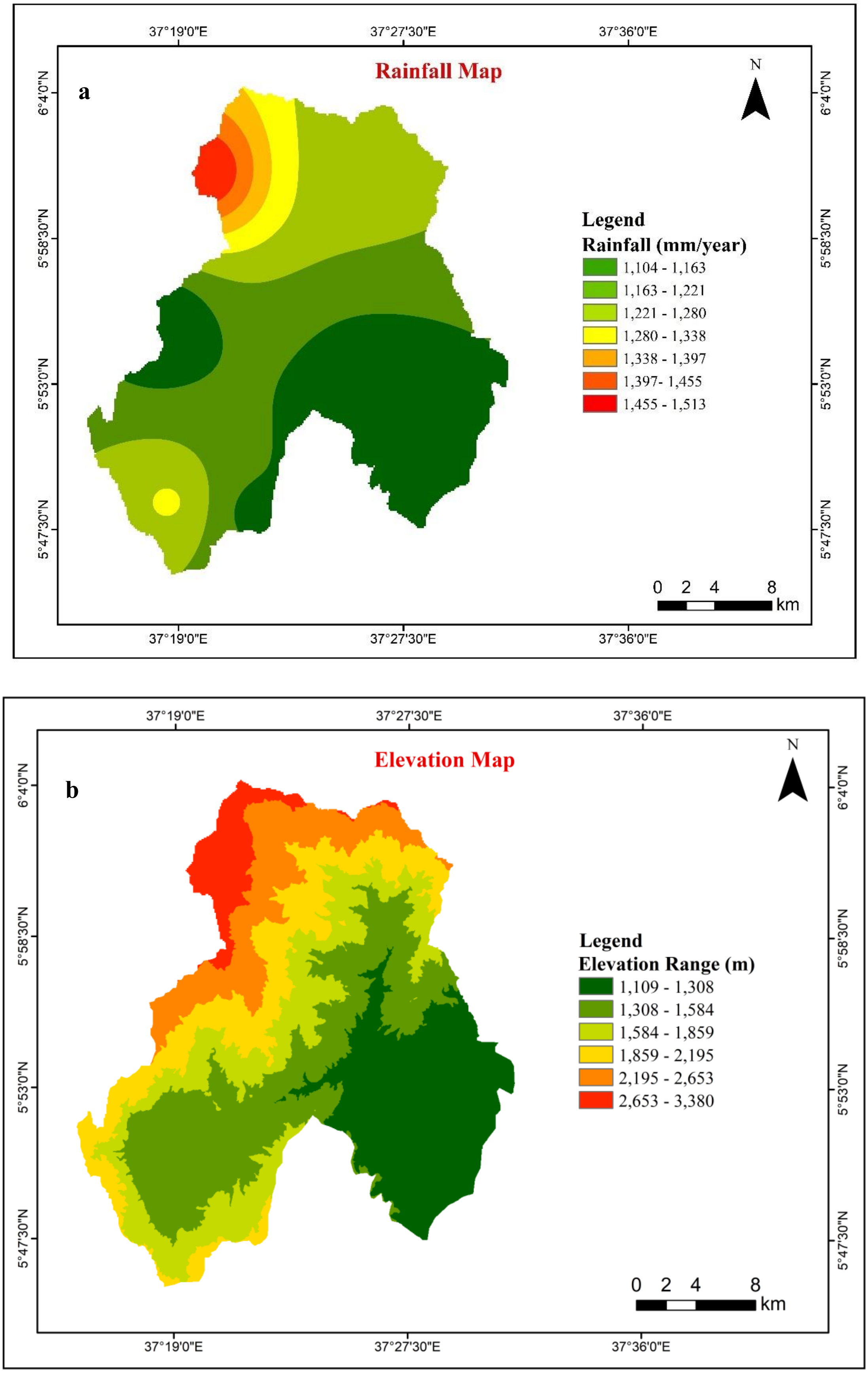


Figure 14: (a) Rainfall map of the study area. (b) Elevation map of the study area.

4.3 Factors contribution

A crucial stage in creating a landslide susceptibility map is assessing each factor's significance in the likelihood of landslides and eliminating those with low predictive ability (Hamedi et al., 2022). The correlation between factors was identified based on computing variance inflation factor (VIF), tolerance (TOL), Pearson correlation coefficient, and Frequency ratio.

4.3.1 Variance Inflation Factor (VIF) and Tolerance (TOL)

In the current study, data filtering was done before modeling was done. It is required before modeling since inconsistent findings could arise from using duplicate data. Therefore, components with collinearity difficulties should be eliminated before modeling to produce high-accuracy findings (Mallick et al., 2022).

The collinearity increases with increasing VIF. When the variance inflation factor (VIF) is greater than 10 or the tolerance is less than 0.10, multicollinearity problems arise, which should be excluded from the process of landslide susceptibility mapping.

The result showed that all tolerance values were more than 0.10, proving that the ten parameters determining landslide susceptibility didn't show collinearity. The ten independent variables, as shown in Table (6) demonstrates, have VIF values of less than 10. This implies that the independent variables do not have a collinearity problem. The VIF value varied from 1.0198 to 2.1289, where the lowest and highest values were observed.

Table 6: VIF and TOL value of causative factors.

NO	Variables	Tolerance	VIF
1	Slope	0.670084	1.492351
2	Aspect	0.954048	1.048165
3	Curvature	0.980514	1.019873
4	DTS	0.736235	1.358262
5	DTL	0.469719	2.128933
6	Soil Type	0.558628	1.790099
7	Lithology	0.566788	1.764328
8	LULC	0.90367	1.106598
9	Rainfall	0.659004	1.51744
10	Elevation	0.6181	1.617862

4.3.2 Pearson correlation

Pearson's coefficient can be used to examine the correlation between factors (Hamedi et al., 2022). To find the proper independent causal elements, Pearson correlation analysis was used to eliminate variables that correlated value >0.7, a sign of strong collinearity (Mallick et al., 2022).

The correlation value does not influence the performance of the classification models. As a result, there is an association among these ten landslide causative factors. Therefore, all ten factors can be included in the landslide susceptibility assessment.

In Figure 15 a positive correlation is shown in shades of red, and a negative correlation is shown in shades of blue. No correlation values close to 0. The correlation matrix in the figure provides valuable insights into the relationships between various factors and their potential influence on landslide occurrence. The strongest observed negative correlation is between Distance to Lineament (DTL) and slope, with a correlation coefficient of -0.51. The strongest positive correlation was also observed between lithology and soil type with a correlation coefficient of 0.54.

In addition, the correlation between elevation and slope is 0.42, suggesting that areas with higher elevations tend to have steeper slopes. These correlations underscore the importance of aspect, slope, and soil type in landslide risk assessment and suggest that these factors should be prioritized in predictive models and mitigation strategies to enhance the accuracy of landslide hazard mapping and to inform effective land-use planning and disaster preparedness initiatives.

The results indicate that there is no issue with covariance among the influencing factors. The findings are consistent with the validation outcomes of VIF and TOL, confirming that all selected causative factors are suitable for landslide susceptibility modeling, thereby ensuring the model's accuracy and precision.

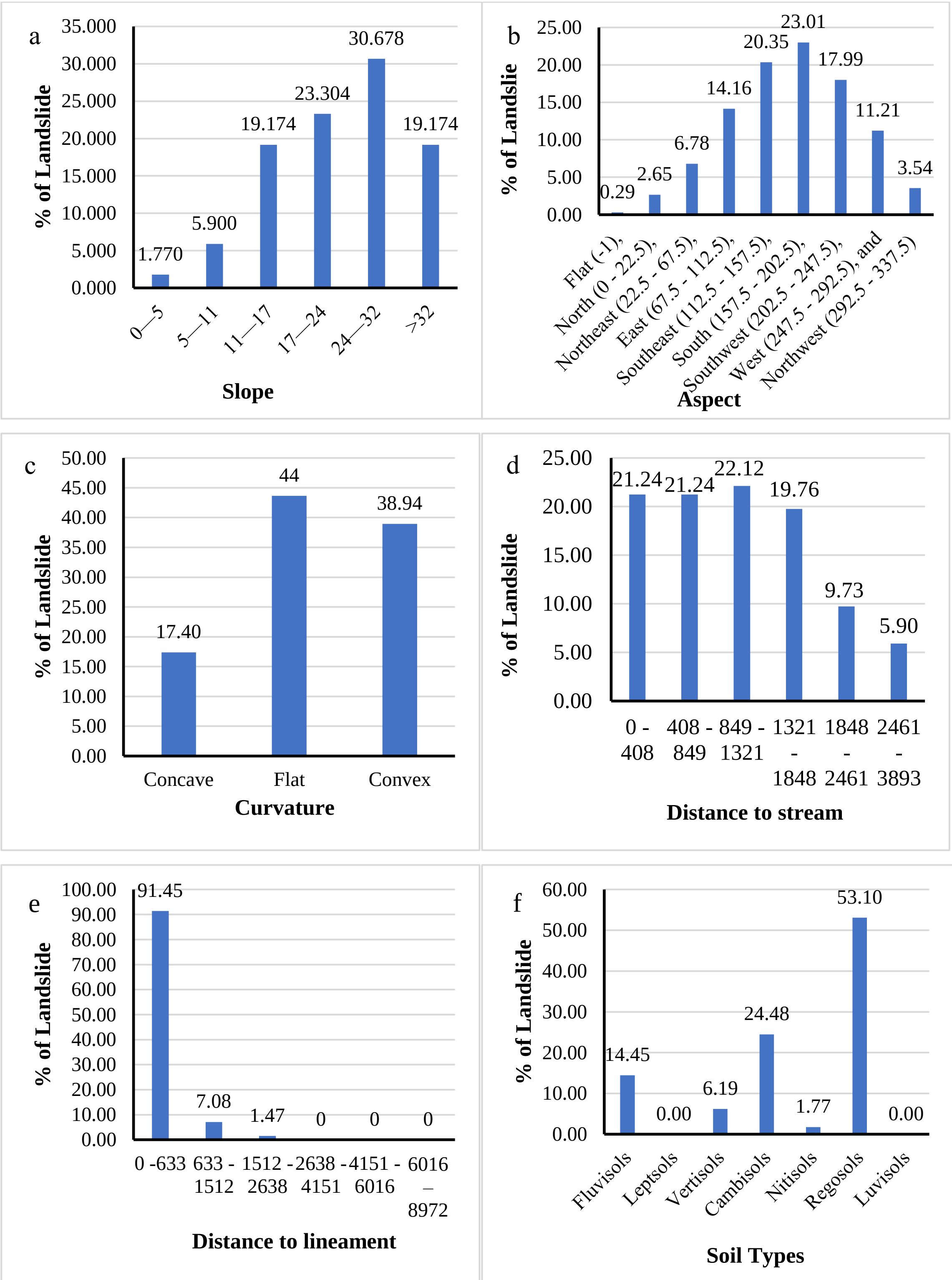
Figure 15: Pearson correlation between causative factors.

4.3.3 Correlation between landslide and causative factor using frequency ratio

The non-linear association between the occurrence of landslides and the chosen causative factors is analyzed using the FR approach (Z. Huang et al., 2023).

The FR value characterizes the various grading intervals of cauasaive factors and the likelihood of a landslide. The grading interval is favorable for landslide hazard occurrence when FR > 1; the grading interval is unfavorable for landslide hazard occurrence when FR < 1; and the probability of landslide hazard occurrence in the grading interval is equal to the average level in the study area when FR = 1.

In this study, the relationship between landslides and each class of causative factors (slope, aspect, curvature, distance to stream, distance to lineaments, soil type, lithology, lULC, Rainfall, and elevation was examined and shown in Figure 16 as follows.


a
% of Landslide
35.000
30.000
25.000
20.000
15.000
10.000
5.000
0.000
1.770
5.900
19.174
23.304
30.678
19.174
0—5
5—11
11—17
17—24
24—32
>32
Slope
b
% of Landslie
25.00
20.00
15.00
10.00
5.00
0.00
0.29
2.65
6.78
14.16
20.35
23.01
17.99
11.21
3.54
Flat (-1),
North (0 - 22.5),
Northeast (22.5 - 67.5),
East (67.5 - 112.5),
Southeast (112.5 - 157.5),
South (157.5 - 202.5),
Southwest (202.5 - 247.5),
West (247.5 - 292.5), and
Northwest (292.5 - 337.5)
Aspect
c
% of Landslide
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
17.40
44
38.94
Concave
Flat
Convex
Curvature
d
% of Landslide
25.00
20.00
15.00
10.00
5.00
0.00
21.24
21.24
22.12
19.76
9.73
5.90
0 - 408
408 - 849
849 - 1321
1321 - 1848
1848 - 2461
2461 - 3893
Distance to stream
e
% of Landslide
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
91.45
7.08
1.47
0
0
0
0 -633
633 - 1512
1512 - 2638
2638 - 4151
4151 - 6016
6016 – 8972
Distance to lineament
f
% of Landslide
60.00
50.00
40.00
30.00
20.00
10.00
0.00
14.45
0.00
6.19
24.48
1.77
53.10
0.00
Fluvisols
Leptsols
Vertisols
Cambisols
Nitisols
Regosols
Luvisols
Soil Types

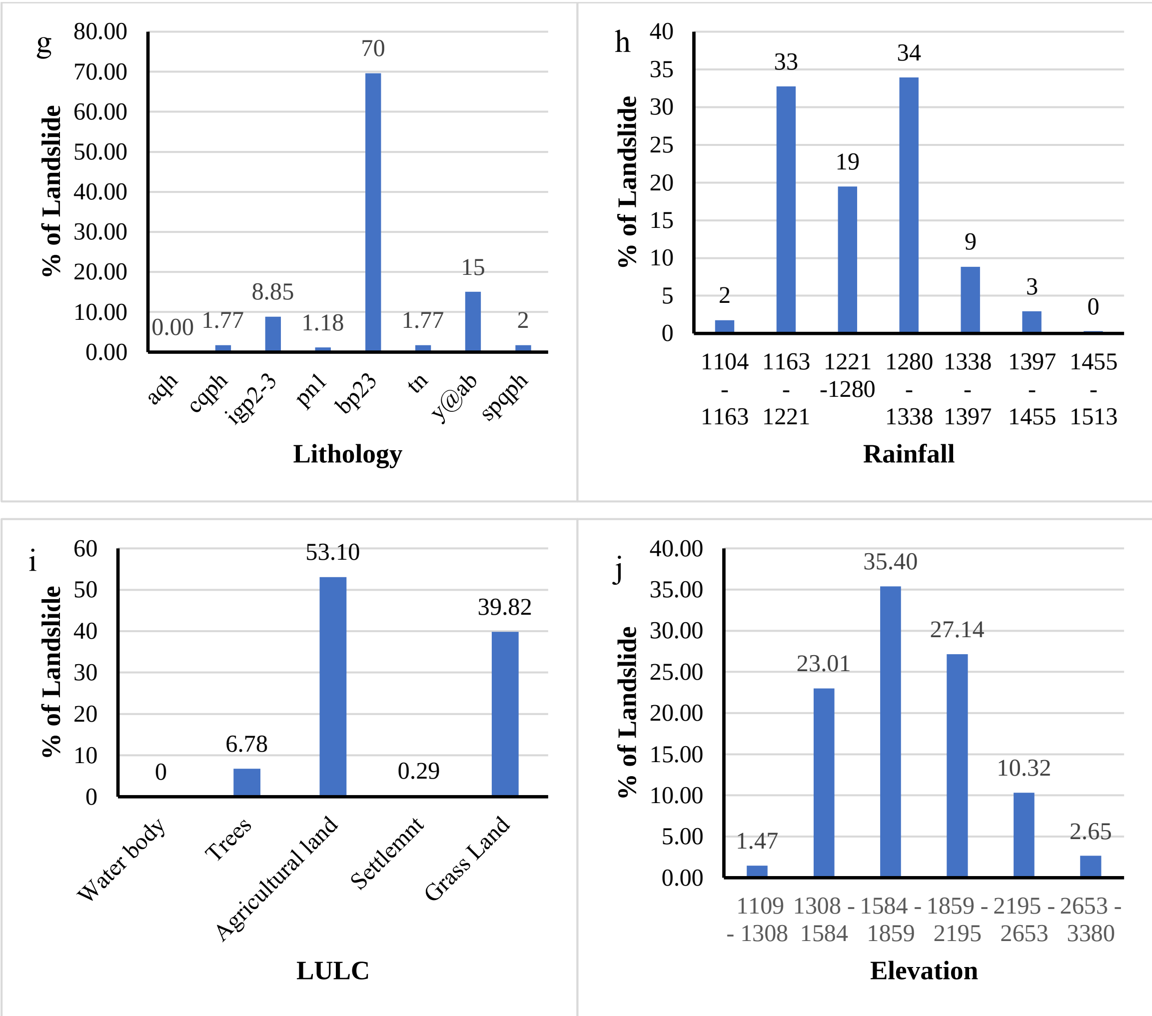


Figure 16: The percentage of landslides in (a) Slope, (b) Aspect, (c) Curvature, (d) Distance to stream, (e) Distance to lineament, (f) Soil type, (G) lithology, (h) LULC, (i) Rainfall, (j) Elevation causative factor classes.

As the slope angle increases, the FR values also increase. FR values are greater than 1, indicating a higher likelihood of landslides occurring at slope intervals between 17 and 32°. For example, the slope classes of (17-24), (24-32), and >32 have the highest FR values of 1.45, 3.11, and 5.46, respectively (Table 7). This suggests that the probability of landslides rises with increasing slope gradient, consistent with findings by Genene and Meten (2021), who reported a high probability of landslides in slope classes greater than 15°.

The aspect analyses showed that the maximum FR value (1.6) was for the southwest, followed by south (1.51), west (1.25), and southeast (1.1), indicating a higher probability of landslide occurrence. The aspect class of flat has zero FR value, which indicates a lower likelihood of landslide occurrence.

The curvature of slopes significantly impacts the Frequency Ratio (FR) values, which indicate the likelihood of landslides. Concave slopes exhibit a higher FR of 1.81, suggesting they are highly susceptible to landslides. This is because their shape collects rainwater, increasing soil saturation and instability. Convex slopes have a moderate FR of 0.285, which is also prone to landslides, though less so than concave slopes. This is due to the outward curvature contributing to gravitational forces that can cause instability. Flat slopes, with a low FR of 0.18, are much less likely to experience landslides since they do not accumulate water in the same way and are less affected by gravity. These observations align with Mersha and Meten (2020), who noted that both concave and convex slopes are more likely to experience landslides due to their shape's influence on rainwater accumulation and gravity.

The FR for distance to stream clearly showed that the classes of (0 - 408m), (408 - 849m), (849 – 1321m), (1321– 1848m), (1848 – 2461m), and 2461 – 3893m have the highest influence on landslide occurrence, with values of 0.85, 0.09, 0.6, 0.47, 0.85, and 0.223, respectively.

Landslides are more likely to occur closer to lineaments. The probability is higher within 0 to 633 meters and 633 to 1512 meters from the lineaments, with respective values of 1.97 and 0.31. Beyond 1512 meters, the probability of landslides decreases.

The FR values for different soil types indicate their susceptibility to landslides. Regosols, Vertisols, and Cambisols have FR values of 36.44, 14.03, and 5.21, respectively, suggesting a higher probability of landslide occurrence in these soil types. Regosols, in particular, have the highest probability of landslide occurrence due to their loose, unconsolidated nature, which makes them more prone to erosion and instability. In contrast, other soil types, such as Leptosols and Luvisols, have an FR value of 0, indicating no probability of landslide occurrence.

As Table 7 indicates, aqh is represented by volcaniclastic deposits, while bp23 corresponds to basalt to trachybasalt lava with subordinate basaltic volcaniclastic deposits. Additionally, cqph represents colluvial sediments. The symbol igp2-3 refers to rhyolitic ignimbrite with subordinate rhyolite lava, and pn1 designates rhyolite lava with subordinate volcanic breccia and ignimbrite. Moreover, spqph is associated with trachyte containing subordinate trachyte volcaniclastic deposits, and y@ab is used to represent biotite.

Frequency analysis shows that the Biotite and volcanoclastic deposits have the highest susceptibility to landslides with FR values of 2.96 and 1.77, respectively.For land use and land cover, grassland and agricultural land have higher FR values of 2.05 and 0.99, respectively.

The FR values for the rainfall intervals (1280-1338mm) indicate a particularly high probability of landslides.

Landslide geological hazards in the study area are primarily found at elevations between 1308 and 2653 meters. The majority, 35.4%, occur between 1584 and 1859 meters as shown in Figure 4.9j with the FR values of 1.92. Though less common, landslides have also been recorded at elevations near or above 3380 meters.

Generally, (slope class > 17º), (slope aspect classes of southwest and south), (curvature classes of concave and convex), (distance to stream classes of 849-1321and >2461m), (distance to Lineament classes of < 1512), (soil type of regosols), (lithology classes of Biotite and Rhyolitic lava), (LULC grassland and agricultural land), (Rainfall class of 1104-1163mm and 1338-1397), (Elevation classes of 1645–1863 m) showed a higher probability of landslide occurrence.

As indicated in Table 4 the most influencing factor in landslide occurrences of ten influencing factors is “Soil Type," specifically the "Regosols" class, with a Frequency Ratio (FR) of 36.447, making it the strongest determinant among all factors analyzed. Following this, the "slope" in the ">33^0" and the "Land use land cover grassland with Fr value of 2.05 also show significant influence, with FR values5.46 and 2.05, respectively. These high FR values indicate that these factors strongly correlate with landslide occurrences, with the "Regosols" soil type being the most critical.

Table 7: The input data used in the analyses for the Frequency ratio.

Parameter	Classes	Class pixels	%class Pixels	Landslide pixels	%Landslide pixels	FR
Slope	0-5	177576	28.502	6	1.770	0.062
	5-11	131712	21.140	20	5.900	0.279
	11-17	130813	20.996	65	19.174	0.913
	17-24	99800	16.018	79	23.304	1.455
	24-32	61274	9.835	104	30.678	3.119
	>32	21863	3.509	65	19.174	5.464
Total		623038	100.000	339	100.000	11.293
Aspect	flat	6312	1.013	1	0.295	0.291
	north	46694	7.496	9	2.65	0.354
	northeast	78932	12.671	23	6.78	0.535
	east	108716	17.453	48	14.16	0.811
	southeast	114900	18.446	69	20.35	1.103
	south	94898	15.235	78	23.01	1.510
	southwest	69848	11.213	61	17.99	1.605
	west	55478	8.906	38	11.21	1.259
	north	47133	7.567	12	3.54	0.468
Total		622911		339		7.937
DTS	0 - 408	163209	26.0172801	75	22.124	0.850
	408 - 849	140073	233.377208	72	21.239	0.091
	849 - 1321	121660	35.1200305	72	21.239	0.605
	1321 - 1848	92886	42.0428005	67	19.764	0.470
	1848 - 2461	71288	11.3631002	33	9.735	0.857
	2461 - 3893	38194	26.4638836	20	5.900	0.223
Total		627310		339		3.096
Curvature	Concave	60020	9.56701373	59	17.404	1.819
	Flat	346412	240.022172	148	43.658	0.182
	Convex	220932	136.407248	132	38.938	0.285
Total		627364		339		3.242
Elevation	1109 - 1308	144325	23.0003299	5	1.474926	0.064
	1308 - 1584	161965	25.8115256	78	23.00885	0.891
	1584 - 1859	115346	18.3820963	120	35.39823	1.926
	1859 - 2195	97063	15.468429	92	27.13864	1.754
	2195 - 2653	74449	11.8645526	35	10.32448	0.870
	2653 - 3380	34343	5.47306655	9	2.654867	0.485
Total		627491		339		5.991
Rainfall	1104-1163	107783	17.1775044	6	2	0.103
	1163-1221	187425	29.8701444	111	33	1.096
	1221-1280	101578	16.1886062	66	19	1.203
	1280-1338	139051	22.1607227	115	34	1.531
	1338-1397	49552	7.89716096	30	9	1.121
	1397-1455	25547	4.07145566	10	3	0.725
	1455-1513	16530	2.63440569	1	0	0.112

Total		627446		339		5.890
LULC	Water boady	9976	1.59206902	0.000	0.000	0.000
	Tress	42656	6.80746753	23	6.785	0.997
	Agricultural land	434304	69.3105396	180	53.097	0.766
	Settelment	18202	2.90485568	1	0.295	0.102
	Grass land	121468	19.3850681	135	39.823	2.054
Total		626606		339.000		3.919
Lithology	aqh	106019	16.8636926	0	0	0.000
	cqph	51023	42.3395763	6	1.769912	0.218
	igp23	51098	47.0372723	30	8.84955	1.089
	pn1	5657	4.04921764	4	1.179941	1.311
	bp23	336680	367.374107	236	69.61652	1.300
	tn	6834	8.39918884	6	1.769912	1.628
	y@ab	31940	60.3347312	51	15.04425	2.961
	spqph	39431	122.788279	6	1.7699	1.770
Total		628682		339		1.185
Soil type	Fluvisols	106224	16.939	49	14.45428	0.853
	Leptsols	62192	9.917	0	0	0.000
	Vertisols	2768	0.441	21	6.19469	14.034
	Cambisols	29456	4.697	83	24.48378	5.212
	Nitisols	140736	22.442	6	1.769912	0.079
	Regosols	9136	1.457	180	53.09735	36.447
	Luvisols	276592	44.106	0	0	0.000
Total		627104		339		56.626
DTL	0-633	291217	46.4231401	310	91.445	1.970
	633-1512	139603	22.2542284	24	7.080	0.318
	1512-2638	87626	13.9685323	5	1.475	0.106
	2638-4151	51945	8.28059492	0	0.000	0.000
	4151-6016	33613	5.35827581	0	0.000	0.000
	6016-8972	23306	3.71522852	0	0.000	0.000
Total		627310		339		2.394

4.4 Landslide susceptibility map

The integration of ten causative factor maps was executed utilizing QGIS software, incorporating training datasets. A comprehensive analysis of the entire region was conducted using the Excel dataset. Support Vector Machine (SVM) and, K-Nearest Neighbors (KNN) models were employed to ascertain Landslide Susceptibility Indices (LSI) for each pixel within the study area's landslide map. The analysis phase was conducted using the RStudio programming language.

The culmination of this process was the generation of the final Landslide Susceptibility Map (LSM). This involved calculating the weights of various layers representing distinct factors influencing landslide occurrences. The weight determination was facilitated by data mining algorithms generated through RStudio. The results were then reclassified using ArcGIS software. In ArcGIS, the Jenks natural breaks classification method was utilized to categorize the landslide susceptibility map into five discernible classes: "very low," "low," "moderate," "high," and "very high." These maps visually represent the varying degrees of landslide susceptibility across different locations within the research area. The very high susceptibility zones are mainly located in the fault concentration area.

4.4.1 Landslide Susceptibility Mapping Using Support Vector Machine (SVM)

In this study, the Support Vector Machine (SVM) model was utilized to classify the study area into five distinct landslide susceptibility classes: very low (0 – 0.164), low (0.164 – 0.341), moderate (0.341 – 0.522), high (0.522 – 0.689), and very high (0.689 – 0.919) indicated in (Figure 17) the resulting distribution of areas across these classes was 15.8% very low, 16.2% low, 17.2% moderate, 22.3% high, and 28.4% very high (Table 8). Comparatively, other studies have demonstrated varying distributions based on regional characteristics and methodological approaches. For instance, Zhang et al. (2021) applied a Random Forest algorithm in a different geographical setting and found a higher proportion of areas classified under the moderate susceptibility class (25%), with lower percentages in the very high category (20%). This may be attributed to differences in input variables, such as the inclusion of additional geotechnical factors in the current study, which could enhance the model's sensitivity to high-risk areas.

Conversely, Kumar et al. (2023) utilized a Logistic Regression model and reported a more balanced distribution across all classes, with each category representing approximately 20% of the study area. The variation between their findings and the present study highlights the impact of model selection on susceptibility mapping outcomes. SVM's capability to handle complex, non-linear relationships likely contributed to the higher classification of areas into the very high susceptibility category in this research.

Overall, the SVM-based landslide susceptibility map in this study corroborates findings from Daviran et al. (2023) extends the understanding of susceptibility distribution by highlighting a significant portion of the area falling under very high susceptibility. This alignment with existing research validates the effectiveness of SVM in accurately identifying high-risk zones, thereby contributing valuable insights for regional landslide management and prevention strategies.

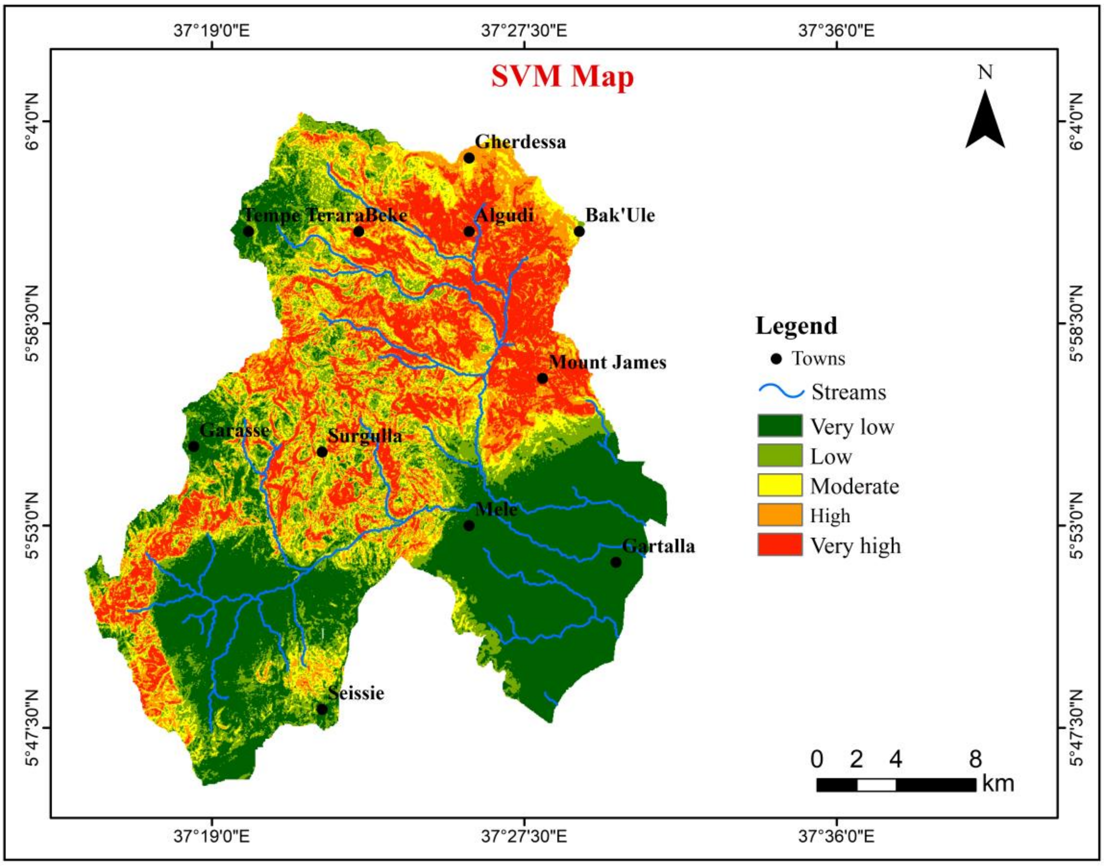


Figure 17: Landslide susceptibility map of the study area using SVM.

4.4.2 Landslide Susceptibility Mapping Using K-Nearest Neighbors (KNN)

The natural breaks approach was used to categorize the landslide susceptibility map into five groups. This method is preferred for uneven data distribution because it may classify the landslide susceptibility map into a separate category based on the inherent data value similarity (Mallick et al., 2022).

Therefore, the KNN Map of the LSI values was classified into five susceptibility areas (Figure 5.10): very low (0 - 0.156), low (0.156 - 0.403), moderate (0.403 - 0.556), high (0.556 - 0.690), very high (0.690 - 1). Based on the model-generated map, the distribution of areas across the five classes is as follows: 18.1% in "very low," 21.4% in "low," 15.8% in "moderate," 24.9% in "high," and 19.7% in "very high" (Table 8). Consequently, it can be inferred that approximately 44.6% of the study area falls into the "high" and "very high" categories, indicating a significant susceptibility to landslides in these regions.

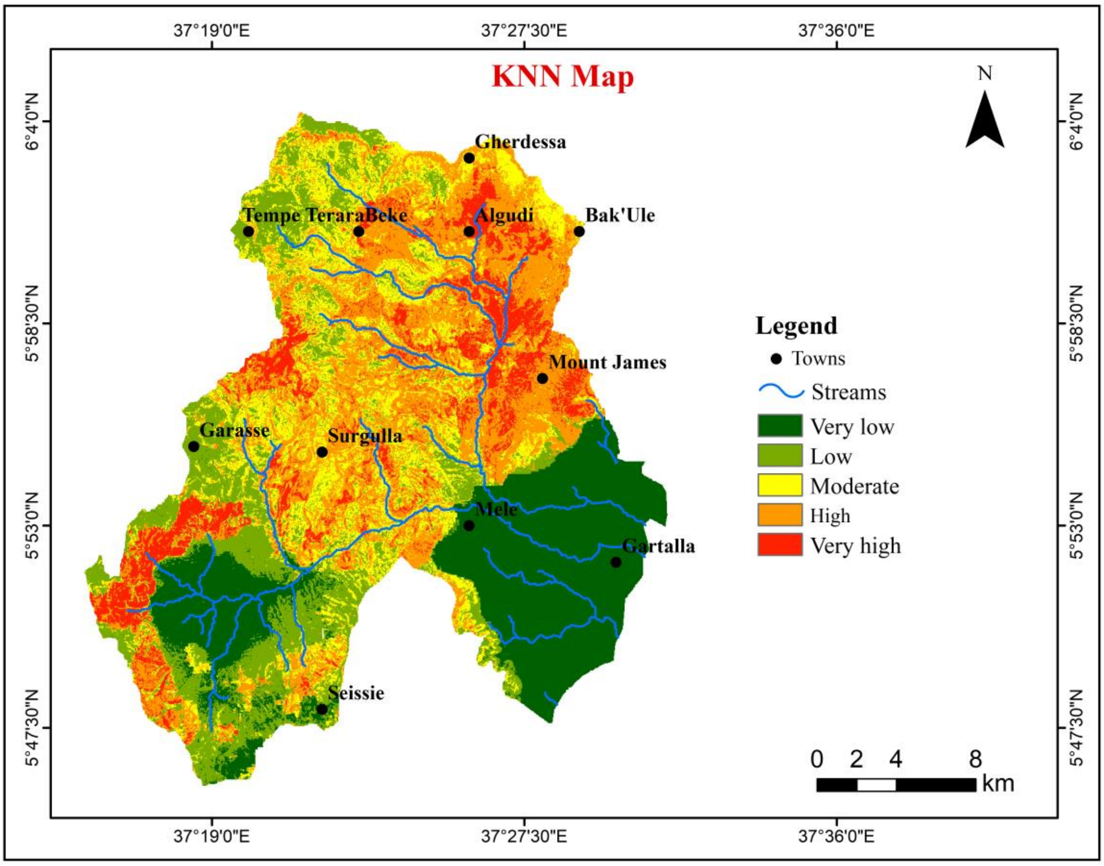


Figure 18: Landslide susceptibility map of the study area using KNN.

Table 8: Area and percentage of landslides for each of the susceptibility classes.

Classes	**SVM**		**KNN**	
	Area (Km^2)	**Percentage**	**Area (Km^2)**	**Percentage**
Very low	87.61	15.8%	100.73	18.1%
Low	90.17	16.2%	119.41	21.4%
Moderate	95.86	17.2%	87.85	15.8%
High	124.21	22.3%	138.6	24.9%
Very high	158.01	28.4%	109.27	19.7%

4.5 Validation of susceptibility maps

An essential step in landslide susceptibility modeling is assessing a model's performance using statistical measures. The most often used statistical method for determining the reliability and competency of a binary prediction model is the Receiver Operating Characteristic (ROC) curve. The prediction performance of a model is measured by the Area Under Curve (AUC) (Mallick et al., 2022). It plots sensitivity as a function of false positive rate (1-specificity). The ratio of true positives to the total of true positives and false negatives indicates a model's sensitivity. The ratio of true negatives to the total of false positives and true negatives is known as specificity. Plotting sensitivity on the y-axis against the false positive rate's cumulative distribution function on the x-axis will create the ROC curve. The accuracy in the ROC approach can be characterized by Four categories: poor (0.5–0.6), average (0.6–0.7), good (0.7–0.8), very good (0.8–0.9), and excellent (0.9–1) that can be applied to the predicted AUC value.

The results showed that the KNN model achieved an AUC of 79.9% for the success rate curve and 78.5% for the predictive rate curve. The SVM model attained an AUC of 81% for both the success and predictive rate curves (Figure 19). Therefore, with an AUC success rate of 81%, the SVM model is identified as the optimal model for the study area.

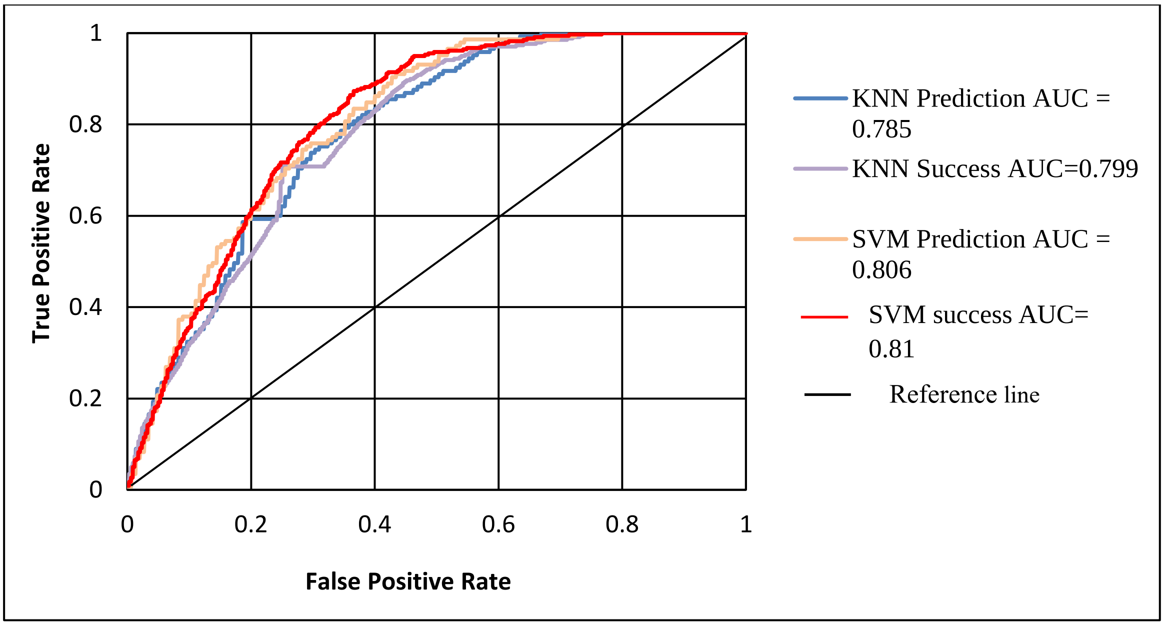


Figure 19: Validation for the Susceptibility Map.

CHAPTER FIVE

5. CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This study comprehensively analyses landslide susceptibility in the Sile and Elgo catchments using K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) machine learning models. Ten key causative factors were evaluated: slope, aspect, distance to streams, distance to lineaments, soil type, lithology, land use/land cover (LULC), rainfall, elevation, and curvature. A total of 534 landslide points and 534 non-landslide points were systematically analyzed. The dataset was divided, with 70% used for training the models and the remaining 30% used for validation.

The frequency ratio analysis revealed that the soil "Regosols" class, with a Frequency Ratio (FR) of 36.447was the most influential factor in landslide prediction. The susceptibility results indicated that the SVM model classified 50.7% of the study area as high or very high susceptibility, compared to 44.6% by the KNN model. Additionally, the AUC values demonstrated the success accuracy of both models, with SVM achieving 80.0% and KNN 80.6%.

These findings highlight the effectiveness of machine learning models for mapping landslide susceptibility. Future research could enhance these models with deep learning techniques to better assess landslide damage. Engaging residents through surveys may also provide insights into the economic impacts, helping to link damage costs to land-use changes driven by local activities and investments. Additionally, strategies to minimize landslide risks include reinforcing slopes, reforestation, and improving drainage systems. Public awareness, resettlement, retaining walls, and vegetation management, are essential for reducing hazards and enhancing resilience.

5.2 Recommendation

Strengthening crucial slopes is highly recommended to ensure stability and prevent mass wasting and erosion. Effective slope stabilization not only mitigates the risk of landslides but also preserves the integrity of the landscape and infrastructure. Implementing proper drainage systems, vegetation cover, and engineering solutions are essential strategies for enhancing slope strength and reducing erosion. Consequently, adopting these measures around gerese contributes to the long-term sustainability and safety of the area.

Reforestation: Due to activities such as logging, agriculture, and natural events. Preventing soil erosion and maintaining water cycles. This process can involve natural regeneration, where trees can regrow.

Efficient management of drainage systems: Drainage is a crucial element in depicting the slope evolution of a region and serves as an indicator of mass wasting and associated erosion characteristics. Throughout the year, the area experiences a significant spring with a high potential discharge, accompanied by a series of low-yield springs at the base of the hills. Therefore, effective management of drainage systems is highly recommended especially around Gina Village, Mele, and Mercha Maro Kebele.

Resettlement: The areas located under unstable zones are likely to experience future landslide activity. Hence, it is essential to relocate the residents currently living in these zones, particularly in Surgulla, Dambile Health Center, and Gedeb Village, which are situated at the base of a highly unstable slope.

Implement public awareness: Implement public awareness campaigns to educate communities about the risks of landslides, recognize early warning signs, and understand appropriate response measures. This effort should also encompass the development of emergency plans and evacuation procedures.

REFERENCES

Abebe, B., Dramis, F., Fubelli, G., Umer, M., and Asrat, A. (2010). Landslides in the Ethiopian highlands and the Rift margins. Journal of African Earth Sciences, 56(4–5), 131–138.

Adnan, M. S. G., Rahman, M. S., Ahmed, N., Ahmed, B., Rabbi, M. F., and Rahman, R. M. (2020). Improving spatial agreement in machine learning-based landslide susceptibility mapping. Remote Sensing, 12(20), 3347.

Adnan, M. S. G., Rahman, M. S., Ahmed, N., Ahmed, B., Rabbi, M. F., and Rahman, R. M. (2020). Improving spatial agreement in machine learning-based landslide susceptibility mapping. Remote Sensing, 12(20), 1–23.

Ahmed, Z., Hussain, A. H. M. B., Ambinakudige, S., Ahmed, M. N. Q., Alam, R., Rezoan, H.-A.-, Das Dola, D., Rahman, M. M., Hassan, R., and Mahmud, S. (2022). Perceived human-induced causes of landslide in Chattogram metropolitan area in Bangladesh. Earth Systems and Environment, 6(2), 499–515.

Akinci, H., and Zeybek, M. (2021). Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey. Natural Hazards, 108(2), 1515–1543.

Ayalew, L., and Yamagishi, H. (2005). The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. Geomorphology, 65(1–2), 15–31.

Ayele, A., and Arvidsson, R. (1997). Fault mechanisms and tectonic implication of the 1985–1987 earthquake sequence in south-western Ethiopia. Journal of Seismology, 1(4), 383–394.

Belayneh, L., Dewitte, O., Gulie, G., Poesen, J., O'Hara, D., Kassaye, A., Endale, T., and Kervyn, M. (2022). Landslides and Gullies Interact as Sources of Lake Sediments in a Rifting Context: Insights from a Highly Degraded Mountain Environment. Geosciences, 12(7), 274.

Chimidi, G., Raghuvanshi, T. K., and Suryabhagavan, K. V. (2017). Landslide hazard evaluation and zonation in and around Gimbi town, western Ethiopia—a GIS-based statistical approach. Applied Geomatics, 9(4), 219–236.

Chowdhury, M., and Sadek, A. W. (2012). Advantages and limitations of artificial intelligence. Artificial Intelligence Applications to Critical Transportation Issues, 6(3), 360–375.

Corominas, J., Moya, J., Ledesma, A., Lloret, A., and Gili, J. A. (2005). Prediction of ground displacements and velocities from groundwater level changes at the Vallcebre landslide (Eastern Pyrenees, Spain). Landslides, 2(2), 83–96.

Daviran, M., Shamekhi, M., Ghezelbash, R., and Maghsoudi, A. (2023). Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm. International Journal of Environmental Science and Technology, 20(1), 259–276.

Đurić, U., Marjanović, M., Radić, Z., and Abolmasov, B. (2019). Machine learning based landslide assessment of the Belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept. Engineering Geology, 256, 23–38.

George, R. M., and Rogers, N. W. (2002). Plume dynamics beneath the African plate inferred from the geochemistry of the Tertiary basalts of southern Ethiopia. Contributions to Mineralogy and Petrology, 144(3), 286–304.

Getachew, N., and Meten, M. (2021). Weights of evidence modeling for landslide susceptibility mapping of Kabi-Gebro locality, Gundomeskel area, Central Ethiopia. Geoenvironmental Disasters, 8, 1–22.

Ghorbanzadeh, O., Blaschke, T., Aryal, J., and Gholaminia, K. (2020). A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. Journal of Spatial Science, 65(3), 401–418.

Hamedi, H., Alesheikh, A. A., Panahi, M., and Lee, S. (2022). Landslide susceptibility mapping using deep learning models in Ardabil province, Iran. Stochastic Environmental Research and Risk Assessment, 36(12), 4287–4310.

Huang, Y., and Zhao, L. (2018). Review on landslide susceptibility mapping using support vector machines. Catena, 165, 520–529.

Huang, Z., Peng, L., Li, S., Liu, Y., and Zhou, S. (2023). GIS-based landslide susceptibility mapping in the Longmen Mountain area (China) using three different machine learning algorithms and their comparison. Environmental Science and Pollution

Research, 30(38), 88612–88626.

Hussain, M. A., Chen, Z., Kalsoom, I., Asghar, A., and Shoaib, M. (2022). Landslide susceptibility mapping using machine learning algorithm: A case study along Karakoram Highway (KKH), Pakistan. Journal of the Indian Society of Remote Sensing, 50(5), 849–866.

Hussain, M. A., Chen, Z., Wang, R., Shah, S. U., Shoaib, M., Ali, N., Xu, D., and Ma, C. (2022). Landslide susceptibility mapping using machine learning algorithm. Civ. Eng. J, 8, 209–224.

Ibrahim, M. B., Mustaffa, Z., Balogun, A. B., Indra, S. H. H., and Nur Ain, A. (2022). Landslide's analysis and Hazard mapping based on ANALYTIC HIERARCHY PROCESS (AHP) using GIS, in Lawas, Sabah-Sarawak. IOP Conference Series: Earth and Environmental Science, 1064(1).

Jaboyedoff, M., Michoud, C., Derron, M.-H., Voumard, J., Leibundgut, G., Sudmeier-Rieux, K., Nadim, F., and Leroi, E. (2018). Human-induced landslides: toward the analysis of anthropogenic changes of the slope environment. In Landslides and engineered slopes. Experience, theory and practice (pp. 217–232). CRC Press.

Jennifer, J. J. (2022). Feature elimination and comparison of machine learning algorithms in landslide susceptibility mapping. Environmental Earth Sciences, 81(20).

Kavzoglu, T., Colkesen, I., and Sahin, E. K. (2019). Machine learning techniques in landslide susceptibility mapping: a survey and a case study. Landslides: Theory, Practice and Modelling, 283–301.

Kebeba, O., Shano, L., Chemdesa, Y., and Jothimani, M. (2024). Integration of Geospatial Analysis, Frequency Ratio, and Analytical Hierarchy Process for Landslide Susceptibility Assessment in the Maze Catchment, Omo Valley, Southern Ethiopia. Quaternary Science Advances, 100203.

Kumar, A., Singh, V. K., Misra, R., Singh, T. N., and Choudhury, T. (2023). Machine learning and IoT-based approaches to detect and predict rainfall-triggered landslides. Revue d'Intelligence Artificielle, 37(5).

Lewis, D. P., Jebara, T., and Noble, W. S. (2006). Support vector machine learning from heterogeneous data: an empirical analysis using protein sequence and structure.

Bioinformatics, 22(22), 2753–2760.

Liu, Z., Gilbert, G., Cepeda, J. M., Lysdahl, A. O. K., Piciullo, L., Hefre, H., and Lacasse, S. (2021). Modelling of shallow landslides with machine learning algorithms. Geoscience Frontiers, 12(1), 385–393.

Loche, M., Alvioli, M., Marchesini, I., Bakka, H., and Lombardo, L. (2022). Landslide susceptibility maps of Italy: Lesson learnt from dealing with multiple landslide types and the uneven spatial distribution of the national inventory. Earth-Science Reviews, 232, 104125.

Mahmoud, S. H., and Gan, T. Y. (2018). Multi-criteria approach to develop flood susceptibility maps in arid regions of Middle East. Journal of Cleaner Production, 196, 216–229.

Mallick, J., Alqadhi, S., Talukdar, S., Sarkar, S. K., Roy, S. K., and Ahmed, M. (2022). Modelling and mapping of landslide susceptibility regulating potential ecosystem service loss: an experimental research in Saudi Arabia. Geocarto International, 37(25), 10170–10198.

Mandal, B., and Mandal, S. (2018). Analytical hierarchy process (AHP) based landslide susceptibility mapping of Lish river basin of eastern Darjeeling Himalaya, India. Advances in Space Research, 62(11), 3114–3132.

Martínek, K., Verner, K., Hroch, T., Megerssa, L. A., Kopačková, V., Buriánek, D., Muluneh, A., Kalinová, R., Yakob, M., and Kassa, M. (2021). Main Ethiopian Rift landslides formed in contrasting geological settings and climatic conditions. Natural Hazards and Earth System Sciences, 21(11), 3465–3487.

Mebrahtu, T. K., Hussien, B., Banning, A., and Wohnlich, S. (2021). Predisposing and triggering factors of large-scale landslides in Debre Sina area, central Ethiopian highlands. Bulletin of Engineering Geology and the Environment, 80(1), 365–383.

Miner, A., Vamplew, P., Windle, D. J., Flentje, P., and Warner, P. (2010). A comparative study of various data mining techniques as applied to the modeling of landslide susceptibility on the Bellarine Peninsula, Victoria, Australia.

Montgomery, D. R., and Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. In WATER RESOURCES RESEARCH

(Vol. 30, Issue 4).

Montrasio, L., and Valentino, R. (2007). Experimental analysis and modelling of shallow landslides. Landslides, 4(3), 291–296.

Mulugeta, T., Shano, L., and Jothimani, M. (2024). Landslide susceptibility modeling in the Kulfo river catchment, rift valley, Ethiopia: An integrated geospatial and statistical analysis. Quaternary Science Advances, 14, 100191.

Noviyanto, A., Sartohadi, J., and Purwanto, B. H. (2020). The distribution of soil morphological characteristics for landslide-impacted Sumbing Volcano, Central Java-Indonesia. Geoenvironmental Disasters, 7(1), 25.

Oyda, Y., Jothimani, M., and Regasa, H. (2024). Assessing landslide susceptibility in Lake Abya catchment, Rift Valley, Ethiopia: A GIS-based frequency ratio analysis. Journal of Degraded and Mining Lands Management, 11(3), 5885–5895.

Oyda, Y., and Regasa, H. (2023). Landslide Assessment and Hazard Zonation in the Birbir Mariam District, Gamo Highlands, Rift Valley Escarpment, Ethiopia. In Natural Hazards-New Insights. IntechOpen.

Raghuvanshi, T. K., Ibrahim, J., and Ayalew, D. (2014). Slope stability susceptibility evaluation parameter (SSEP) rating scheme–an approach for landslide hazard zonation. Journal of African Earth Sciences, 99, 595–612.

Rebala, G., Ravi, A., and Churiwala, S. (2019). An introduction to machine learning. Springer.

Rooney, T. O. (2010). Geochemical evidence of lithospheric thinning in the southern Main Ethiopian Rift. Lithos, 117(1–4), 33–48.

Segoni, S., Piciullo, L., and Gariano, S. L. (2018). A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides, 15(8), 1483–1501.

Shano, L., Gisila, B., Jerene, W., Ekaso, D., Baye, T., Degife, T., Chameno, G., Dosa, Z., and Jothimani, M. (2024). Fatal landslides in Kencho, Shacha and Gozdi villages, Gofa zone, Ethiopia: A detailed investigation (Geological, Geotechnical, geophysical and geospatial) of the July 22, 2024 catastrophe and its socioeconomic repercussions. Quaternary Science Advances, 16(September).

Shano, L., Raghuvanshi, T. K., and Meten, M. (2020). Landslide susceptibility evaluation and hazard zonation techniques – a review. In Geoenvironmental Disasters (Vol. 7, Issue 1). Springer.

Shano, L., Raghuvanshi, T. K., and Meten, M. (2021). Landslide hazard zonation using logistic regression model: The case of Shafe and Baso catchments, Gamo highland, Southern Ethiopia. Geotechnical and Geological Engineering, 1–19.

Shano, L., Raghuvanshi, T. K., and Meten, M. (2022). Landslide Hazard Zonation using Logistic Regression Model: The Case of Shafe and Baso Catchments, Gamo Highland, Southern Ethiopia. Geotechnical and Geological Engineering, 40(1), 83–101.

Sim, K. Ben, Lee, M. L., and Wong, S. Y. (2022). A review of landslide acceptable risk and tolerable risk. In Geoenvironmental Disasters (Vol. 9, Issue 1). Springer Science and Business Media Deutschland GmbH.

Singh, K., Arya, A. K., and Agarwal, K. K. (2020). Landslide occurrences along lineaments on NH-154A, Chamba, Himachal Pradesh; extracted from satellite data Landsat 8, India. Journal of the Indian Society of Remote Sensing, 48, 791–803.

Singh, P. K., Kar, A. K., Singh, Y., Kolekar, M. H., and Tanwar, S. (2019). Proceedings of ICRIC 2019: Recent Innovations in Computing (Vol. 597). Springer Nature.

Taalab, K., Cheng, T., and Zhang, Y. (2018). Mapping landslide susceptibility and types using Random Forest. Big Earth Data, 2(2), 159–178.

Tadesse, L., Uncha, A., and Toma, T. (2024). Landslide vulnerability mapping using multi-criteria decision-making approaches: in Gacho Babba District, Gamo Highlands Southern Ethiopia. Discover Applied Sciences, 6(2), 1–15.

Tadesse, L., Uncha, A., and Toma, T. (2024). Landslide vulnerability mapping using multi - criteria decision - making approaches : in Gacho Babba District , Gamo Highlands Southern. Discover Applied Sciences.

Teffera, F. E. (2016). The ecology of the major Ethiopian Rift Valley lakes Abaya and Chamo, with special reference to water quality and food web structure.

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

Tesfa, C., and Woldearegay, K. (2021). Characteristics and susceptibility zonation of landslides in Wabe Shebelle Gorge, south eastern Ethiopia. Journal of African Earth Sciences, 182, 104275.

Thi Ngo, P. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Kariminejad, N., Cerda, A., and Lee, S. (2021). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. Geoscience Frontiers, 12(2), 505–519.

Valdes Carrera, A. C., Mendoza, M. E., Allende, T. C., and Macías, J. L. (2023). A review of recent studies on landslide hazard in Latin America. Physical Geography, 44(3), 243–286.

Verner, K., Buriánek, D., Svojtka, M., Peřestý, V., Megerssa, L., Tadesse, T., Kussita, A., Alemayehu, D., and Hroch, T. (2021). Tectonometamorphic evolution and U–Pb dating of the high-grade Hammar Domain (Southern Ethiopian Shield); implications for the East-African Orogeny. Precambrian Research, 361, 106270.

Whipp, D. M., and Ehlers, T. A. (2019). Quantifying landslide frequency and sediment residence time in the Nepal Himalaya. Science Advances, 5(4), eaav3482.

Woldearegay, K. (2013). Review of the occurrences and influencing factors of landslides in the highlands of Ethiopia: With implications for infrastructural development. In Momona Ethiopian Journal of Science (MEJS) (Vol. 5, Issue 1).

Wubalem, A. (2021). Landslide susceptibility mapping using statistical methods in Uatzau catchment area, northwestern Ethiopia. Geoenvironmental Disasters, 8(1), 1.

Wubalem, A., Getahun, B., Hailemariam, Y., Mesele, A., Tesfaw, G., Dawit, Z., and Goshe, E. (2022). Landslide susceptibility modeling using the index of entropy and frequency ratio method from Nefas-Mewcha to Weldiya Road Corridor, Northwestern Ethiopia. Geotechnical and Geological Engineering, 40(10), 5249–5278.

Yalcin, A. (2005). An investigation on Ardesen (Rize) region on the basis of landslide susceptibility. Karadeniz Technical University Trabzon, Turkey.

Yalcin, A., and Bulut, F. (2007). Landslide susceptibility mapping using GIS and digital photogrammetric techniques: A case study from Ardesen (NE-Turkey). Natural Hazards, 41(1), 201–226.

Zhang, C., Yin, Y., Yan, H., Li, H., Dai, Z., and Zhang, N. (2021). Reactivation characteristics and hydrological inducing factors of a massive ancient landslide in the three Gorges Reservoir, China. Engineering Geology, 292, 106273.

Zheng, X., He, G., Wang, S., Wang, Y., Wang, G., Yang, Z., Yu, J., and Wang, N. (2021). Comparison of machine learning methods for potential active landslide hazards identification with multi-source data. ISPRS International Journal of Geo-Information, 10(4), 253.

APPENDIX

Appendix 1: A literature review from different Machine learning Articles.

	Title	Author Name and Year	Journal Name	Influencing factors	Method
1	A Comparative Assessment of Machine Learning Models for Landslide Susceptibility Mapping in the Rugged Terrain of Northern Pakistan	Naeem Shahzad et al.,2022	Applied science	• Elevation, • Slope, • Aspect, • Topographic wetness index (TWI), • Topographic position index (TPI), • Distance to drainage, • Distance to fault, • Distance to road, • Normalized difference vegetation index (NDVI), • Rainfall,	• Random forest • Support vector machine • Maximum entropy • Gradient-boosting machine • Logistic regression

				• Land cover/land use (LCLU), • Geological map	
2	Landslide susceptibility assessment and mapping using state-of-the art machine learning techniques	Hamid Reza Pourghasem et al.,2021	Earth Science Informatics	• Altitude • Slope degree • Slope aspect • Plan curvature • Profile curvature • Distance from rivers • Distance from road • Distance from fault • Geology features • Land-use features	• Generalized linear mode (GLM) • Mixture discriminant analysis (MDA) • Boosted regression tree (BRT) • Functional discriminant analysis (FDA)
3	Spatiotemporal landslide susceptibility mapping using machine learning models: A case study from district Hattian Bala, NW Himalaya, Pakistan	Ahmad Hammad Khaliq et al.,2022	Ain Shams Engineering Journal	• Slope • Aspect • elevation, • curvature, • plane curvature, • profile curvature, • topographic wetness index (TWI), • lithology, • distance to faults,	• Random Forest (RF) • Logistic Regression (LR)

				• distance to streams, • distance to roads, • normalized difference vegetation index (NDVI) • land use/ land cover (LULC)	
4	GIS-based landslide susceptibility mapping of the Meghalaya-Shillong Plateau region using machine learning algorithms	Navdeep Agrawal and Jagabandhu Dixit 2023	Bulletin of Engineering Geology and the Environment	• Aspect • Curvature • Distance from faults • Distance from rivers • Distance from roads • Elevation • Geomorphology • Lithology • LULC • NDVI • Rainfall • Slope • Topography wetness index (TWI) • Soil texture • Stream power index (SPI)	• artificial neural network (ANN), • extreme gradient boosting (XGBoost), • K-nearest neighbors (KNN), • random forest (RF), • support vector machine (SVM)
5	Landslide hazard risk modeling in north-west	Ali Jamali (2020)	Modeling Earth Systems	• Elevation • Slope	• Random Forest,

	of Iran using optimized machine learning models		and Environment	• Aspect • Curvature • Rainfall • Distance to stream • Lithology	• eXtreme Gradient Boosting, • Salp Swarm Algorithm Multi-layer Perceptron (SSAMLP), • Multi-verse Optimization Multi-layer Perceptron (MVOMLP)
6	An Integrated Approach of Machine Learning, Remote Sensing, and GIS Data for the Landslide Susceptibility Mapping	Israr Ullah et al.,2022	Land	• elevation, • slope degree, • slope aspect, • general curvature, • plan curvature, • profile curvature, • landcover classification system, • Normalized Difference Water Index (NDWI), • Normalized Difference Vegetation Index (NDVI), • soil, • lithology,	• Linear Regression (LiR), • Logistic Regression (LoR), • Support Vector Machine (SVM)

				• fault density, • topographic roughness index, • road density	
7	A Comparative Analysis of Certainty Factor-Based Machine Learning Methods for Collapse and Landslide Susceptibility Mapping in Wenchuan County, China	Xinyue Yuan et al.,2022	Remote Sensing	• Slope • Aspect • Curvature • Terrain relief • Topographic wetness index (TWI) • Lithology • Soil type • Distance to fault • Rainfall • Distance to river • Peak ground acceleration (PGA) • Normalized difference vegetation index (NDVI) • Land use	• Logistic regression • Support vector machine • Random forest

8	Landslide Susceptibility Mapping Using Machine Learning Algorithm: A Case Study Along Karakoram Highway (KKH), Pakistan	Muhammad Afaq Hussain et al.,2021	Journal of the Indian Society of Remote Sensing	• Distance to River • Curvature • NDVI • TWI • Landcover • Roughness	• Random Forest (RF), • Extreme Gradient Boosting (XGBoost), • k-Nearest Neighbors (KNN)
9	Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia	Ahmed Mohamed Youssef and Hamid Reza Pourghasemi 2021	Geoscience Frontiers	• Altitude, • Lithology, • Distance to faults, • Normalized difference vegetation index (NDVI), • Landuse/landcover (LULC), • Distance to roads, • Slope angle, • Distance to streams, profile curvature, • Plan curvature, • Slope length (LS), • Slope-aspect.	• SVM • RF • Multivariate Adaptive Regression Spline (MARS), • Artificial Neural Network (ANN), • Quadratic Discriminant Analysis (QDA), • Linear Discriminant Analysis (LDA), • Naive Bayes (NB)

10	Landslide Susceptibility Mapping using Machine Learning Algorithm	Muhammad Afaq Hussain et al.,2022	Civil Engineering Journal	• Aspect, • Elevation, • Slope, • NDVI, • Curvature, • SPI, • TWI, • Lithology, • Landcover.	• Extreme Gradient Boosting (XGBoost), • Random Forest (RF), • k-Nearest Neighbors (KNN)
11	Landslide Susceptibility Mapping Using Machine Learning	Ageenko et al., 2022	Geo-Information	• Elevation, • Terrain Roughness Index (TRI), • Standard Deviation of Slope, • Distance from the Coast, • Distances from Roads, • Distances from Railroads, • Distances from Quarries	• RF • SVM • LR

12	GIS-based landslide susceptibility mapping in the Longmen Mountain area (China) using three diferent machine learning algorithms and their comparison	Ziyan Huang et al.,2023	Environmental Science and Pollution Research	• Elevation • Slope • Altitude difference • Rainfall • NDVI • Distance to road • Distance to river • Distance to fault • Land use • PGA(Peak ground acceleration) • Lithology	• Random forest (RF), • Support vector machine (SVM), • Decision tree (DT) regression
13	Landslide Susceptibility Mapping in Three Upazilas of Rangamati Hill District Bangladesh: Application and Comparison of GIS-based Machine Learning Methods	Yasin Wahid Rabby et al.,2020	Geocarto International	• TWI), • Stream Power Index (SPI), • NDVI, • LULC, • land use/land cover change, • distance from the road, • Elevation, • slope, • aspect, • plan, • profile curvatures	• K-Nearest Neighbour (KNN), Random Forest (RF) • Extreme Gradient Boosting (XGBoost)

14	A Comparison Study of Landslide Susceptibility Spatial Modeling Using Machine Learning	Nurwatik et al.,2022	Geo-Information	• Slope, • Elevation, • Aspect, • NDVI, • Geological type, • Soil type, • Distance from the fault • Distance from the river, • River density, • TWI, • Land cover, • Annual rainfall	• RF, • NB (naïve Bayes), • KNN (k-nearest neighbor)
15	Landslide susceptibility mapping of mountain roads based on machine learning combined model	Dou HQ et al.,2023	J. Mt. Sci.	• Elevation • Slope angle • Curvature • Topographic relief • Lithology • Distance to fault • Distance to river • Rainfall • NDVI • Distance to road • FVC	• SVM, • RF • Artificial neural network (ANN)

16	Factors Affecting Landslide Susceptibility Mapping: Assessing the Influence of Different Machine Learning Approaches, Sampling Strategies and Data Splitting	Minu Treesa Abraham et al.,2021	Land	• elevation • slope • aspect • NDVI • SPI • TWI • Rainfall	• RF • KNN • SVM • LR • NB
17	Improving Spatial Agreement in Machine Learning-Based Landslide Susceptibility Mapping	Mohammed Sarfaraz Gani Adnan et al.,2020	Remote sensing	• Aspet • Elevation • Curvature • Slope angle • Stream Power Index (SPI) • Distance to stream • Landcover • NDVI • Geology • Soil type • Soil texture • Distance to road	• KNN) • Multi-Layer Perceptron (MLP), • RF , • SVM

18	Mapping Landslide Susceptibility Using Machine Learning Algorithms and GIS: A Case Study in Shexian County, Anhui Province, China	Zitao Wang et al.,2020	Symmetry	• Slope aspect, • Slope degree, • Plane curvature, • Profile curvature, • Topographic relief, • Surface roughness • Distance from fault • Lithology • Distance from river • SPI • TWI • Flow length • Rainfall • NDVI • Distance from roads	• LR, • SVM, • RF, • Gradient Boosting Machine (GBM), • Multilayer Perceptron (MLP)
19	Landslide susceptibility assessment using SVM machine learning algorithm	Miloš Marjanović et al.,2011	Engineering geology	• Elevation • Slope • Aspect • Slope length • TWI • Plan curvature • Profile curvature • Distance from stream • Lithology	• SVM, • Decision Trees • Logistic Regression

				• Distance from fault • NDVI • Distance from geo-boundary	
20	Machine learning for high-resolution landslide susceptibility mapping: case study in Inje County, South Korea	Xuan-Hien Le et al.,2023	Frontiers	• Soil depth • Slope • Azimuth • Curvature • TWI • Forest condition • Catchment length • Catchment area • Tree diameter • Bed rock	• LR • KNN • RF • GNB • SVM • XGB
21	Comparison of machine learning methods for potential active landslide hazards identification with multi-source data.	Zheng et al., 2021	Geo-Information	• Slope • Aspect • Curvature • Distance from river • Drainage desnsity • Topographic index • NDVI • Landcover	• Naive Bayes (NB), • Decision Tree (DT), • Support Vector Machine (SVM) and

				• Distance from road	• Random Forest (RF) algorithms
22	Feature elimination and comparison of machine learning algorithms in landslide susceptibility mapping	Jennifer, 2022	Environmental Earth Sciences	• Rainfall • Elevation • Slope • Curvature • TWI • SPI • NDVI • Soil • Geology • Geomorphology • Distance to lineaments • LULC • Distance to streams	• , Adaptive Boosting (AdaBoost), • Naïve Bayes (NB), • Neural Network (NNET), • RF, • SVM and • eXtreme Gradient Boosting (XGBoost)
23	Comparing classical statistic and machine learning models in landslide susceptibility mapping in Ardanuc (Artvin), Turkey	Akinci and Zeybek, 2021	Natural hazards	• Altitude • Aspect • Land cover • Curvature • Distance to drainage • Distance to fault • Lithology • Slope • TWI • Distance to road	• SVM, • RF, and • LR

24	Machine learning based landslide assessment of the Belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept	Đurić et al., 2019	Engineering geology	• Aspect • TWI • TRI • Plan curvature • Lithology • Hydrology • Land cover •	• SVM • RF
25	Improving Spatial Agreement in Machine Learning-Based Landslide Susceptibility Mapping	(Adnan et al., 2020b)	MDPI	• Aspect • Elevation • Slope • Curvature • SPI • Distance to stream • Land cover • NDVI • Geology • Soil texture • Distance to road	• KNN, • Multi-Layer Perceptron (MLP), • RF, and • SVM,

Appendix 2a: Annual rainfall in each station.

	2013	2014	2015	2016	2017	2018	2019	2020	2021	20202
Arbaminch	1061.1	866.2	1121.3	795.1	752.6	787.2	1866.9	1241	860.35	502.2
Arfaide	628.1	976.1	106.5	636.5	890.2	980.2	933.95	1391.7	887.35	528.2
Chencha	1183.3	1254.9	978.6	950.3	413.5	413.5	750.55	453	374.35	268.1
Kemba	869.35	1498.9	1576.7	1108.3	1694.35	2160.4	1501.5	1806.6	1392.05	1432.3
Malanaha	1224	1079.25	930.05	1019.4	1038.3	1005.5	1213.65	1484.4	827.95	781.3
Gerese	1318	1207.63	1531.48	1139	1165.42	1207.6	1434.37	2030.26	1416.54	1104.3
Arguba	1722.7	914.2	983.4	774.6	907.8	959.95	845.7	832.65	358.55	408.6

Appendix 2b: Maximum monthly rainfall.

	Arbaminch	Arfaide	Chencha	Kemba	Malanaha	Gerese	Arguba
April	169.64	183.6	89.168	286.735	201.805	282.88	176.68
May	145.42	98.51	167.44	256.975	130.73	213.84	113.95
September	67.86	89.28	63.935	182.71	181.65	92.95	76.09
October	151.33	100.43	111.55	176.48	147.5	186.2	107

Appendix 3: Landslide inventory data of the Study area.

NO	Easting	Northing	Elevation(m)	Mode of failure
1	37.38693	5.88374	1393	Debris slide
2	37.45119	5.928996	1277	Earth slide
3	37.44957	5.926443	1154	Earth slide
4	37.30626	5.893119	1099	Earth slide
5	37.30658	5.894169	1011	Earth slide
6	37.31296	5.897017	1021	Earth slide
7	37.31795	5.89318	1352	Earth slide
8	37.3037	5.881814	1188	Earth slide
9	37.35411	5.944254	1286	Earth slide
10	37.34295	6.009693	1156	Earth slide
11	37.34601	5.914277	2900	Rock slide
12	37.34048	5.90324	1338	Rock slide
13	37.41923	6.01396	1332	Earth slide
14	37.39786	5.902428	1159	Earth slide
15	37.29683	5.807805	3671	Earth slide
16	37.46477	5.934805	2693	Earth slide
17	37.46127	5.936959	1944	Earth slide
18	37.41932	5.939573	1591	Earth slide
19	37.45409	5.937912	1307	Earth slide
20	37.40596	5.926305	787	Earth slide
21	37.41318	5.935597	1516	Earth slide
22	37.32555	5.891395	1578	Rock slide
23	37.27184	5.853327	418	Earth slide
24	37.29994	5.803133	1748	Earth slide
25	37.34394	5.893198	3493	Earth slide
26	37.34646	5.88538	2563	Earth slide
27	37.39968	5.901685	2249	Earth slide
28	37.43141	6.026397	1588	Rock slide
29	37.36584	5.877806	1288	Earth slide
30	37.3938	5.984187	1455	Earth slide
31	37.44968	5.990549	997	Earth slide
32	37.34289	5.951044	1534	Rock slide
33	37.42398	6.018338	746	Earth slide
34	37.44924	5.995014	1680	Earth slide
35	37.34055	5.95965	2887	Earth slide
36	37.44194	5.997546	1069	Earth slide
37	37.44133	6.002323	1127	Earth slide
38	37.44116	5.998994	982	Debris Slide
39	37.35925	5.943498	1240	Earth slide
40	37.2895	5.871221	933	Rock slide
41	37.28869	5.870938	978	Earth slide
42	37.29171	5.877258	2496	Earth slide
43	37.30056	5.881567	1234	Earth slide

44	37.29959	5.881462	1436	Earth slide
45	37.30123	5.881348	1348	Earth slide
46	37.30402	5.88296	1548	Earth slide
47	37.30505	5.884296	1288	Earth slide
48	37.30492	5.885373	1264	Earth slide
49	37.30567	5.885256	2480	Earth slide
50	37.30196	5.887003	1719	Earth slide
51	37.3031	5.886117	915	Earth slide
52	37.29706	5.877863	995	Earth slide
53	37.30811	5.877233	1182	Earth slide
54	37.31646	5.882937	1110	Earth slide
55	37.3176	5.882423	1104	Earth slide
56	37.31374	5.883853	1450	Earth slide
57	37.31102	5.885941	1776	Earth slide
58	37.3164	5.884424	1610	Earth slide
59	37.32093	5.887585	1802	Earth slide
60	37.3342	5.885514	974	Earth slide
61	37.3363	5.883844	1262	Earth slide
62	37.33622	5.885082	1668	Earth slide
63	37.33334	5.883217	1309	Earth slide
64	37.32994	5.883782	1258	Earth slide
65	37.32905	5.885232	2930	Earth slide
66	37.32547	5.887732	1296	Earth slide
67	37.30215	5.881997	939	Earth slide
68	37.30092	5.88476	1216	Earth slide
69	37.29233	5.882527	746	Earth slide
70	37.29576	5.882203	1954	Debris slide
71	37.2922	5.870674	1220	Debris slide
72	37.2956	5.872124	1377	Earth slide
73	37.29762	5.873652	981	Earth slide
74	37.30284	5.875357	1273	Earth slide
75	37.29915	5.871619	1460	Earth slide
76	37.33314	5.925251	1436	Transitional slide
77	37.33385	5.927128	1503	Earth slide
78	37.33255	5.921084	1783	Earth slide
79	37.34624	5.924005	2910	Earth slide
80	37.34415	5.928649	2657	Earth slide
81	37.34509	5.926365	1518	Earth slide
82	37.34609	5.953064	3290	Earth slide
83	37.35005	5.958516	996	Earth slide
84	37.34876	5.957808	1988	Earth slide
85	37.34918	5.892012	3259	Earth slide
86	37.3499	5.908206	1461	Earth slide
87	37.35422	5.903855	2222	Earth slide
88	37.41859	5.936288	1984	Earth slide
89	37.41671	5.938431	2942	Earth slide

90	37.40706	5.937972	1162	Earth slide
91	37.40918	5.935478	2936	Earth slide
92	37.4585	5.995506	3462	Earth slide
93	37.46231	5.981327	3334	Earth slide
94	37.48745	5.93596	1366	Earth slide
95	37.47347	5.941599	3819	Earth slide
96	37.40009	5.907804	1546	Earth slide
97	37.39396	5.916258	1265	Earth slide
98	37.39228	5.916843	1812	Earth slide
99	37.3871	6.05475	1298	Transitional slide
100	37.37182	6.060578	2554	Earth slide
101	37.41958	6.024965	2991	Earth slide
102	37.41273	6.028843	1948	Earth slide
103	37.46868	6.011185	1764	Earth slide
104	37.46368	6.013376	1093	Earth slide
105	37.44932	6.053675	2423	Earth slide
106	37.40601	5.970799	1979	Earth slide
107	37.40357	5.972877	1266	Transitional slide
108	37.39695	5.98155	2056	Earth slide
109	37.40483	5.980522	1743	Debris slide
110	37.38505	6.019261	1462	Earth slide
111	37.389	6.022781	2760	Earth slide
112	37.38719	6.019657	2859	Transitional slide
113	37.40529	6.003365	2747	Earth slide
114	37.40347	6.005593	1322	Earth slide
115	37.40293	6.004955	2360	Debris slide
116	37.38239	5.991547	2373	Earth slide
117	37.36897	5.995284	789	Debris slide
118	37.35099	5.980128	1226	Earth slide
119	37.3514	5.981536	918	Earth slide
120	37.35056	5.974415	1625	Earth slide
121	37.36194	5.948915	1635	Earth slide
122	37.33186	5.936481	1281	Debris slide
123	37.34301	5.95625	1466	Earth slide
124	37.34284	5.95679	1922	Debris slide
125	37.3602	5.994077	1443	Earth slide
126	37.36041	5.99521	2321	Earth slide
127	37.36227	6.017063	1869	Earth slide
128	37.38182	6.059741	978	Earth slide
129	37.34234	5.818415	1676	Earth slide
130	37.36017	5.899176	1121	Debris slide
131	37.3785	5.896633	1205	Debris slide
132	37.3769	5.896727	1344	earth slide
133	37.37468	5.897384	1266	Earth slide
134	37.42734	5.977948	1251	Earth slide
135	37.42146	5.995883	1165	Debris slide

136	37.4311	5.995177	1712	Earth slide
137	37.36464	6.059578	929	Earth slide
138	37.36388	6.058712	1880	Earth slide
139	37.40092	5.999405	1213	Earth slide
140	37.39521	6.008887	1909	Earth slide
141	37.40295	6.032505	1697	Earth slide
142	37.4017	6.03334	2563	Earth slide
143	37.3966	5.962399	1146	Earth slide
144	37.41456	5.991558	1460	Earth slide
145	37.41704	5.99081	1569	Earth slide
146	37.38428	5.941357	1135	Earth slide
147	37.38127	5.942498	927	Earth slide
148	37.37469	5.947548	899	Earth slide
149	37.38727	5.973255	951	Earth slide
150	37.4483	5.932722	1036	Earth slide
151	37.44288	5.933436	1776	Earth slide
152	37.44707	5.951941	1380	Earth slide
153	37.45508	5.981961	1283	Earth slide
154	37.43222	6.014277	1988	Earth slide
155	37.43576	6.019214	2074	earth slide
156	37.44985	6.007035	1458	earth slide
157	37.42482	6.043426	984	Rock slide
158	37.42575	6.04259	963	Rock slide
159	37.45751	6.0436	1025	Rock slide
160	37.362	5.92645	2476	Rock slide
161	37.3504	5.920538	1548	Earth slide
162	37.35221	5.879002	2276	Earth slide
163	37.34732	5.88211	1815	Earth slide
164	37.40129	5.890252	2590	Earth slide
165	37.39796	5.892185	1293	Earth slide
166	37.34296	5.90204	917	Earth slide
167	37.45897	6.041186	948	Earth slide
168	37.45444	6.047583	1606	Rock slide
169	37.45893	6.04011	956	Earth slide
170	37.3952	5.97922	1421	Transitional slide
171	37.41459	5.977511	1207	Debris slide
172	37.40459	5.968503	1524	Debris slide
173	37.41126	5.966573	1809	Earth slide
174	37.42518	5.957676	566	Earth slide
175	37.42208	5.964641	1213	Earth slide
176	37.42955	5.971334	2431	Earth slide
177	37.45798	5.953138	2571	Rock slide
178	37.4562	5.945824	2475	Rock slide
179	37.45974	5.94524	3011	Rock slide
180	37.45667	5.949658	3402	Rock slide
181	37.45924	5.958349	2894	Rock slide

182	37.45484	5.970073	2569	Rock slide
183	37.46071	5.970655	8063	Debris slide
184	37.45675	5.928362	2626	Earth slide
185	37.34621	5.910754	2352	Earth slide
186	37.34744	5.90821	2148	Earth slide
187	37.35368	5.936139	1744	Earth slide
188	37.3557	5.942136	1109	Earth slide
189	37.35679	5.947621	1571	Earth slide
190	37.36126	5.887958	1615	Earth slide
191	37.36898	5.886663	1164	Earth slide
192	37.37033	5.885464	1133	Earth slide
193	37.36467	5.888319	2496	Earth slide
194	37.36666	5.886408	1875	Earth slide
195	37.35405	5.914631	2315	Earth slide
196	37.3908	5.923866	3140	Earth slide
197	37.39394	5.923017	2201	Earth slide
198	37.38428	5.937775	1994	Earth slide
199	37.37777	5.945612	1366	Earth slide
200	37.35109	5.943798	2602	Earth slide
201	37.33827	5.947175	1152	Earth slide
202	37.39963	5.942286	2041	Earth slide
203	37.43118	5.965031	2126	Earth slide
204	37.44262	6.037971	1833	Earth slide
205	37.42164	6.012486	1263	Debris slide
206	37.36041	6.058987	2005	Earth slide
207	37.3797	5.925808	1854	Earth slide
208	37.36929	6.059053	1760	Earth slide
209	37.37574	6.060026	975	Earth slide
210	37.45031	5.99796	1615	Rock slide
211	37.46439	5.994099	1873	Rock slide
212	37.27779	5.852412	1789	Rock slide
213	37.27514	5.854835	1581	Earth slide
214	37.27758	5.857944	1128	Earth slide
215	37.27473	5.857166	1985	Transitional slide
216	37.2723	5.851658	1265	Earth slide
217	37.27339	5.849433	1882	Transitional slide
218	37.26626	5.844718	2873	Earth slide
219	37.27673	5.841286	1341	Debris slide
220	37.27374	5.841555	1359	Debris slide
221	37.28107	5.860704	2171	Rock slide
222	37.28668	5.862636	2069	Earth slide
223	37.27066	5.84827	1923	Earth slide
224	37.39107	6.053893	1179	Earth slide
225	37.28625	5.8293	1051	Debris slide
226	37.28527	5.842762	1989	Debris slide
227	37.28104	5.84297	1340	Earth slide

228	37.2717	5.846502	967	Debris slide
229	37.27514	5.851183	1272	Debris slide
230	37.26741	5.844938	994	Debris slide
231	37.31006	5.897558	1055	Earth slide
232	37.34418	5.818062	959	Earth slide
234	37.34417	5.814868	2051	Earth slide
235	37.2845	5.847745	1615	Earth slide
236	37.28438	5.868273	1332	Earth slide
237	37.28774	5.877902	1793	Earth slide
238	37.29375	5.876256	1246	Earth slide
239	37.47312	5.914713	1828	Debris slide
240	37.47435	5.919365	2604	Earth slide
241	37.47558	5.928346	4724	Earth slide
242	37.46633	5.937158	1878	Transitional slide
243	37.4698	5.941162	2694	Earth slide
244	37.48292	5.949569	3804	Rock slide
245	37.47675	5.953454	2808	Debris slide
246	37.47618	5.950649	2411	Debris slide
247	37.46897	5.96554	2186	Earth slide
248	37.47529	5.9675	1950	Earth slide
249	37.474	5.958282	914	Debris slide
250	37.48134	5.934388	3209	Debris slide
251	37.45144	5.932776	1922	Debris slide
252	37.45109	5.95217	1215	Debris slide
253	37.45418	5.9533	1154	Debris slide
254	37.44793	5.948283	1156	Earth slide
255	37.44538	5.946746	1778	Earth slide
256	37.40878	5.988847	1756	Earth slide
257	37.40834	5.964913	1967	Earth slide
258	37.39369	5.962179	755	Earth slide
259	37.39886	5.957031	1521	Earth slide
260	37.39408	5.953672	1547	Earth slide
261	37.38826	5.952632	711	Earth slide
262	37.46119	5.909383	1461	Earth slide
263	37.41241	5.938602	1671	Earth slide
264	37.4051	5.942917	2594	Earth slide
265	37.40001	5.937379	956	Transitional slide
266	37.4147	5.931458	1104	Earth slide
267	37.37067	5.892384	2551	Earth slide
268	37.3638	5.907395	989	Earth slide
269	37.36293	5.916903	1279	Earth slide
270	37.35727	5.826159	1664	Earth slide
271	37.35011	5.814806	1806	Debris slide
272	37.28983	5.821592	2920	Debris slide
273	37.28824	5.819349	1019	Earth slide
274	37.29459	5.817079	1822	Earth slide

275	37.28737	5.80891	1504	Earth slide
276	37.29298	5.810384	3021	Earth slide
277	37.28768	5.812786	1540	Transitional slide
278	37.28295	5.81357	934	Earth slide
279	37.29034	5.797829	1619	Earth slide
280	37.29926	5.795721	1542	Earth slide
281	37.30776	5.794737	1288	Earth slide
282	37.30715	5.787704	1914	Earth slide
283	37.29348	5.823593	1866	Earth slide
284	37.28231	5.83918	1172	Earth slide
285	37.29126	5.81672	2378	Transitional slide
286	37.38365	5.882699	1535	Debris slide
287	37.38464	5.885358	1079	Debris slide
288	37.38573	5.88755	830	Earth slide
289	37.38636	5.887845	616	Earth slide
290	37.38666	5.889609	949	Earth slide
291	37.3878	5.890992	858	Earth slide
292	37.39001	5.901261	1682	Earth slide
293	37.38605	5.904914	1149	Earth slide
294	37.38365	5.906829	973	Earth slide
295	37.38108	5.907626	1082	Earth slide
296	37.37302	5.898806	2943	Earth slide
297	37.35888	5.909414	1194	Earth slide
298	37.37239	5.990487	1270	Earth slide
299	37.37387	5.989843	1159	Earth slide
300	37.37422	5.988163	1204	Earth slide
301	37.36562	6.001499	1918	Earth slide
302	37.37183	6.043469	2798	Transitional slide
303	37.40963	6.029461	1898	Earth slide
304	37.40626	6.029802	923	Earth slide
305	37.408	6.0413	1354	Earth slide
306	37.42156	6.046108	1286	Earth slide
307	37.4238	6.044199	1376	Earth slide
308	37.42267	6.021714	925	Earth slide
309	37.42307	6.01933	1789	Earth slide
310	37.46094	6.028601	1042	Earth slide
311	37.46157	6.026838	1640	Earth slide
312	37.4587	6.026441	1334	Earth slide
313	37.4573	6.026545	951	Earth slide
314	37.45956	6.01468	2389	Earth slide
315	37.45809	6.005007	1118	Transitional slide
316	37.45778	6.007397	1044	Debris slide
317	37.46163	6.009143	1891	Debris slide
318	37.47431	5.996741	1880	Earth slide
319	37.42696	6.016571	1360	Transitional slide
320	37.42851	6.025978	2637	Earth slide

321	37.4293	6.028248	1490	Earth slide
322	37.43452	6.03387	1093	Earth slide
323	37.42882	6.048841	2092	Earth slide
324	37.44355	6.046482	1028	Earth slide
325	37.36222	5.970514	1570	Rock slide
326	37.35634	5.969326	994	Rock slide
327	37.36277	5.971954	1249	Earth slide
328	37.35766	5.995884	1785	Earth slide
329	37.3884	6.021158	2868	Earth slide
330	37.43287	6.017114	1796	Earth slide
331	37.43402	6.026546	1005	Earth slide
332	37.43334	6.025231	1209	Earth slide
333	37.37112	5.919602	1082	Earth slide
334	37.30437	5.8915	1012	Earth slide
335	37.35102	5.95915	1279	Earth slide
336	37.39924	5.914854	1596	Earth slide
337	37.39492	5.968344	1581	Earth slide
338	37.41865	5.950819	1361	Earth slide
339	37.43333	5.971911	2594	Debris slide
340	37.45129	5.991541	1423	Debris slide
341	37.45165	5.997049	1099	Debris slide
342	37.44742	5.995678	1508	Earth slide
343	37.42499	5.978133	1197	Earth slide
344	37.32426	5.890919	1177	Earth slide
345	37.31928	5.888857	1248	Earth slide
346	37.33311	5.91844	1853	Earth slide
347	37.41883	5.968091	2192	Earth slide
348	37.44681	6.001216	1469	Earth slide
349	37.32033	5.893463	1919	Earth slide
350	37.43408	5.960864	2767	Earth slide
351	37.37058	5.879828	2833	Earth slide
351	37.33378	5.953499	1452	Rock slide
352	37.33869	5.95272	1809	Earth slide
353	37.34268	5.953821	1683	Rock slide
354	37.34356	5.954453	1917	Rock slide
355	37.34131	5.956941	1381	Earth slide
356	37.33496	5.956784	1062	Earth slide
356	37.35122	5.968137	1389	Earth slide
357	37.35325	5.971061	1253	Rock slide
358	37.35145	5.976675	1344	Rock slide
359	37.35178	5.982711	1963	Earth slide
360	37.35078	5.983915	1008	Earth slide
361	37.35147	5.992657	1533	Rock slide
362	37.35047	5.991664	2112	Earth slide
363	37.35199	5.987681	1020	Earth slide
364	37.34988	5.990008	1342	Earth slide

365	37.35675	5.994727	1371	Earth slide
366	37.36055	5.992423	920	Earth slide
367	37.36571	5.9934	2599	Earth slide
368	37.36966	5.976422	1239	Earth slide
369	37.37829	5.981044	901	Earth slide
370	37.37485	5.980514	1039	Earth slide
371	37.36828	5.980427	987	Earth slide
372	37.37899	5.984541	935	Earth slide
373	37.3981	5.983205	1094	Earth slide
374	37.39783	5.975721	1115	Earth slide
375	37.40114	5.97745	1892	Earth slide
376	37.39052	5.972068	1328	Earth slide
377	37.37959	5.973378	1054	Earth slide
378	37.37601	5.976444	1414	Earth slide
379	37.38048	5.972467	1233	Earth slide
380	37.39044	5.969592	1632	Earth slide
381	37.41116	5.954449	1055	Debris slide
381	37.40457	5.952905	1818	Debris slide
382	37.40317	5.954878	1802	Debris slide
383	37.39731	5.959209	2285	Earth slide
384	37.37596	5.979747	1642	Earh slide
385	37.36424	5.966768	1344	Earth slide
386	37.36282	5.968389	2685	Earth slide
387	37.30451	5.800691	3268	Earth slide
388	37.26543	5.842453	1284	Rock slide
389	37.2667	5.848407	978	Rock slide
390	37.4241	5.989828	1825	Debris slide
391	37.43152	5.988519	1056	Earth slide
392	37.42553	5.997713	1711	Earth slide
393	37.4182	5.994924	1104	Earth slide
394	37.4155	6.021301	1228	Earth slide
395	37.43776	6.033704	1355	Earth slide
396	37.43651	6.032585	1192	Earth slide
397	37.4353	6.028168	1180	Earth slide
398	37.44154	6.049967	1809	Earth slide
399	37.41812	5.978845	1065	Earth slide
400	37.41974	5.978961	1582	Earth slide
401	37.42135	5.977899	836	Earth slide
402	37.4247	6.014199	1726	Earth slide
403	37.42073	6.014824	1259	Earth slide
404	37.37718	5.958196	946	Earth slide
405	37.41457	6.01176	1417	Earth slide
406	37.41738	6.018278	1133	Earth slide
407	37.41829	6.01619	1159	Earth slide
408	37.36694	5.913359	1797	Earth slide
409	37.32923	5.786068	2741	Earth slide

410	37.32333	5.791829	2028	Earth slide
411	37.33122	5.79218	1502	Earth slide
412	37.32312	5.79732	2220	Earth slide
413	37.35243	5.796056	1563	Earth slide
414	37.36823	5.878526	1875	Earth slide
415	37.36952	5.911336	1104	Earth slide
416	37.37645	5.927729	957	Earth slide
417	37.40944	5.939635	2218	Earth slide
418	37.43245	6.007629	1201	Earth slide
419	37.36047	5.972882	2680	Rock slide
420	37.4341	5.980149	3532	Earth slide
421	37.4108	5.902651	5800	Earth slide
422	37.41716	5.901874	3645	Debris slide
423	37.38473	5.956215	3714	Debris slide
424	37.34636	5.938225	1905	Earth slide
425	37.35987	5.934597	4750	Earth slide
426	37.44328	6.034607	1837	Debris slide
427	37.37915	6.037797	1765	Earth slide
428	37.4077	5.936317	1144	Debris slide
429	37.40406	5.936823	1467	Debris slide
430	37.32992	5.928169	3481	Earth slide
431	37.36286	5.901909	2016	Earth slide
432	37.38397	6.055539	1313	Earth slide
433	37.34071	5.919987	2034	Rock slide
434	37.28258	5.877666	2327	Earth slide
435	37.33475	5.917881	2184	Earth slide
436	37.34079	5.895722	2296	Earth slide
437	37.34758	5.893997	3705	Earth slide
438	37.36042	5.880293	1762	Earth slide
439	37.35956	5.875924	1945	Earth slide
440	37.38574	6.03331	1907	Earth slide
441	37.38187	6.049495	1971	Earth slide
442	37.38849	5.885583	1654	Earth slide
443	37.39129	5.884802	1430	Earth slide
444	37.40971	5.883176	2305	Earth slide
445	37.40767	5.88505	1860	Earth slide
446	37.40554	5.881195	2370	Earth slide
447	37.38719	5.877261	1251	Earth slide
448	37.41649	5.885737	1648	Earth slide
449	37.39626	5.908098	2067	Debris slide
470	37.42832	5.937261	1166	Earth slide
471	37.3833	6.018255	1207	Earth slide
472	37.47582	6.023108	953	Transitional slide
473	37.45433	5.97572	2848	Earth slide
474	37.4498	5.957648	2052	Earth slide
475	37.4516	5.916203	2297	Earth slide

476	37.37206	5.823411	3099	Earth slide
477	37.37566	5.827761	2793	Transitional slide
478	37.35503	5.898653	4326	Earth slide
479	37.34386	5.885126	1587	Earth slide
480	37.30036	5.878428	1966	Earth slide
481	37.41454	5.952568	2052	Earth slide
482	37.31006	5.893362	2553	Earth slide
483	37.43584	6.009849	930	Earth slide
484	37.42999	6.009536	1624	Earth slide
485	37.43889	5.974163	3609	Earth slide
486	37.29027	5.801194	2446	Earth slide
487	37.29143	5.811419	1811	Earth slide
488	37.28594	5.810956	2040	Earth slide
489	37.27933	5.828329	2632	Earth slide
470	37.2778	5.83897	1419	Earth slide
471	37.28179	5.850349	2053	Rock slide
472	37.32261	5.887299	1356	Rock slide
473	37.31586	5.89342	2646	Earth slide
474	37.37329	6.05954	1314	Earth slide
475	37.37758	5.927914	1201	Earth slide
476	37.37222	5.917728	2527	Earth slide
477	37.3744	5.921695	3022	Earth slide
480	37.38412	5.930863	3043	Debris slide
481	37.37744	5.941828	2342	Earth slide
482	37.37681	5.951574	3524	Earth slide
483	37.38417	5.958679	2984	Earth slide
484	37.38503	5.963872	1437	Earth slide
485	37.38507	5.971651	2403	Earth slide
486	37.39769	6.00296	1781	Earth slide
487	37.39322	5.889974	1656	Earth slide
488	37.39323	5.888233	2182	Earth slide
489	37.37662	5.923779	2950	Earth slide
490	37.36363	5.985114	1788	Earth slide
491	37.29612	5.795723	2528	Earth slide
492	37.30244	5.806771	2169	Earth slide
493	37.29079	5.828621	1586	Earth slide
494	37.28718	5.833025	2277	Rock slide
495	37.37993	6.013046	2610	Earth slide
496	37.37751	5.996176	3445	Earth slide
497	37.33152	5.884298	1463	Rock slide
498	37.32685	5.886098	1835	Rock slide
499	37.41181	6.026289	1799	Earth slide
500	37.41708	6.027872	1755	Earth slide
501	37.44417	6.012243	3490	Earth slide
502	37.34898	5.951523	3082	Earth slide
503	37.36262	5.879258	1304	Rock slide

504	37.38704	5.988697	2301	Earth slide
505	37.28436	5.85313	2528	Rock slide
506	37.28367	5.838317	2654	Rock slide
507	37.269	5.845895	1127	Rock slide
508	37.26911	5.840971	1537	Rock slide
509	37.46595	5.950165	3645	Earth slide
510	37.46921	5.958001	2116	Earth slide
511	37.42345	5.936747	2453	Earth slide
512	37.41821	5.934176	1469	Earth slide
513	37.42839	5.956883	2227	Earth slide
514	37.42169	5.961816	2110	Earth slide
515	37.35653	5.967949	1964	Rock slide
516	37.35534	5.811756	2619	Earth slide
517	37.35995	5.824222	1408	Debris slide
518	37.37401	6.041192	1816	Debris slide
519	37.35576	5.910336	1078	Earth slide
520	37.41744	5.928623	2129	Earth slide
521	37.41377	5.9143	1175	Earth slide
522	37.33897	5.904286	1751	Earth slide
523	37.3398	5.907598	918	Earth slide
524	37.44717	6.032703	1542	Earth slide
525	37.43832	6.013132	1072	Earth slide
526	37.43814	6.025581	999	Earth slide
527	37.44004	6.028254	1013	Transitional slide
528	37.38709	6.031232	1572	Earth slide
529	37.46432	6.007092	1720	Earth slide
530	37.44602	5.994964	1520	Earth slide
531	37.29238	5.794537	1894	Earth slide
532	37.30685	5.887897	2066	Rock slide
533	37.30637	5.886472	1788	Rock slide
534	37.44055	6.018193	1731	Earth slide

Appendix 4: Recommendation chart.