Landslide Susceptibility Mapping Using GIS-based Information Value and Frequency Ratio Methods in Gindeberet area, West Shewa Zone, Oromia Region, Ethiopia

Alembante Genene
Addis Ababa Science and Technology University
Matebie Meten (matebe21@gmail.com)
Addis Ababa Science and Technology University https://orcid.org/0000-0002-1529-190X

Research

Keywords: GIS, Frequency Ratio, Information Value, Landslide Susceptibility, Gindeberet, Ethiopia

Posted Date: February 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-219331/v1

Landslide Susceptibility Mapping Using GIS-based Information Value and Frequency Ratio Methods in Gindeberet area, West Shewa Zone, Oromia Region, Ethiopia

Alembante Genene (MSc) [1*] (First author)

[1] Department of Geology, College of Applied Sciences, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia; P.O.Box: 16417
Telephone: +251-967231973; E-mail: alembantegenene5064@gmail.com

Matebie Meten (PhD) [1, 2] (Corresponding author)

[1] Department of Geology, College of Applied Sciences,
Addis Ababa Science and Technology University, Ethiopia
P.O.Box: 16417; Addis Ababa Ethiopia
Telephone: +251-911899279;
E-mail: matebe21@gmail.com

[2] Mineral Exploration, Extraction and Processing Center of Excellence;
Addis Ababa Science and Technology University, Ethiopia;
P.O.Box: 16417; Addis Ababa Ethiopia

Abstract

The study area is found in Gindeberet district of West Shewa zone in Oromia Regional State of Ethiopia.This area is highly susceptible to active surface processes due to the presence of rugged morphology with steep scarps, sharp ridges, cliffs, deep gorges and valleys. This study aimed to identify and evaluate the causative factors and to prepare the landslide susceptibility maps (LSMs) of the study area. Two bivariate statistical models i.e. Information value(IV) and the Frequency ratio(FR), were used. First, active, reactivated and passive landslides and scarps were identified using Google Earth image interpretation and extensive field survey for landslide inventory. A total of 580 landslide were randomly selected into two datasets in which (80%)460 landslides were used for modeling and (20%)116 landslidesfor validation. conditioning factors (slope, aspect, curvature, distance from stream, distance from lineaments, lithology, rainfall and land use) were combined with a training landslide dataset in a ArcGIS to generate LSMs which weredivided into verylow, low, moderate, high and veryhigh susceptibility zones. LSMs for IV and FR models were validated using the Area under(ROC) curve showing a success rate of 0.836 and 0.835 respectively and a predictive rate of 0.817 and 0.818 respectively wich showed a good performance of both models. The resulting LSMs can be used for land use planning and management.

Keywords*:GIS,Frequency Ratio,Information Value, Landslide Susceptibility,Gindeberet,Ethiopia*

1. Introduction

Surface of the earth is always in a dynamic change as a result of different mass wasting processes such as landslide and these changes are more common in mountainous terrains. Landslide is defined as a mass movement of rock, debris or earth down a slope resulting in a geomorphic change of the Earth's surface (Glade 1997).It can be triggered by various natural external stimuli such as intense and/or prolonged rainfall, volcanic eruption, earthquake, snow melts and also human actions that affect the drainage or groundwater condition.

Globally, every year landslide disasters cost lives of thousands and billions of USD in property damage. These phenomena cause lots of damage affecting people, organizations, industries and the environment (Glade 1997). Ethiopia is one of the mountainous countries which are characterized by a steep slope, deep gorge; frequent fault escarpment, highly weathered rock, and intensive and prolonged rainfall. All those conditions are favorable for the occurrence of slope movements after rainy seasons in most parts of the country (Woldearegay 2013). According to many studies conducted in Ethiopia (Woldearegay 2008; Ibrahim 2011 cited in Meten *et al*. 2015), landslides in Ethiopia had resulted in the loss of human and animal lives, damages in infrastructures and properties in the last five decades.

The current study area is found in Gindeberet District, West Shewa Zone, Oromia Region in Central Ethiopia. This area is highly susceptible to landslide problems with the occurrence of many active landslides in the last few decades which extensively damaged gravel roads, houses and farmlands in which the magnitude of the problem is alarming and the vulnerability of the people's lives and property from landslides needs special attention and so far the area has not been studied before in a greater level of detail.

The development of more advanced qualitative and quantitative methods and GIS data-processing techniques have allowed conducting numerous studies effectively and as a result, landslide susceptibility, hazards and risk assessment have attracted attentions of many researchers (VanWesten et al. 2006). According to Aleotti and Chowdhury (1999), there are two general approaches that can be used to study landslide susceptibility. These are the field-based qualitative- and data-driven quantitative approaches. Field-based, qualitative (heuristic) approach directly depends on the researcher's expertise because all of the decision rules to prepare the landslide susceptibility maps are evaluated by the researcher. Statistical methods are used to estimate the relative contributions of the factors responsible for slope instability and to make some predictions

based on these factors (Suzen 2002). Accordingly, the data analysis in statistical approaches can be grouped into two i.e. bivariate and multivariate methods (Soeters and vanWesten 1996). The aim of this paper is to generate the landslide susceptibility maps of the study area using GIS-based information value and frequency ratio models.

2. Study Area

The study area is found in Gindeberet District which is located in West Shewa Zone of Oromia Region in Central Ethiopia (Figure 1). The area is located 182Km northwest of Addis Ababa and geographically bounded between 37° 40′ to 37°57′E longitudes and 9°30′ to 9°40′N latitudes. This area can be classified into two main physiographic regions of plateau area and rugged terrain with an elevation ranging from 1386 to 2600m a.s.l. Due to the presence of sharp ridges, cliffs, rugged slope faces, steep scarps, deep gorges and valleys, the area is highly susceptible to active surface processes. The drainage pattern of the area is dendritic type in which the river channel follows the slope of the terrain. As observed in physiographic map of the study area in Figure 2, the slopes of the study area are facing towards the north and south directions.The study area is divided into three agro-ecological zones of the highland, midland and lowland. The wet season starts on May and extends up to October and the peak wettest season occurs during July and August. This area receives rainfall twice in a year with a heavy precipitation from June to September and light to moderate precipitation from mid-March to mid-April. In July and August, the peak rainfall is in between 300 and 400 mm per month respectively in which more than 50% of the annual precipitation is accumulated during these months. The study area is moderately vegetated and most parts of the ridges are covered by eucalyptus trees, wild grass and scattered bushes. Most of the gentle slopes and flat land in the study area are intensively used for agricultural purpose and grazing land for livestock production. The most dominant agricultural crops are “teff”, sorghum, barley, wheat, maize, pulses and small scale production of vegetables.

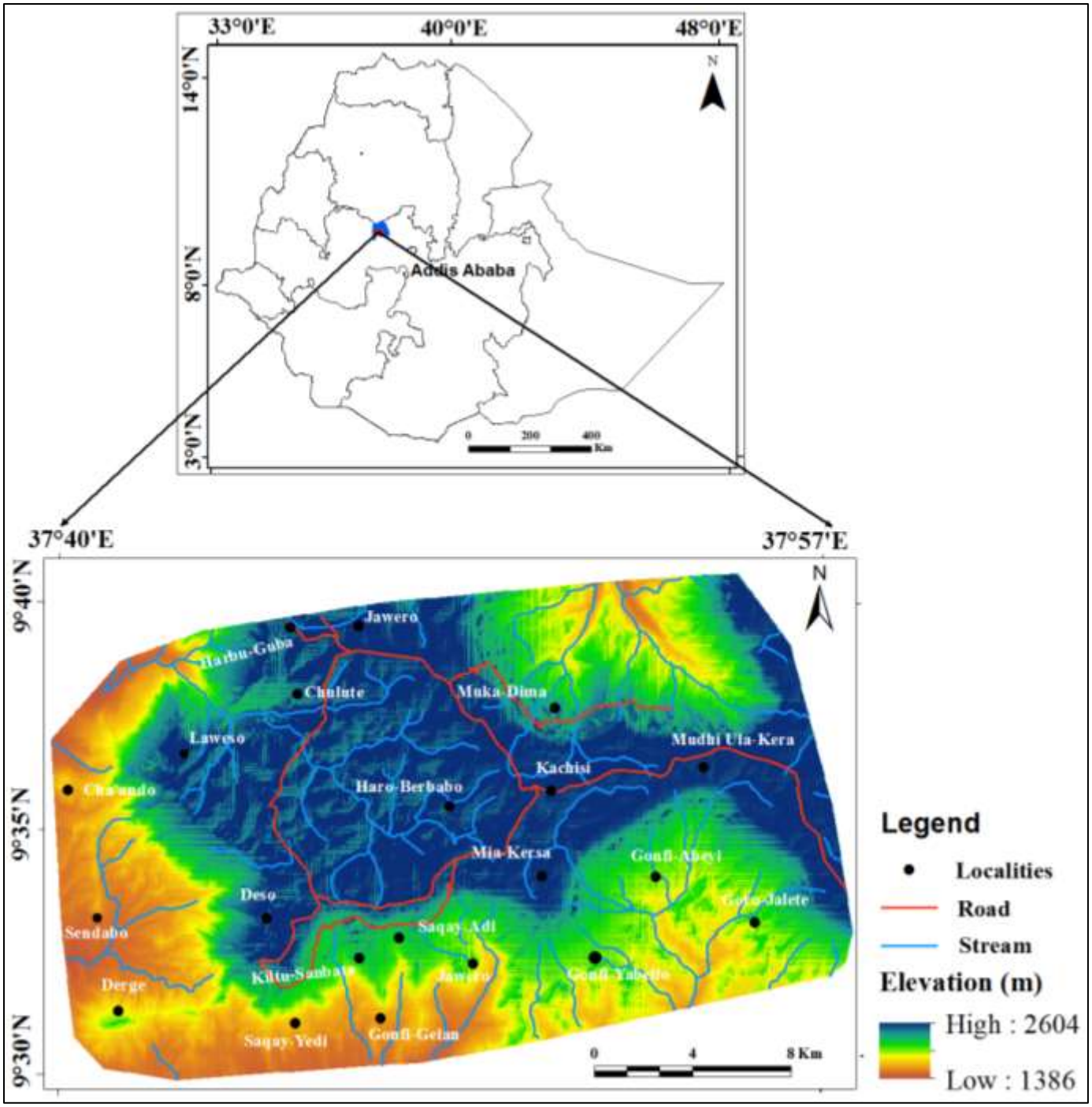


Fig. 1 Location Map of the Study area

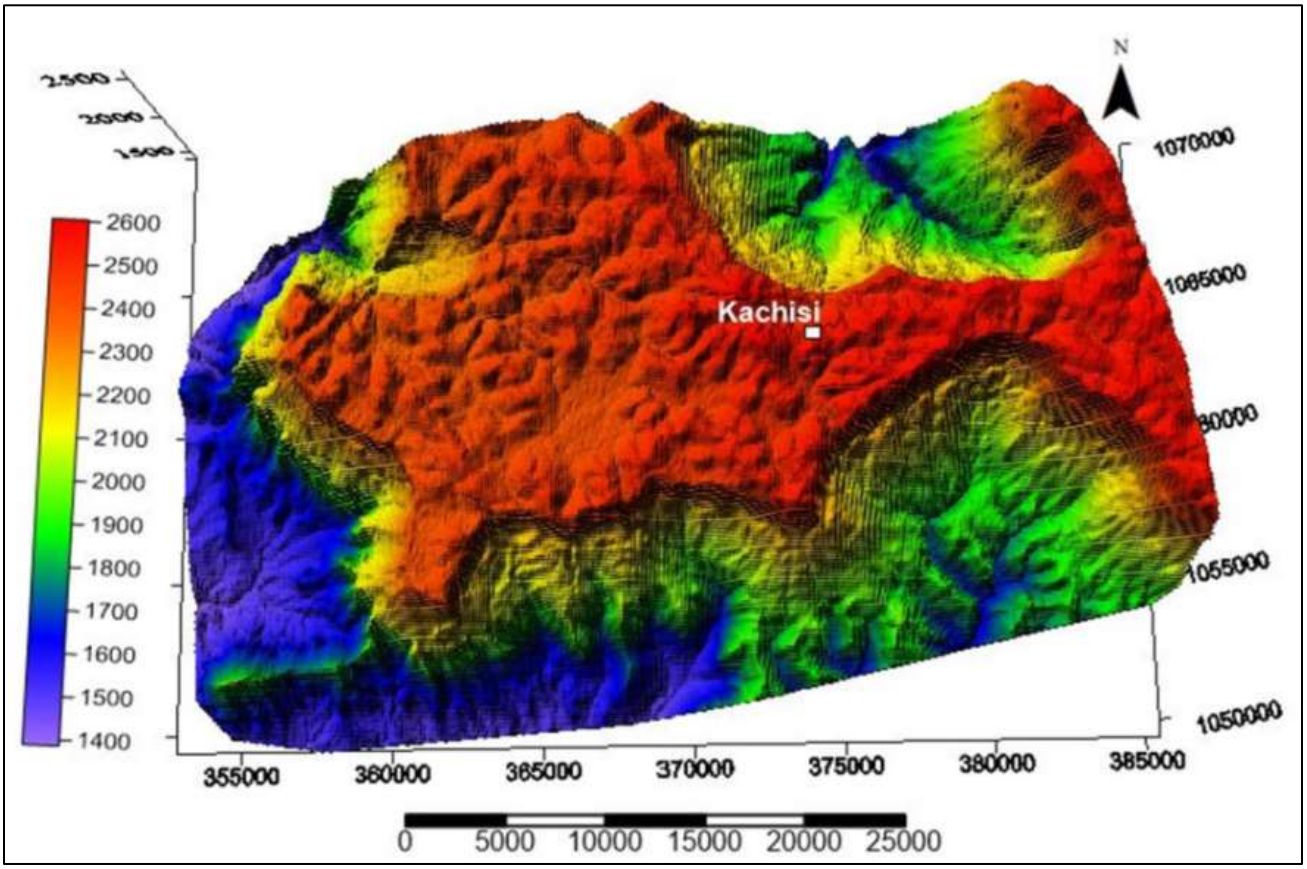


Fig. 2 Physiography of the study area

3. Data used and Methodology

In order to achieve the objective of this research work four main stages were followed including data acquisition and organization, database construction and analysis, preparing the landslide susceptibility maps and finally verification and comparison of the landslide susceptibility maps.

3.1. Data Acquisitionand Organization

The data acquisition part of this research work was based on primary and the available secondary data. Secondary data includes both published and unpublished papers. Topographic map, regional geological map and meteorological data were collected from the Ethiopian Geospatial Information Agency, Geological Survey of Ethiopia and National Metrology Agency of Ethiopia respectively. DEM of the study area was used to generate topographic parameters such as slope, aspect and curvature. Google Earth images were used to identify the land cover and geomorphologic features of the study area. After a comprehensive and thorough literature review and preliminary site

investigation, a detailed field work was carried out to collect the primary data such as description of different types of lithological units and identifying their relative degree of weathering, visually inspection of slope steepness, collection of available spring location, mapping both active and landslide scar locations and identification of materials involved, failure mechanisms, state of activity (active, dormant etc.) and measuring of their size and shape and identifying land use and human activities and their role for landslide occurrence in the study area.

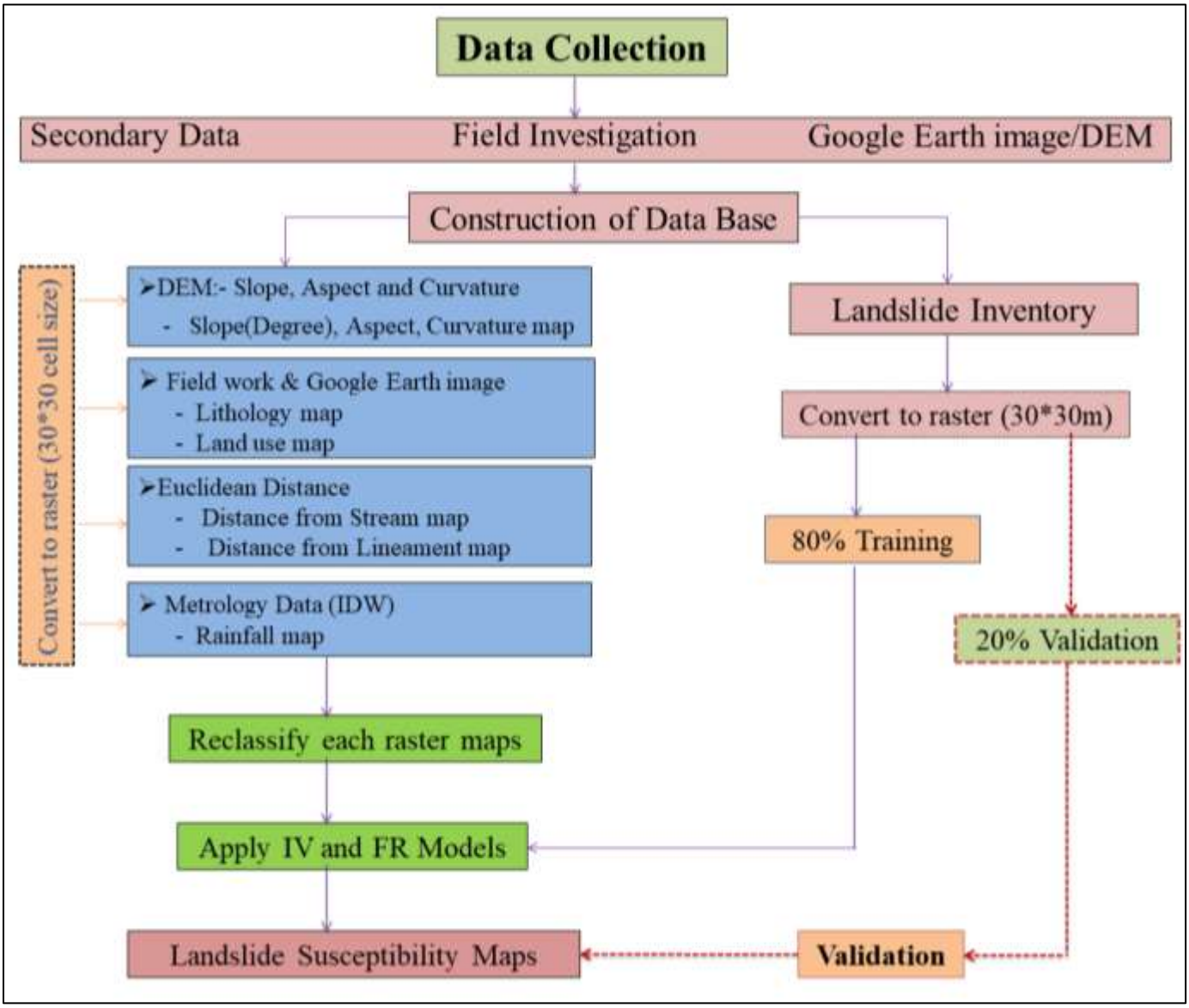


Fig. 3 Flowchart of the research work

3.2. Landslide inventory map

The landslide inventory map (Figure 4) was produced through intensive field survey and Google Earth image interpretation with time series data. Extensive field studies conducted in mid-November to mid-December 2019 was used to map known landslides using GPS locations and check the size and shape of the landslides to identify the type of movements and the materials involved and to determine the landslides state of activity (active, reactivated, dormant etc.). In this study, a landslide inventory with a total of 580 (24356 pixels) landslides were identified and mapped as vector-based polygons then converted to the raster format with a pixel size of 30m by 30m in ArcGIS software and they were randomly subdivided into two data sets i.e. 80%(20336 pixels) or (464 landslides) used for the susceptibility model building and 20%(4020 pixels) or (116 landslides) used for model validation (Figure 4). From a total 580 landslides, 98 landslides were identified during the field work and the remaining landslide polygons were identified and collected from Google Earth image. Landslide distributions in the study area are dominantly found in northern and southern parts of the study area which are characterized by rugged and steep slopes, deep gorges, highly fractured and weathered rocks. The inventory is composed of debris slide, rotational and translational slide, progressive creep movement, rock slides and rock fall and complex types of landslide. All landslide events are represented by polygon features and these landslides affect a total area of 21.88Km^2 and the maximum area of the landslide polygon was 0.7576Km^2 and the minimum area was 0.000389Km^2. In this study, the landslide classification system developed by Varnes (1978) was used.

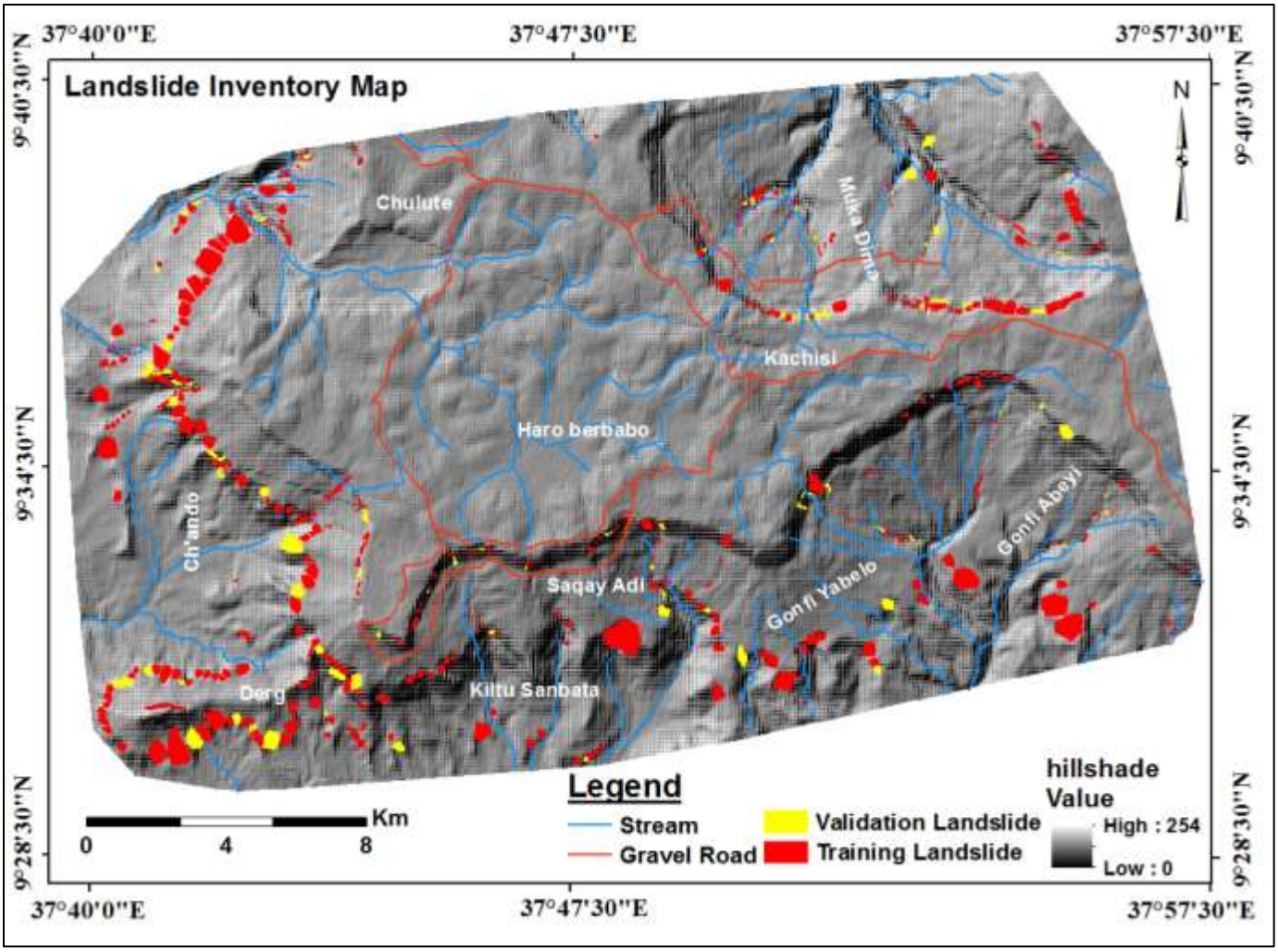


Fig. 4 Landslide inventory map

3.3. Landslide Causative Factors

The selected factors for landslide susceptibility assessment in a GIS-based study must be operational, represented over the entire area, variable, fundamental and measurable (Ayalew and Yamagishi 2005). Based on the above criteria eight predisposing factors were selected including slope, aspect, curvature, lithology, landuse/landcover, rainfall, distance from rivers and distance from lineaments. These eight landslide causative factors were prepared in a raster format using the same spatial projection (WGS 1984 UTM zone 37N) and cell size of 30×30 m using ArcGIS 10.4 which is used to evaluate the spatial relationship between them and the landslides in the study area. Topographic parameters such as slope, aspect and curvature maps were extracted from digital elevation model(DEM). The lithological map was prepared from data collected in the field and from

existing geological map while the landuse/landcove map was generated from Google Earth image interpretation with field survey and verification. The lineament map was prepared from 3D Google Earth image interpretation with field survey and the distance from lineament was prepared through GIS based buffering analysis. Distance from stream map was extracted from drainage map of the study area and constructed through Euclidean distance buffering. The rainfall map was prepared by interpolating the 30-year average rainfall data of the five nearest rain gage stations of the study area using IDW interpolation in the spatial analyst tool in order to get spatially distributed rainfall raster map.

3.3.1. Aspect

Aspect is defined as the slope direction, it have an essential influence on landslide occurrence as it controls the exposure of slope to sunlight, cold and hot winds and rainfall(Huang et al 2015). The aspect map of the study area is derived from DEM using ArcGIS software and classified into nine classesi.e. Flat, North,Northeast,East, Southeast, South, Southwest, West and Northwest (Figure 5a).

3.3.2. Slope

As slope angle increases gravity's parallel component increases and the perpendicular component decreases, so slope failure occurs more frequently on the steeper slope (Huang et al 2015). The slope of the study area was derived from DEM using GIS software. Slope of the study area range between 0° - 69.97° and it was classified into 8 classes of 0° – 5°, 5° – 10°, 10° – 15°, 15° – 20°, 20° – 25°, 25° – 35°, 35° – 45°, 45° – 69.97° (Figure 5b).

3.3.3. Curvature

Curvature is defined as the rate of change of slope gradient in a particular direction and controls the hydraulic condition and gravity. Curvature may refer to the convex, concave and flatness of a slope. According to Pradhan (2010), Alkhasawneh et al. (2013) as cited in Meten et al. (2015) the negative value refers valley, positive value refers hill slope and zero/ near zero values refer to flat land area. Curvature of the study area was derived from DEM by ArcGIS software and classified in to three classes of convex, concave and flat (Figure 5c).

3.3.4. Distance from Stream

The probability of landslide occurrence increases as distance to the stream decreases (Cellek 2019). In this study, a drainage network was extracted from stream network shape file and then the distance to drainage was generated by Euclidean distance in ARCGIS 10.4. Finally, the distance to drainage was reclassified into fourteen classes: 0 – 50, 50 – 100, 100 – 150, 150 – 200, 200 – 300, 300 – 400, 400 – 500, 500 – 900, 900 -1400, 1400 – 1800, 1800 – 2000, and > 2000 meter (Figure 5d).

3.3.5. Distance from Lineament

The potential of landslide increases as the distance from lineaments decreases, (Pradhan and Lee 2007). Lineaments of the study area were prepared from Google Earth image interpretations and field survey. Then the distance to lineament was generated by Euclidean distance in ARCGIS 10.4 software which was reclassified into ten classes of 0 – 100, 100 – 200, 200 – 300, 300 – 400, 400 – 500, 500 – 800, 600 – 1200, 1200 – 1700, 1700 – 2200 and >2200meter (Figure 5e).

3.3.6. Lithology

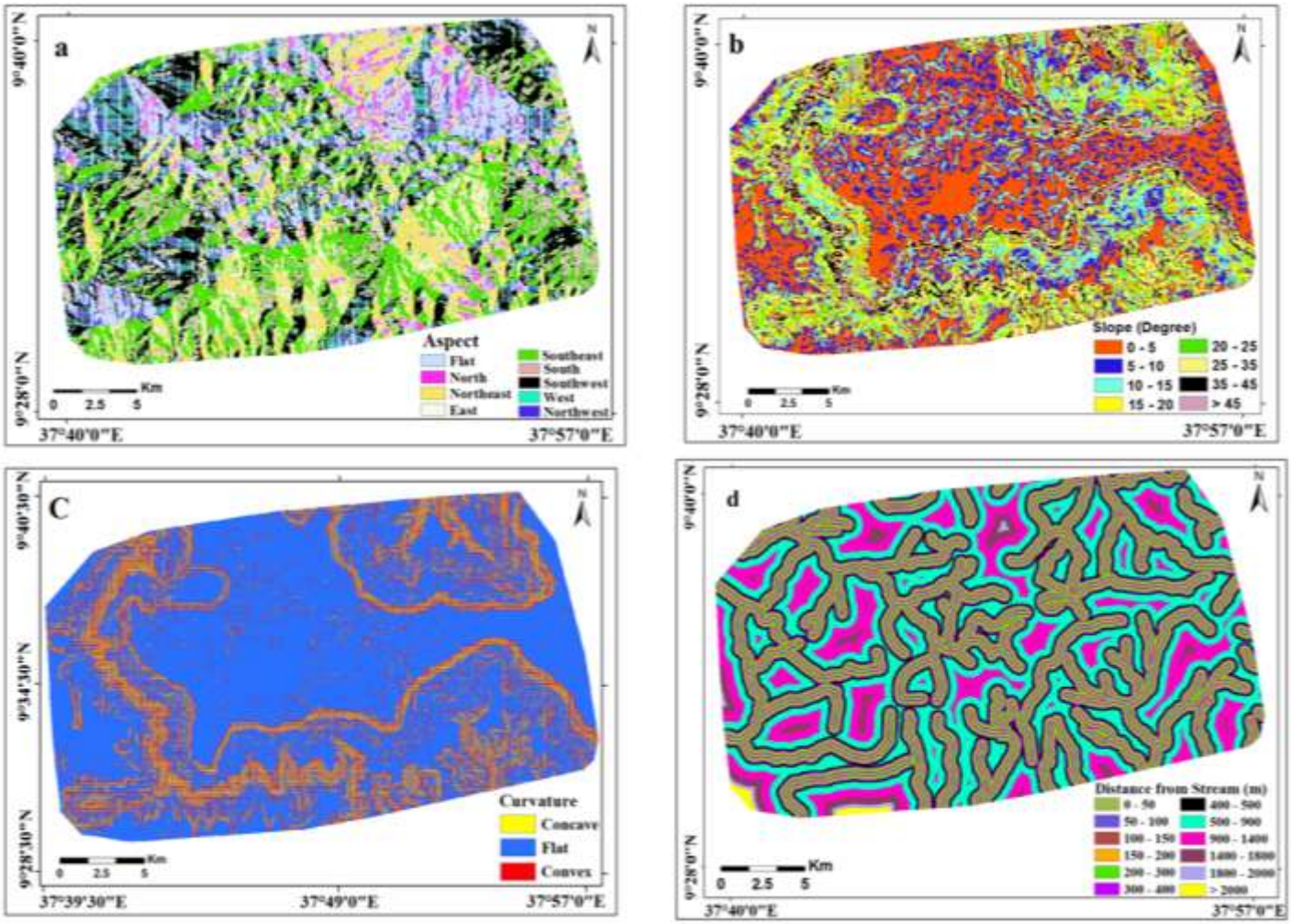
Lithological variations often lead to a difference in strength and permeability of rocks and have a significant role in landslides occurrence (Sarkar et al. 2013). In the current study, lithological map of the study area (Figure 5f) was prepared from existing regional geological map and detail field survey. The study area contains six lithological units namely Quaternary Superficial sediment, alluvial deposits, Residual soil, Basalt, Limestone and Sandstone.

3.3.7. Land use

Land cover is also one of the key factors responsible for the occurrence of landslides, since, barren slopes are more prone to landslides (Gomes et al 2005). The land use map of the study area (Figure 5g) was prepared from Google Earth image interpretations and the analysis of this factor with landslide was done using ArcGIS tools. Seven land-use types were identified including dense forest, moderate forest, sparse forest, shrubs, bare land, agricultural land and settlement.

3.3.8. Rainfall

Rainfall plays an important role in reducing the shear strength and increasing pore pressure (Yalcin, 2007). The available data from five rain gage stations surrounding the study area was obtained from Ethiopian National Meteorological Agency(Kachisi, Yejube, Alge, Hareto and Fincha). These rain gage stations were interpolated using IDW interpolation technique that can be used to compute the unknown spatial rainfall data from the known data of sites that are adjacent to the unknown site (Chen and Liu, 2012) in ArcGIS and were classified in to five classes i.e., 1621 – 1645, 1645 – 1670, 1670 – 1695, 1695 – 1719 and 1719 – 1744mm/year (Figure 5h).



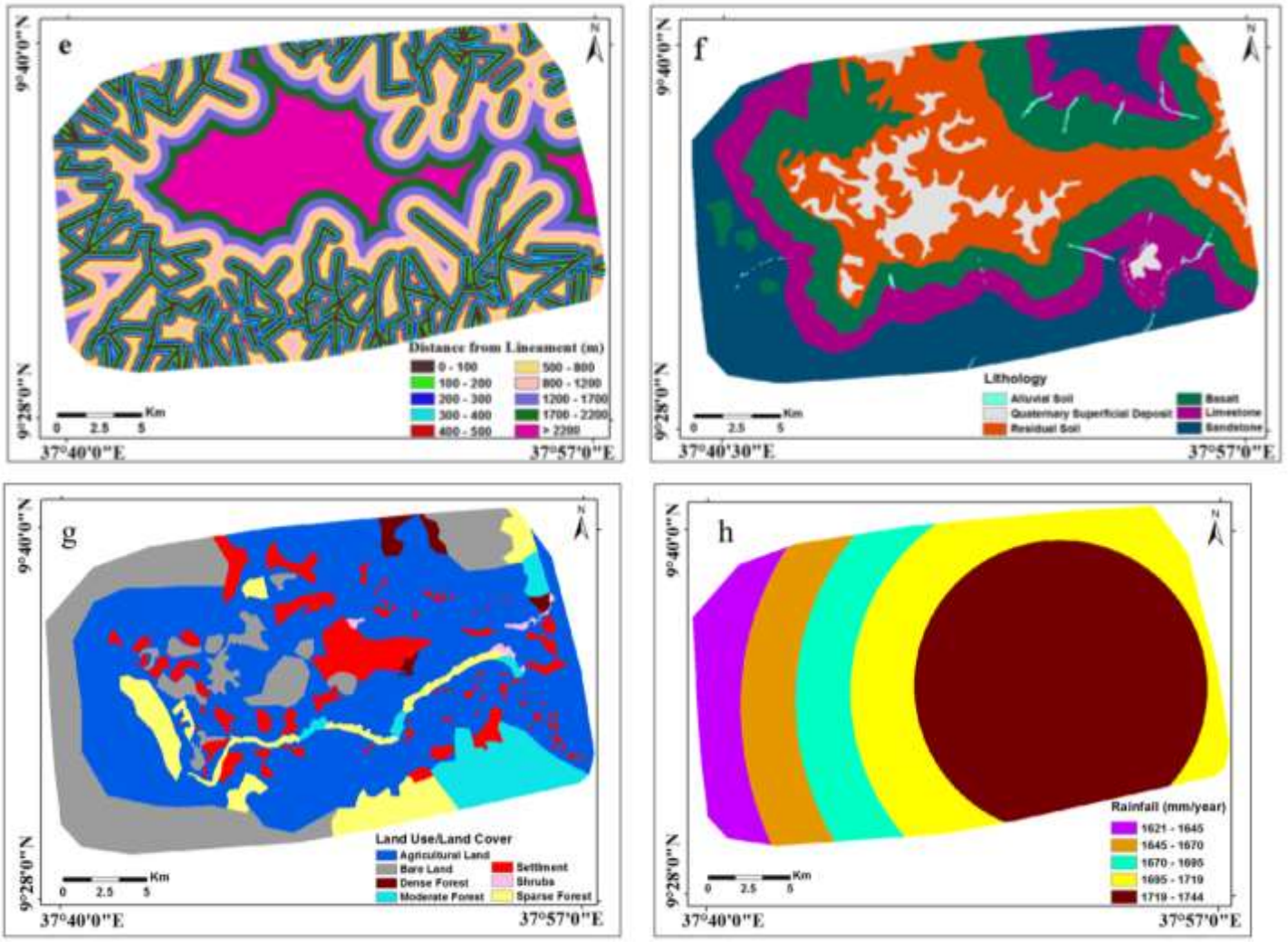


Fig. 5 Landslide causative factors used in this study a) Aspect, b) Slope, c) Curvature, d) Distanc from Stream, e) Distance from Lineament, f) Lithology, g) Land Use, h) Rainfall

3.4. Information Value Method

The information value(IV) model is a bivariate statistical analysis method that was developed from information theory and developed by Yin and Yan (1988) and a little bit modified by (VanWesten 1993). The objective of the information value model was to find the combination of significant factors by determining the probability of a landslide event based on the comprehensive information available. This method is important to determine the degree of influence of individual causative factors responsible for landslide occurrence (Kanungo et al. 2009). Information value for each factor class can be calculated using:

$$IV = \log(Conditional\ probability/prior\ probability) \quad (1)$$

Where a conditional probability is the ratio of landslide pixels in a class to the class pixels and prior probability is the ratio of the total number of landslide pixels to the total number of study area pixels. In a practical sense, the information value of each factor class is calculated as:

$$I(H, X_i) = \log \frac{Npix(Si)/Npix(Ni)}{\sum Npix\,(Si)/\sum Npix\,(Ni)} \quad (2)$$

whereas I(H, Xi) is the information value of a factor class; Npix(Si) is the number of pixels of a landslide within class i; Npix(Ni) is the number of pixels within class i; Σ Npix(Si) is the number of pixels of a landslide within the entire study area; Σ Npix(Ni) is the number of pixels within the entire study area. Therefore, the landslide susceptibility index (LSI) for each pixel was computed by summing the information values of each factor class as follows:

$$LSI = \sum_{i=0}^{n} I(H, Xi) = \sum_{i=0}^{n} (\ln) \frac{Npix(Si)/Npix(Ni)}{\sum Npix(Si)/\sum Npix(Ni)} \quad (3)$$

When LSI < 0, the likelihood of a landslide is less than average; when LSI = 0, the likelihood of a landslide is average; when LSI > 0, the likelihood of landslide is greater than average. This means that the greater the information value, the greater the possibility of a landslide.

3.5. Frequency Ratio Method

The frequency ratio approach is based on the observed relationships between the distribution of landslides and each landslide related factor to correlate between the landslide locations and the landslide factors in the study area (Lee and Pradhan 2007). To calculate the frequency ratio, the area ratio of landslide occurrence and non-occurrence was calculated for different classes or types of each factor after which an area ratio for the class or type of each factor of the total area was calculated.

$$\text{Frequency ratio (FR)} = \frac{\text{Slide ratio}}{\text{Class ratio}} \quad (4)$$

$$\text{Slide ration} = \frac{\text{Number of Landslide pixel in class}}{\text{Total number of Landslide pixels}} \quad (5)$$

$$\text{Class ratio} = \frac{\text{Number of pixels in individual class}}{\text{Total number of pixels in whole class}} \quad (6)$$

$$\text{Frequency Ratio} = \frac{\frac{\text{Landslide pixels in each class}}{\text{Landslide pixel in the whole area}}}{\frac{\text{Area pixels in each class}}{\text{Area pixels in the whole area}}} \quad (7)$$

The FR value represents the degree of the correlation between landslide and the concerned class of the conditioning factors. So a value of 1 means an average value. If the value is > 1, there is a high correlation and FR < 1 means a lower correlation.Then, the landslide susceptibility index was calculated using $\text{LSI} = \sum FR = \text{FR}_1 + \text{FR}_2 + \text{FR}_3 + .. + \text{FR}_8$. The higher the LSI value, the greater will be the risk (Huang et al. 2015).

4. Result and Discussion

4.1. Relationship between Landslide Occurrence and Causative Factors

To evaluate the contribution of each factor towards landslide susceptibility in this study, the landslide distribution data layer has been compared to various thematic data layers separately. The number of landslide pixels falling on each class of the thematic data layers has been recorded and weights have been calculated on the basis of both Information value and Frequency ratio methods (Table 1 & 2).

For both models, (slope class > 45º), (aspect classes of northwest, west, southwest, south and southeast facing slopes), (curvature classes of concave and convex), (distance from stream classes of < 300 m), (distance from lineament classes of < 800m), (lithology classes of alluvial soil, basalt, limestone and sandstone), (rainfall classes of 1621 – 1645, and 1645 – 1670 mm/year) and (land use/land cover classes of sparse forest, moderate forest and bare land) showed a higher probability of landslide occurrence.

Table 1 Rating of each factor class from a spatial relationship between each factor class and landslide using information value model

Data Layers	Class	Npix(Ni)	% Npix(Ni)	Npix(Si)	%Npix(Si)	Con Prob(X)	Prior Prob(Y)	(X/Y)	IV
Slope(degree)	0 - 5	216161	35.42	3680	18.10	0.017	0.033	0.51	-0.29
	5 - 10	138995	22.78	1257	6.18	0.009	0.033	0.27	-0.57
	10 - 15	96285	15.78	1767	8.69	0.018	0.033	0.55	-0.26
	15 - 20	57043	9.35	2108	10.37	0.037	0.033	1.11	0.04
	20 - 25	33386	5.47	2170	10.67	0.065	0.033	1.95	0.29
	25 - 35	37079	6.08	3953	19.44	0.107	0.033	3.20	0.51
	35 - 45	19043	3.12	2697	13.26	0.142	0.033	4.25	0.63
	> 45	12267	2.01	2704	13.30	0.220	0.033	6.61	0.82
	Flat	149049	24.42	4860	23.90	0.033	0.033	0.98	-0.01
Aspect	N	24130	3.95	389	1.91	0.016	0.033	0.48	-0.32
	NE	86591	14.19	2117	10.41	0.024	0.033	0.73	-0.13
	E	30771	5.04	947	4.66	0.031	0.033	0.92	-0.03
	SE	113819	18.65	3956	19.45	0.035	0.033	1.04	0.02
	S	44034	7.22	1560	7.67	0.035	0.033	1.06	0.03
	SW	113661	18.63	4539	22.32	0.040	0.033	1.20	0.08
	W	24932	4.09	1108	5.45	0.044	0.033	1.33	0.13
	NW	23272	3.81	860	4.23	0.037	0.033	1.11	0.04
Curvature	Concave	55844	9.15	5444	26.77	0.097	0.033	2.93	0.47
	Flat	463404	75.94	8243	40.53	0.018	0.033	0.53	-0.27
	Convex	91011	14.91	6649	32.70	0.073	0.033	2.19	0.34
Distance from Stream(m)	0 - 50	41163	6.75	2628	12.92	0.064	0.033	1.92	0.28
	50 - 100	38932	6.38	2512	12.35	0.065	0.033	1.94	0.29
	100 - 150	39192	6.42	2377	11.69	0.061	0.033	1.82	0.26
	150 - 200	28757	4.71	1548	7.61	0.054	0.033	1.62	0.21
	200 - 300	72724	11.92	3222	15.84	0.044	0.033	1.33	0.12
	300 - 400	60345	9.89	1973	9.70	0.033	0.033	0.98	-0.01
	400 - 500	61492	10.08	1548	7.61	0.025	0.033	0.76	-0.12
	500 - 900	170262	27.90	2828	13.91	0.017	0.033	0.50	-0.30
	900 - 1400	79100	12.96	1256	6.18	0.016	0.033	0.48	-0.32
	1400 - 1800	13956	2.29	268	1.32	0.019	0.033	0.58	-0.24
	1800 - 2000	2063	0.34	66	0.32	0.032	0.033	0.96	-0.02
	> 2000	2273	0.37	110	0.54	0.048	0.033	1.45	0.16

Distance from Lineament(m)	0 - 100	68305	11.19	2361	11.61	0.035	0.033	1.04	0.02
	100 - 200	58332	9.56	2618	12.87	0.045	0.033	1.35	0.13
	200 - 300	61697	10.11	3237	15.92	0.052	0.033	1.57	0.20
	300 - 400	47562	7.79	2955	14.53	0.062	0.033	1.86	0.27
	400 - 500	44865	7.35	2422	11.91	0.054	0.033	1.62	0.21
	500 - 800	88972	14.58	3550	17.46	0.040	0.033	1.20	0.08
	800 - 1200	67273	11.02	1664	8.18	0.025	0.033	0.74	-0.13
	1200 - 1700	53148	8.71	1094	5.38	0.021	0.033	0.62	-0.21
	1700 - 2200	40920	6.71	435	2.14	0.011	0.033	0.32	-0.50
	> 2200	79185	12.98	0	0.00	0.000	0.033	0.00	0.00
Lithology	QSD	45960	7.53	306	1.50	0.007	0.033	0.20	-0.70
	Basalt	152777	25.03	5860	28.82	0.038	0.033	1.15	0.06
	Residual Soil	152864	25.05	372	1.83	0.002	0.033	0.07	-1.14
	Alluvial Soil	4088	0.67	218	1.07	0.053	0.033	1.60	0.20
	Limestone	106874	17.51	6497	31.95	0.061	0.033	1.82	0.26
	Sandstone	147695	24.20	7083	34.83	0.048	0.033	1.44	0.16
Rainfall(mm)	1621 - 1645	55457	9.09	4315	21.22	0.078	0.033	2.33	0.37
	1645 - 1670	70735	11.59	4513	22.19	0.064	0.033	1.91	0.28
	1670 - 1695	75689	12.40	2179	10.71	0.029	0.033	0.86	-0.06
	1695 - 1719	143296	23.48	1453	7.14	0.010	0.033	0.30	-0.52
	1719 - 1744	265082	43.44	7876	38.73	0.030	0.033	0.89	-0.05
Land use	Bare land	129304	21.19	6257.00	30.77	0.048	0.033	1.45	0.16
	Shrubs	2146	0.35	32.00	0.16	0.015	0.033	0.45	-0.35
	Dense Forest	6839	1.12	0.00	0.00	0.000	0.033	0.00	0.00
	Settlement	54390	8.91	39.00	0.19	0.001	0.033	0.02	-1.67
	Moderate Forest	35616	5.84	2385.00	11.73	0.067	0.033	2.01	0.30
	Sparse Forest	44078	7.22	3104.00	15.26	0.070	0.033	2.11	0.32
	Agricultural Land	337887	55.37	8519.00	41.89	0.025	0.033	0.76	-0.12

Where; Npix(Si) is the number of pixels of a landslide within classes i; Npix(Ni) is the number of pixels within classes, QSD is quaternary superficial sediments.

Table 2 Rating of each factor class from a spatial relationship between each factor class and landslide using frequency ratio model

Data Layers	Class	Class Pixels	% class pixels	Landslide Pixels	% Landslide pixels	Slide Ratio(a)	Class Ratio(b)	FR =a/b
Slope(degree)	0 to 5	216161	35.42	3680	18.10	0.181	0.354	0.511
	5 to 10	138995	22.78	1257	6.18	0.062	0.228	0.271
	10 to 15	96285	15.78	1767	8.69	0.087	0.158	0.551
	15 to 20	57043	9.35	2108	10.37	0.104	0.093	1.109
	20 to 25	33386	5.47	2170	10.67	0.107	0.055	1.950
	25 to 35	37079	6.08	3953	19.44	0.194	0.061	3.199
	35 to 45	19043	3.12	2697	13.26	0.133	0.031	4.250
	> 45	12267	2.01	2704	13.30	0.133	0.020	6.615
Aspect	Flat	149049	24.42	4860	23.90	0.239	0.244	0.978
	N	24130	3.95	389	1.91	0.019	0.040	0.484
	NE	86591	14.19	2117	10.41	0.104	0.142	0.734
	E	30771	5.04	947	4.66	0.047	0.050	0.924
	SE	113819	18.65	3956	19.45	0.195	0.187	1.043
	S	44034	7.22	1560	7.67	0.077	0.072	1.063
	SW	113661	18.63	4539	22.32	0.223	0.186	1.198
	W	24932	4.09	1108	5.45	0.054	0.041	1.334
	NW	23272	3.81	860	4.23	0.042	0.038	1.109
Curvature	Concave	55844	9.15	5444	26.77	0.268	0.092	2.925
	Flat	463404	75.94	8243	40.53	0.405	0.759	0.534
	Convex	91011	14.91	6649	32.70	0.327	0.149	2.192
Distance from Stream(m)	0 - 50	41163	6.75	2628	12.92	0.129	0.067	1.916
	50 - 100	38932	6.38	2512	12.35	0.124	0.064	1.936
	100 - 150	39192	6.42	2377	11.69	0.117	0.064	1.820
	150 - 200	28757	4.71	1548	7.61	0.076	0.047	1.615
	200 - 250	72724	11.92	3222	15.84	0.158	0.119	1.330
	250 - 300	60345	9.89	1973	9.70	0.097	0.099	0.981
	300 - 500	61492	10.08	1548	7.61	0.076	0.101	0.755
	500 - 900	170262	27.90	2828	13.91	0.139	0.279	0.498
	900 - 1400	79100	12.96	1256	6.18	0.062	0.130	0.476
	1400 - 1800	13956	2.29	268	1.32	0.013	0.023	0.576

Distance from Lineament(m)	1800 - 2000	2063	0.34	66	0.32	0.003	0.003	0.960
	>2000	2273	0.37	110	0.54	0.005	0.004	1.452
	1 - 100	68305	11.19	2361	11.61	0.116	0.112	1.037
	100 - 200	58332	9.56	2618	12.87	0.129	0.096	1.347
	200 - 300	61697	10.11	3237	15.92	0.159	0.101	1.574
	300 - 400	47562	7.79	2955	14.53	0.145	0.078	1.864
	400 - 500	44865	7.35	2422	11.91	0.119	0.074	1.620
	500 - 800	88972	14.58	3550	17.46	0.175	0.146	1.197
	800 - 1200	67273	11.02	1664	8.18	0.082	0.110	0.742
	1200 - 1700	53148	8.71	1094	5.38	0.054	0.087	0.618
	1700 - 2200	40920	6.71	435	2.14	0.021	0.067	0.319
	> 2200	79185	12.98	0	0.00	0.000	0.130	0.000
Lithology	QSD	45960	7.53	306	1.50	0.015	0.075	0.200
	Basalt	152777	25.03	5860	28.82	0.288	0.250	1.151
	Residual Soil	152864	25.05	372	1.83	0.018	0.250	0.073
	Alluvial Soil	4088	0.67	218	1.07	0.011	0.007	1.600
	Limestone	106874	17.51	6497	31.95	0.319	0.175	1.824
	Sandstone	147695	24.20	7083	34.83	0.348	0.242	1.439
Rainfall(mm)	1621 - 1645	55457	9.09	4315	21.22	0.212	0.091	2.335
	1645 - 1670	70735	11.59	4513	22.19	0.222	0.116	1.915
	1670 - 1695	75689	12.40	2179	10.71	0.107	0.124	0.864
	1695 - 1719	143296	23.48	1453	7.14	0.071	0.235	0.304
	1719 - 1744	265082	43.44	7876	38.73	0.387	0.434	0.892
Land use	Bare land	129304	21.19	6257.00	30.77	0.308	0.212	1.452
	Shrubs	2146	0.35	32.00	0.16	0.002	0.004	0.447
	Dense Forest	6839	1.12	0.00	0.00	0.000	0.011	0.000
	Settlement	54390	8.91	39.00	0.19	0.002	0.089	0.022
	Moderate Forest	35616	5.84	2385.00	11.73	0.117	0.058	2.010
	Sparse Forest	44078	7.22	3104.00	15.26	0.153	0.072	2.113
	Agricultural Land	337887	55.37	8519.00	41.89	0.419	0.554	0.757

4.2. Landslide susceptibility mapping

4.2.1. Information Value Method

The calculated information value for each parameter classes is converted into raster map in ArcGIS, then the landslide susceptibility index was prepared by summing the information values of all parameters corresponding to each pixel in the map using raster calculator.

$$LSI = IV_{Slo} + IV_{Asp} + IV_{Curv} + IV_{Dis\text{-}Str} + IV_{Dis\text{-}Lin} + IV_{Litho} + IV_{Rain} + IV_{LU} \quad (8)$$

Where IV_{Slo}, IV_{Asp}, IV_{Curv}, $IV_{Dis\text{-}Str}$, $IV_{Dis-Lin}$, IV_{Litho}, IV_{Rain} and IV_{LU} are the information values for slope, aspect, curvature, distance from stream, distance from lineament, lithology, rainfall and landuse respectively.

when LSI < 0, the likelihood of a landslide is less than average; when LSI = 0, the likelihood of a landslide is equal to average; when LSI > 0 is greater than average with a greater information value which means a higher probability of landslide occurrence.

For IV model, the final LSI values of the study area range from -5.29 to 2.67 which was classified into five classes of verylow (-5.29 to -3.02), low (-3.02 to -1.62), moderate (-1.62 to -0.5), high (-0.5 to 0.53) and very high (0.53 to 2.67) using the natural breaks method. LSI, landslide percentage, landslide density and area coverage of landslide susceptibility class were shown in Table 3.

4.2.2. Frequency Ratio Method

The fundamental concept of this method is to calculate the ratio between the density of the phenomena in a given class and the density of the same class (Lee and Talib 2005). In the present study with the help of ArcGIS 10.4, the landslide factors were converted into raster maps with a pixel size of 30*30m. The spatial relationship between landslide locations and each landslide factor was analyzed and the number of landslide pixels in each class has been evaluated and the frequency ratio for each factor class is calculated. Then the Frequency Ratio ratings of factors in the form of raster maps were summed in ArcGIS using raster calculator to prepare the landslide susceptibility index (LSI) as follows.

$$LSI = \sum FR = FR_1 + FR_2 + FR_3 + \cdots + FR_8 \quad (9)$$

A higher LSI means a higher susceptibility to landslide while a lower one indicates a lower susceptibility (Bui et al. 2012). LSI values range from 2.19 to 20.01 in the FR model. Using the reclassify function, the LSI map was reclassified into five classes of very low (2.19 – 5.74), low (5.74 – 7.97), moderate (7.97 – 10.33), high (10.33 – 13.26) and very high (13.26 – 20.01) by using the natural breaks method. LSI, landslide percentage, landslide density and area coverage of landslide susceptibility class were shown in Table 4.

Table 3 LSI, landslide percent, and landslide density of IV

Landslide Susceptibility Class	LSI	LSP	LSP(%)	TLSP	TLSP [a] (%)	LSM Area(Km^2)	LSM Area [b] (%)
VeryLow	-5.29 - - 3.02	43734	7.17	23	0.11	39	7.1
Low	-3.02 - -1.62	136995	22.45	81	0.4	123	22.4
Moderate	-1.62 - -0.5	145615	23.86	1339	6.58	131	23.9
High	-0.5 - 0.53	179599	29.43	6499	31.96	162	29.5
VeryHigh	0.53 - 2.67	104315	17.09	12394	60.95	94	17.1
Total		610258		20336		549	

Table 4 Landslide susceptibility classes index and landslide percent, of FR

Landslide Susceptibility Class	LSI	LSP	LSP(%)	TLSP	TLSP[a] (%)	LSM Area (Km^2)	LSM Area [b] (%)
VeryLow	2.19 – 5.74	159321	26.11	55	0.27	143	26.1
Low	5.74 – 7.97	169780	27.11	1496	7.36	153	27.8
Moderate	7.97 – 10.33	162592	26.64	5606	27.57	146	26.7
High	10.33 –13.26	77772	12.74	5418	26.95	70	12.7
Very High	13.26 –20.01	40793	6.68	7698	37.85	37	6.7
Total		610258		20336		549	

LSP = Landslide susceptibility pixels and TPLSP = Training landslide pixels

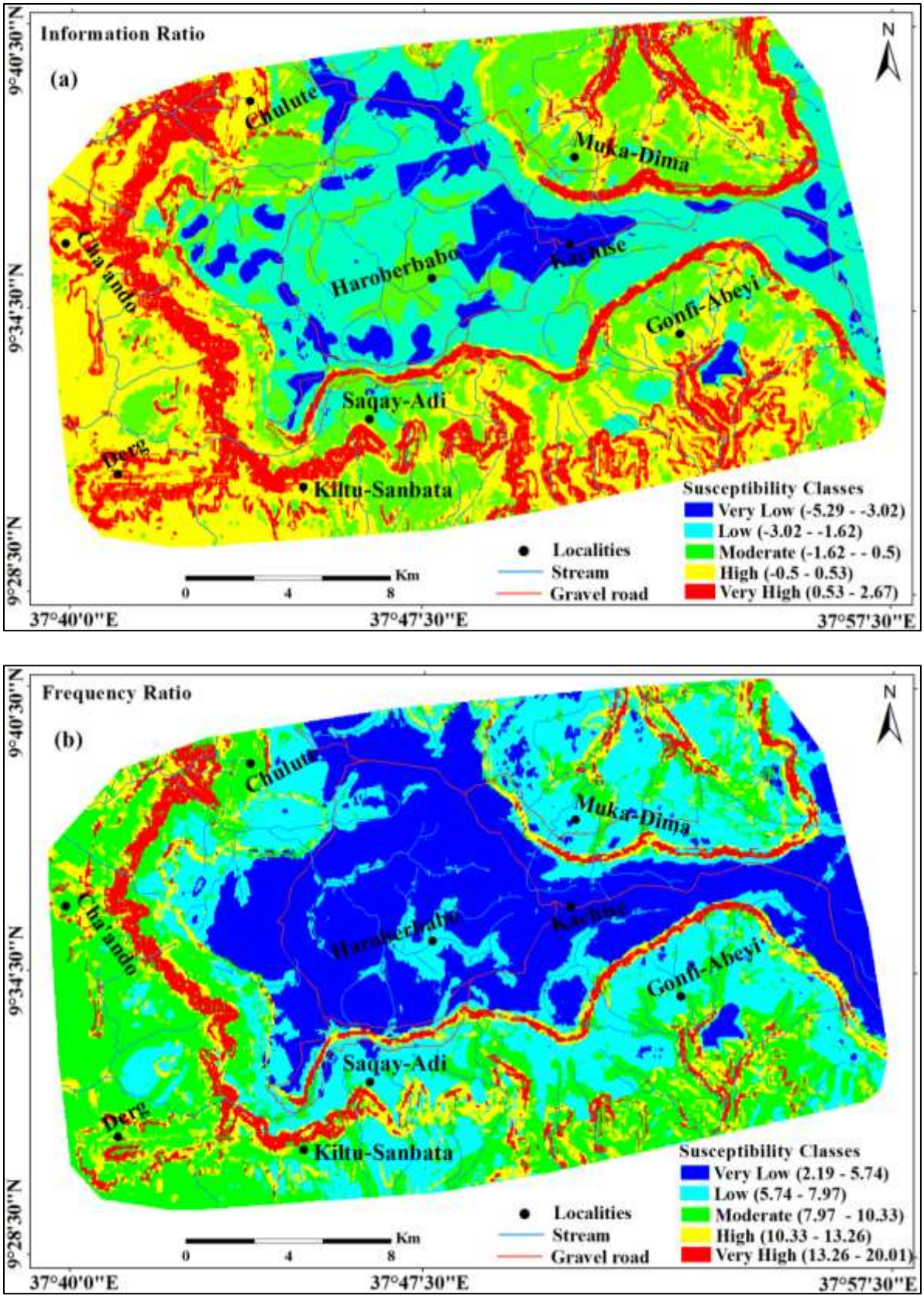


Fig. 6 Landslide Susceptibility Map of the study area using

4.3. Verification of the susceptibility maps

4.3.1. Area under the Curve (AUC)

The area under the curve(AUC) is the measure that indicates the accuracy of the landslide susceptibility maps by creating success and prediction rate curves. The resulting area under the curve indicates the probability that more pixels were correctly labeled than incorrectly labeled (Ghorbanzadehet al. 2018). Therefore, the greater AUC values indicate a higher accuracy of the resulting susceptibility map. The success rate assesses how many landslides sites, which are used in the model, are successfully captured by the susceptibility map and consequently represents a measure of model efficiency (Neuhäuser 2012).The predictive rate calculates the percentage of the independent landslides captured with the susceptibility map. Therefore, it can be assessed how many "unknown" landslides could be "predicted" (Neuhäuser 2012). In the ROC method, the area under the curve(AUC) values, ranging from 0.5 to 1.0, are used to evaluate the accuracy of the model. AUC close to 1 suggest a higher model reliability while near or less than 0.5 suggest that the model is invalid (Chung and Fabbri 2003).

In this study, to obtain the success and predictive rate, the calculated susceptibility index values were sorted in descending order and classified into 100 classes and combined with training and validating landslide raster. Success rate was obtained by comparing the 80% (training landslide pixels) with the landslide susceptibility maps of both models and the predictive rate was obtained by comparing the 20% (validation landslide pixels) which were not included in the models producing the landslide susceptibility maps of both models. The results showed that the AUC of the success and predictive rate curves for the IV model are 0.836 and 0.817while for FR model, it was 0.835and 0.818 respectively (Figure 6). The AUC of the success and predictive rate curves range between 0.8 and 0.9 indicating a good performance of both models.

4.3.2. Landslide Density Index (LDI)

In this study, the LDI value was obtained by combining the training landslides and validation landslide with landslide susceptibility and calculated using equation 10. The higher LDI on the high and very high landslide susceptibility regions further confirmed that the model is reliable and accurate (Fazye et al. 2018). As a result the LDI value for both models increased from very low to very high susceptibility classes as presented in Table 5.

$$\text{LDI}=\frac{Percent\ of\ observed\ Landslide}{Percent\ of\ Predicted\ Landslide} \quad (10)$$

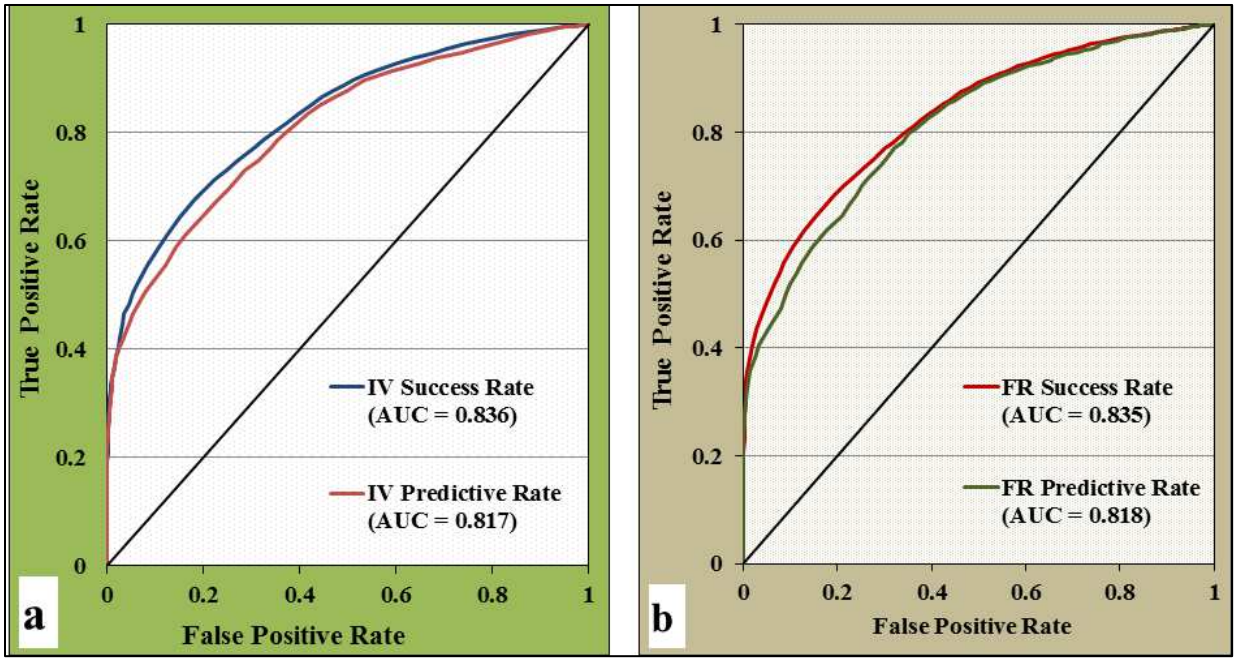


Fig. 6 Success and Predictive rate Curves for a) IV and b) FR models

Table 5 LDI value for IV and FR

	LS Class	LSI	LSP	%of LSP(a)	TLP	% of TLSP(b)	LDI (b/a)	VLP	%of VLP(c)	LDI (c/a)
Information value	Verylow	-5.29 to -3.02	43734	7.17	23	0.11	0.016	0	0.00	0.00
	Low	-3.02 to -1.62	136995	22.45	81	0.40	0.018	31	0.77	0.03
	Moderate	-1.62 to -0.50	145615	23.86	1339	6.58	0.276	383	9.53	0.40
	High	-0.50 to 0.53	179599	29.43	6499	31.96	1.086	1231	30.62	1.04
	Veryhigh	0.53 to 2.67	104315	17.09	12394	60.95	3.565	2375	59.08	3.46
		Total	610258	100	20336	100		4020	100	
	LS Class	LSI	LSP	% of LSP(a)	TLP	% of TLSP(b)	LDI (b/a)	VLP	% of VLP(c)	LDI (c/a)
Frequency Ratio	Verylow	2.19 – 5.74	159321	26.11	55	0.27	0.01	7	0.17	0.01
	Low	5.74 – 7.97	169780	27.82	1496	7.36	0.26	450	11.19	0.4
	Moderate	7.97 – 10.33	162592	26.64	5606	27.57	1.03	1010	25.12	0.94
	High	10.33 – 13.26	77772	12.74	5481	26.95	2.11	1159	28.83	2.26
	Veryhigh	13.26 –20.01	40793	6.68	7698	37.85	5.66	1394	34.68	5.19
		Total	610258	100	20336	100		4020	100	

5. Conclusion

The main aim of this research work was the application, testing and comparison of a bivariate statistical models which are capable of describing the relationship between landslides and a number of causative factors to create reliable susceptibility maps for land use planning and management. To achieve this objective, two bivariate statistical methods i.e. information value and frequency ratio models have been used. A total of 580 landslides have been identified from Google Earth image interpretations and field survey. Then, 80%(464 landslides) were used as

training landslides to build the models while the remaining 20%(116 landslides) were used as testing landslides to evaluate the performance of the models eight predisposing factors (slope, aspect, curvature, distance from stream, distance from lineament, lithology, annual rainfall and land use) were used for the analysis and evaluation of the spatial relationship between these factors and landslides. The resulting LSMs from both models were subdivided into five susceptibility classes. The information value that have positive value and frequency ratio greater than 1 were found in the factor classes of a slope greater than 15°, curvatures of concave and convex and aspect of southeast, south, southwest, west and northwest facing slopes. In case of distance from stream, the five factor classes in between 0 and 300m and in case of distance from lineaments, the six factor classes in between 0 and 800m showed the highest probability of landslide occurrence. Lithological units (basalt, alluvial soil, limestone and sandstone), two rainfall classes in between 1621 and 1670mm/year and land use(bare land, sparse and moderate forest) had a high probability for landslide occurrence. LSI map of the study area was prepared based on IV and FR methods in ArcGIS10.4 using the spatial analyst tools of raster calculator and reclassified into five susceptibility classes of very low, low, moderate, high and very high using the natural breaks method of classification to produce the final landslide susceptibility maps. Finally, the models was validated using area under ROC curve and Landslide Density Index (LDI). In particular, the fitting performance and the prediction capability of the resulting landslide susceptibility models have been ascertained using the same landslide data which is used to obtain the model itself (training 80%) and independent landslide information which was not used to construct the model (validation 20%) respectively. In this study, AUC of the success rate and predictive rate curves range between 0.8-0.9 indicating a good performance of both models. The information value and frequency ratio models were validated using the Area under the curve(AUC) of the receiver operating curve(ROC) curve with a success rate of 0.836 and 0.835 respectively and a predictive rate of 0.817 and 0.81 respectively.

Generally, for a regional-scale map(1:50,000) of the study area, the bivariate statistical models (Information value and frequency ratio models) were found to be reliable. Besides this, the procedures of producing LSMs were relatively simple and cost-effective. The findings of this study can help land use planners, geologists and civil engineers to identify areas that are susceptible to landslides. Since the results are of regional-scale, the LSMs may be less useful for a site-specific development that requires large-scale maps.

Recommendation

Increasing the vegetation cover of the area by planting the trees, controlling the drainage system found in the study area, applying preventive measures such as retaining and gabion walls and achieving widespread public awareness about landslide hazards can reduce the risk.

Declarations

Availability of data and materials

Rainfall data were collected from the National Metrology Agency of Ethiopia. Topographical Map was purchased from the Ethiopia Geospatial Information Agency. DEM data'sare freely available from http://gdex.cr.usgs.gov/gdex/ website.

Competing Interests

There was no any competing interest.

Funding

There was no any funding received from any governmental or non-governmental organization as the first author was a self-sponsored MSc student.

Authors' contributions

AG as a first author has mostly participated in the whole process of this research work including the field-work, data collection, database preparation and compling the results. MM as an advisor, participated from the inception stage of this research and commented in each phase for further improvement. AG addressed the comments given from his advisor in terms of its scientific justification, methodological aspect and English correction before submission and finally the two authors approved the submission to this journal.

Acknowledgments

The first author would like to thank Mr. Leulalem Shano and Mr.Nega Getachewu for their useful comments and suggestions to improve this research work.

Authors' Information

Alembante Genene (MSc) [1*] (First author)

[1] Department of Geology, College of Applied Sciences, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia; P.O.Box: 16417
Telephone: +251-967231973; E-mail: alembantegenene5064@gmail.com

Matebie Meten (PhD) [1, 2] (Corresponding author)

[1] Department of Geology, College of Applied Sciences,
Addis Ababa Science and Technology University, Ethiopia
P.O.Box: 16417; Addis Ababa Ethiopia
Telephone: +251-911899279;
E-mail: matebe21@gmail.com

[2] Mineral Exploration, Extraction and Processing Center of Excellence;
Addis Ababa Science and Technology University, Ethiopia;
P.O.Box: 16417; Addis Ababa Ethiopia

Reference

Aleotti, P.; Chowdhury, R (1999). Landslide hazard assessment: Summary review and new perspectives. Bull. Int. Assoc. Eng. Geol. 58, 21–44.

Ayalew, L., & Yamagishi, H. (2005). The application of GIS based logistic regression for landslide susceptibility mapping in the KakudaYahiko Mountains Central Japan. Geomorphology 65(1):15–31.

Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012). Landslide susceptibility mapping at HoaBinh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. Computers & Geosciences.45:199-211.

Chung, C.J.F.; Fabbri, A.G. Validation of Spatial Prediction Models for Landslide Hazard Mapping. Nat. Hazards 2003, 30, 451–472.

Fayez, L., Pazhman, D., Binh, T.P., Dholakia, M.B., Solank, H.A., and Khalid, M. (2018). Application of frequency ratio model for the development of landslide Susceptibility Mapping at Part of Uttarakhand State, India. International Journal of Applied Engineering, Vol. 13 (9):6846-6854.

Glade, T., 1997.Establishing the frequency and magnitude of landslide-triggering rainstorm events in New Zealand. Environmental Geology 35, 160–174.

Gomez H, Kavzoglu T (2005) Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. J EngGeol 78(1 2):11–27.

Ibrahim J. (2011). Landslide assessment and hazard zonation in Mersa and Wurgessa, North Wollo, Ethiopia. Unpublished Master Thesis, School of Graduate Studies, Addis Ababa University, Addis Ababa, Ethiopia, 1-10.

Junyi Huang, Qiming Zhou & Fenglong Wang (2015). Mapping the landslide susceptibility in Lantau Island, Hong Kong, by frequency ratio and logistic regression model, Annals of GIS, 21:3, 191-208.

Kanungo D.P., Arora, M.K., Sarkar, S. and Gupta, R.P. (2009). Landslide susceptibility zonation mapping a review. J South Asia Diaster stud., Vol. 2: 81 – 105.

KifleWoldearegay. (2008). Characteristics of a large-scale landslide triggered by heavy rainfall in Tarmaber area, central highlands of Ethiopia. Geophysical Research Abstracts, vol.10.

KifleWoldearegay. (2013). Review of the occurrences and influencing factors of landslides in the highlands of Ethiopia with implications for infrastructural development.Mekelle University, Mekelle, Ethiopia. Journal of Muna, Vol. 5 : 331.

Lee, S. & Pradhan, B. (2007). Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides. 4:33–41.

Meten, M., Bhandary, N. P., &Yatabe, R. (2015). GIS-based frequency ratio and logistic regression modelling for landslide susceptibility mapping of Debre Sina area in central Ethiopia. Journal of Mountain Science, 12(6), 1355–1372. doi:10.1007/s11629-015-3464-3.

Neuhäuser B, Damm B, Terhorst B. GIS-based assessment of landslide susceptibility on the base of the weights-of-evidence model. Landslides 2012;9:511-28.

Pradhan B, Oh H-J, Buchroithner M. Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. Geomatics, Natural Hazards and Risk 2010;1:199-223.

Seda C (2019). Effects of Stream Distance to landslide. Engineering and Natural science. 4(1); 268-275.

Shraban S, Archana K,Tapas R(2013). Landslide Susceptibility Assessment Using Information value methods in parts of the Darjeeling Himalays. Geological society of India, 82, 351-362.

Soeters R, van Westen CJ (1996) Slope instability recognition, analysis and zonation. In: Turn AK, Schuster RL (eds.) Landslides: Investigation and Mitigation. Transportation Research Board Special Report 247, National Academy Press, Washington DC, pp.129–177.

Suzen, M.L., 2002. Data driven landslide hazard assessment using geographical information systems and remote sensing, M.E.T.U. PhD Thesis, Unpublished. 196 pp.

vanWESTEN, C.J. (1993). Application of Geographical Information System to landslide hazard zonation. ITC, International institute for aerospace and earth res. surv. Enschede, The Netherlands: ITC Publication.

vanWesten CJ, van Asch TWJ, Soeters R (2006) Landslide hazard and risk zonation – why is it still so difficult? Bulletin of Engineering Geology and the Environment 65:167–184.

Yalcin A. The effects of clay on landslides: A case study. Applied Clay Science 2007; 38:77-85.

YIN, K.J. and YAN, T.Z (1988). A statistical prediction model for slope instability of metamorphosed rocks. Proceedings of the 5th international symposium on landslides, Lausanne, Switzerland, Vol. 2: 1269–1272.

Figures

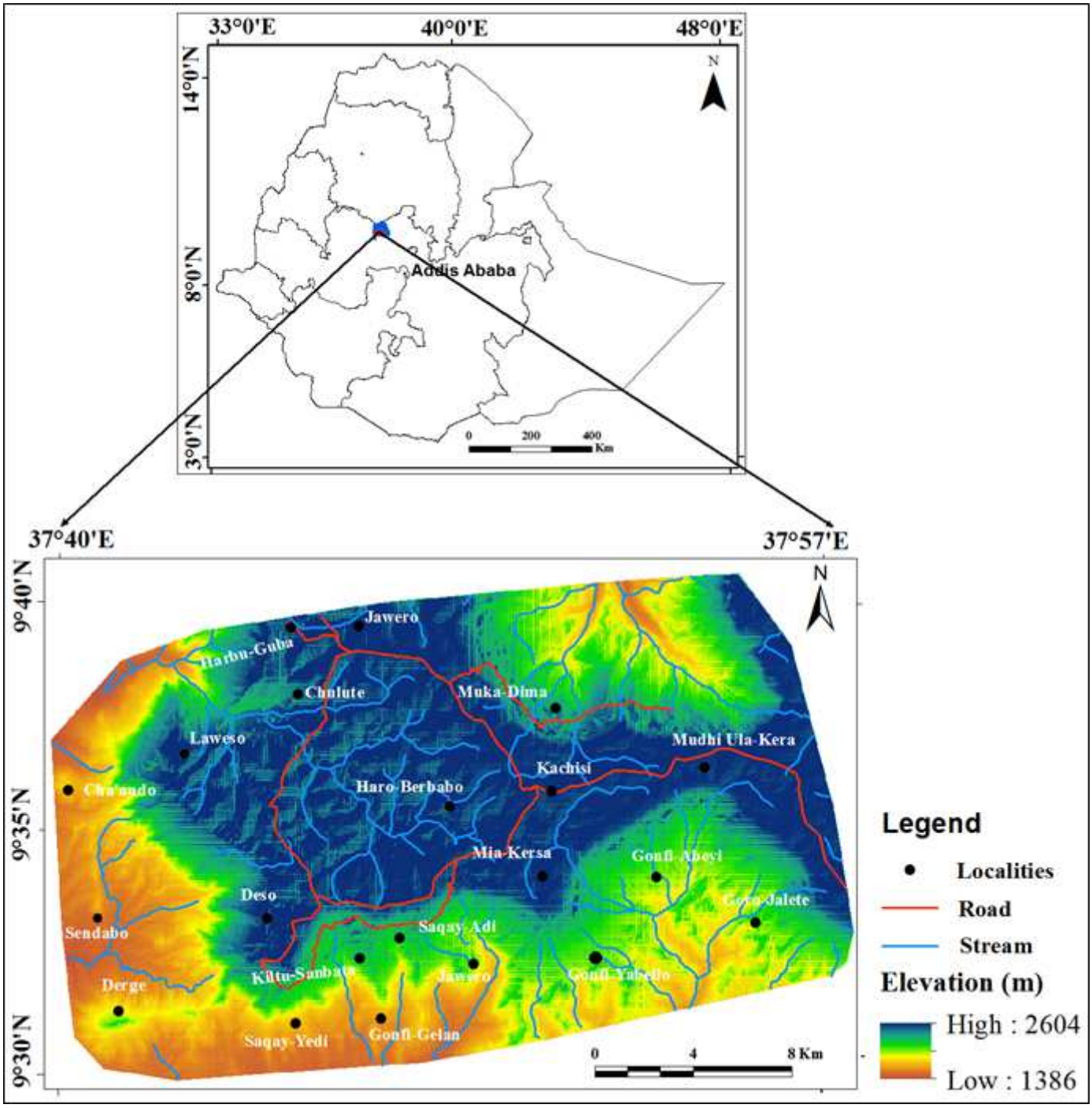


Figure 1

Location Map of the Study area Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

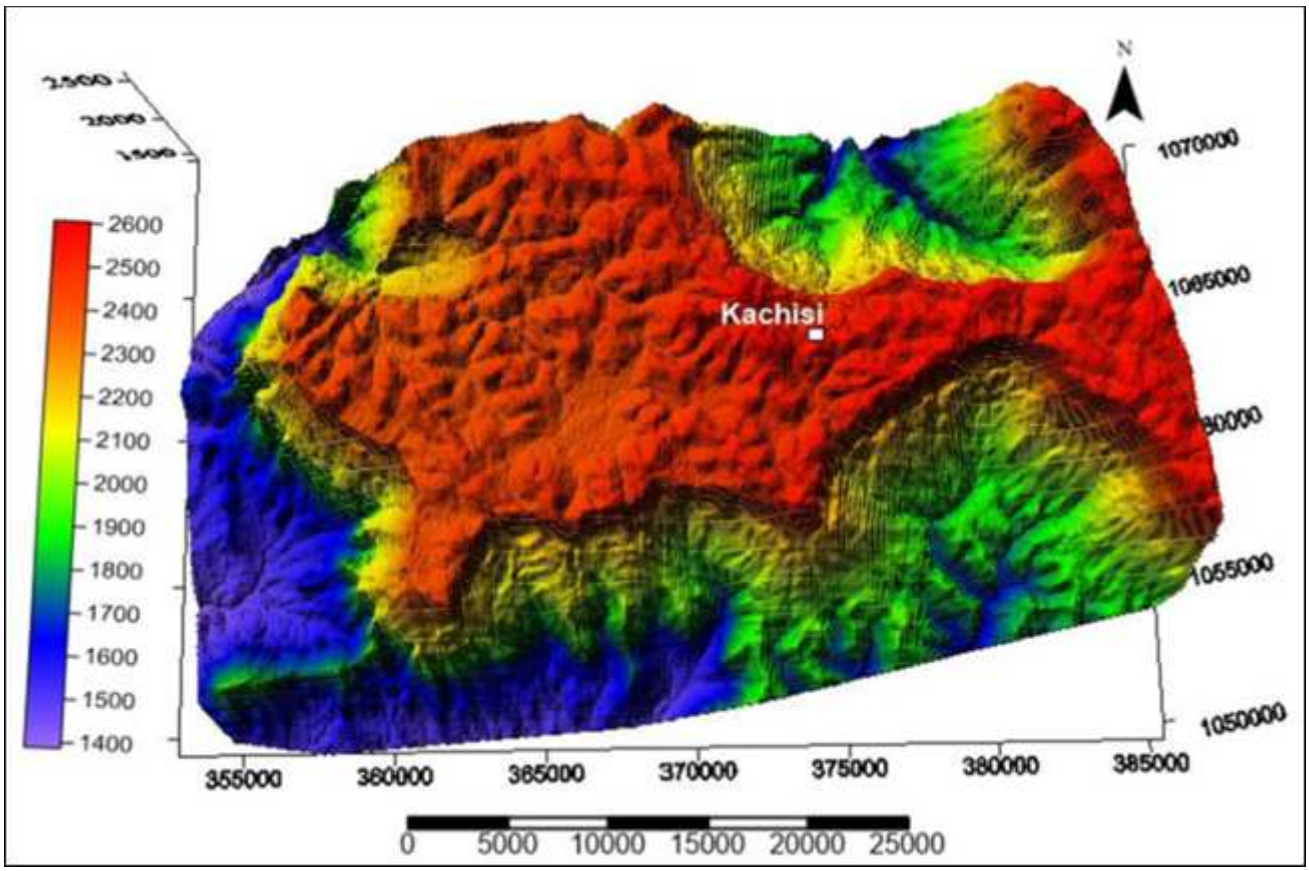


Figure 2

Physiography of the study area Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

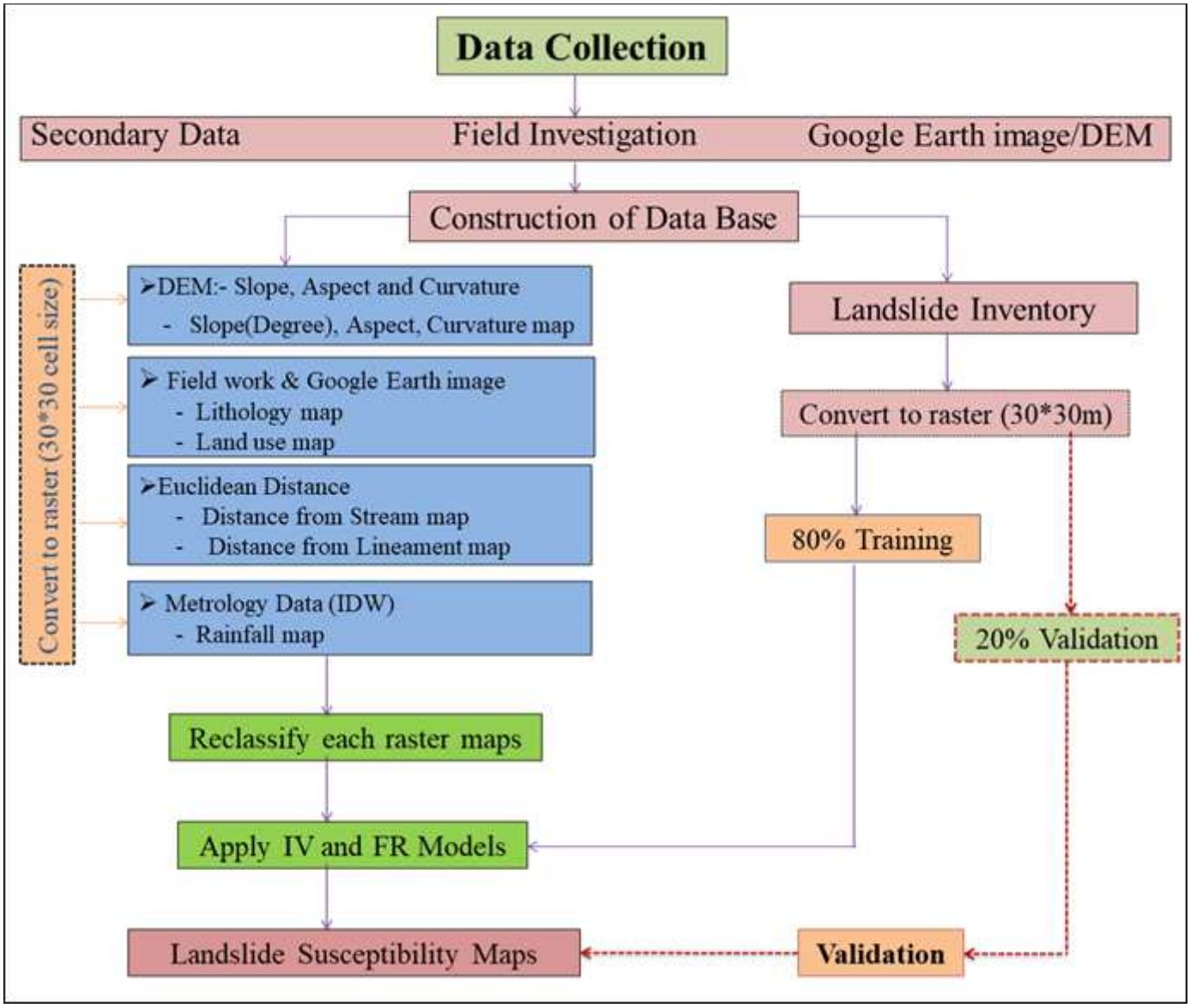


Figure 3

Flowchart of the research work

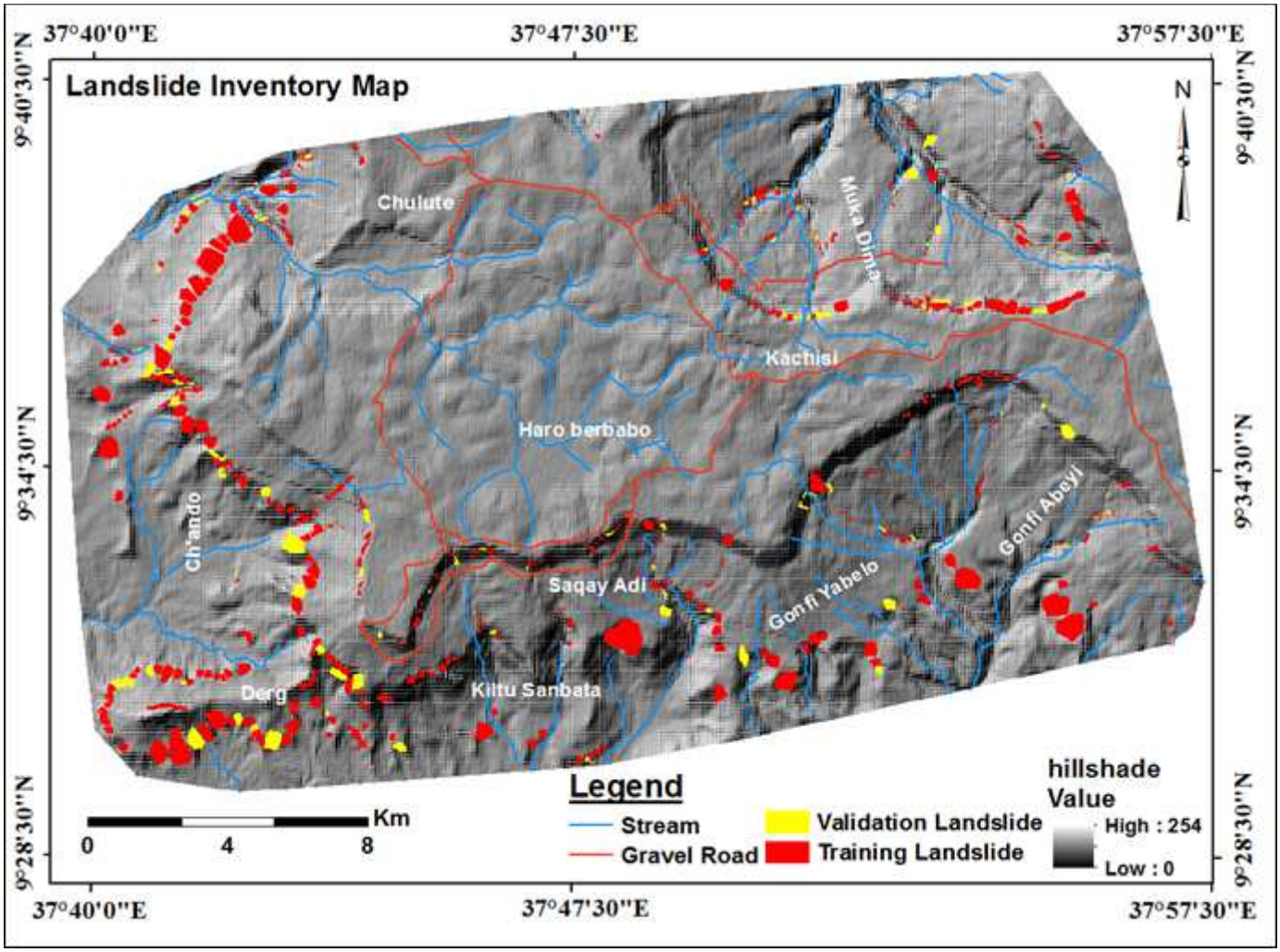


Figure 4

Landslide inventory map Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

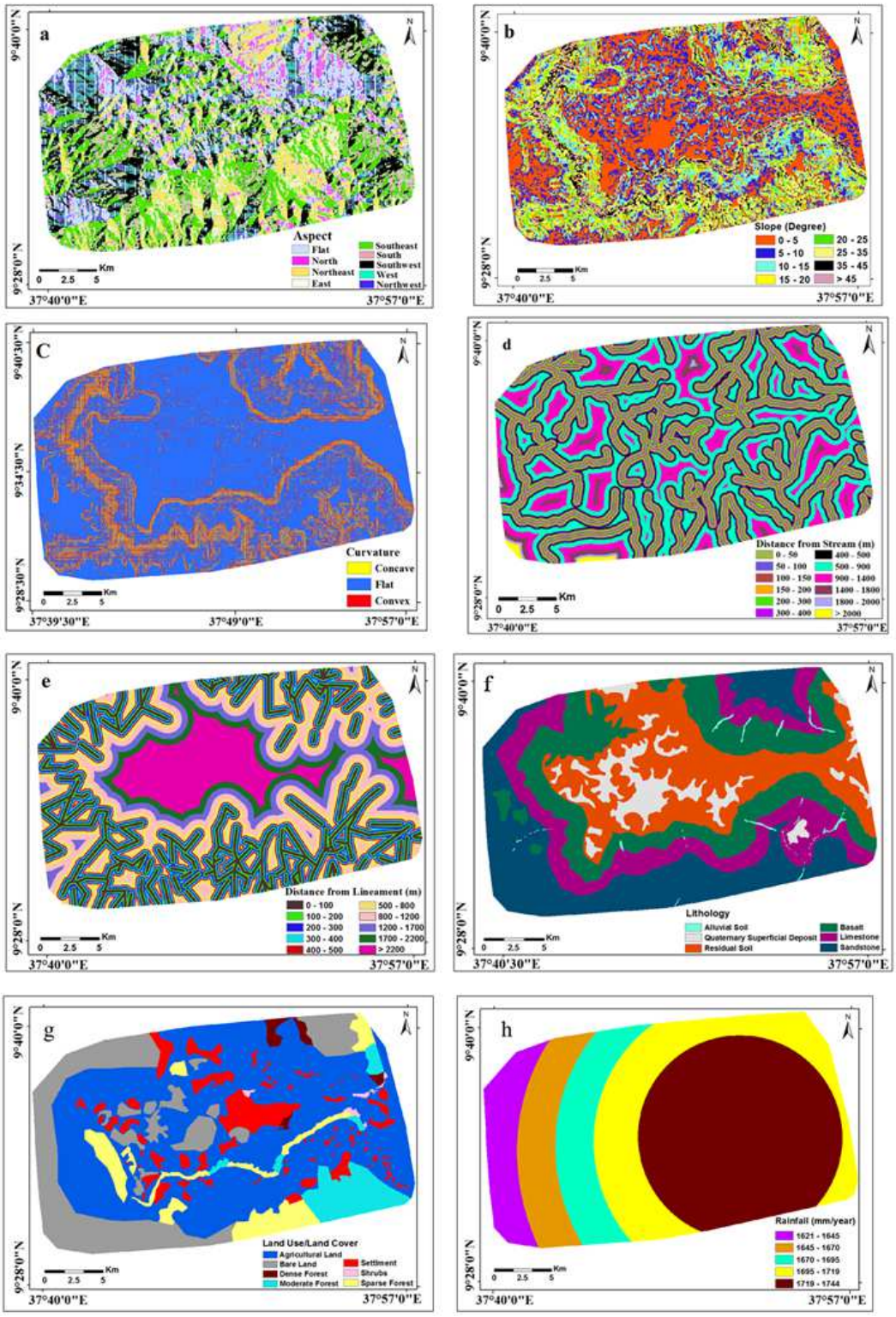


Figure 5

Landslide causative factors used in this study a) Aspect, b) Slope, c) Curvature, d) Distanc from Stream, e) Distance from Lineament, f) Lithology, g) Land Use, h) Rainfall Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its

authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

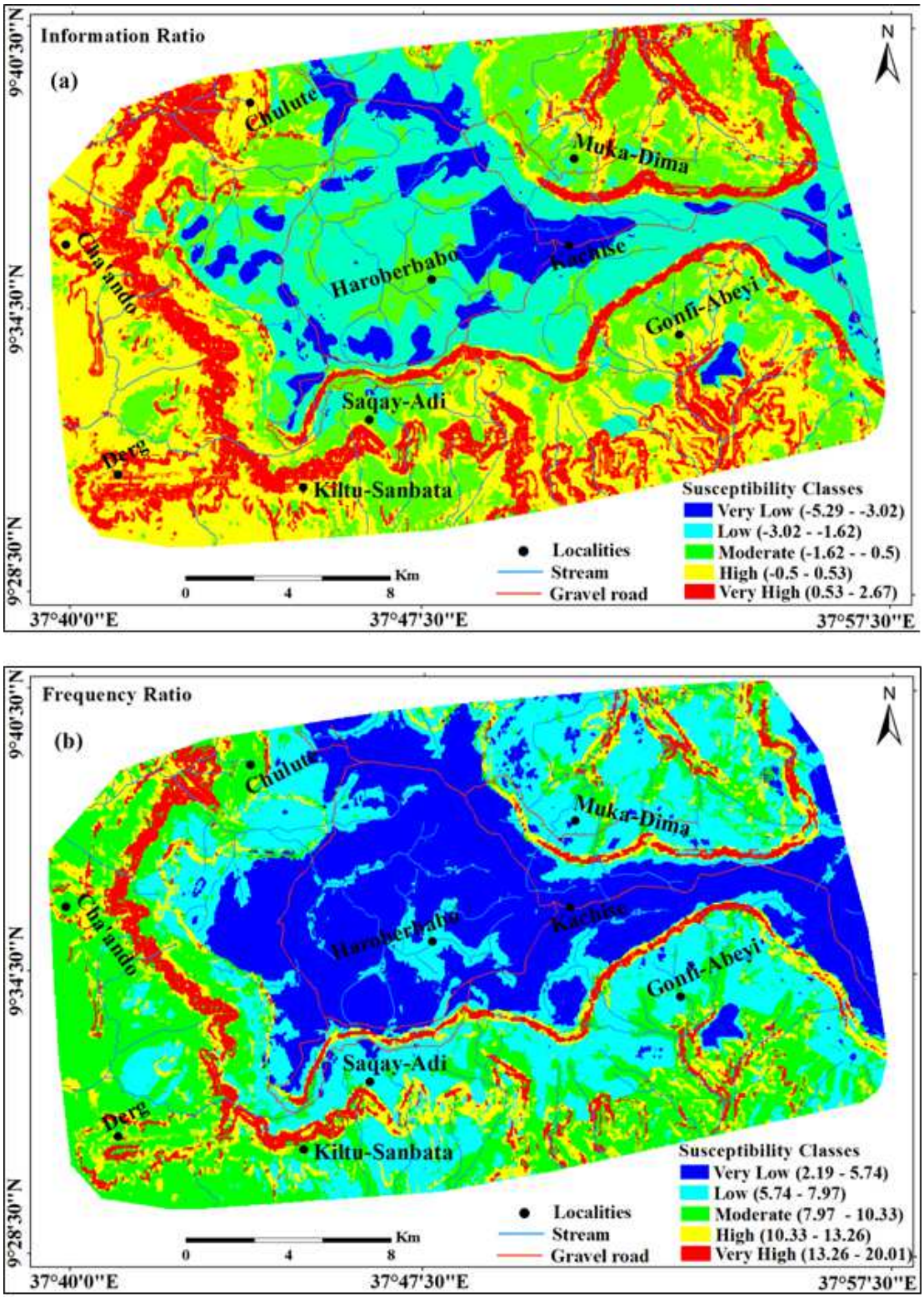


Figure 6

Landslide Susceptibility Map of the study area using Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its

authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

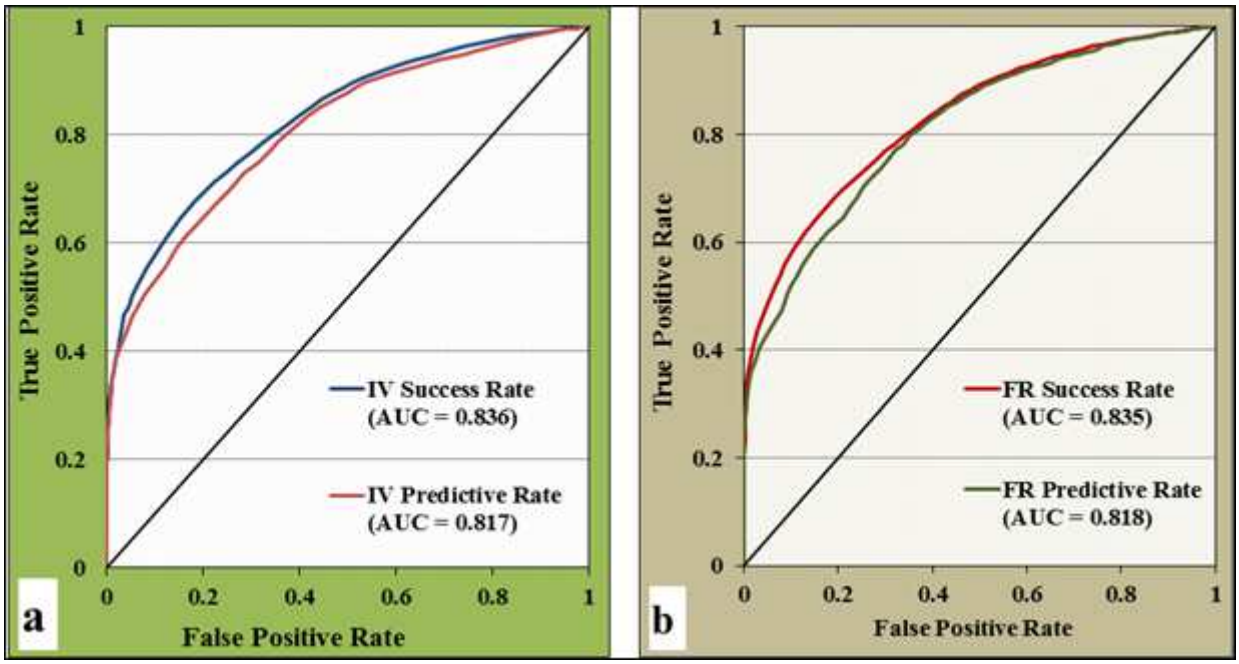


Figure 7

Success and Predictive rate Curves for a) IV and b) FR models