# MAPPING OF LANDSLIDE SUSCEPTIBLE ZONES IN MALAPPURAM DISTRICT, KERALA USING MACHINE LEARNING ALGORITHM AND GEOINFORMATICS.

काशी हिन्दू विश्वविद्यालय BANARAS HINDU UNIVERSITY

**THESIS**

**SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF**

**Master of Technology**

**In**

**Agricultural Engineering**

**( SOIL AND WATER CONSERVATION ENGINEERING)**


**Submitted by**

**SHUCHISMITA GIRI**

**Supervisor**

*Dr. V.K. Tripathi*

**Co-Supervisor**

*Dr. Thenmozhi M*

**DEPARTMENT OF FARM ENGINEERING**

**INSTITUTE OF AGRICULTURAL SCIENCES**

**BANARAS HINDU UNIVERSITY**

**VARANASI-221005**

**INDIA**



**I.D. No. 20412AEN004** **2022** **Enrolment No. 429389**

*Dedicated to*

*Baba, Mamma & Tiki*

काशी हिन्दू विश्वविद्यालय

BANARAS HINDU UNIVERSITY

## CERTIFICATE

**Ref. No.: ……** **Dated: ……**

To,
The Registrar (Academic)
Banaras Hindu University,
Varanasi–221005,
U.P., India

**Through:** The Head, Department of Farm Engineering, I. Ag. Sc., BHU, Varanasi

Dear Sir,

We have great pleasure in forwarding the thesis entitled **"Mapping of landslide susceptible zones in Malappuram district, Kerala using machine learning algorithm and geoinformatics"** submitted by **Miss. Shuchismita Giri, I.D. No.: 20412AEN004, Enrolment No.: 429389** in partial fulfillment of the requirements for the award of the degree of **Master of Technology in Agricultural Engineering (Soil and Water Conservation Engineering)**, Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi.

This is to certify that the work has been carried out solely by **Miss. Shuchismita Giri** under my supervision and guidance and her findings and data presented herein are genuine and original to the best of my knowledge and belief and no part of the work has been submitted for any other degree or distinction.

Thanking you,

Yours faithfully,

**FORWARDED BY**

**Dr. V. K. Tripathi**
**Assistant Professor**
**(Supervisor)**

**(HEAD)**

**Dr. Thenmozhi M**
**Scientist-B**
**(Co-supervisor)**

# Mapping Of Landslide Susceptible Zones in Malappuram District, Kerala Using Machine Learning Algorithm and Geoinformatics


by

**SHUCHISMITA GIRI**


**THESIS SUBMITTED IN THE PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF**

**MASTER OF TECHNOLOGY**

**IN**

**AGRICULTURAL ENGINEERING**

**(SOIL AND WATER CONSERVATION ENGINEERING)**


**DEPARTMENT OF FARM ENGINEERING**
**INSTITUTE OF AGRICULTURAL SCIENCES**
**BANARAS HINDU UNIVERSITY**
**VARANASI-221 005**



**I.D. No.: 20412AEN004** **2022** **Enrolment No. - 429389**


## THESIS APPROVED BY ADVISORY COMMITTEE

**Chairman** : **Dr. V. K. Tripathi**
**Assistant Professor**
**Department of Farm engineering**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**Co-Supervisor** : **Dr. Thenmozhi M**
**Scientist-B**
**KSCSTE-Center for Water Resources Development and Management**
**Kozhikode, Kerala.**

**Member** : **Dr. Amitava Rakshit**
**Associate Professor**
**Department of Soil Science and Agricultural Chemistry**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**Member** : **Dr. Reema Sharma**
**Assistant Professor**
**Department of Farm engineering**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**External Examiner** :

# C E R T I F I C A T E

We, the undersigned members of the advisory committee of **Miss. Shuchismita Giri**, **Id. No.: 20412AEN004**, a candidate for the degree of **Master of Technology in Agricultural Engineering (Soil and Water Conservation Engineering),** agree that the thesis entitled "**Mapping of landslide susceptible zones in Malappuram district, Kerala using machine learning algorithm and geoinformatics**" may be submitted in partial fulfillment of the requirements for the degree.

**Chairman** : **Dr. V. K. Tripathi**
**Assistant Professor**
**Department of Farm engineering**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**Co-Supervisor** : **Dr. Thenmozhi M**
**Scientist-B**
**KSCSTE-Center for Water Resources Development and Management**
**Kozhikode, Kerala.**

**Member** : **Dr. Amitava Rakshit**
**Associate Professor**
**Department of Soil Science and Agricultural Chemistry**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**Member** : **Dr. Reema Sharma**
**Assistant Professor**
**Department of Farm engineering**
**Institute of Agricultural Sciences**
**Banaras Hindu University, Varanasi, U.P.**

**External Examiner:**

# ACKNOWLEDGEMENT

*At the outset, I bow my head with in great reverence to the lotus feet of **Bharat Ratna Mahamana Pandit Madan Mohan Malviya Ji**, the founder of Banaras Hindu University for his lifetime sacrifice and efforts in establishing such a great temple of learning for the course of millions of students like me.*

*I consider myself fortunate and greatly privileged to have worked under the supervision and guidance of my beloved guru and chairperson of my advisory committee **Dr. V. K. Tripathi**, Assistant Professor, Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University for his perpetual guidance, wise counsels, easy approachability, constructive criticism, persistent encouragement as well as parental affection during the course of research.*

*I would like to express my deep sense of reverence and sincere gratitude to **Dr. Thenmozhi M**, Scientist-B, KSCSTE-CWRDM, Kozhikode for her invaluable inputs, guidance, corporation, and constant encouragement from beginning to the end of the thesis.*

*I am grateful to the esteemed member of my advisory committee **Dr. Amitava Rakshit,** Associate Professor, Department of Soil Science and Agricultural Chemistry and **Dr. Reema Sharma,** Assistant Professor, Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University for their constant encouragement and moral boost-up during the course of experimentation.*

*I owe my gratefulness to **Dr. A. K. Nema**, Professor and Head, Institute of Agricultural Sciences, Banaras Hindu University, for his constant support and inspiration throughout my study period.*

*I extend my indebtness to **Prof. Dr. V. K. Chandola, Prof. Dr. R. M. Singh, Dr. Shrinivasa D. J., Dr. Abhishek Singh, Dr. Shashi Shekhar** and **Mr. Rajan Kumar** of the Department of Farm Engineering, Institute of Agricultural Sciences, B.H.U., for their discerning comments, valuable suggestions, and helpful attitude towards me during the course of investigation.*

*I am eternally grateful to **Dr. Manoj P. Samuel,** Executive Director, KSCSTE-CWRDM, Kozhikode for allowing me to carry out my M.Tech thesis project under such an esteemed institute. I would like to extend my gratitude to **Er. Ambili G.K**, Head, and **Dr. Naveen K**, Scientist-B, Training and Outreach Programme Group, KSCSTE-CWRDM, Kozhikode for providing their valuable support during the thesis preparation. With immense pleasure, I want to thank **Er. Dinu Maria Jose, Mr. Royal Tata, and Mr. Murugavel**, Project Fellow, KSCSTE-CWRDM.*

*I will remain grateful to all the non-teaching staff members, especially **Mr. O. P. Singh, Mr R. Viswakarma, Mr. Anil, Mr. C. P. Singh,** and **Mr. Uma** of Department Farm Engineering, for their co-operation throughout the course of my study.*

*My heartfelt and special thanks to my senior **Mr. Kanhu Charan Panda** for his kind guidance and co-operation during my research.*

*It is a pleasure for me to offer thanks to my seniors* ***Mr. Saswat Kumar Kar, Mr. Pushpendra, Ms. Liansangpuii, Ms. Samiksha Panda, Ms. Snehil Dubey, Mr. Manish, Mr. Vikas Singh, Mr. Subodh Hanwat, Mr. Chandra Kishor,*** *and* ***Mr. Kuldeepak Paul .***

*I want to thank my M.Tech friends* ***Ms. Madhu, Mr. Achal Singh, Mr. Anoop Dongre, Mr. Shivam kumar Dwivedi, Mr. Rahul Kumar, Mr. Gaurav Maurya, Mr. Sarvadanand Tiwari, Mr. Dheeraj Sonker, Ms. Akansha Maravi, Ms. Sagorika Chakraborty, Ms. Nisthha Sharnagat, Mr. Ram Thakur,*** *and* ***Ms. Anuradha Sahu.***

*A word of appreciation is extended to my juniors* ***Mr. Kumar Saurabh, Mr. Santosh Kumar Vishwakarma, Mr. Rohit Kumar, Mr. Shringar Mishra, Ms. Akanksha Sharma, Mr. Ankit Kumar Yadav, Mr. Deepak kumar, Mr. Yugojyoti Barik, Ms. Aishwarya Mansingh, Mr. Jitendra Kumar Verma, Mr. Darshan M, Mr. Khujde Raju Gangaram, Mr. Jageshvar Prasad, Mr. Anuj Kumar, Mr. Om Prakash Mishra,*** *and* ***Mr. Aditya Kumar .***

*The graces of the God are always blessed to me and give me patience and power to overcome the difficulties which came in my accomplishment of this endeavour. I cannot dare to say thanks but only pray to bless me always.*

*Last but not the least, I record my sincere thanks to my beloved sister and respected parents for their encouragement and blessings during my ups and down.*

*Date:*

*Place: Varanasi*

*Shuchismita Giri*

*Department of Farm Engineering,*
*Institute of Agricultural Sciences,*
*Banaras Hindu University, Varanasi-221005*

# CONTENT

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF ABBREVIATIONS AND SYMBOLS

| | | |
|---|---|---|
| ° | : | Degree |
| % | : | Per cent |
| °C | : | Degree Celsius |
| $\Sigma$ | : | Summation |
| & | | And |
| m | : | Meter |
| mm | : | Milli meter |
| sq. km | : | Square Kilo meter |
| et al. | : | et alii (and others) |
| No. | : | Number |
| amsl | : | Above Mean Sea Level |
| SVM | : | Support Vector Machine |
| KNN | : | K-Nearest Neighbour |
| RF | : | Random Forest |
| LHEF | : | Landslide hazard evaluation factor |
| TMC | : | Terrain Mapping Complexes |
| TMU | : | Terrain Mapping Units |
| LULC | : | Land Use and Land Cover |
| TMS | : | Terrain Mapping Subunits |
| GIS | : | Geographical Information System |
| AHP | : | Analytical Hierarchy Process |
| AUC | : | Area Under Curve |
| ROC | : | Receiver Operating characteristic |
| RES | : | Rock Engineering System |
| NDVI | : | normalized difference vegetation Index |
| STI | : | Sediment Transport Index |
| WoE | : | Weight of Evidence |
| mLRF | : | modified Landslide Relative Frequency |
| BPNN | : | backpropagation neural network |
| SSRE | : | Sum Squared Relative Error |

| | | |
|---|---|---|
| TWI | : | Topographic Wetness Index |
| TRI | : | Topographic Ruggedness Index |
| LR | : | Logistic Regression |
| GSI | : | Geological Survey of India |
| UTM | : | Universal Transverse Mercator |
| WGS | : | World Geodetic System |
| DTs | : | Decision Trees |
| TPR | : | True Positive Rate |
| FPR | : | False Positive Rate |
| TNR | : | True Negative Rate |
| MCC | : | Matthews correlation coefficient |
| PPV | : | Positive Prediction Value |
| AUROC | : | Area Under ROC curve |
| TP | : | True Positive |
| TN | : | True Negative |
| FP | : | False Positive |
| FN | : | False Negative |
| LBE | : | LogitBoost Ensemble |
| DB | : | Deep Boost |
| SGB | : | Stochastic Gradient Boosting |
| BRT | : | Boosted Regression Tree |
| LDA | : | Linear Discriminant Analysis |
| GNB | : | Gaussian Naïve Bayes |
| LB | : | Logit Boost |
| RoF | : | Rotation Forest |
| FLDA | : | Fisher's Linear Discriminate Analysis |
| NB | : | Naïve Bayes |
| BN | : | Bayesian network |
| GARP | : | Genetic Algorithm for Rule-set Prediction |
| DEM | : | Digital Elevation Model |
| ASTER | : | Advanced spaceborne Thermal Emission and Reflection |
| ML | : | Machine Learning |

# 1. INTRODUCTION

Landslide is one of the most dangerous and unpredictable natural hazards, that has affected the social and economic development of hill ranges. Landslide is defined as the downward movement of debris, rock, organic matter, or earth along the slope under the influence of gravity (Cruden, 1991). Landslide causes a devastating situation in an area, that leads to loss of lives, property, agriculture, settlement, infrastructures, and loss of connectivity. Natural hazards like landslides also affect the emotional well-being of sufferers.

According to research conducted by the National Institute of Disaster Management in 2011, India endures a financial loss of Rs 150-200 crore per annum due to incidents of landslides (NIDM, 2011). As per a study conducted by academics at Sheffield University in the United Kingdom, India was one of the countries, which is the most hit by human-caused deadly landslides between 2004 and 2016. The study looked at 5,031 deadly landslides throughout the world and found that India had roughly 10,900 deaths from 829 landslides, accounting for 18% of total global mortality.

As the Government of India spends a gigantic portion of its GDP on the relief, rescue, and mitigation projects for disasters, it is a great economic drawback for the nation. Recourses that could have been utilized for the development of an area, are used for stabilization and restoration of the landslide-hit regions. The catastrophic character of such multi-hazards was provoked by irresponsible slope changes, mismanagement of water drainage, unplanned residential constructions across river channel courses and inside active floodplains, and a lack of landslide early warning systems (Wadhawan et al., 2020). Therefore, disaster preparedness is much needed by identifying the susceptible area and developing disaster mitigation techniques according to the condition of the area.

Kerala is a state that is extremely sensitive to natural calamities including growing climatic variations, as it is located along the seaside and the steep slope of the Western Ghats hills. Kerala is also one of the most densely inhabited Indian states

(860 people per square kilometer), making it even more sensitive to disaster-related damages and losses. The western portion of the Western Ghats mainly comprises the eastern portion of Kerala and is one of the country's most susceptible places to landslides. 1500 square kilometers of the Western Ghats are susceptible, and landslides are reported every year when the monsoon season sets in. Occurrences of landslides in mountain regions during the monsoon season, causing road crashes, blockage of river beds due to silting, and significant damage to properties and lives (Kuriakose, 2010). Rainfall-induced landslides are one of the most prevalent forms of risk in tropical climatic settings (Oommen et al., 2019). Thus, an awareness of rainfall patterns will assist in providing a comprehensive picture of an impending landslide and can design methods to mitigate and reduce its impacts.

## 1.1. Landslide

Landslide can be best described as a complex correlative action of geological, geomorphological, and meteorological factors. Due to Slope failure caused by the inability to support its weight, soil mass including debris, rock, and organic material starts to move downward under gravitational effect. Although landslide is a natural phenomenon, when it occurs in populated or well-developed areas, it causes immense physical, social, and economical damages. Various parts of a landslide are graphically represented in Figure 1.1.

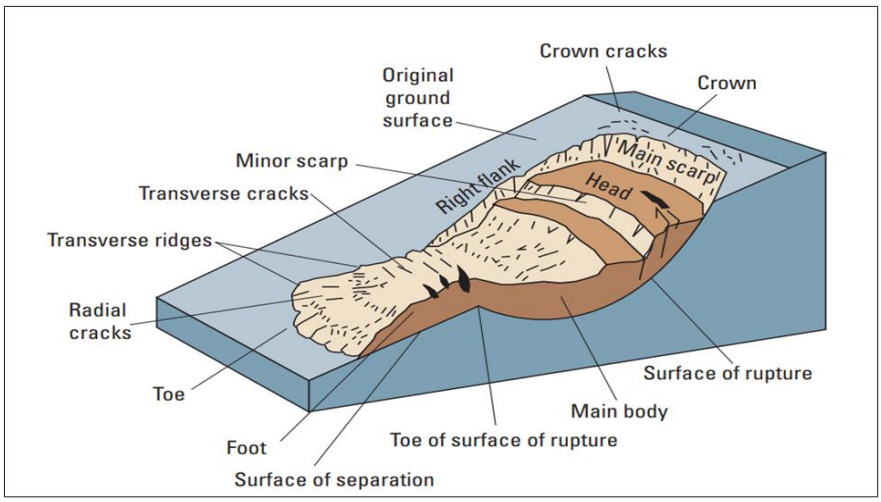


Figure 1.1: Illustration of commonly used labels for the parts of a landslide (Source: Varnes, 1978)

## 1.2. Types of landslides

D. Cruden (1996) classified landslides focusing on the type of material involved and the type of movement occurring on the surface of rupture. Figure 1.2 picturizes the type of landslides taking place in nature.

### *1.2.1. Fall* :

When soil or rock detached from a steep slope with little to no interaction of shear displacement between moving units and descends to the lower region in the air free-falling, bouncing, rolling, or leaping, the process called fall.

### *1.2.2. Topple* :

A displacement of a soil mass or rock, which starts to have a forward rotation around a point of axis below its center of gravity, is classified as toppling movement. The presence of water or ice in cracks of the displaced mass causes an increase in the weight of upslope material and the driving force of gravity results in the toppling of rock, debris (coarse material), or earth material (fine-grained).

### *1.2.3. Slide* :

It is a downward movement of soil or rock primarily due to intense shear failure along the surface of rupture. Slides can be divided into two groups according to the mechanism of movement -

(a) Rotational Slide

(b) Translational Slide

In the rotational slide, the surface of rupture is concave towards the direction of movement forming a spoon-shaped. The slide movement is rotational about an axis that is parallel to the contour of the slope. The moving mass does not get highly deformed while undergoing rotational sliding and moves as a relatively coherent mass along the rupture surface. While the head of the displaced material moves vertically downward, the upper surface of the displaced material may tilt backward towards the scarp. A slump is formed when there are several numbers of parallel curved planes are

involved in a rotational slide. In translational slide, the soil mass or detached material moves down and outward, along a relatively planar surface. The sliding material gets disintegrated into smaller parts. Rotational slide tends to increase the shear resistance and decrease the driving force to gain stability. Whereas in translational slides, the material tends to progress if the surface is sufficiently inclined and if low shear resistance prevails. failure along pre-existing structural features like faults, joints, and bedding planes results in translational slides.

### *1.2.4. Spread* :

Sometimes slide is resulted from the lateral spreading of soft clay present under firmer material. When pore water pressure increases in homogeneous clay or softer underlying material, it loses its cohesivity through liquefaction or plastic deformation and cannot support the stronger material above it. Block spread, liquefaction spread, and lateral spread is examples of various kinds of spread. This kind of movement generally takes place on a gentle slope or flat terrain.

### *1.2.5. Flow* :

The component velocities of the displaced mass have similarities to that of viscous liquid. The shear surfaces are short-lived, closely spaced, and not usually preserved. The landslide type changes from slide to flow, based on water content and velocity of progression. Debris flow is one of the examples of flow-type landslides, in which rotational and transitional slide gains velocity due to increased water quantity and internal mass loses cohesion. Debris flows can be deadly due to the high velocity and uncertainty of happening.

### *1.2.6. Complex* :

When a landslide shows various types of movement at different parts and at different times during the progress of the landslide, it is called a complex landslide. So, most landslides are the combination of more than one type of movement and are classified as complex type landslides.

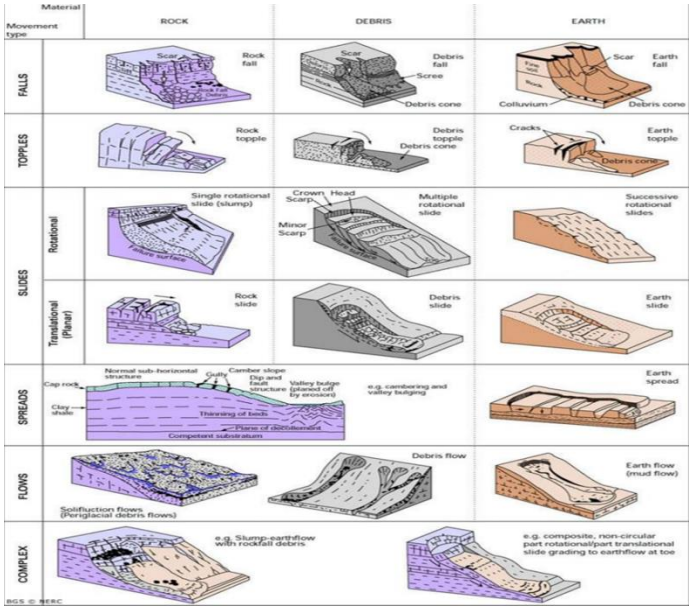


Figure 1.2: Classification of types of landslides according to movement and material involved.

## 1.3. Causes of landslide

Varnes (1978) had broadly grouped the causes of landslides into 3 groups-

(A) Increase in shear stress- Removal of lateral support, imposition of surcharges, transitory stresses, and uplifting are the reasons for increasing shear stress. Erosion by water streams, excavations for quarries or canals, piping, or mining may lead to the reduction of lateral support. Again, the addition of water through precipitation or seepage from an irrigation system, construction, and waste dumps, growth of glaciers and vegetation increase both the length and the height of the slope, thus increasing the chances of landslide. Change in the local stress field due to earthquakes, explosions, or passage of heavy vehicles are instances of increased shear stress. Uplifting steepens the slope in the area as drainage responds by increased incision leads to landslide.

(B) Low strength- Materials having low strength (clay and organic matter) or may become weak after saturation with water, clay produced from chemical weathering of certain rocks, and weak arrangement of particles are the causative agents of low

strength in an area. Faults, bedding surfaces, foliations, cleavages, joints, fissures, shears, and sheared zones contribute to discontinuity and weakening of soil mass.

(C) Reduction of material strength- Various physiochemical reactions, weathering, thermal expansion and contraction, desiccation cracking due to dry weather, dissolution of natural rock cement that holds particles together, reduction of intergranular pressure, friction, and capillary tension due to saturation also reduces the strength of the material.

Table1.1: List of triggering factors of landslide

| Sl. No. | Factors | Description |
|---|---|---|
| 1. | Geological | Weak and weathered material, Jointed or fissured material, Contrast in stiffness, |
| 2. | Morphological | Tectonic or volcanic uplift, Fluvial or wave erosion of slope toe, Vegetational removal, Deposition loading slope, Glacial rebound |
| 3. | Physical | High-intensity rainfall, Earthquake, Rapid snowmelt, Rapid flood, Thawing |
| 4. | Anthropogenic | Excavation of slope, Mining, Artificial vibration, Deforestation |

## 1.4. Landslides in India

About 15% of the landmass of India is a landslide-prone area comprised of the Himalayas, the North-eastern hill ranges, the Nilgiris, the Western Ghats, the Eastern Ghats, and the Vindyas. North-east Himalayan region of Darjeeling and Sikkim, North-West Himalayan states like Uttarakhand, Himachal Pradesh, and Jammu & Kashmir, the Western Ghats and Konkan hills, and the Eastern Ghats of Aruku area in Andhra Pradesh comprise about 0.18 million sq. km, 0.14 million sq. km, 0.09 million sq. km, and 0.01 million sq. km respectively (India Today, 2015). Therefore, a total of

0.42 million sq. km or 12.6% of land area excluding snow-covered area are susceptible to landslide.

The presence of tectonically unstable younger geological formations in the region of the Himalayan Mountain belt makes it susceptible to severe seismic activity. The Indian plate moves in the north direction under the Eurasian plate at the speed of 4.5cm per year, thus creating shearing stress around the region. Being characterized by a lateritic cap, vigorous mining, deforestation, and erosion due to high-intensity rainfall, the Western Ghats along with the Konkan coast, Nilgiris pose an uncertain landslide hazard. India is the country, where the number of deaths due to landslide stands at number 11,000, which is the highest globally, within 12 years (2006-17). In June 2013, 5000 deaths were reported in the landslide at Kedarnath, which was the most tragic landslide due to flash floods. Some fatal landslide incidents in India are listed below-

1. On 18$^{th}$ September 1948 over 500 people died in a landslide that happened in Guwahati, Assam caused by heavy rain.
2. A disastrous Landslide, triggered by a flood on 4$^{th}$ October (1968), took the lives of thousands of people in Darjeeling, West Bengal. The 60 km highway was cut into 91 parts.
3. Entire village was washed out in Malpa, Uttarakhand and 380 people died in consecutive landslides that occurred between 11$^{th}$ August to 17$^{th}$ August 1998.
4. Heavy rainfall followed by land erosion resulted in a landslide in July 2000, in Mumbai. Around 67 people died in that unfortunate incident.
5. In the Amboori region of Kerala, 40 people died due to a landslide on 9$^{th}$November 2001.
6. Malin landslide of Maharashtra occurred on 30$^{th}$ July 2014, took the innocent lives of 151 people and 100 people went missing.
7. Death of 400 people, as well as the destruction of 106 households, took place in a disastrous landslide in the Kuwari region of Uttarakhand on 10$^{th}$ March 2018.

## 1.5. Landslides in Kerala

Being close to the seashore of the Arabian sea along with steep slope formations of the Western Ghats, Kerala is physically vulnerable to landslide hazards. Again, due to a dense population of 860 persons per square kilometer, the state is highly vulnerable to loss and damage socio-economically.

The average annual rainfall of Kerala is about 3000 millimeters and sometimes it could exceed 5500 millimeters in a mountainous region. 1500 sq. km. of the Western Ghats of the eastern part of Kerala is highly vulnerable with the onset of monsoon season and experiences several landslides, which leads to the collapsed road, river basin silting, heavy damage to properties of public and agricultural losses.

The most common type of landslide in Kerala is debris flow, which is called 'UrulPottal' in the local Malayalam language. It is a rapidly moving slurry of water, rock, vegetation, fine soil particle, and debris. This is the most destructive type of slope failure among all the types of landslides.

Death of 295 people had occurred due to 85 disastrous landslides, within 55 years (1961-2016), which was reported by the Kerala State Disaster Management Authority. The document "disaster management plan" published by the Kerala government in 2016, describes that 14.4% of the state, or 5607 sq. km. of the area is a landslide vulnerable zone.10 taluks, 25 taluks, 14 taluks, and 28 taluks of Kerala are high-risk zone, moderate-risk zone, low-risk zone, and stable zone respectively.

Due to the migration of population from plains to high lands, and subsequent clearing of forests for infrastructure, the pressure increases on vulnerable weak land. An increasing number of landslides are being reported during the monsoon period (June to September) when the driving force of landslides is the highest. On account of climate change, high-intensity rainfall and unmonitored anthropogenic activities, the peak of landslide hazards will go up even more in Kerala.

## 1.6. Landslides in Malappuram

Malappuram district is a geographically diverse region with the presence of dense forest land, coastal land, Western Ghats mountainous range, and well-developed human settlements. But this highly eco-rich state can't save itself from natural disasters like floods and landslides during monsoon. On account of High-intensity rainfall during the monsoon season, increasing cases of calamitous events of landslides and floods have been reported every year.

In 2019, the Western Ghats grama panchayat of Pothukallu near Nilambur was the terrible epicenter of landslides, damage, and human deaths. Over 50 residents of 18 houses were buried alive for more than a week when a landslide slammed Kavalappara into the panchayat, leveling dozens of hectares of valuable agriculture and destroying homes and public facilities.

The number of deaths by drowning in floodwaters was, however, few. Mudslides turned out to be the real villain. And, it was not raining alone that triggered such a mammoth mudslide as the one that took away 18 families of Kavalappara. The overall number of human lives lost in the district during the 2019 monsoon storm is nearly 70 and continuing. The majority of victims of landslides and debris flows in the district were socio-economically disadvantaged. They lived on the vulnerable outskirts because they had few other options.

Because of the widespread destruction, the loss of property and harm to means of livelihood remains incalculable. Paathar, a bustling bazaar in Pothukallu, was destroyed, leaving a massive scar behind. Across the district, hundreds of houses got mauled. The business was hit like never before as unexpected inundation destroyed merchandise worth crores of rupees in towns like Nilambur, Kondotty, and Malappuram.

Besides the district's fragile rural economy, already deeply hurt by demonetization and reverse migration from the Gulf, is in ruins. Nature has responded furiously to human vandalism. The district is now paying the price for mindless development carried out in recent decades by destroying hills, forests, water bodies, and wetlands.

For instance, large portions of Oorakamala and Cheruppadimala, two of the highest hills in the district, were ruthlessly carved up for granite and laterite. Hills have been grotesquely wounded by quarrying; forests have been shaved off by commercial agriculture; paddy fields and ponds that in the past served as a buffer for floodwaters have turned into modern mansions; rivers have shrunken and streams vanished.

## 1.7. Objective of the study

Landslide study of Malappuram and western ghats region had been carried out before by many scientists. A recent Investigation and study, which was carried out by Savith et al.(2021), Ajmal & Saud S (2022), and Jones et al.(2021)show that adaptation of geomorphic analysis is most popular and although they have stated rainfall as a triggering agent, inclusion of annual rainfall amount during analysis has not been noticed yet for present study area. The major benefits of probabilistic and machine learning methods are their objective statistical foundation, reproducibility, capability to quantitatively examine the contribution of elements to landslide development, and possibility for continual update (Youssef & Pourghasemi, 2021). Despite this significant advantage, very few studies had carried out using machine learning techniques for the Malappuram region. Based on the above limitations of previous studies, we have framed the following objectives:

1. To develop a landslide susceptibility map by the qualitative technique of the weighted overlay method (WoM)

2. To delineate landslide susceptibility map by the quantitative technique of machine learning models such as Artificial Neural Network (ANN), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), and Random Forest (RF).

3. To compare the accuracy of developed maps, which can be useful for future identification and delineation of landslide-vulnerable areas by using the model of higher accuracy.

# 2. REVIEW OF LITERATURE

Several methods have been developed by the researchers in the attempt to create a landslide zonation map. Each method has its advantages, limitations, and uncertainty owing to the different factors considered and the evaluation technique adopted (Fall et al., 2006; Kanungo et al., 2006; Casagli et al., 2004).

So, for reviewing of literature, we can classify all the methods and techniques mainly into two groups, which are (a) qualitative and (b) quantitative. The qualitative analysis includes a direct approach of geomorphological analysis and landslide inventory methods and an indirect approach of heuristic (expert) evaluation techniques. Quantitative analyses are statistical methods, deterministic methods, probabilistic approaches, and methods of artificial intelligence.

## 2.1. Qualitative analysis of landslide

**Anbalagan (1992)** had considered geology, the morphology of slope, relative relief, LULC, and groundwater scenario as contributory factors of past landslides and had carried out an extensive study about their effect on landslides in the Kathgodam-Nainital area. In addition to this, more detailed information was gathered during the field survey. Each factor was given a maximum value of LHEF (Landslide hazard evaluation factor) and the subcategories of that factor were assigned values within the range of LHEF. The final map was delineated by summing up the ratings of all the factors for each facet.

**Ibsen & Brunsden (1996)** carried out a landslide study along the British south coast with the purpose to assess the frequency of landslide incidents and forecasting any future mishaps regarding slope failures. Through their assessment, they found that temporal occurrence study is extremely dependable on the analysis of past landslide events. Therefore, good quality and sufficient data should be provided for the mapping of landslide inventory maps. Considering the research-based analysis, acquiring, and organizing the dataset is a strenuous task. After analyzing the inventory landslide data as well as predicted rainfall events, which were forecasted by the Hadley Centre climate model, researchers had concluded that landslide events will

rise with the increasing changeability and the total number of rainfall events for the region.

**Panikkar & Subramanyan (1997)** conducted an analytical study for landslide disasters in a study area, which comprises 445 sq. km of the Dehradun and Mussoorie region. Researchers had considered the geomorphic parameters for the delineation of landslide zones of the study area. Three Terrain mapping complexes (TMC) were recognized such as structural hills, residual hills, and plains. 14 terrain mapping units (TMU) were identified by overlaying the TMC map and lithology map. 45 homogenous terrain mapping subunits (TMS) were recorded from the combination of TMU map, drainage density map, and relief map. Landslide hazard zonation map was formed after the assignment of four different scores (very high, high, low, and very low) to TMS, as per the experience and knowledge collected from the field survey. After analyzing the hazard zonation map, it was found that 5.1%, 14.1%, 18.4%, and 62.4% of the study area were very high, high, low, and very low respectively. Moreover, land use and land cover changes were analyzed for a period of 60 years (1930-1990) by the researchers to demarcate the effect of anthropogenic activities on the frequency of landslides. Due to the gradual invasion of human settlement into forest regions, forest cover had reduced from 45% (1930) to 18.74% (1990). Around 60% of the landslide had taken place in deforested land and only 3% of the landslide were noticed in the unchanged forest area. This data has confirmed that deforestation plays a major role in landslide incidents. When the LULC map and landslide distribution map were overlaid, it was found that 9%, 65%, 4%, and 10% of the total landslides were taken placed in the forested area, sparsely-vegetated area, cultivated land, and abandoned quarries respectively.

**Turrini & Visintainer (1998)** had developed a numerical-cartographical method that assigns a rate to each parameter, which was considered as landslide causative agents of areas containing dolomite in Italy. For this purpose, thematic maps were delineated for (a) erodibility of rocks and quaternary deposits, (b) permeability of soil, (c) geometric ratio of discontinuities and slope, and thickness of quaternary deposits (d) angle of the slope and (e) land use. Overlaying these thematic maps of different rating

values had resulted in a final map of landslide hazards, which could be used to address the slope failure problem of the referred study area.

**Dai & Lee (2002)** combined GIS a database with digital maps and historical airborne pictures to establish the correlation between various physical factors of the study area and landslide occurrences on Lantau Island in Hong Kong. After analyzing the statistical significance of factors, slope gradient, lithology, elevation, slope aspect, and land use were considered for extraction of the landslide vulnerability map using GIS, whereas slope morphology and distance from drainage lines were dismissed due to statistical insignificance. Researchers had also studied the physical features of landslides.

**Ray et al. (2007)** had applied a fuzzy-based method for the demarcation of landslide susceptibility map in the tectonically active region of Garhwal Himalaya. As a landslide causative agent, he had considered lithology, weathering process, geomorphology, lineaments, drainage, land use land cover, man-made causes, soil type, soil depth, slope gradient along with slope aspect. All the raster layers of stated factors were combined by a fuzzy gamma operator in the GIS platform and the final reclassified map was extracted of five different classes. For validation purposes, a landslide inventory map was also prepared based on a comparison between the satellite image before landslide and after the landslide. This research gave the result that, around 37% of the total land of the study area is highly susceptible to landslides and around 72% of past landslides fall into the category of the high and very high susceptible zone. A broad look into the role of precipitation patterns in the respective study area showed that high rainfall intensity, prolonged rainfall duration, and higher seasonal cumulative rainfall may become the reason for landslide initiation.

**Ruff & Czurda (2008)** applied the heuristic method for their study area, located within the Northern Calcareous Alps to analyze the data layers of geotechnical class, condition of bedding, layouts of tectonics, slope angle, the orientation of slope, vegetation cover and erosion. Before carrying out this analysis, the required data layers of causative factors were marked by a comparison of grided geological maps with past landslide activities. Using bivariate statistics, the vulnerability of each layer

was determined and accordingly weighted with different indices. The final map of slope failure was outlined by combining these pre-weighted data layers.

**Pourghasemi et al. (2012)** had derived the landslide susceptibility map of north Iran by incorporating fuzzy logic and analytical hierarchy process (AHP) models. For validation purposes, a 70:30 ratio was taken for preparing the training and testing dataset out of 78 past landslides. Slope degree, aspect, plan curvature, altitude, lithology, land use, distance from rivers, distance from roads, distance from faults, stream power index, slope length, and topographic wetness index were considered as causative factors to derive the landslide susceptibility map. After being verified by using the area under curve (AUC) approach, it was concluded that the fuzzy logic model (89.7%) performs better than the AHP model (81.1%) for the study area and data layers considered. ROC result shows a better prediction accuracy of 8.60% (89.70–81.10%) for the fuzzy logic model than the AHP model.

**Raghuvanshi et al. (2014)** had taken various intrinsic parameters and external triggering agents into consideration to formulate a new slope susceptibility evaluation parameter (SSEP) for their study area in Wurgessa Kebelle, northern Ethiopia. This newly developed SSEP rating scheme was completely based on the researcher's experience and ability to decide the comparative significance of intrinsic parameters and external triggering agents in the landslide occurrences. The geometry of slope, slope forming material, geological discontinuities, and LULC of the study area was regarded as intrinsic parameters, whereas seismicity, precipitation characters, and anthropogenic activities were considered as triggering agents. The desired landslide susceptibility map was delineated by adding up the values assigned to each factor by the SSEP method. After validating with inventory landslide map, it was concluded that 8.33%, 83.33%, and 8.34% of the area was moderately hazardous, highly hazardous, and very highly hazardous respectively.

**Girma et al. (2015)** had incorporated GIS technology with a statistical method to assign proper value to various landslide causative parameters such as type of soil, lithology, slope, aspect, elevation, curvature, LULC, and groundwater state. Using geoinformatics and GIS, all the causative layers were prepared and rates were

assigned according to their respective impact on slope failure. The extracted final landslide hazard map was validated by inventory landslides, which had been prepared by locating 45 past landslides around the Ada Berga district of central Ethiopia. 24%, 32%, 17%, 25%, and 2% of the study area were labelled as no-risk zone, low-risk zone, moderate risk zone, high-risk zone, and very high-risk zone respectively.

**Meten et al. (2015)** had picked up the Selelkula area of the Jema River Gorge in Central Ethiopian Highland as a study area for mapping landslide susceptibility. The approach adopted by the researchers was the integration of geoinformatics with the fuzzy logic method and rock engineering system (RES). Using GIS techniques, various raster layers were created such as slope, aspect, profile curvature, plan curvature, lithology, land use, distance from the river, and distance from lineament. For fuzzy logic analysis, the frequency ratio of all subclasses of each class was determined from several pixels of class and no of landslide pixels in that respective class. Normalization of frequency ratio within the range of 0 to 1 was carried out and susceptibility maps were demarcated by applying different fuzzy operators (fuzzy gamma, fuzzy AND, fuzzy OR, fuzzy product, and fuzzy sum). The most appropriate map was given by the fuzzy gamma ($\gamma = 0.8$) operator with maximum prediction accuracy of 87.2%. The validation process of the resulted map showed that 77.98% of landslide pixels were included in the high and very high susceptibility zone of the map. The landslide susceptibility map, which was produced through the RES method gave a prediction accuracy of 88.6% and 83.3% of total landslide pixels were included in high and very high susceptibility classes. Although fuzzy logic analysis is data-driven and RES analysis is based on the researcher's expertise, RES gave a better result than fuzzy logic for this study.

**Shahabi et al. (2015)** had carried out a study to plot the landslide susceptibility map for central Zab basin, Iran by adopting frequency ratio, logistic regression, and fuzzy logic methods. Through this analysis, researchers evaluated various causative factors such as slope, aspect, elevation, lithology, normalized difference vegetation index (NDVI), land cover, precipitation, distance to fault, distance to drainage, and road distance. Three susceptibility maps were developed by using three techniques and the accuracy of performance was determined. During fuzzy logic analysis, 94.64% and

85.11% accuracy were developed by gamma operator (λ = 0.975) and fuzzy OR operator respectively. 96%, 95%, and 94% accuracy were shown by the logistic regression model, fuzzy logic model, and frequency ratio model.

**Wang et al. (2016)** had adopted weight of evidence (WoE) models for the region of Gongliu county, China along with the machine learning approach of ANN for the derivation of landslide susceptibility map. 163 landslides of the past were collected through vigorous field investigation and aerial photographs and divided into training (70%) and testing (30%) data. As a causative factor of landslide, researchers had considered the angle of slope, slope aspect, curvature, plan curvature, profile curvature, altitude, distance to rivers, distance to roads, lithology, rainfall, normalized difference vegetation index (NDVI), and sediment transport index (STI). The weight of evidence model calculates the weight for each landslide predictive factor with respect to the presence or absence of the previous landslides within the area. Both the susceptibility maps were compared using the area under the curve (AUC) method. The success rate of 82.51% and 79.82% was found for the ANN model and WoE method respectively. Moreover, the prediction accuracy of 77.31% was shown by the ANN model and 74.59% was found for the WoE model.

**Myronidis et al. (2016)** had applied the analytic hierarchy process (AHP) for Western Cyprus in the SE Mediterranean region to deduce the landslide susceptibility map. 70% of 1540 total past landslides were taken to calculate the modified Landslide Relative Frequency (mLRF) of each causative factor considered. A pairwise comparison was carried out between the mLRF for each landslide instability factor and all possible combinations by using the AHP method and relative weights were produced for all the causative factors. Extraction of the final susceptibility map was done by calculating the Overall Landslide Susceptible Index (OLSI) value of each pixel of the area in GIS software. This research had revealed that human activities affect significantly the frequency of landslides. In addition to this, this study had also revealed that slope angle of 20°-45°, slope aspect of 22.5° to 157.5° (NE-SE) and elevation of 1000–1200 m, gypsiric-regosols soil type, siltstones rock formations, and higher rainfall range of 551 to 600mm were more vulnerable to landslide incidents. The map of landslide susceptibility of the study area gave the outcome that 31.5, 32.6,

24.5 and 11.4% of the study area were very high, high, moderate, and low susceptibility zones respectively. 30% of the recorded past landslide were used for validation by the ROC curve and found an accuracy of 73.9%, which was good.

**Basharat et al. (2016)** presented their landslide study for Balakot, NW Himalayas, Pakistan. They had taken various landslide contributing factors such as slope angle, aspect, curvature, elevation, lithology of rock formation, land cover, faults, road network, and hydrology of the study area. after creating raster layers in GIS software analytic hierarchy process (AHP) was applied for the assignment of numerical weightage to each factor. With the weighted overlay method (WoM), a landslide susceptibility map was extracted, which showed very high, high, moderate, and low susceptible zones. Following the validation process by ROC curve, it was concluded that 76% of past landslides were correctly classified by using WoM. Furthermore, this study also disclosed that 69% of the study area is a very high to high-risk zone on account of a steep slope, lithological pattern, and presence of faults and 31% of the study area is a low to moderately susceptibility zones.

## 2.2. Quantitative analysis of landslide

**Neaupane & Achet (2004)** applied the machine learning method for the study area of Okharpauwa, Nepal to observe the landslide incidents and predict the magnitude of the landslide. Researchers had made use of MATLAB software for the development of a backpropagation neural network (BPNN). They had considered five contributing factors of landslide such as infiltration coefficient, antecedent rainfall, the gradient of the slope, groundwater depth along with soil shear strength. The model gave the best performance with two hidden layers with five and nine nodes sequentially. Sum Squared Relative Error (SSRE) was found to be 7.168%, which was satisfactory. This study concluded that the availability of better-quality information on contributing factors will help in better prediction of slope failure incidents with the BPNN method.

**Ercanoglu (2005)** evaluated the slope failure vulnerability of SE Bartin of Turkey using a machine learning algorithm, which is an artificial neural network. With the data of 317 past landslides, landslide inventory data was created. DEM and ASTER image was utilized to demarcate the parameter maps of six causative factors, such as

Slope angle, slope aspect, elevation, topographical shape, wetness index, and NDVI. When landslide inventory maps and parameter maps were intersected, the most favourable factors for landslide incidents came to light. Finally, the Landslide susceptibility map, which was brought out by the ANN algorithm revealed that 26.5% of the study area was landslide-prone. This ANN model was good as 82.5% and 86.3% accuracy were found for the validation data set and testing data set respectively.

**Goetz et al. (2015)** had compared the various statistical method of analysis as well as machine learning (ML) models for landslide susceptibility modeling in the Lower Austria region. The statistical methods, which were adopted, were Logistic regression (GLM), generalized additive models (GAM), and weights of evidence (WOE). The ML techniques were the support vector machine (SVM), random forest classification (RF), and bootstrap aggregated classification trees (bundling) with penalized 13 discriminant analysis (BPLDA).This research revealed that the Prediction performance of the RF and BPLDA modeling techniques was the best among all the models for the study area. Although the relative importance of each causative factor was dissimilar in every model, highly ranked variables, which affect the landslide the most, were slope angle, surface roughness, and plan curvature.

**Pham & Prakash (2017)** carried out a comparative study on the landslide vulnerability mapping capability of four machine learning techniques. Researchers had adopted four ML techniques such as LogitBoost Ensemble (LBE), Fisher's Linear Discriminate Analysis (FLDA), Logistic Regression (LR), and Support Vector Machines (SVM) for the study area of Tehri Garhwal district, Uttarakhand state. As input factors, slope angle, elevation, curvature, slope length, soil type, slope aspect, land use, depth of valley, lithology, Normalized Difference Vegetation Index (NDVI), Stream Power Index (SPI), Sediment Transport Index (STI), Topographic Ruggedness Index (TRI), Topographic Wetness Index (TWI), distance to lineaments, distance to rivers, distance to roads, and rainfall were taken. Validation of these models revealed that LBE had the highest prediction ability with an AUC of 0.972. SVM, LR, and FLDA models had the AUC values of 0.945,0.873,and 0.870respectively.

**Djuric et al. (2019)** used two machine learning techniques of Support Vector Machines and Random Forest to delineate landslide susceptibility maps for the region of Belgrade, Serbia. 72 raster data layers of 24 less frequently used causative factors were created including different resolutions such as 25, 50, 100, and 200 m. Cross-Scaling (CrSc) method was implemented by taking different resolution training and testing datasets. Again, an adaptation of multi-class learning like stable zone, dormant landslide zone, and active landslide zone, had made this study the leading-edge research. This study revealed that RF creates the results of better accuracy than SVM at a lower resolution and at a higher resolution, and the difference in accuracy decreases. Improvement in performance of both models was noticed by cross-scaling of 100m training data set. Resolution and causative factor selection have less effect on the performance of Random Forest than SVM models.

**Kavzoglu et al. (2019)** had compared the usual statistical method with novel machine learning algorithms such as bagging, random forest (RF), rotation forest (RotFor), and support vector machines (SVM) for landslide vulnerability mapping of Macka region of Trabzon, Turkey. The base of all machine learning techniques is past and present landslides help in predicting the spatial distribution of landslides in any region. The significant role of geology, geomorphology, hydrology and landcover was considered in the initiation of landslide. Polygons of 54 past landslides and 23 non-landslides represented 5144 pixels and 1188 pixels respectively. The overall accuracy was found to be 83.08%, 87.23%, 85.31%, 84.85%, and 78.46% for bagging, RF, RotFor, SVR, and logistic regression (LR) methods respectively. Evaluation of performance was done by using the ROC curve and AUC value.0.931, 0.963, 0.959, 0.955, and 0.868 were the AUC value of bagging, RF, RotFor, SVR and LR models respectively. From the above study, it was concluded that the Random Forest algorithm was the best-fitted model for the prediction of spatial landslides in the study area.

**Chen et al. (2020)** carried out a study to apply advanced machine learning algorithms such as support vector machines (SVM) and artificial neural network (ANN) to deduce the landslide vulnerability map for Achaia Regional Unit, Greece. For the analysis purpose, 335 landslides of past and spatial data of 12 landslide influencing factors such as elevation, slope angle, slope aspect, curvature, plan curvature, profile

curvature, topographic wetness index, stream power index, distance to faults, distance to a river, lithology, and hydrological cover were collected. The result of the models revealed that the highest learning accuracy and prediction accuracy was of the SV M model (AUC =0.977) and ANN (AUC =0.800) respectively.

**Nhu et al. (2020)** conducted a study to plot the landslide susceptibility map of Bijar City, Iran. They had assessed 111 shallow past landslides data and twenty causative factors. This study aimed to compare the efficacy of maps produced by different machine learning algorithms such as Logistic Model Tree (LMT), Logistic Regression (LR), Naïve Bayes Tree (NBT), Artificial Neural Network (ANN), and Support Vector Machine (SVM). Various statistical indices such as root mean square error (RMSE), mean absolute error (MAE), sensitivity, specificity, accuracy, ROC, and AUC were used to determine the suitability and accuracy of models. This assessment culminated with the result that the LMT model was the best-fitted model for the study area. LMT model has the highest AUC both for the training (0.938) and validation (0.932) dataset. MAE and RMSE value of the training dataset were 0.207 and 0.304 respectively and for the validation dataset value of MAE and RMSE was 0.216 and 0.313 respectively.

**Zhao & Zhao (2021)** had applied support vector machine (SVM) and particle swarm optimization coupled with support vector machine (PSO-SVM) models for a comparative study of accuracy in the landslide mapping process. 186 past landslide database was collected from the study area, which was in Luoyang County, China. 70:30 ratio was taken for training and testing of the model. 16 landslide causative elements such as latitude, slope angle, slope, aspect, plan curvature, profile curvature, SPI, TPI, TRI, lithology, distance to faults, TWI, distance to rivers, NDVI, distance to roads, LULC, and precipitation pattern were analyzed. Grid units and slope units were combined with ML techniques to derive the respective maps. PSO-SVM model based on slope units was performed well with the corresponding area under the curve (AUC) values of training and testing datasets were 0.945 and 0.9245.

**Hussain Gardezi et al. (2021)** had tried to evaluate the landslide susceptibility prediction potentiality of six different methods, which include the analytical hierarchy

process (AHP), the weight of evidence (WoE), index of entropy (IoE), multi-layer perceptron (MLP), support vector machine (SVM) and binary logistic regression (BLR). In an effort to train the models, 224 past landslides data (70:30) of Muzaffarabad-Neelum trunk road, Kashmir Himalaya along with a spatial database of 13 landslide influencing elements were collected. Collinearity between independent variables was checked by a variance inflation factor (VIF) and Pearson correlation statistical tests and no collinearity was found. As a means to assess the correlation between the triggering factors, Gain Ratio (GR) and Relief-F methods were performed. Both, the method, lithology, LULC, and slope angle were found to have a major impact on a landslide. Six landslide susceptibility maps were demarcated. This study had given a conclusive statement that the BLR method gave better performance than other machine learning models and the AHP method gave the least accuracy after the validation by various statistical indices such as AU-ROC, ACC, F-score, MCC, Kappa, and confusion matrices.

**Ghasemian et al. (2022)** had designed a deep learning (DP) model to map the various degree of landslide susceptibility of Kamyaran city, Iran. Landslide inventory data was created by using 118 past landslide locations. Out of 25 triggering factors of landslide, only 16 factors were considered after applying the information gain ratio (IGR) technique. Three machine learning algorithms such as support vector machine (SVM), Reduced Error Pruning (REPTree), and Naive Bayesian tree **(**NBTree) were also applied and their corresponding performance was evaluated by sensitivity, specificity, accuracy, F1-measure, and area under-the-receiver operating characteristic curve (AUC). Along with the desired landslide susceptibility map, this research also revealed that the DP model had the highest accuracy of 92.6% among all the models and had the highest AUC value of 0.870.

# 3. MATERIALS AND METHODS

Narration of geomorphological and meteorological characters of study area such as location, climate, rainfall pattern, geology, geomorphology, humidity, temperature, and soil texture were done under this section. Data required for analysis, data sources and methodology, which is used for the present study has also been described in detail.

## 3.1. Description of the study area

Malappuram is one of Kerala's 14 districts in India. The district lies between 75° and 77° east longitude and 10° and 12° north latitude. Malappuram is divided into three natural regions: lowland, midland, and highland. The lowland region spans along the sea coast, the midland region in the center, and the highland region in the east and northeast. It has an area of around 3546.20 $km^2$, making it Kerala's third-largest district by area (Biju et al., 2021). The location of the study area in India is demonstrated in Figure3.1.

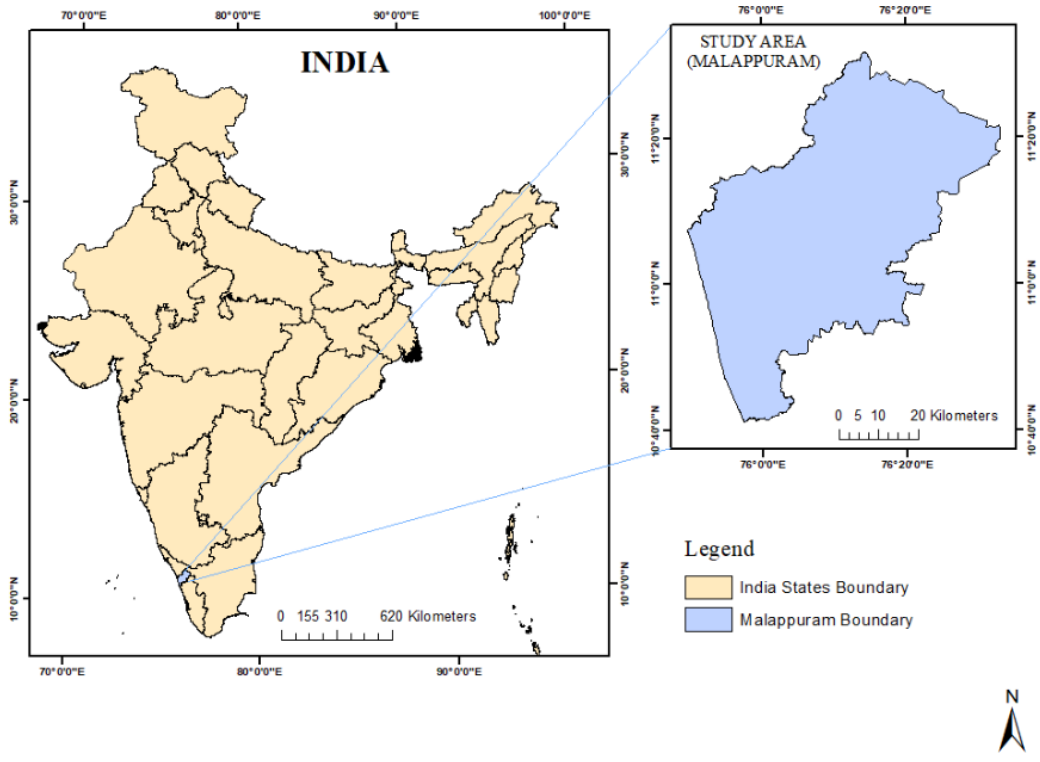


Figure 3.1:Location of the study area

Malappuram district is surrounded to the east by the Nilgiri Hills, to the west by the Arabian Sea, and the south by the districts of Thrissur and Palakkad. The research region is primarily drained by the Kadalundi, Chaliyar, and Bharathapuzha rivers (locally known as the Ponnani River).Only the Chaliyar and Bharathapuzha rivers are perennial, while the others dry up in the summer, making Malappuram district very drought-prone.

The river Kadalundi is formed by the convergence of two of its largest tributaries, the Olipuzha and the Veliyar. The Olipuzha gets its name from the forest of the 'Cherakkobban Mala' (1160m AMSL), while the Veliyar gets its name from the woodland of the 'Erattakomban Mala' (1190m AMSL). The Kadalundi river is 130 kilometers long and drains an area of 1274 square kilometers. The river empties into the Lakshadweep Sea about 5 kilometers south of the Chaliyar river mouth. The Chaliyar river, one of the state's major rivers, flows from the Ilambalari Hills in Tamil Nādu's Nilgiris district. The river passes through the Malappuram district's northern boundary towns of Nilambur, Mambad, Edavanna, Areakode, and Feroke. It merges with the Lakshadweep Sea at Beypore. In Kerala State, the river is 169 kilometers long and drains an area of 2535 square kilometers. The Bharathapuzha or Ponnani river is Kerala's second-longest river, flowing from the Anamalai Hills in the Western Ghats. The Ponnani river flows below the confluence of the Bharathapuzha and Gayathripuzha rivers. It runs through the districts of Palakkad, Malappuram, and Trichur before emptying into the Lakshadweep Sea at Ponnani in Malappuram.

The drainage pattern of the district's three rivers is mainly dendritic. Tidal impacts may be felt in towns 6 to 8 kilometers from the shore, such as Vallikkunnu and Tirurangadi. The drainage features analysis says that the Kadalundi river, the Ponnani river, and the Chaliyar river are of the fourth order, the fifth-order, and the seventh order respectively.

### *3.1.1. Rainfall and climate :*

The study area's climatic characteristics include a dry season (December-February) and a hot season (March-May), as well as the South-West monsoon (June-September) and the North-East monsoon (October-December). The district's average rainfall is

2793.3mm. The SW monsoon contributes the most rainfall, followed by the NE monsoon. The southwest monsoon is often exceptionally heavy, accounting for roughly 73.5% of total rainfall.16.4% of the total precipitation comes from the NE monsoon. Nearly 9.9% and 0.2% of annual rainfall are obtained from summer showers (March-May) and winter period rain events (January-February) respectively.

### *3.1.2. Temperature :*

The weather is often hot and humid. The months of March and April are the warmest, while January and February are the coldest. Maximum temperatures vary from 28.9°C to 36.2°C, while minimum temperatures range from 17.0°C to 23.4°C. The temperature begins to rise in January and reaches a high in March and April, then fall during the monsoon season and begins to rise again in September.

### *3.1.3. Relative humidity and wind speed :*

During the early hours, the relative humidity varies from 84 to 94 percent. During the peak monsoon months of June to September, humidity levels are higher. During the morning and evening hours, the wind is predominantly from the east as well as the west. From December to February, the wind speed is higher. It varies from 2.9 to 7.2 km/h.

### *3.1.4. Geomorphology :*

Geomorphologically, the area is separated into three sections: coastal plain (less than 7.5 m above sea level), midland (7.5 – 75 m above sea level), and highland (above 75 m AMSL). The coastal plains are a small strip of land that stretches along the coast from KadalundiNagaram in the north to Ponnani in the south. It gets quite narrow to the north of Tirur, with the greatest breadth visible at the Chauravallam-Tirurangadi stretch. The midlands cover the territory between the coastal plain in the west and the high hills in the east. This is the district's most visible physiographic unit. This has flat-topped hillocks with steep 'U' shaped troughs and ridges.The valley is a prospective agricultural location for paddy, areca nut, vegetable, banana, and coconut cultivation. The net cropped area of Malappuram is 174931 hectares. The larger area is used for mixed plantation than paddy. The hilltops are mainly barren, with thick

and dense laterite covering them. The district's eastern regions are characterized by high hills, gorges, and escarpments. The height of the hill ranges reaches 1127 meters above sea level. Forests cover many of the highlands.

### *3.1.5. Geology :*

Malappuram District is home to two significant geological bands. The charnockite group of rocks covers much of the region. Mimetite makes up the eastern sections. Charnockite gneiss, granitic gneiss, and hornblende boitite gneiss are all members of the Charnockite group. Banded magnetite quartzite, pyroxene granulite, amphibolites, hornblende granulite, and pyroxenite are listed in decreasing order of abundance in the areal distribution in the Malappuram district (Sreekumar & Aslam, 2017). Biotite hornblende gneiss and quartz of eldspathic gneiss are examples of migmatitic complexes. Acid intrusive include pegmatite and quartz veins, whereas basic intrusive include dolerite and gabbro. Lateritisation is widespread, with the isolated capping of Neogene Warakali sediments reaching a thickness of more than 10m along the shore. Unconsolidated Quaternary sediments are confined to coastal planes.

### *3.1.6. Soil texture :*

The soils of the district can be broadly divided into the following categories (i) Soils of the low lands (Alluvial soil)- Those are mainly seen along the coastal plains and valleys. In nature, the soils extend from completely drained sand to moderately/well-drained sand to sandy clay. (ii) Soils of the Mid/Uplands (Lateritic Soil)- These are predominantly lateritic soils found in the district's midlands. These are deep to extremely deep, gravelly to clayey, and well-drained. (iii) Central Sahyadri soils (Hydromorphic soil)- These are deep moderately, well-drained clayey soils with a significant gravel concentration. Erosion ranges from moderate to severe. There are hard laterites with rock-out crops. (iv) Soils of Malappuram's eastern region (forest loamy soil)- These soils are deep or extremely deep, well-drained, with loamy to clayey characteristics and a high gravel concentration.

## 3.2. Data collection of physical factors

For the development of the landslide susceptibility map of the Malappuram district, the study area boundary shape file was collected from the DIVA-GIS website. We have considered 14 causative factors, which have a direct and indirect impact on the occurrence of landslides. DEM (Digital Elevation Model) of SRTM (Shuttle Radar Topography Mission) of 30m spatial resolution has been downloaded from the USGS earth explorer website. Information about slope, aspect, elevation, curvature, and relative relief has been derived from DEM. Both geological maps of scale 1:2000000 and geomorphology map of scale 1:250000 for the study area have been downloaded from the GSI website (Geological Survey of India). Soil map was collected from FAO Digital soil map of scale 1:5000000. For the derivation of NDVI (Normalized Difference Vegetation Index) information, Sentinel-2 satellite imagery of 10m spatial resolution was collected. 30 years of daily rainfall data starting from 1991 to 2020 was collected from IMD (India Meteorological Department) for six rain gauge stations such as Nilambur, Manjeri, Anakkayam, Angadipuram, Perinthalmanna, Ponnani. For land use land cover (LULC) classification, a Sentinel-2 satellite image of spatial resolution of n 10m was collected for the year 2020. Using GIS platforms, 2018 and 2019 landslides were mapped by visually interpreting satellite photos of pre-landslide situations and post-landslide situations. Mainly, Google earth pro, survey data from GSI (Geological Survey of India),and previous inventory maps of the landslide from Kerala State Disaster Management Agency (KSDMA) were utilized to overview the landslide area in the study area. 200 past landslide points and an equal number of points of non-landslide points are chosen for analysis purposes. All the maps created were projected according to UTM Zone-34N with a datum of WGS 84.

## 3.3. Methods of analysis

### *3.3.1. Weighted overlay method (WoM) :*

The weighted overlay method is a semiqualitative method. By this method, multi-criteria analysis (MCA) can be done based on various causative factors. Raster layer databases of factors are created using GIS software. Then weightage value is assigned to each sub-classes of the factor layer using the weighted overlay tool in ArcGIS10.8.

Weightage, which has been assigned to factors, indicates the relative significance of influencing elements in landslide development. Assignment of weightage is carried out based on the knowledge acquired from past landslide incidents and vigorous field surveys (Sarkar & Kanungo, 2004). The overall weight of all the thematic layers should be 100. The sub-classes of thematic layers are again reclassified into a common reference scale of 1 to 9, where 9 indicates the most favorable factor and 1 indicates the least favorable factor. A single layer is generated from the reclassified layers, which shows the degree of landslide risk (high, medium, low) in any region. Only spatial distribution of vulnerability can be shown through this method.

### *3.3.2. Physical factors of the study area :*

Physical factors which contribute to landslide initiation has been considered such as slope, elevation, soil texture, relative relief, geology, geomorphology, rainfall, drainage density, aspect, lineament density and LULC.

*3.3.2.1. Slope* :

Being the most significant natural element in the process of landslide, the slope is expressed in terms of percentage or degree. The slope is directly linked to the likelihood of landslide occurrences (Clerici et al., 2002). As the angle of slope increases, so does the soil shear stress, making the region vulnerable to landslides (T. Raghuvanshi et al., 2015). This is not to say that gentler slopes are not vulnerable to landslides. Other conditions might cause even a gradual slope to be at risk of landslides. A steeper slope increases the velocity of water flowing down and the slope material that supports the slope is being eliminated, exacerbating the slope instability problem. The amount of runoff and hence the extent of erosion are influenced by the slope's degree and height (Azeze, 2021).The angle of slope steepness is calculated using the basic trigonometric equation 3.1of young’s method (P. Abraham & E., 2013), which is given as

$$Tan\ U = \frac{V}{d} \tag{3.1}$$

where U is the angle of slope, V is the vertical contour interval (altitude difference between two points on the same line of slope), and d is the contour separation (horizontal distance between the two points)**.**

*3.3.2.2. Aspect* :

Aspect identifies the downslope direction of the maximum rate of change in value from each cell to its neighbors. Aspect can be thought of as the slope direction. The values of the output raster are the compass direction of the aspect. Some climatic occurrences, such as intensity and amount of rainfall, amount of sunlight, and the morphologic nature of the terrain, influence the soil strength and subsequently tendency of landslides**.** As an aspect of a slope impacts the moisture retention and plant cover, it indirectly affects the landside occurrence (T. Raghuvanshi et al., 2015). Hillsides that get copious rainfall attain saturation faster; nevertheless, this is also connected to the slope's filtering capacity, which is governed by a variety of criteria such as slope topography, soil type, permeability, porosity, humidity, organic components, land cover, and climatic season (H. R. Pourghasemi et al., 2012). As a result, the slope-forming material's pore water pressure varies**,** which affects the stability potential of the land.

*3.3.2.3. Elevation* :

The elevation of a geographic location is its height above a fixed reference point, often the mean sea level. Mountain peaks at very high elevations are often made of worn rocks with a high shear strength (R S et al., 2016). At intermediate heights, however, slopes are coated by a thin layer of colluvium, making them more prone to landslides. At extremely low elevations, the landslide risk is limited since the terrain is gentle and covered with a thick layer of colluvium and (or) residual soils, and a higher elevated water table will be required to trigger slope collapse (Dai & Lee, 2001).

*3.3.2.4. Relative relief* :

The highest difference in altitude within a cell is called relative relief (Oguchi, 1997). Relief is an important landscape feature that reflects the effects of uplift and erosion

(Schmidt & Montgomery, 1995). Landslides restrict and react to mountain relief, and relief is a crucial geometric factor that enhances landslide development (Bui et al., 2012; Roering, 2012; Bucci et al., 2016). Relative relief has a substantial influence on the frequency of landslides, with higher relative relief resulting in more landslides (Pareek et al., 2010). Relative relief is directly related to the degree of dissection. Therefore, an increase in the degree of dissection causes increases in relative relief. As a result, this factor is essential in slope failure research.

*3.3.2.5. Lineament density* :

Lineament represent the lines of the earth's surface, which are created by joints and faults. It reflects a relationship with subsurface features such as rock basement architecture (Hobbs, 1904). Joints, faults, and shear zones are considered topographically negative lineament and have more influence on landslide susceptibility. Faults and joints are regarded as discontinuous structural weaknesses, as the presence of these features reduces the shear strength of the mass (Hencher, 1987). Faults are defined by fractures along which displacement of the earth's surface has occurred vertically previously. Low shear strength, relatively high permeability, and containment of high groundwater flow are major characteristics of faults that initiates slope failure. Accumulation of water along the fault lines increases the pore water pressure and chemical weathering of rocks. Moreover, joints are defined as brittle fractures on the surface of rocks, along which very little displacement in the horizontal direction to the plane of fracture has occurred or no displacement at all. Less shear strength of joints reveals their vulnerability towards slope failure under triggering conditions. Therefore, analysis of distance from lineament and lineament density carries importance in the study of landslide susceptibility.

*3.3.2.6. Rainfall* :

Rainfall is one of the natural causative factors, which triggers slope failure. As the volume, duration, and intensity of rainfall regulate the rainfall erosivity, soil erosion, infiltration, and runoff, the scenario of rainfall pattern in a region should be studied to get the probability of slope failure. The saturation of soil with water increases the pore pressure in soil and decreases the soil strength (Abramson et al.,2002). This

phenomenon is more common in clay soil which tends to store water in its pores. A combination of anthropogenic interferences in land use patterns in a vulnerable area and high duration or high-intensity rainfall leads to slope failure. The construction of roads, houses, pipeline structures, and mixed plantation disturbs the natural slope of the region. Water being the important factor in the chemical weathering process, soil and rock of higher rainfall regions undergo molecular and structural change more often than other regions.

*3.3.2.7. Land use and land cover (LULC)* :

Land use and land cover are important factors in the analysis process of landslide, as it affects this hazard indirectly. The considerable change in land use pattern creates disequilibrium in the environment and shows an increment in the frequency of landslide occurrences. Improper agricultural practices like the Plantation of shallow and weak rooted plants on a steep slope for commercial benefits, the building of houses by diverting the path of stream flow, construction of roads on an unfavorable slope, excessive mining and quarrying with inconsideration of slope conditions can create disturbances in natural arrangements of rock, soil mass or path of stream flow. Deforestation in steep slope regions causes acceleration of soil erosion, reduction of soil cohesion, increase runoff, and subsequent occurrence of debris flow.

*3.3.2.8. Drainage density* :

The ratio of the total length of streams to the whole area is known as drainage density. High drainage densities are associated with impermeable strata, high rainfall, minimal vegetation, and vigorous stream incision, all of which could be linked to mass movement (Whalley et al.,1991). Slope stability can be affected by proximity to hydrological features such as streams or rivers. Due to erosion and undercutting at the foot of slopes, streams may have an unfavorable effect on slope stability. Slopes become more susceptible to landslides during periods of heavy rainfall, or they can produce catastrophic breakdown along weak lineaments such as joints, bedding, and exfoliation plains (R S et al., 2016).

*3.3.2.9. Soil texture* :

The physio-chemical character of soil plays a vital role in the landslide process. For the prediction of slope failure in any region, the spatial distribution of soil, the thickness of soil on the earth's surface, and hydrological aspects of soil should be studied thoroughly. Upon saturation with rainfall, the water content of soil increases and subsequently the pore pressure. The causes washing away of soil by decreasing its cohesiveness and the material above it tends to slide down too. Clay soil has a higher water holding capacity compared to other soil types and is having high tendency to be a causative agent in a landslide.

*3.3.2.10.Geology and geomorphology* :

The analysis of landforms, their genesis processes, and the spatial connections between forms and processes is known as geomorphology. Slope failures were also recurring in other tectonically dormant and older mountainous places, notably when heavy rains and major anthropogenic intrusions were associated. Later, the investigations were expanded to include other mountainous regions, revealing that other geological, climatic, and anthropogenic characteristics were shown to be responsible for landslides, regardless of age or tectonic state (Skilodimou et al., 2018). Due to complex tectonic geomorphology and the resulting significant influx of precipitation and pore pressure increase, the Western Ghats of Kerala are prone to repeated landslides. Furthermore, the vast deforestation permits rainwater access directly. This hypothesis of ours about landslides caused by pore pressure is backed up by previous research (Galeandro et al., 2014).

*3.3.2.11.Curvature* :

A part of the surface can be concave, convex, or plan, this can be decided by using curvature value. The Curvature function's output can be used to define the physical properties of a drainage basin for a better understanding of erosion and runoff processes. The curvature value can be used to determine soil erosion patterns and water distribution on land. There are the three types of curvature, such as profile, planform, and standard. Profile curvature is parallel to the direction of the maximum

slope, whereas planform curvature is perpendicular to the direction of the maximum slope. The combination of profile and planform curvature is called standard curvature. The profile curvature affects flow acceleration and deceleration, and so erosion and deposition. The planform curvature affects flow convergence and divergence of landslide material and water in the direction of landslide motion (Carson & Kirkby, 1972). When you consider both the planform and the profile curvature, we may gain a better understanding of the flow through a surface. The vulnerability to landslides is affected by both profile and plan curvatures. In Figure 3.2, the columns show the planform curves, and the rows show the profile curve. The planform columns are positive, negative, and 0, going from left to right. The profile curves are negative, positive, and 0, going from top to bottom.

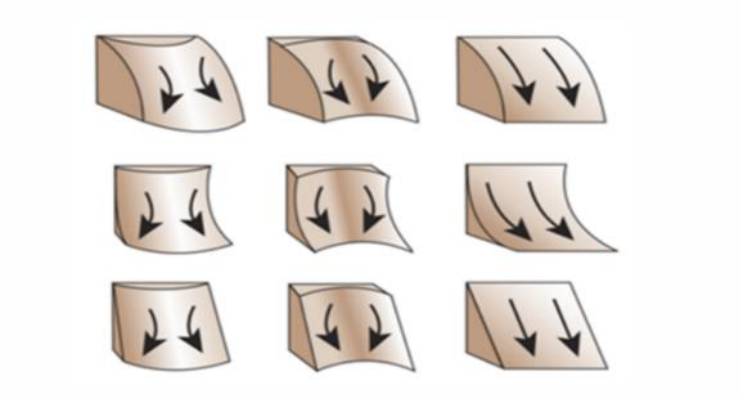

Figure 3.2: Profile and planform curvatures

### *3.3.3. Artificial Neural Network (ANN) :*

The ANN technique (Pradhan & Lee, 2010) is a generic nonlinear function approximation method that has been frequently used for regression applications such as landslide susceptibility modeling. An artificial neural network is a "computational mechanism able to acquire, represent, and compute a mapping from one multivariate space of information to another, given a set of data representing that mapping".

The input layer, output layer, and hidden layer (Figure 3.3) are used to make the multi-layer perceptron (MLP) neural network, which is the most common and widely used ANN architecture (Csáji, 2001). Each layer contains a sufficient number of

neurons. Each connection between neurons in one layer and neurons in the next layer has a weight attached to it. The input layer is a passive layer that just receives data (e.g., the data from various sources). As a result, the number of input layer neurons equals the number of sources to be utilized. Unlike the input layer, the data is actively processed by both the hidden and output layers. The output layer, as the name implies, is responsible for the neural network's output. The trial-and-error method is adopted to determine the number of buried layers and their neurons.

The back-propagation (BP) algorithm is frequently utilized to apply the MLP neural network for landslide susceptibility modeling problems (Wang et al., 2016). The initial weights of the inputs are chosen at random using the BP method, and the mean-square error is used to summarise the differences between the calculated and expected values over all observations. This algorithm iteratively minimizes an error function (E), which is calculated from a sample of the target (known) and network-derived outputs. The process is repeated until E converges to a minimum value, at which point the updated weights are calculated by the equation 3.2, which was used by Li et al., (2012).

$$E = 0.5( \sum_{i=1}^{n} T_i - O_i )^2 \qquad (3.2)$$

where *Ti* is the target output vector, *Oi* is the network output vector and *n* is the number of training samples. The network is supposed to be trained once the suitable weights for all connections have been identified. To evaluate the network's performance, these weights are used to determine the class outputs of testing (known) datasets. If the testing dataset's accuracy is low, it suggests that even though the network was trained correctly, it has yet to achieve generalization capacity and needs to be retrained. The updated weights are used to determine the output of the full dataset after the network has been retrained and tested to the appropriate accuracy.

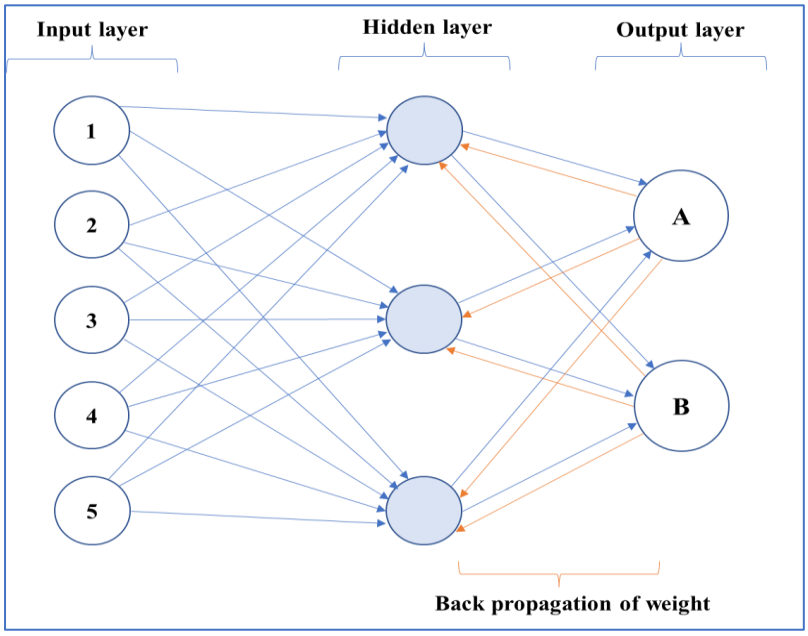


Figure 3.3: Graphical representation of back propagation algorithm in ANN network

### *3.3.4. Support Vector Machine (SVM) :*

SVM is a statistical learning theory-based supervised machine learning classification algorithm (Vapnik, 2000). SVM has been widely and successfully utilized for a range of classification and regression applications because it can handle linearly non-separable and high-dimensional data sets. SVM uses landslide training data to map data into a high-dimensional feature space and determine an ideal hyperplane with a maximum margin to distinguish between landslide and non-landslide classes (Pourghasemi et al., 2013). These maximum separating margins are known as support vectors. Figure 3.4 has shown the classification principle of the SVM model. The SVR algorithm presupposes that each collection of input features (for example, landslide conditioning factors) has a distinct relationship with its target variable (i.e., landslide susceptibility index). As a result, the SVR algorithm determines the rules for estimating unknown test data samples' target values from a group of inputs. The linear (LN), polynomial (PL), sigmoid (SIG), and radial basis functions (RBF) are the four most common kernel functions used in landslide mapping.

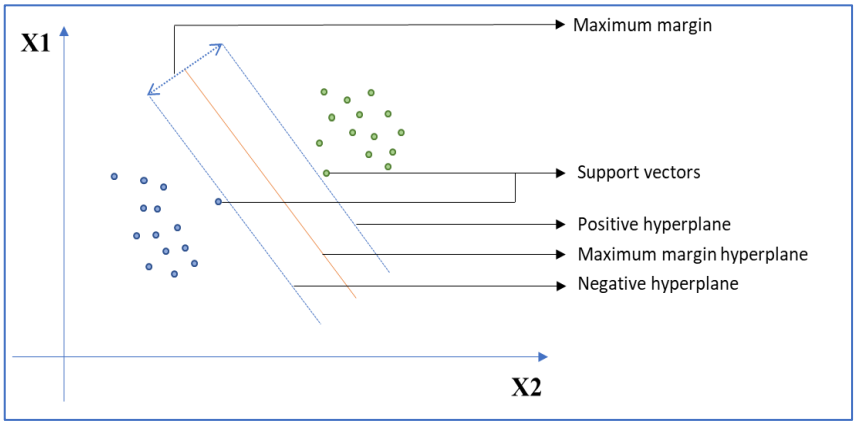


Figure 3.4:Graphical representation of SVM-model

### *3.3.5. Random Forest (RF) :*

Breiman had developed Random Forest (RF) technique in order to carry out classification and regression operation of the nonparametric multivariate (Catani et al., 2013; Breiman, 2001). Through the random forest approach, several relatively uncorrelated decision trees (DT) are created and finally, a random forest is formed by combining all the individual DTs. This method eliminates the high variance or bias of a single DT while doing prediction for a dataset (Erener et al., 2016).As the RF technique makes use of cross-validation, the problem of overfitting the model or improper randomization of the dataset can be easily avoided. By employing the bootstrap bagging procedure, several subsets of original data are created and used as a training dataset for Decision Trees. The result will be decided based on the highest number of votes of DTs. The fundamental principle of a random forest model is exhibited in Figure 3.5.

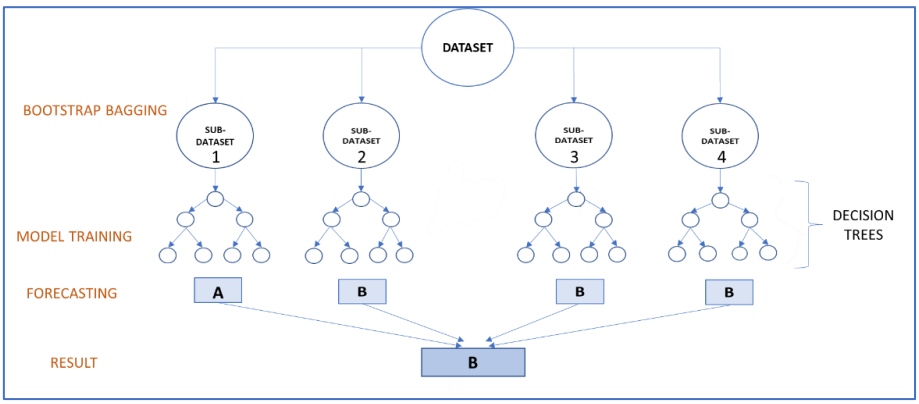


Figure 3.5: Graphical representation of RF model

### *3.3.6. K- Nearest Neighbour (KNN)*

KNN is a supervised machine learning algorithm, which can be used for both classification and regression operations. Being a non-parametric model, it is able to analyze real data without any assumptions. KNN is best-known as a lazy algorithm, as it does not approximate the target function globally. By this approach, a new unclassified data is categorized into any target class based on the similarity with the neighboring data points. K stands for several nearest neighbors and the choice of K-value is very crucial in KNN analysis. For instance, if K is taken as 4, then the four nearest neighbors will be chosen by the algorithm and the new data point is classified after inspecting its similarity of it with the four known data points. The concept of a KNN model is well explained in Figure 3.6.

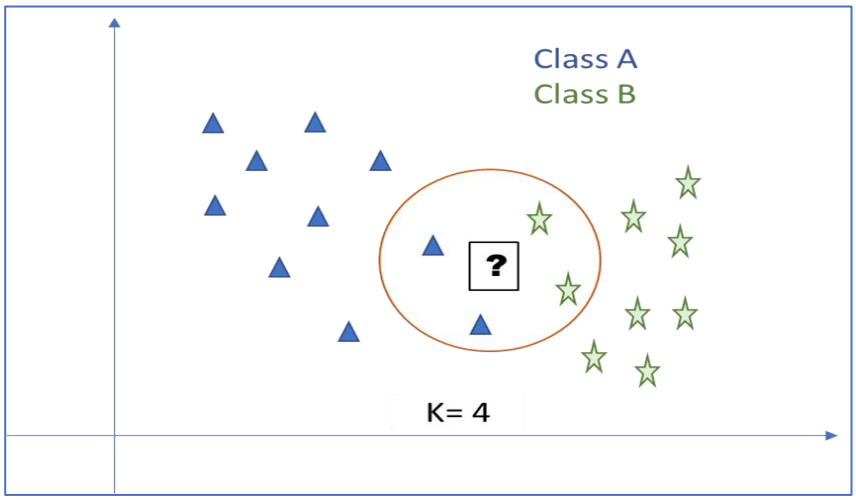


Figure 3.6: Graphical representation of the KNN model

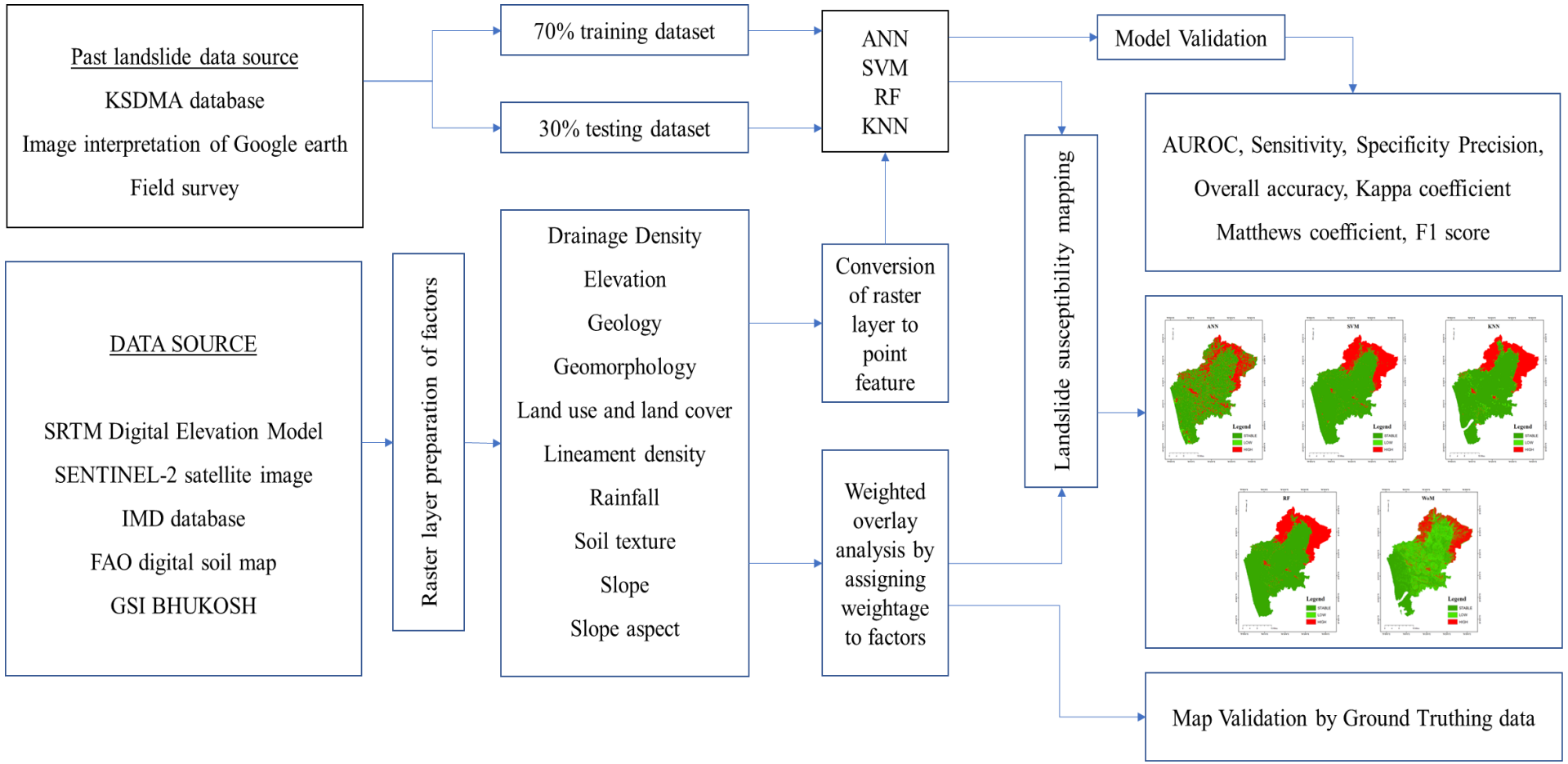


Figure 3.7: Flow chart of the methodology

# 4. RESULTS AND DISCUSSION

In this chapter, we have presented the results, which were obtained from the Weighted overlay method (WoM) and four machine learning models. The results were also compared with the previous research work of scholars in the corresponding field.

## 4.1. Composition of raster layer of physical factors

### *4.1.1. Slope* :

Using ARC GIS 10.8 software, the slope map of the study area was extracted from the downloaded DEM of SRTM and is given in Figure 4.1 (a). The range of slope was classified into eight groups, such as less than 1°, 1° to 3°, 3° to 5°, 5° to 10°, 10° to 15°, 15° to 20°, 20° to 30°, more than 30°. From the prepared map, we observed that the northern part of the district is a high-slope region. 196 $km^2$ of the Malappuram district are having a slope of more than 30° and 301$km^2$ of the area falls under a slope range of 20° to 30°. Although this high slope area comprises a small part of the whole district but may affect the low-lying area the most during debris flow kind landslide.

### *4.1.2. Aspect* :

The aspect map, which was derived for the study area in Figure 4.1 (b), was categorized into ten groups. These groups were Flat, North, Northeast, East, Southeast, South, Southwest, West, Northwest, and North. As North facing slopes are shielded from drained wind and sunrise mare than south-facing slopes, they absorb more moisture. The surficial process working on the drier slope flattens it more easily, the moister slope is constantly steeper than the drier Southwest slope.

### *4.1.3. Elevation*:

Because high elevation locations receive more rainfall than lower elevation places, they are more susceptible to landslides (Coe et al., 2013). The elevation of the study area was categorized into more than 2000m, 1500m to 2000m, 1000m to 1500m, 400m to 1000m, and less than 400m elevations, and the map were presented in Figure 4.1 (c).

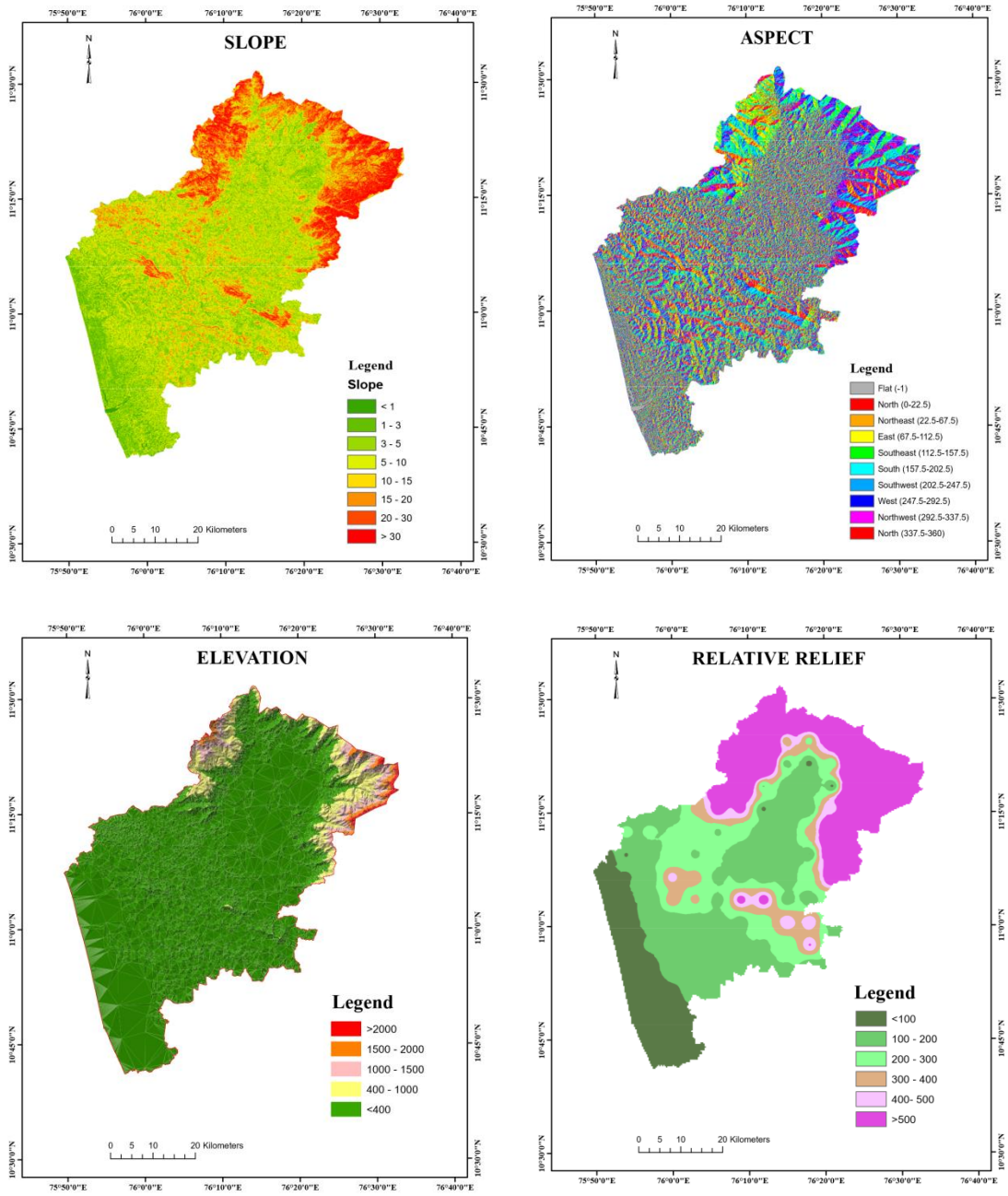


Figure 4.1: (a) Slope map, (b) Aspect map, (c) Elevation map, (d) Relative relief map of study area

### *4.1.4. Lineament density* :

Number of lineaments per kilometer square area was grouped into five groups. The high lineament density region is more susceptible to the landslide physically. This can be well explained due to the weak geological formation of the region, which is making it more susceptible compared to other regions. The lineament density of the study area is shown in Figure 4.1 (e).

### *4.1.5. Relative relief* :

The study area was classified into six groups concerning the ranges of relative aspects in Figure 4.1 (d) . These groups are such as less than 100m, 100m to 200m, 200m to 300m, 300m to 400m, 400m to 500m, and more than 500m. It was observed that maximum landslides had happened around the relative relief above 400m.

### *4.1.6. Land use and land cover (LULC)* :

Malappuram region is covered with mainly by plantations and mixed plantations and dense forest in the northern part. Due to population growth into high slope lands, cutting of slope for building, farming and plantation make this area more unfavorable. During the field survey, it was observed that hilly slope regions were cleared and was converted into rubber plantation zone. The weak and shallow root system of rubber plants is another reason for the low strength to hold the soil mass during the high rainfall season. Invasion of settlement zone into adverse slope area was taken note during field survey. Barren lands, which were developed due to the quarrying process, can be observed in Malappuram District easily. Quarrying disturbs the geomorphology of the region. All these minor incidents, yet having a huge impact on debris landslide flow. Land use and land cover of the study area are shown in Figure 4.2 (c),which is grouped into nine groups Dense Forest, Plantation, Forest, Settlement, Waterbody, Wasteland, Barren Land, Cropland, and Paddy land. LULC map was produced by supervised classification and accuracy assessment (Table 4.1) shows the accuracy of 83. 70 % after ground-truthing had been carried out.

Table4.1 : Accuracy assessment of LULC classification

| Classes | Random points | Correct | Incorrect | Accuracy |
|---|---|---|---|---|
| Dense forest | 20 | 18 | 2 | 90 % |
| Forest | 20 | 17 | 3 | 85 % |
| Plantation | 20 | 16 | 4 | 80 % |
| Settlement | 20 | 18 | 2 | 90 % |
| Waterbody | 10 | 9 | 1 | 90 % |
| Wasteland | 10 | 7 | 3 | 70 % |
| Barren Land | 15 | 13 | 2 | 86.67 % |
| Cropland | 10 | 8 | 2 | 80 % |
| Paddy land | 10 | 7 | 3 | 70 |

### *4.1.7. Rainfall* :

Past 30 years' daily rainfall data was collected from IMD to plot the rainfall map for the study area in Figure 4.2 (b). Though water is one of the reasons for the initiation of landslides, we observed that regions having rainfall within 2648 mm and 2734.63mm of annual precipitation, show more frequency of landslides than regions, which are having an annual rainfall of more than 2800mm. This exception can be explained by the rainfall erosivity of the region. Due to the high intensity of rainfall around the high slope region, the rainfall erosivity is more in the north part of the study area than in the low slope area of the southern part.

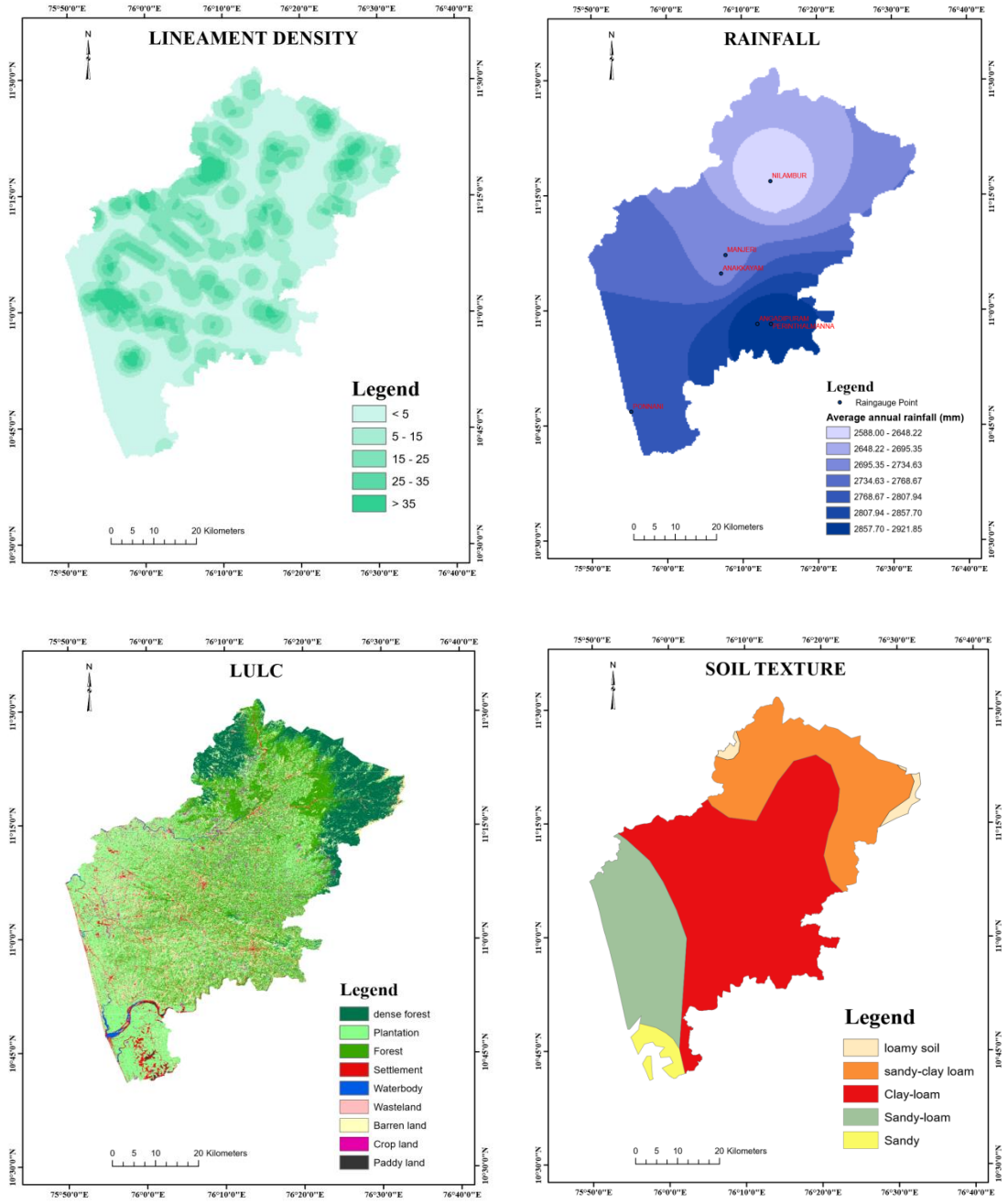


Figure 4.2: (a) Lineament density map, (b) Rainfall map, (c) Land use and land cover, (d) Soil map of the study area

### *4.1.8 Geomorphology* :

Eco-rich Malappuram geomorphology map in Figure 4.3 (a), consists of the coastal plain, denudational hill, pediment-pediplain complex, structural hills, piedmont alluvial plain, active flood plain, waterbody, and anthropogenic terrain. Past landslides show that denudational hills and structural hills are major landslide-prone regions. Denudational hills are having the typical character of being steeply sloping and highly weathered. These kinds of hills comprise sandstone, shales, and boulder beds. Structural hills are originated due to the tectonic process and highly dissected due to the drainage lines. The presence of sandstones, shales, and limestones with prominent joints and fractures are the features of structural hills. Both the structural hills and denudational hills are highly dissected. They have a low capability to store moisture as groundwater and mostly rainwater runs down as surface runoff from this region.

### *4.1.9. Geology* :

Malappuram geology is given in Figure 4.3 (b) and has been divided into five types of rock classes such as Charnockite gneissic complex, Migmatite gneissic complex, Peninsular gneissic complex, Coastal, and glacial sediments, and Granite. The parent rocks in the Malappuram district are charnockite rocks, and the weathering process of the rocks produces reddish-brown colored laterite soil (Sajinkumar et al., 2013). Charnockite is a felsic rock found in the lower crust that contains orthopyroxene and garnet, among other minerals (Touret & Huizenga, 2012). Hillslope instability is induced is induced by this weathered granite, charnockite, and gneissic rocks (Kannan et al., 2011).

### *4.1.10. Curvature* :

Fora better understanding of the runoff and erosion process of the study area, a curvature map, which is given in Figure 4.3 (c), was derived from DEM. The range of curvature is between 50.56 to 59.69. A high positive value of curvature was observed in the northern part of Malappuram. Profile and planform curvature map was also derived to observe the concavity, convexity, and flatness of the surface along with convergence and divergence of flow across a surface. Upon Observing these three

maps in an integrated manner, we concluded that the high slope region of Malappuram is sideward convex and the surface is upwardly concave. Being sideward convexity of surface, streams are diverged rather than converged into one stream. Upwardly concavity of the surface accelerates the stream flow and increases the kinetic energy of flow, which causes massive soil erosion.

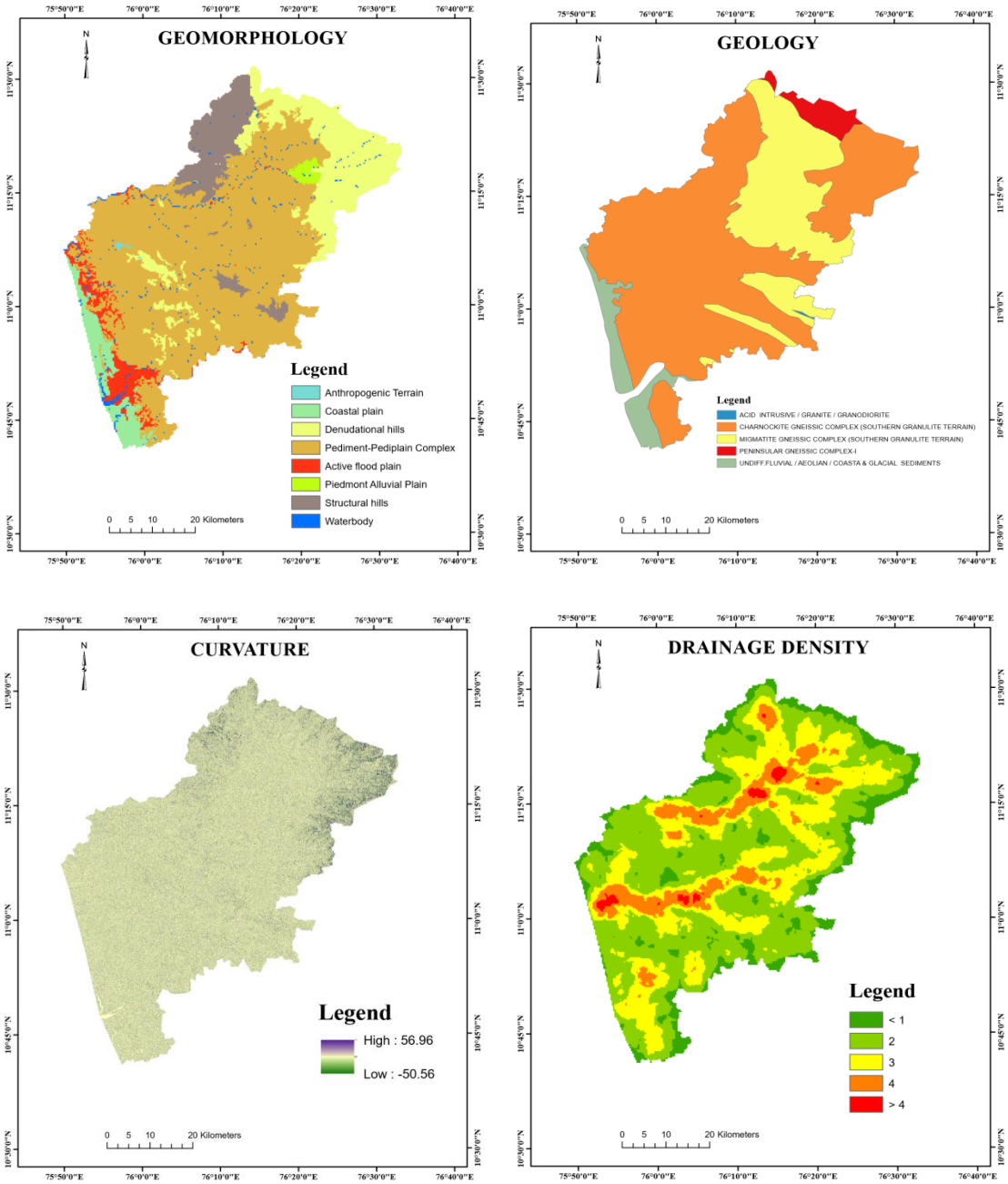


Figure 4.3: (a) Geomorphology (b) Geology (c) Curvature (d) Drainage density map

### *4.1.11. Drainage density* :

The drainage map of the study area is illustrated in Figure 4.3 (d) and the analysis of density shows a stream order of up to six. Drainage density is an indirect indicator of groundwater conditions, which play a key influence in landslide activity. Kanungo et al. (2006) discovered an inverse association between landslides and drainage density, which they believe is due to the increased infiltration in weathered gneisses, which implies high instability in the area. Low drainage density implies low surface runoff and higher infiltration to groundwater, which may increase slope instability in a region. This can be well explained subsequent development of high pore pressure in the soil. Again, the occurrence of landslides in a high drainage density region may be due to the seepage of water into the adjacent region. As both seepage and infiltration rate of water depend on the texture of the soil, the role of drainage density depends on the soil texture and other geological conditions of the experimental area. The drainage density of the current study area is classified into five distinct groups such as less than 1, 1 to 2, 2 to 3, 3 to 4, and more than 4.

### *4.1.12. Soil texture* :

Soil has significant variances depending on geology, relief, vegetation type, and climatic conditions. One of the most essential characteristics is soil texture. The slower the flow of soil fluids and rainfall infiltration, the finer the soil texture. The soil of Malappuram is categorized as loamy soil, sandy-clay-loam soil, clay-loam soil, sandy-loam soil, and sandy soil. The high slope region is dominated by sandy-clay loam soil. The presence of clay reduces the soil shear strength after being saturated. As clay has the least transmissibility through it, the increased pore water pressure will result in the washing away of clay and other cementing material between rocks. High clay content along with high slope has worsened the landslide frequency of the region. The distribution of different soil textures in the study area can be studied from Figure 4.2 (d).

## 4.2. Generation of landslide susceptibility maps

Various raster layers and their corresponding data , which were produced before, were used as an input for the generation of final landslide susceptibility map through Weighted Overlay Method (WoM) and machine learning models.

### *4.2.1. Mapping through Weighted overlay Method (WoM) :*

Weighted overlay analysis was carried out using the geospatial analyst tool of "weighted overlay". For this analysis purpose, several landslide-controlling parameters are derived from the aforementioned geospatial data sets including slope, aspect, geology, percentage of soil types (including clay, loam, silt, and sand, among others), land use land cover classification, lineament density and hydrological factors like drainage density. To the SRTM elemental horizontal scale of 30 m, all parameters have been downscaled or interpolated. Using the "reclassify" spatial analyst tool, these maps were again categorized with a suitable range for a convenient weightage assignment process. Each factor and its subclasses were assigned (Table 4.2) with a numerical value according to export opinion subjectively. The basic equation 4.1, which is used for weighted overlay analysis, was given by Li et al. ( 2012).

$$S = \frac{\sum Wi\ Sij}{\sum Wi} \qquad (4.1)$$

Where $S$ is the spatial unit value in the output map, $W_i$ is the weight of the $i^{th}$ factor, and $S_{ij}$ is the weight of the $j^{th}$ subfactor of the $i^{th}$ factor. For the field survey of past landslides, data were collected from Kerala State Disaster Management Authority (KSDMA) and various landslide locations were marked carefully using satellite imagery and Google earth pro software. Landslide type, plantation of the affected area, the slope of the surface, presence of any major stream or direction of the stream during monsoon period along with preventive measures of Government and awareness of local people had been noted down and observed meticulously. The final map of landslide susceptibility was derived by overlaying the raster layer with weightage through software. The produced map was categorized into three classes high-risk zone, low-risk zone, and stable zone in Figure4.4.

The final map generated shows that denudational structural hills are a high-risk zone and similar result was obtained by Hariharan et al. (2016),when a landslide case study was carried out in Thenmala Sub-Watershed of the Western Ghats. Moreover, high slope and elevation region is marked as high risk zone in the resulted map which shows that these two factors play a vital role in landslide susceptibility. Machireddy (2020) had also concluded that elevation and slope have major effect on landslide initiation in the Tirumala hill ranges of Andra-Pradesh. As per the case study of Aneesah et al. (2019), the dense forest region should be less vulnerable to landslide hazards, but our analyzed map shows that dense forest area is also susceptible to landslides. This contradiction can be well explained by the removal of ancient trees by the local people for rubber, coconut, pineapple plantation, etc, which reduces the soil holding capacity of the respective area.

Table 4.2: List of weightages given to various factors

| **Sl. No.** | **Factors** | **Influence (%)** | **Sub-classes** | **Weightage** |
|---|---|---|---|---|
| 1. | Slope | 30 | < 5° | 1 |
| | | | 15° - 15° | 2 |
| | | | 15° - 25° | 7 |
| | | | 25° - 35° | 8 |
| | | | > 35° | 9 |
| 2. | Lineament density | 10 | <5 | 1 |
| | | | 5 – 15 | 9 |
| | | | 15 – 25 | 8 |
| | | | 25 - 35 | 7 |
| | | | >35 | 6 |
| 3. | Aspect | 7 | Flat | 1 |
| | | | North | 2 |
| | | | Northeast | 3 |
| | | | East | 4 |

| Sl. No. | Factors | Influence (%) | Sub- classes | Weightage |
|---|---|---|---|---|
| | | | Southeast | 5 |
| | | | South | 6 |
| | | | Southwest | 7 |
| | | | South | 8 |
| | | | West | 9 |
| | | | Northwest | 2 |
| | | | North | 2 |
| 4 | Drainage density | 5 | < 1 | 5 |
| | | | 1 - 2 | 6 |
| | | | 2 - 3 | 7 |
| | | | 3 - 4 | 8 |
| | | | > 4 | 9 |
| 5. | Geology | 10 | Migmatite GC | 8 |
| | | | CharnockiteGC | 9 |
| | | | Undiff fluvial/ Aeolian/ Coastal sediments | 2 |
| | | | Acidintrusive/granite/ Granodiorite | 3 |
| | | | Peninsular gneissic complex | 7 |
| 6. | Land use and land cover | 15 | Dense forest | 9 |
| | | | Forest | 9 |
| | | | Plantation | 8 |
| | | | Settlement | 1 |
| | | | Waterbody | 1 |
| | | | Wasteland | 1 |

| | | | | |
|---|---|---|---|---|
| | | | Barren land | 2 |
| | | | Cropland | 2 |
| | | | Mixed Plantation | 9 |
| | | | Paddy field | 1 |
| 7. | Soil | 23 | Sandy-Clay- Loam | 7 |
| | | | Clay- Loam | 8 |
| | | | Sandy- Loam | 3 |
| | | | Sandy | 4 |
| | | | Loam | 9 |

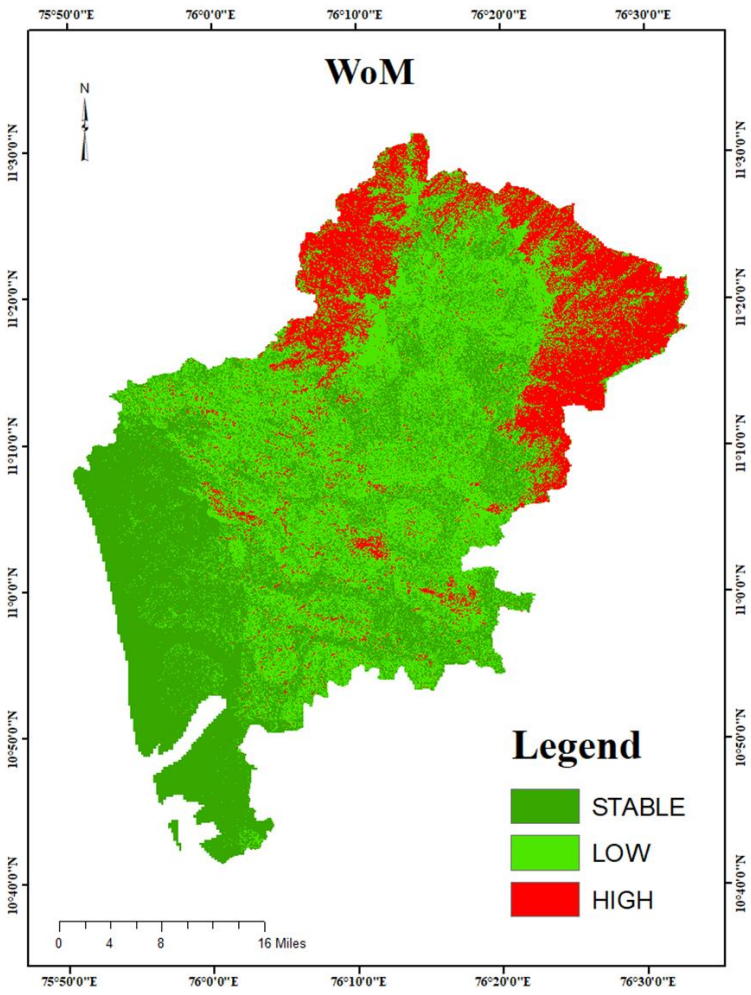


Figure 4.4: landslide susceptibility map derived from WoM method

### *4.2.2. Mapping through ANN model :*

ANN analysis was carried out using the "neuralnet" package. "Logistic" as activation function and "sse" or "sum of squared error" as error function was used for analysis. The model was run several times with a training dataset to obtain a neural network withthe least error and high accuracy. Input independent datasets of our model were slope, elevation, geology, geomorphology, soil texture, LULC, rainfall amount, and drainage density. The dependent variables were landslide and non-landslide. The best model (Figure 4.5) was obtained as 8 input layers, one hidden layer with 5 nodes, and one output layer. The map, generated from ANN is given in Figure 4.11 (a).

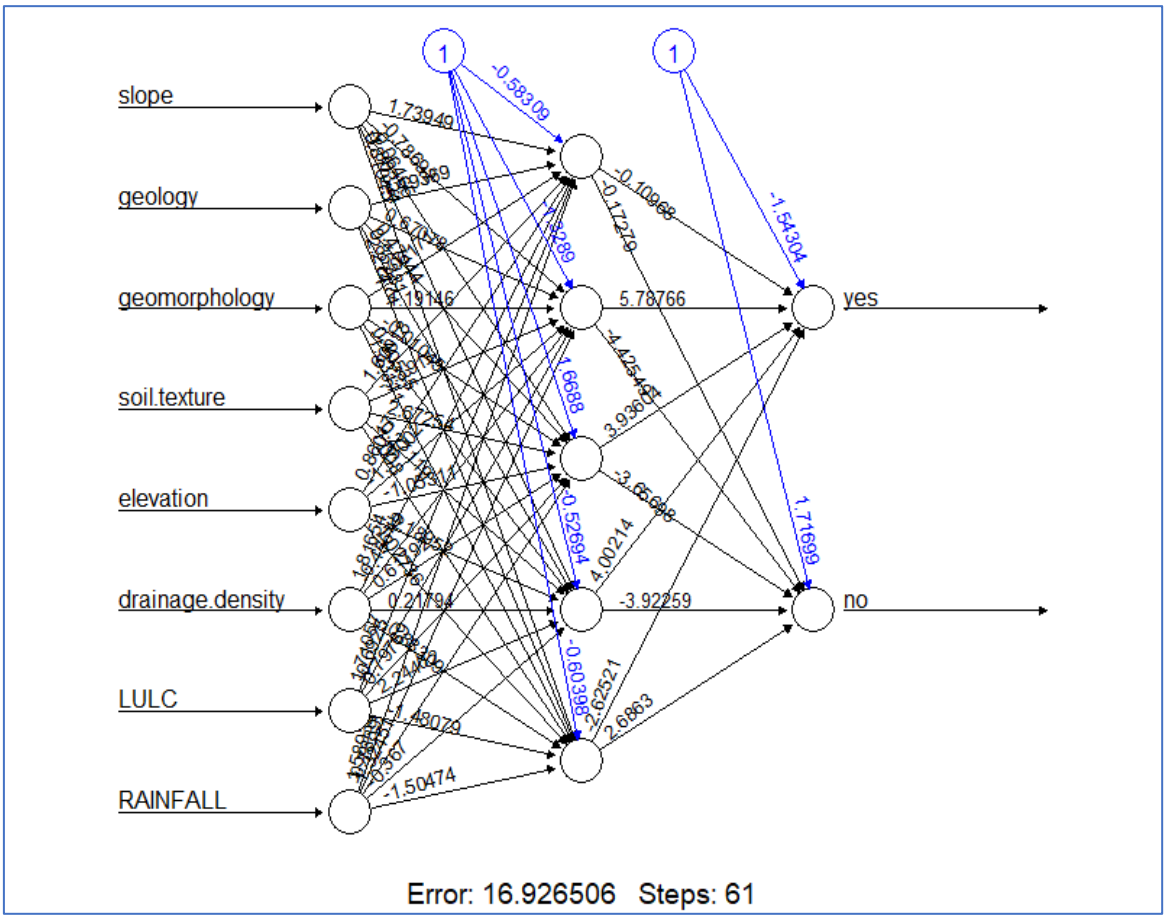


Figure 4.5: Artificial Neural Network Model

The generalized weights were obtained from the model were again plotted in Figure 4.6 to check the independent effect of each covariate on the output. If a covariate shows a large variance in generalized weight distribution, then it has a non-linear effect on output. Moreover, the high concentration of generalized weights near zero

indicates no effect of the covariate on the outcome(Intrator, 2001).After plotting, we got to know that all covariates such asslope, geology, geomorphology, soil texture, elevation, drainage density, LULC and rainfall have a non-linear effect on landslides.

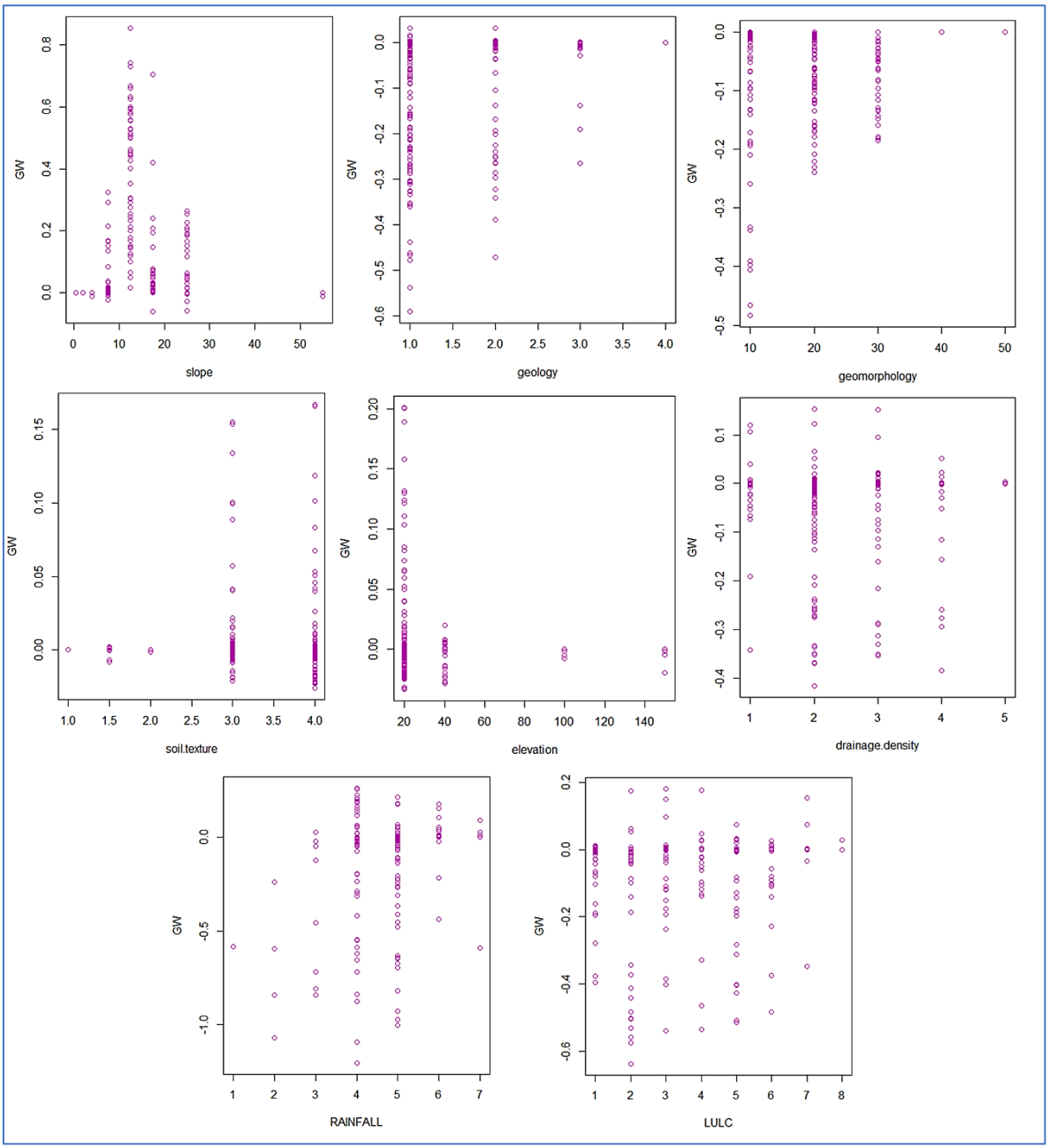


Figure 4.6: General weights for covariate (a) slope, (b) geology, (c) geomorphology, (d) soil texture, (e)elevation, (f) drainage density, (g) rainfall, and (h) LULC

Excellent result of ANN model was found by Jacinth Jennifer & Saravanan (2021)for the Idukki district of Kerala with prediction rate of 0.83 while considering 15 different triggering factors for analysis. Again Mandal & Mondal (2019) had analysed the landslide susceptibility of the Darjeeling Himalaya region and 81.5% accuracy level was found with 15 input layer, one hidden layer of 30 hidden neurons, and one output layer of 2 neurons. Ali et al. (2022) had combined the RF tress with ANN model to look over the landslide susceptibility of Himalayan region of Sikkim and mapped the vulnerability of the region with 97.4% prediction accuracy. Contrast to these above studies, poor accuracy result was obtained from our study.

### *4.2.3. Mapping through SVM model :*

The Support Vector Machine (SVM) model was developed using package "e1071". For our present dataset, kernel function, and radial basis function **(**RBF) showed the highest accuracy and were accepted for analysis. The same training data set and input factors were considered as in ANN modeling and a landslide susceptibility map was derived.A confusionmatrix was established to investigate the statistical parameters. The map, which was generated by prediction is shown in Figure4.11 (b).

SVM model shows an excellent performance with a high accuracy result for our study area and with corresponding data. Similar result was also produced by B. T. Pham et al. (2016), who had derived the susceptibility map for Uttarakhand area by using the SVM, FLDL, LR, NB, and BN model and obtained the highest AUC OF 0.95 by SVM model. Although both SVM and GARP model were adopted by Ramachandra et al. (2010) for landslide susceptibility study of locations along the Western Ghats, SVM had higher accuracy of 96%. Moreover Rana & Babu, (2022) had found that SVM (accuracy 87.45%) gives as excellent result as RF (accuracy 88.64%), while carrying out a case study of landslide susceptibility along the Western Ghats.

### *4.2.4. Mapping through RF model :*

For establishing Random Forest (RF) model, the R-package of "randomForest" was used. Before training the model, we had to set two parameters, these are "mtry" and "ntree".the 1st parameter "mtry" represents a number of variables randomly sampled as candidates at each split and the 2nd parameter ntree represents number of trees to grow. The value of "mtry" was selected as "2" due to the minimum "Out-Of-Bag" (OOB) error, obtained from the Figure 4.7 (b). After training the model with the training dataset, the relative importance of each factor was derived and it was found that slope, soil texture, land use, and land cover pattern, and geomorphology have the maximum contribution to the landslide initiation process. The relative importance of slope, geology, geomorphology, soil texture, elevation, drainage density, LULC and rainfall is graphically illustrated in Figure 4.7 (a).

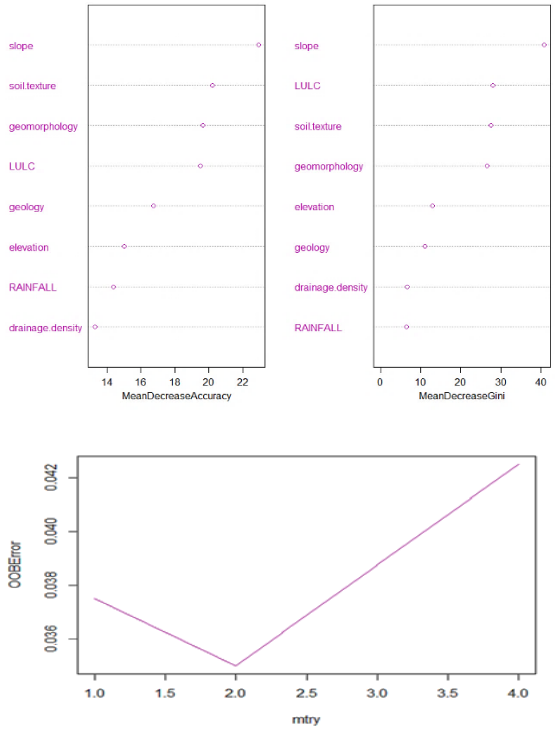


Figure 4.7: (a) Relative importance of variable (b) graph of OOB error VS mtry

Trained model of RF was used to predict the landslide susceptibility map of our study area and is given in Figure 4.11 (c).

Model validation shows that RF models give the best prediction result for our study with the highest accuracy among all the models. Ali et al. (2022) had got the best result using RF with accuracy of 90.4% among all the models of machine learning, which he had adopted to analyses lope failure probability of a highway corridor situated in Himalayan region. Furthermore, highest accuracy of 81% was obtained from RF model by Prasad et al. (2021), when he applied RF, DB, SGB, RoF, BRT, and LB for landslide vulnerability mapping of mountainous region of western India.

### *4.2.5. Mapping through KNN model :*

K-Nearest Neighbour (KNN) method was adopted for the prediction of landslides for our dataset. For this purpose, R-package "caret" was imported. The most important task in KNN modeling is to set the value of K. Decreasing the value of K to 1 will unstable the model while increasing the K value will stabilize the model due to majority voting. Moreover, increasing the value of k risks increasing errors. Therefore, the K value is increasing up to an optimal point. For the optimal model ,the K value was taken as "10" after repeated cross-validation as shown in the graph of Figure4.8. Euclidean distance function was used for analysis, although there are several distance functions present such as Hamming distance, Minkowski distance, Kullback-Leiber (KL) divergence, and BM25. The model was trained with the same dataset as used before with the same input features and finally, the prediction was done to get the map of landslide susceptibility Figure 4.11(d).

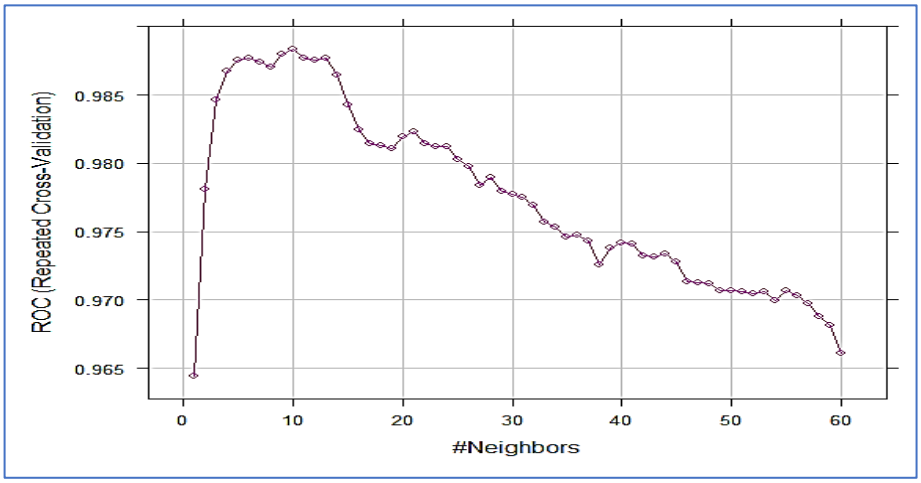


Figure 4.8:Optimal value of K

Although the accuracy of the KNN model is less than the RF and SVM model, it gives a better result than the ANN model. Many alike results were also derived by various researchers. M. T. Abraham et al. (2021) had applied SVM, NB,LR, RF along with KNN model for vulnerability mapping of slope failure along the Western Ghats and found satisfactory accuracy of 88% by KNN model with K value of 9. From the study of landslide susceptibility mapping of Idukki district of Kerala, which was carried out by Bhargavi et al. (2011), It was clearly evident that KNN model gives satisfactory prediction accuracy of 98.22% along with LR, LDA, DT, and GNB model.

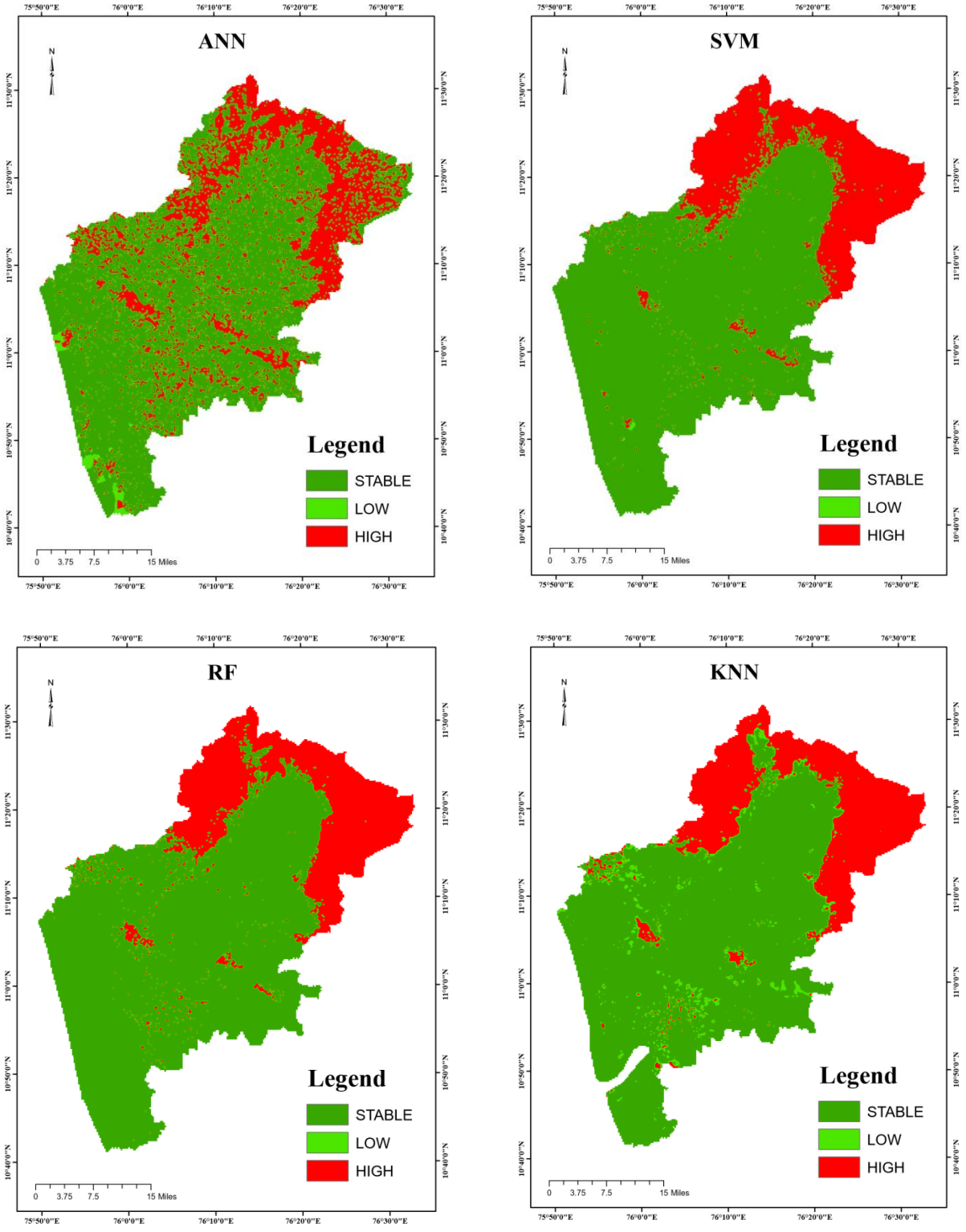


Figure 4.9: Landslide susceptibility map by (a) ANN, (b) SVM, (c) RF, and (d) KNN model

## 4.3. Validation of the models

Validation of machine learning models was done by Area Under Curve method (AUC), Receiver Operating Characteristics (ROC), specificity, sensitivity, Kappa Coefficient,Matthew's correlation coefficient (MCC), precision, accuracy, and F1-score.

ROC curve and AUC curve are one of the most used validation algorithms for machine learning models (Nguyen et al., 2020). ROC is a probability curve that represents the graph of true positive rate (TPR) or sensitivity against false positive rate (FPR) at different classification thresholds in a binary classification problem. TPR is the percentage of positive points correctly classified under positive class and FPR is the percentage of negative points incorrectly classified under positive class.

$$TPR = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4.2}$$

$$FPR = \frac{False\ Positive}{False\ Positive + True\ Negative} \tag{4.3}$$

AUC measures the area that comes under the ROC curve and is bounded by the x-axis. The value of AUC ranges from 0 to 1.At the AUC value is equal to 0, all the negative values will be grouped under the positive class and all the positive values will be grouped under the negative class. When the value of AUC is below 0.5, the model is unable to differentiate the positive and negative data points and it shows the model's low efficiency. High AUC indicates that the model has better capability to classify the data into correct respective classes. Thus combinedly, AUC and ROC are represented as AUROC and this is a significant index to measure the prediction ability of a model. The AUROC graph of different models for their corresponding training and the testing dataset was plotted in Figure 4.10. AUROC graph of the training dataset and testing dataset represents the success rate and prediction rate respectively. The success rate of the ANN model, SVM model, RF model, and KNN model is97.13%, 99.89%, 99.56%, and 99.65% respectively.

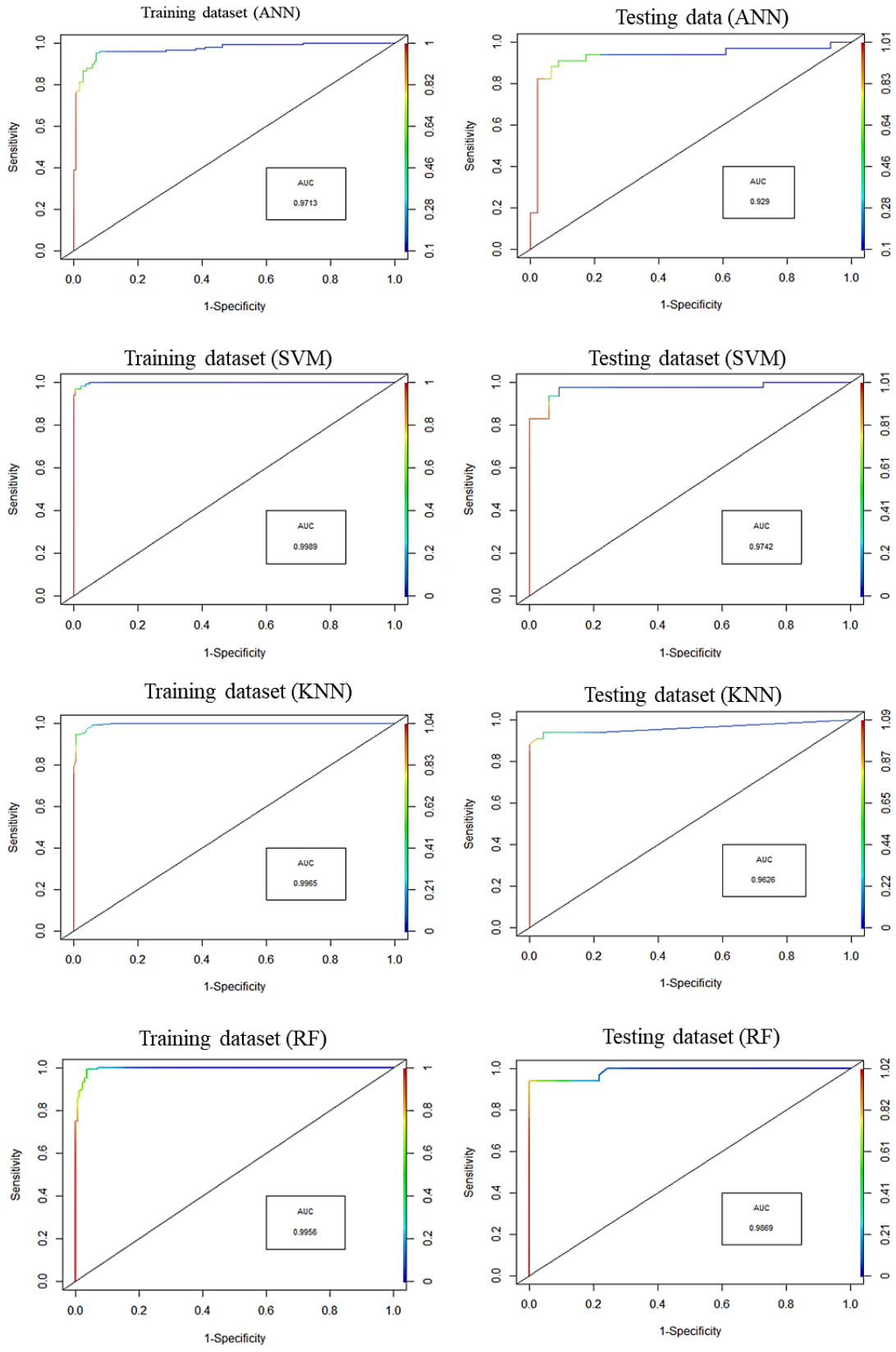


Figure 4.10: AUROC graph of training and testing data ofmachine learning models

This shows that the SVM model has the highest success rate during the training the model. Among all the models. Moreover, the prediction rate of the ANN model, SVM model, RF model, and KNN model is 92.2%, 97.42%, 98.69%, and96.26% respectively. The random forest has the highest prediction rate followed by the SVM model. Other statistical indices like sensitivity, specificity, precision, overall accuracy, F1 score, kappa coefficient, and Matthews coefficient were calculated from the confusion matrix elements (Table 4.3) such as true negative, true positive, false negative, and false positive. if a positive value is wrongly classified to negative class, it is termed as false negative (FN) and when it is classified correctly to positive class, it is called true positive (TP). A false positive (FP) label is given to a negative point when it is misclassified under a positive class and a true negative is a correctly classified negative point into a negative class.

Sensitivity represents the rate of correctly classified positive values out of total actual positive values present in a dataset. This is also known as recall or true positive rate (TPR). Specificity represents the rate of correctly classified negative values out of total negative values present in the initial dataset. Specificity is also termed a true negative rate (TNR). A higher value of sensitivity and specificity indicates the efficient classification of positive values and negative values into positive class and negative class respectively. The precision of a confusion matrix indicates, the rate of correct prediction of true positive values out of all the predicted positive values, and for this reason, precision is also called a positive prediction value (PPV). Overall accuracy gives a view of the rate of correct prediction of both positive and negative values out of the total dataset. F1 score is the harmonic mean of precision and recall. High precision and recall are achieved when the F1 score is near 1. Matthews correlation coefficient (MCC) depicts the rate of agreement between initial observed data and final predicted data. MCC ranges from -1 to +1 (perfect prediction). Unless the value of MCC is -1, 0, or +1, it is not a reliable indicator. A high kappa coefficient is good for prediction(Tang et al., 2015).

$$Specificity = \frac{TN}{TN + FP} \tag{4.4}$$

$$Precision = \frac{TP}{TP + FP} \quad (4.5)$$

$$F1\ Score = \frac{2 \times Precision \times Sensitivity}{Precision + Recall} \quad (4.6)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (4.7)$$

During the training of the model, the highest specificity (98.66%), precision (98.82%), and overall accuracy (98.44%)were found in the RF model, which is followed by the SVM model. The highest sensitivity was shown by the KNN model, followed by the RF model. F1 score (0.98), kappa coefficient (0.97), and Matthews coefficient (0.97) are the highest in the case of the RF model. The performance of the ANN model is very poor compared to the rest three models.

In the course of validation of test data, the highest sensitivity (97.14), precision (95.45), specificity (94.44), and accuracy (96.20) are found for the SVM model followed by the RF model. Moreover, the value of the F1 score, Kappa coefficient, and Matthews coefficient was the highest for the SVM model, which was 0.96, 0.92, and 0.92 respectively. An SVM model is followed by an RF model.

Table 4.3: statistical indices of training and testing dataset of different models

| Models | ANN | | SVM | | RF | | KNN | |
|---|---|---|---|---|---|---|---|---|
| Indices | Training | Testing | Training | Testing | Training | Testing | Training | Testing |
| Sensitivity | 5.71 | 4.44 | 97.22 | 97.14 | 98.00 | 88.89 | 99.281 | 93.939 |
| precision | 4.14 | 8.57 | 96.05 | 95.45 | 98.82 | 95.45 | 93.923 | 93.617 |
| specificity | 6.71 | 5.88 | 95.24 | 94.44 | 98.66 | 94.12 | 92.617 | 91.176 |
| accuracy | 5.00 | 6.25 | 96.57 | 96.20 | 98.44 | 92.50 | 96.250 | 93.750 |
| F1 score | 0.06 | 0.05 | 0.96 | 0.96 | 0.98 | 0.91 | 0.958 | 0.925 |
| Matthews coefficient | -0.90 | -0.87 | 0.93 | 0.92 | 0.97 | 0.85 | 0.926 | 0.872 |
| kappa coefficient | -0.89 | -0.84 | 0.93 | 0.92 | 0.97 | 0.85 | 0.924 | 0.872 |

# 5. SUMMARY AND CONCLUSION

Mudslides or debris flow is the most common type of landslide and the most destructive in nature. Unfortunately, it is a frequently occurring event along the Western Ghats of Kerala during the peak rainfall period of monsoon. With ever-changing demography and climatic factors, it is high time to carry out studies regarding early detection of landslide susceptible zone. The invention of warning systems, investigation of triggering agents along with mitigation plans will prevent or bring down the socio-economic damaging effect of landslides.

In the present study, an attempt has been made to delineate the landslide susceptible zones for landslide-stricken Malappuram district of Kerala by using both qualitative method and quantitative methods. As input, we have considered geomorphological characteristics such as slope, elevation, soil texture, geology, geomorphology, aspect, land use land cover (LULC), curvature, drainage density, lineament density, relative relief along with the annual average precipitation of the study area. Data for these characteristics were collected from Digital Elevation Map (DEM), Geological Survey of India (GSI) and India Meteorological Department (IMD).

Past landslide points were analysed and field surveys were done for examining the factors, which affect the initiation of landslides. Based on knowledge acquired from the survey, a landslide susceptibility map was derived by adopting the weighted overlay method (WoM) in the ArcGIS environment.

Landslides occurred in past (at 200 data points) and 200 non-landslide points were randomly divided into the ratio of 70:30 for training and testing purposes. Various statistical packages were used in R software V 4.1.2 to deduce, classify, and predict landslide susceptibility map of study are through machine learning models of ANN, SVM, RF, and KNN. Four landslide susceptibility maps were predicted by models, which were again validated by AUROC method and statistical indices such as sensitivity, specificity, precision, accuracy, F1 score, Matthews coefficient, and Kappa coefficient.

Based on the above analysis, which was conducted on the comparative study of landslide susceptibility mapping using geoinformatics and machine learning algorithms, the following conclusions can be drawn-

1. The raster layers of physical factors and final landslide maps were analysed and it was observed that areas having higher slope such as greater than 35° are more susceptible to landslide. This can be well explained by the higher rate of runoff flowing down the slope with a scouring effect during peak rainfall period.

2. High risk susceptibility zones were observed in the area having higher drainage density. This indicates the landslide of Malappuram is initiated due to higher pore water pressure in soil during monsoon.

3. While executing a field survey for LULC accuracy assessment, it was noticed that dense forest area with high slope is cleared to establish rubber plantation. There is an urgent need to protect these lands through soil conservation measures.

4. Geological survey of study area reveals that high-risk zone area is dominated by clay soil to clayey-loam soil, which again confirms that higher water holding capacity of clay decreases the shear strength of soil to bear external weight. This is one of the major causes of landslides.

5. Accuracy shown by training dataset under RF, SVM, KNN, and ANN is found to be 98.44%, 96.57%, 96.25%, and 5% respectively. Again, 96.20%, 93.75%, 92.50%, and 6.25% accuracy were shown under SVM, KNN, RF, and ANN model during testing respectively. The performance of the SVM model was better in comparison with other four machine learning models for predicting landslide susceptibility in study area.

6. SVM, RF, and KNN gives better accuracy compared to weighted overlay method (WoM) due to less human interference, while generating results through ground truthing .

7. The predicted maps can be utilized for the selection of safe zone for the purpose of developmental and constructive works by engineers and planners.

8. As landslide susceptibility mapping depends on the quality of data, size of data, availability of data as well as the skill of an expert to identify the environmental factors, there is a scope to enhance the quality of study with the combination of more datasets in future.

9. There is a need to create awareness in the region susceptible for higher landslides during a high precipitation period of a year. Therefore, this study can serve as an initiative to design mitigation measures by inspecting high-risk zones on a map, and develop an alert system to save lives and property.

# REFERENCES


Abraham, M. T., Satyam, N., Lokesh, R., Pradhan, B., & Alamri, A. (2021). Factors Affecting Landslide Susceptibility Mapping: Assessing the Influence of Different Machine Learning Approaches, Sampling Strategies and Data Splitting. In *Land* (Vol. 10, Issue 9). https://doi.org/10.3390/land10090989

Abraham, P., & E., S. (2013). Landslide hazard zonation in and around Thodupuzha-Idukki-Munnar road, Idukki district, Kerala: A geospatial approach. *Journal of the Geological Society of India*, *82*, 649–656. https://doi.org/10.1007/s12594-013-0203-7

Abramson, L. W., Macdonald, H. M., Parsons, T. S. L., Douglas, Q. &, Francisco, S., & Parsons, G. M. B. (2002). *SLOPE STABILITY AND STABILIZATION METHODS Second Edition*. 1–54.

Ajmal, P., & Saud S, J. (2022). *Estimating Landslide Risk Management Index for Nilambur Taluk, Kerala Using Analytic Hierarchy Process—A Case Study* (pp. 205–218). https://doi.org/10.1007/978-3-030-80312-4_17

Ali, S. A., Parvin, F., Pham, Q. B., Khedher, K. M., Dehbozorgi, M., Rabby, Y. W., Anh, D. T., & Nguyen, D. H. (2022). An ensemble random forest tree with SVM, ANN, NBT, and LMT for landslide susceptibility mapping in the Rangit River watershed, India. In *Natural Hazards* (Issue Sharpe 1938). Springer Netherlands. https://doi.org/10.1007/s11069-022-05360-5

Anbalagan, R. (1992). Landslide hazard evaluation and zonation mapping in mountainous terrain. *Engineering Geology*, *32*(4), 269–277. https://doi.org/10.1016/0013-7952(92)90053-2

Aneesah, R., Suresh, V. M., Varshini, S. V, & Nijamir, & K. (2019). *Zoning and Mapping of Landslide Hazard in the Nilgiris District, Tamil Nadu, India*. *November*, 978–955.

Azeze, A. (2021). *Landslide Inventory, Susceptibility, Hazard and Risk Mapping* (p.

32). https://doi.org/10.5772/intechopen.100504

Basharat, M., Shah, H. R., & Hameed, N. (2016). Landslide susceptibility mapping using GIS and weighted overlay method: a case study from NW Himalayas, Pakistan. *Arabian Journal of Geosciences*, *9*(4), 292. https://doi.org/10.1007/s12517-016-2308-y

Bhargavi, G., Arunnehru, J., Road, J. N., & Nadu, T. (2011). *Landslide vulnerability Assessment in Devikulam Taluk , Idukki District , Kerala using GIS and Machine Learning algorithms Department of Computer Science and Engineering , SRM Institute of Science and Technology , Vadapalani campus 1 Introduction 2 Study* . https://www.indiatoday.in/india/story/kerala-rains-all-5-gates-idukkidam-.

Biju, A. K., Mathew, J., George, A. M., Raj, D., & Thampy, A. (2021). *A Study in Correlation of Landslides and Increase in Areal Extent of Quarries in Malappuram District*. *9*(9), 25–30.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bucci, F., Santangelo, M., Cardinali, M., Fiorucci, F., & Guzzetti, F. (2016). Landslide distribution and size in response to Quaternary fault activity: the Peloritani Range, NE Sicily, Italy. *Earth Surface Processes and Landforms*, *41*(5), 711–720. https://doi.org/https://doi.org/10.1002/esp.3898

Bui, D., Pradhan, B., Löfman, O., Revhaug, I., & Dick, Ø. (2012). Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): A comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. *Catena*, *96*, 28–40. https://doi.org/10.1016/j.catena.2012.04.001

Carson, M. A. (Michael A. ., & Kirkby, M. J. (1972). *Hillslope form and process*. University Press.

Casagli, N., Catani, F., Puglisi, C., Delmonaco, G., Ermini, L., & Margottini, C. (2004). An Inventory-Based Approach to Landslide Susceptibility Assessment and its Application to the Virginio River Basin, Italy. *Environmental and*

*Engineering Geoscience*, *10*, 203–216. https://doi.org/10.2113/10.3.203

Catani, F., Lagomarsino, D., Segoni, S., & Tofani, V. (2013). Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. *Nat. Hazards Earth Syst. Sci.*, *13*(11), 2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Chen, W., Chen, Y., Tsangaratos, P., Ilia, I., & Wang, X. (2020). Combining Evolutionary Algorithms and Machine Learning Models in Landslide Susceptibility Assessments. *Remote Sensing*, *12*, 3854. https://doi.org/10.3390/rs12233854

Clerici, A., Perego, S., Tellini, C., & Paolo, V. (2002). A procedure for landslide susceptibility zonation by the conditional analysis method. *Geomorphology*, *48*, 349–364. https://doi.org/10.1016/S0169-555X(02)00079-X

Coe, J., Godt, J., Baum, R., Bucknam, R., & Michael, J. (2013). *Landslide susceptibility from Topography in Guatemala*.

Cruden, D. (1996). Cruden,D.M.,Varnes, D.J.,1996, Landslide Types and Processes, Transportation Research Board, U.S. National Academy of Sciences, Special Report, 247: 36-75. *Special Report - National Research Council, Transportation Research Board*, *247*, 36–57.

Cruden, D. M. (1991). A simple definition of a landslide. *Bulletin of the International Association of Engineering Geology - Bulletin de l'Association Internationale de Géologie de l'Ingénieur*, *43*(1), 27–29. https://doi.org/10.1007/BF02590167

Csáji, B. C. (2001). Approximation with artificial neural networks. *Faculty of Sciences, Etvs Lornd University, Hungary*, *24*(48), 7.

Dai, F. C., & Lee, C. F. (2001). Terrain-based mapping of landslide susceptibility using a geographical information system: A case study. *Canadian Geotechnical Journal*, *38*, 911–923. https://doi.org/10.1139/t01-021

Dai, F. C., & Lee, C. F. (2002). Landslide characteristics and slope instability

modeling using GIS, Lantau Island, Hong Kong. *Geomorphology*, *42*, 213–228. https://doi.org/10.1016/S0169-555X(01)00087-3

Djuric, U., Marjanović, M., Radić, Z., & Abolmasov, B. (2019). Machine learning based landslide assessment of the Belgrade metropolitan area: Pixel resolution effects and a cross-scaling concept. *Engineering Geology*, *256*. https://doi.org/10.1016/j.enggeo.2019.05.007

Ercanoglu, M. (2005). Landslide susceptibility assessment of SE Bartin (West Black Sea region, Turkey) by artificial neural networks. *Natural Hazards and Earth System Science*, *5*. https://doi.org/10.5194/nhess-5-979-2005

Erener, A., Mutlu, A., & Sebnem Düzgün, H. (2016). A comparative study for landslide susceptibility mapping using GIS-based multi-criteria decision analysis (MCDA), logistic regression (LR) and association rule mining (ARM). *Engineering Geology*, *203*, 45–55. https://doi.org/https://doi.org/10.1016/j.enggeo.2015.09.007

Fall, M., Azzam, R., & Noubactep, C. (2006). A multi-method approach to study the stability of natural slopes and landslide susceptibility mapping. *Engineering Geology*, *82*, 241–263. https://doi.org/10.1016/j.enggeo.2005.11.007

Feby, B., Achu, A. L., Jimnisha, K., Ayisha, V. A., & Reghunath, R. (2020). Landslide susceptibility modelling using integrated evidential belief function based logistic regression method: A study from Southern Western Ghats, India. *Remote Sensing Applications: Society and Environment*, *20*, 100411. https://doi.org/https://doi.org/10.1016/j.rsase.2020.100411

Galeandro, A., Doglioni, A., Simeone, V., & Šimůnek, J. (2014). Analysis of infiltration processes into fractured and swelling soils as triggering factors of landslides. *Environmental Earth Sciences*, *71*(6), 2911–2923. https://doi.org/10.1007/s12665-013-2666-7

Ghasemian, B., Shahabi, H., Shirzadi, A., Al-Ansari, N., Jaafari, A., Kress, V., Renoud, S., Ramadhan, A., & Geertsema, M. (2022). A Robust Deep-Learning

Model for Landslide Susceptibility Mapping. *Sensors*, *22*, 1–28. https://doi.org/10.3390/s22041573

Girma, F., Raghuvanshi, T. K., Ayenew, T., & Hailemariam, T. (2015). Landslide hazard zonation in Ada Berga district, Central ethiopia- A GIS based statistical approach. *Journal of Geomatics*, *9*(1), 25–38.

Goetz, J., Brenning, A., Petschko, H., & Leopold, P. (2015). Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. *Computers & Geosciences*. https://doi.org/10.1016/j.cageo.2015.04.007

Hariharan, S., V, S. A., & Reghunath, R. (2016). *Landslide Susceptibility Mapping : A Case Study in Thenmala Sub-Watershed ,. 2*, 50–53.

Hencher, S. R. (1987). The implications of joints and structures for slope stability. *Slope Stability. John Wiley*, 145–186.

Hobbs, W. H. (1904). Lineaments of the Atlantic border region. *Bulletin of the Geological Society of America*, *15*(1), 483–506.

Hussain Gardezi, S. A., Usmani, N., Chen, X., Ikram, N., Ahmad, S., & Wani, S. (2021). *Application of Data-Driven Techniques For Landslide Susceptibility Prediction Along An Earthquake Affected Road Section in Kashmir Himalaya*. https://doi.org/10.21203/rs.3.rs-1105939/v1

https://www.indiatoday.in/education-today/gk-current-affairs/story/landslides-india-283441-2015-07-20

Ibsen, M.-L., & Brunsden, D. (1996). The nature, use and problems of historical archives for the temporal occurrence of landslides, with specific reference to the south coast of Britain, Ventnor, Isle of Wight. *Geomorphology*, *15*. https://doi.org/10.1016/0169-555X(95)00073-E

Intrator, N. (2001). Interpreting Neural-Network Results : *Methods*, *37*, 1–17.

Jacinth Jennifer, J., & Saravanan, S. (2021). Artificial neural network and sensitivity

analysis in the landslide susceptibility mapping of Idukki district, India. *Geocarto International*, 1–23. https://doi.org/10.1080/10106049.2021.1923831

Jones, S., Kasthurba, A., Bhagyanathan, A., & BV, B. (2021). Impact of anthropogenic activities on landslide occurrences in southwest India: An investigation using spatial models. *Journal of Earth System Science*, *130*. https://doi.org/10.1007/s12040-021-01566-6

Kannan, M., Saranathan, E., & Anbalagan, R. (2011). Macro Landslide Hazard Zonation Mapping - Case Study from Bodi – Bodimettu Ghats Section, Theni District, Tamil Nadu - India. *Journal of the Indian Society of Remote Sensing*, *39*(4), 485–496. https://doi.org/10.1007/s12524-011-0112-4

Kanungo, D., Arora, M., Sarkar, S., & Gupta, R. P. (2006). A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. *Engineering Geology*, *85*, 347–366. https://doi.org/10.1016/j.enggeo.2006.03.004

Kavzoglu, T., Colkesen, I., & Sahin, E. (2019). Machine Learning Techniques in Landslide Susceptibility Mapping: A Survey and a Case Study. In *Advances in Natural and Technological Hazards Research* (pp. 283–301). https://doi.org/10.1007/978-3-319-77377-3_13

Kuriakose, S. L. (2010). Physically-based dynamic modelling of the effect of land use changes on shallow landslide initiation in the Western Ghats of Kerala, India. In *Faculty of Geo-information Science and Earth Observation (ITC),: Vol. PhD* (Issue TC dissertation number 178). https://dspace.library.uu.nl/handle/1874/188136

Li, J., Cheng, J., Shi, J., & Huang, F. (2012). *Brief Introduction of Back Propagation (BP) Neural Network Algorithm and Its Improvement BT - Advances in Computer Science and Information Engineering* (D. Jin & S. Lin (eds.); pp. 553–558). Springer Berlin Heidelberg.

Machireddy, S. R. (2020). *Geospatial Technologies in the Mapping of Landslide Hazard Zonation*. *11*(1), 1082–1093.

Mandal, S., & Mondal, S. (2019). *Artificial Neural Network (ANN) Model and Landslide Susceptibility BT - Statistical Approaches for Landslide Susceptibility Assessment and Prediction* (S. Mandal & S. Mondal (eds.); pp. 123–133). Springer International Publishing. https://doi.org/10.1007/978-3-319-93897-4_5

Meten, M., Bhandary, N., & Yatabe, R. (2015). Application of GIS-based fuzzy logic and rock engineering system (RES) approaches for landslide susceptibility mapping in Selelkula area of the Lower Jema River Gorge, Central Ethiopia. *Environmental Earth Sciences*, *74*. https://doi.org/10.1007/s12665-015-4377-8

Myronidis, D., Papageorgiou, C., & Theophanous, S. (2016). Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). *Natural Hazards*, *81*(1), 245–263. https://doi.org/10.1007/s11069-015-2075-1

Neaupane, K., & Achet, S. H. (2004). Use of backpropagation neural network for landslide monitoring: A case study in the higher Himalaya. *Engineering Geology*, *74*, 213–226. https://doi.org/10.1016/j.enggeo.2004.03.010

Nguyen, H.-D., Pham, V.-D., Nguyen, Q.-H., Pham, V.-M., Pham, M. H., Vu, V. M., & Bui, Q.-T. (2020). An optimal search for neural network parameters using the Salp swarm optimization algorithm: a landslide application. *Remote Sensing Letters*, *11*(4), 353–362. https://doi.org/10.1080/2150704X.2020.1716409

Nhu, V.-H., Shirzadi, A., Shahabi, H., Singh, S., Al-Ansari, N., Clague, J., Jaafari, A., Chen, W., Miraki, S., Dou, J., Luu, C., Górski, K., Pham, B., Nguyen, H., & Ahmad, B. Bin. (2020). Shallow Landslide Susceptibility Mapping: A Comparisons Between Logistic Model Tree, Logistic Regression, Naïve Bayes Tree, Artificial Neural Network, and Support Vector Machine Algorithms. *International Journal of Environmental Research and Public Health*.

Oguchi, T. (1997). Drainage Density and Relative Relief in Humid Steep Mountains with Frequent Slope Failure. *Earth Surface Processes and Landforms*, *22*, 107–

120. https://doi.org/10.1002/(SICI)1096-9837(199702)22:23.3.CO;2-L

Oommen, T., K S, S., & C L, V. (2019). SAR based assessment of 2018 Kerala Floods. *AGU Fall Meeting Abstracts*, *2019*, NH12A-05. https://ui.adsabs.harvard.edu/abs/2019AGUFMNH12A..05O

Panikkar, S. V, & Subramanyan, V. (1997). Landslide hazard analysis of the area around Dehra Dun and Mussoorie, Uttar Pradesh. *Current Science*, *73*(12), 1117–1123. http://www.jstor.org/stable/24101273

Pareek, N., Sharma, M., & Arora, M. (2010). Impact of seismic factors on landslide susceptibility zonation: A case study in part of Indian Himalayas. *Landslides*, *7*, 191–201. https://doi.org/10.1007/s10346-009-0192-1

Pham, B., & Prakash, I. (2017). Evaluation and Comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, Logistic Regression, and Support Vector Machines Methods for Landslide Susceptibility Mapping. *Geocarto International*, *34*. https://doi.org/10.1080/10106049.2017.1404141

Pham, B. T., Pradhan, B., Tien Bui, D., Prakash, I., & Dholakia, M. B. (2016). A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). *Environmental Modelling & Software*, *84*, 240–250. https://doi.org/https://doi.org/10.1016/j.envsoft.2016.07.005

Pourghasemi, H., Pradhan, B., & Gokceoglu, C. (2012). Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. *Natural Hazards*, *63*, 1–32. https://doi.org/10.1007/s11069-012-0217-2

POURGHASEMI, H. R., JIRANDEH, A. G., PRADHAN, B., XU, C., & GOKCEOGLU, C. (2013). Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. *Journal of Earth System Science*, *122*(2), 349–369. https://doi.org/10.1007/s12040-013-0282-2

Pourghasemi, H. R., Mohammady, M., & Pradhan, B. (2012). Landslide susceptibility

mapping using index of entropy and conditional probability models in GIS: Safarood Basin, Iran. *Catena*, *97*(Complete), 71–84. https://doi.org/10.1016/j.catena.2012.05.005

Pradhan, B., & Lee, S. (2010). Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling and Software*, 747–759. https://doi.org/10.1016/j.envsoft.2009.10.016

Prasad, P., Loveson, V. J., Das, S., & Chandra, P. (2021). Artificial intelligence approaches for spatial prediction of landslides in mountainous regions of western India. *Environmental Earth Sciences*, *80*(21), 720. https://doi.org/10.1007/s12665-021-10033-w

R S, A., Loghin, A.-M., P G, V., K Jacob, M., & Krishnamurthy, R. (2016). Landslide Susceptible Zone Mapping Using ARS and GIS Techniques in Selected Taluks of Kottayam District, Kerala, India. *International Journal of Applied Remote Sensing and GIS*, *3*, 16–25.

Raghuvanshi, T. K., Ibrahim, J., & 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. https://doi.org/https://doi.org/10.1016/j.jafrearsci.2014.05.004

Raghuvanshi, T., Negassa, L., & Kala, P. (2015). GIS based Grid overlay method versus modeling approach – A comparative study for landslide hazard zonation (LHZ) in Meta Robi District of West Showa Zone in Ethiopia. *The Egyptian Journal of Remote Sensing and Space Science*, *18*. https://doi.org/10.1016/j.ejrs.2015.08.001

Ramachandra, T. V, Member, S., Kumar, U., Member, S., Aithal, B. H., Member, S., Diwakar, P. G., & Joshi, N. V. (n.d.). *Theme3_36*. 380–389.

Rana, H., & Babu, G. L. S. (2022). Object-Oriented Approach for Landslide Mapping Using Wavelet Transform Coupled with Machine Learning: A Case Study of

Western Ghats, India. *Indian Geotechnical Journal*, *52*(3), 691–706. https://doi.org/10.1007/s40098-021-00587-8

Ray, P. K. C., Dimri, S., Lakhera, R. C., & Sati, S. (2007). Fuzzy-based method for landslide hazard assessment in active seismic zone of Himalaya. *Landslides*, *4*(2), 101–111. https://doi.org/10.1007/s10346-006-0068-6

Roering, J. (2012). Tectonic geomorphology: Landslides limit mountain relief. *Nature Geoscience*, *5*, 446–447. https://doi.org/10.1038/ngeo1511

Ruff, M., & Czurda, K. (2008). Landslide susceptibility analysis with a heuristic approach in the Eastern Alps (Vorarlberg, Austria). *Geomorphology*, *94*, 314–324. https://doi.org/10.1016/j.geomorph.2006.10.032

Sajinkumar, K. S., Anbazhagan, S., Rani, V., & Muraleedharan, C. (2013). A paradigm quantitative approach for a regional risk assessment and management in a few landslide prone hamlets along the windward slope of Western Ghats, India. *International Journal of Disaster Risk Reduction*, *7*. https://doi.org/10.1016/j.ijdrr.2013.10.004

Sarkar, S., & Kanungo, D. (2004). An Integrated Approach for Landslide Susceptibility Mapping Using Remote Sensing and GIS. *Photogrammetric Engineering and Remote Sensing*, *70*, 617–625. https://doi.org/10.14358/PERS.70.5.617

Savith, S., A.M, S., & .P, V. (2021). Post landslide Investigation of Shallow Landslide: A case study from the Southern Western Ghats, India. *Disaster Advances*, *14*, 52. https://doi.org/10.25303/147da5221

Schmidt, K., & Montgomery, D. (1995). Limits to Relief. *Science*, *270*, 617–620. https://doi.org/10.1126/science.270.5236.617

Shahabi, H., Hashim, M., & Ahmad, B. Bin. (2015). Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin, Iran. *Environmental Earth Sciences*, *73*(12), 8647–8668. https://doi.org/10.1007/s12665-015-4028-0

Skilodimou, H., Bathrellos, G., Koskeridou, E., Soukis, K., & Rozos, D. (2018). Physical and Anthropogenic Factors Related to Landslide Activity in the Northern Peloponnese, Greece. *Land*, *7*, 85. https://doi.org/10.3390/land7030085

Sreekumar, S., & Aslam, A. (2017). Geospatial Approach for Landslide Disaster Management: A Case Study from India. *International Journal of Applied and Advanced Scientific Research (IJAASR) Impact*, *5*(2), 120–127. www.dvpublication.com

Tang, W., Hu, J., Zhang, H., Wu, P., & He, H. (2015). Kappa coefficient: a popular measure of rater agreement. *Shanghai Archives of Psychiatry*, *27*(1), 62–67. https://doi.org/10.11919/j.issn.1002-0829.215010

Touret, J., & Huizenga, J. (2012). Charnockite microstructures: From magmatic to metamorphic. *Geoscience Frontiers*, *3*, 745–753. https://doi.org/10.1016/j.gsf.2012.05.004

Turrini, M., & Visintainer, P. (1998). Proposal of a method to define areas of landslide hazard and application to an area of the Dolomites, Italy. *Engineering Geology*, *50*, 255–265. https://doi.org/10.1016/S0013-7952(98)00022-2

Vapnik, V. (2000). *The Nature of Statistical Learning Theory*. Springer: New York.

Varnes, D. (1978). Slope Movement Types and Processes. *Special Report*, *176*, 11–33.

Wadhawan, S. K., Singh, B., & Ramesh, M. V. (2020). Causative factors of landslides 2019: case study in Malappuram and Wayanad districts of Kerala, India. *Landslides*, *17*(11), 2689–2697. https://doi.org/10.1007/s10346-020-01520-5

Wang, Q., Li, W., Xing, M., Wu, Y., Pei, Y., Yang, D., & Bai, H. (2016). Landslide susceptibility mapping at Gongliu county, China using artificial neural network and weight of evidence models. *Geosciences Journal*, *20*. https://doi.org/10.1007/s12303-016-0003-3

Whalley, B., Cooke, R., & Doornkamp, J. (1991). Geomorphology in Environmental

Management. A New Introduction. *Transactions of the Institute of British Geographers*, *16*, 492. https://doi.org/10.2307/623035

Youssef, A. M., & Pourghasemi, H. R. (2021). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. *Geoscience Frontiers*, *12*(2), 639–655. https://doi.org/https://doi.org/10.1016/j.gsf.2020.05.010

Zhao, S., & Zhao, Z. (2021). A Comparative Study of Landslide Susceptibility Mapping Using SVM and PSO-SVM Models Based on Grid and Slope Units. *Mathematical Problems in Engineering*, *2021*, 1–15. https://doi.org/10.1155/2021/8854606

****