Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

GIS-based landslide susceptibility zonation mapping using frequency ratio and logistics regression models in the Dessie area, South Wello, Ethiopia

Tesfaye Chala Korma (tesf.chala2@gmail.com)
Bahir Dar University

Research Article

Keywords: Dessie area, South wello, landslide, landslide susceptibility, Frequency Ratio, Logistic Regression

Posted Date: May 12th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1633474/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

The purpose of this study was to evaluate landslides influencing factors and prepare landslide susceptibility zonation map, Dessie area is situated in South wello Ethiopia. For this study, lithology, slope, curvature, elevation, land use, slope aspect, distance from road, river and tectonic fault were considered as landslide causative factors. The landslide inventory database was developed from extensive fieldwork and Google Earth imagery and randomly partitioned as training datasets (70%) and testing datasets (30%) using the Subset features Tools in ArcGIS. The training datasets combined with landslide causative factors to produce landslide susceptibility zonation maps. The testing datasets used to validate the maps using Area under the curve (AUC) of the Receiver Operating Characteristic (ROC) method. The spatial relationship between landslide and landslide causative factors analyzed in frequency Ratio and logistic regression models and the results used to prepare landslide susceptibility zonation (LSZ) maps in GIS environments. Landslide susceptibility zonation maps were classified into five categories: very low, low, medium, high and very high zones. Finally, the results of analysis were validated by comparing known landslide location with LSZ maps using AUC of ROC. The prediction accuracy value obtained shows that the logistic regression model (84.4%) is better in prediction than frequency ratio model (80.8%). The analysis shows that the prediction rate of FR model achieved 80.8% while the LR model achieved 84.4%. These maps show areas prone to landslides that will be used for proper land use planning and further sustainable development activities in the area.

1. Introduction

Landslides are one of the most serious geological hazards in the world, especially in mountainous areas, causing widespread destruction on the loss of life, properties, and infrastructure (Parise & Jibson 2000; Alexander 2005; Pan et al. 2008; Mersha & Meten 2020). Globally, landslides cause appropriately 1000 deaths per year and property damage of more than 4 billion (Lee & Pradhan 2007; Solaimani et al. 2013). Over the past 25 years, governmental and international research institutes have invested considerable resources to assess landslides and develop susceptibility maps (Guzzetti et al. 1999; Yalcin et al. 2011). Landslides in mountainous region of Ethiopia are a serious geological hazard, often resulting in loss of life, properties and infrastructure (Ayalew 1999; Abebe et al. 2010; Woldearegay 2013; Mersha & Meten 2020). About 60% of Ethiopia population lives in the highlands (with elevation ranging from 1500 m – 4550 m) with complex geology, heavy rainfall, rough morphology, and active geomorphological processes (Ayalew 1999; Mersha & Meten 2020). In recent years, landslide related hazards are increasing due to rapid expansion of urbanization, improper land use practice and slope excavation over landslide-prone areas. Hence, to reduce the hazard related to landslides, it is important to identify areas susceptible to landslides and predict future landslides for land use management.

Dessie area is the most landslide-prone region of South wello in Ethiopia, which are characterized by frequent landslide occurrences causing damages to infrastructures and properties. Landslides interact with geology, human activities such as land-use change; develop infrastructures without detailed geotechnical investigations are aggravating and reactivating landslides, which caused to damage

infrastructures and property Fig. 2. From 2014 to 2020, there was a significant increase in landslides in the Dessie, especially in the area under new construction (e.g. Agere Gizat site 2016 due to over-excavation) Fig. 2C and 2D. Large scale landslide hazard may be expected in the future due to unexpected urban expansion, infrastructures developments in landslide prone-area. To minimize the damage caused by landslides, it is important to evaluate the factors causing landslides and determine landslide susceptible areas for future suitable land use management.

Hence, main objectives of this study were to evaluate the landslide causing factors and apply bivariate (frequency ratio) and multivariate statistical (logistic regression) analysis within the GIS framework to prepare the landslide susceptibility zonation maps. Landslide susceptibility zonation mapping plays a key role in developing sustainable infrastructure, urban plan and land-use planning. The words used to study landslide related hazard such as landslide hazard mapping, landslide susceptibility zonation, etc., are very confusing among the researchers (Anbalagan 1992; Guzzetti et al. 1999; Ayalew & Yamagishi 2005; Lee & Pradhan 2007; Yilmaz 2009; Anbalagan et al. 2015). According to Varnes (1984), in the context of landslides, the term 'zonation' refers to the general term for dividing the land surface into similar zones and based on the actual or potential hazard of landslides. Since the 1980s, 1990s and early 2000s, landslide hazard zonation mapping has been used by many authors to study landslides (Anbalagan et al. 2015). Also, the researchers describe the term 'landslide susceptibility mapping, in the context of landslides given to the spatial probability of occurrence of landslides based on a set of local geo-environmental conditions (Lee & Sambath 2006; Kanungo et al. 2012). Landslide hazard mapping is also a division of terrain into zones characterized by the spatial and temporal probability of landslide occurrence (Guzzetti et al. 1999; Guzzetti 2006; Anbalagan et al. 2015; Shano et al. 2020).

Nowadays landslide susceptibility zonation combines the terms landslide hazard zonation, landslide susceptibility and is often used to study landslide associated hazards (Kanungo et al. 2012; Anbalagan et al. 2015). Similarly, in this study landslide hazard zonation, landslide susceptibility and landslide susceptibility zonation have been considered as the same. GIS-based landslide studies using frequency ratio (FR) and logistic regression (LR) models are used by several researchers to map landslide susceptibility zonation based on local geo-environmental factors (Aleotti & Chowdhury 1999; Kanungo et al. 2012; Mersha & Meten 2020). Many approaches are used to study landslides such as qualitative, semi-qualitative and quantitative. Qualitative and semi-qualitative methods are subjective and represent susceptibility level by descriptive expressions and expert based. Quantitative methods such as deterministic, probability approaches and statistical methods are closely based on mathematical models, with no personal bias (Aleotti & Chowdhury 1999; Kanungo et al. 2012; Mersha & Meten 2020). Quantitative methods are also used to evaluate landslide events and predict future landslides. Thus, for this study quantitative approaches of FR and LR models were embraced and applied to prepare landslide susceptibility zonation maps in the Dessie area of South Wello, Ethiopia. The output susceptibility zonation maps have been validated using the area AUC of the ROC.

2. Study Area

Dessie area is located in the South Wello, highlands of Ethiopia. It is characterized by a chain of mountains and a technologically fragile zone along the western Afar margin Fig. 1A and B. It is bounded by two N–S striking steep slopes formed by the action of normal faults: the Tossa escarpment to the west and the antithetic Azwa Gedel escarpment to the east (Ayenew & Barbieri 2005; Fubelli et al. 2008; Abebe et al. 2010). Most of the river network and their engagement processes resulting in deep incision emerged from Tossa Chain mountain made the area more susceptible to landslide (Ayalew 1999; Fubelli et al. 2008; Abebe et al. 2010) (Fig. 1B).

The morphology of this area is very gentle to very steep slope; the gentle slope is covered by Quaternary deposit (Tefera et al. 1996; Ayenew & Barbieri 2005). The climate of the Dessie area is characterized as sub-humid to humid with a bimodal rainfall regime and the annual rainfall distribution is characterized by pronounced seasonality, with the heaviest rainfall occurring in July and August (Ayalew 1999; Fubelli et al. 2008). Intensive rainfall has caused disastrous flash floods in river basins and many landslides on the slope (Ayalew 1999).

2.1. Geological and Geomorphological setting of Dessie area

The geology of the Dessie area consists of sequences of basaltics Trap Series ranging between 30 and 25 Ma and Quaternary deposit (Tefera et al. 1996; Ayenew & Barbieri 2005; Fubelli et al. 2008). These rocks are moderate to highly weathered and densely jointed basaltic rock which corresponds to stratified beds of successive lava flows, inter-bedded with weakly degraded volcanic units and several reddish paleosol horizons (Ayenew & Barbieri 2005; Fubelli et al. 2008; Vařilová et al. 2015). The porous and friable vesicular basalt, which overlying the stratified basalt is dominantly covered in the central part of Dessie Town (Ayenew & Barbieri 2005; Vařilová et al. 2015). The floor of Dessie Basin and Boru Meda are filled with Quaternary sediments which reach up to several tens of meters thick (Fubelli et al. 2008). These Quaternary sediments are mainly alluvial-swampy and colluvial deposits likely derived from volcanic products (Ayenew & Barbieri 2005; Fubelli et al. 2008).

The Oligocene continental flood basalt volcanism, which resulted in producing sequences of basalt and ignimbrite, is considered the start of the geomorphological evolution of the Dessie area (Fubelli et al. 2008). In the Pliocene-Quaternary times, the western Afar margin has been affected by E-directed extension (Boccaletti et al. 1998; Chorowicz et al. 1999; Fubelli et al. 2008). This extensional phase was accompanied by an important regional uplift which raised the Ethiopian plateau by 800–1000 m to attain the present elevation (Almond 1986; Fubelli et al. 2008). This area also has been affected by active geomorphological processes and human activities such as land-use change and construction, especially at relatively low elevation and at the foot of the Tosa Chain Mountain. The slopes of the western and eastern areas (from Dessie to Konbolcha Road, Tosa Slopes, Azwa Gedel and Dorommezlia slide) were faced rock falls, and rock slides while the Center of the Dessie was influenced by a rotational landslide (Ayenew & Barbieri 2005; Fubelli et al. 2008; Vařilová et al. 2015).

3. Methods

3.1 Data collection and preparation

To achieve the objective of the study, necessary geo-environmental data were collected and organized from different data sources. These include DEM (Digital Elevation Model) from USGS, Satellite image, Google Earth imagery from Google Earth, relevant literature and extensive fieldwork were major data sources used to map landslide susceptibility zonation maps in the area Table 1. Slope, slope aspect and curvature were derived from DEM using the spatial analyst tool in ArcGIS 10.3 while the land-use map has been extracted from Landsat-8, 2020 satellite image using supervised classification techniques. The distance from drainage was extracted from the DEM database while road networks were extracted from Google Earth. Also, Google Earth is used as references to check the land use types for each class and determines the area of high frequency of landslides. The distance from river and distance from road were calculated by the Euclidean distance tool in spatial analysis of ArcGIS 10.3.

During fieldwork, data collection was carried out on landslide inventory datasets on both active and scarp areas, land-use conditions, man-made activities prone to landslide have been determined. Then, actual field data has been organized in excel format and analyzed and processed in the GIS environment. The spatial database containing landslide causative factors that were influencing to landslide occurrence are elevation, slope, slope aspects, curvature, distance to fault, distance from river, distance from road, lithology and land use were prepared. All data layers have been rasterized and resampled as uniform pixel size (30x30) and thematic maps prepared. The spatial correlation between the landslide location and landslide causative factors were evaluated using FR and LR approaches. Finally, Landslide Susceptibility zonation maps were prepared using FR and LR approach by summing all statistically significant landslide causative factors using map algebra in the ArcGIS environment Eqs. 2 and 3.

Table 1
Sources of data

Data	Description	Source
Landsat 8	Downloaded	https://earthexplorer.usgs.gov/
Aster DEM (Digital Elevation Model DEM) Resolution 30 m	Downloaded	https://earthexplorer.usgs.gov/
Slope angle and Slope aspect	Derived from DEM 30 m	DEM 30 m
Land-use	Derived from Landsat 8	Landsat 8 OLI
Slope aspect	Derived from DEM 30 m	DEM 30 m
Road	Extracted	Google earth
Drainage	Extracted	DEM 30 m
Landslide location	Inventory map	Field data/ literature/ Google earth
Geological and structural data	Literature	Geological Survey of Ethiopia

3.2 Spatial Database

3.2.1. Landslide Inventory database

Preparing a reliable landslides inventory database is the basic step to study landslide susceptibility zonation mapping. According to Lee and Talib (2005), the basic assumption is that future landslides will occur under the same condition factors as past landslides. Hence, it is important to determine the location, spatial distribution, frequency and intensity of past landslides during preparation landslides susceptibility zonation map. In this study, more 210 landslides points were identified by extensive fieldwork and from Google Earth imagery interpretation Fig. 3. For this study, landslide datasets were randomly divided into training datasets (70%) and testing datasets (30%) using the Subset features Tools in ArcGIS (Chung & Fabbri 2003). The training datasets combined with landslide causative factors to produce landslide susceptibility zonation maps using GIS tools and testing datasets used to validate the performance of the maps.

3.2.2 Landslides influencing factors

Landslide influencing factors considered in this study are lithology, slope, slope aspect, curvature, land use, distance from fault, distance from road, and distance from river are described in the following subsections.

Lithology

The lithological units in the study area are basalt, colluvial and alluvial deposits edited from regional geological maps (Fig. 4A). Landslides occurred in all lithological units, but the frequency and spatial distribution of landslides was relatively concentrated in thick colluvial deposits and alluvial deposits. Basalt is cover the highest elevation (Tossa chain ridge) and forming steep slope (Azwa Gedel and Doromezleya) to gentle. Colluvial deposit is affected by numerous landslides, mostly retrogressive rotational slides of different size (Fubelli et al. 2008). Alluvial deposits are restricted to low-lying areas close to Borkena river and have high contribution in occurrences of landslides along riverbanks (Ayenew & Barbieri 2005; Fubelli et al. 2008). Most of the slow-moving and some rotational slides were observed in the colluvial and alluvial deposited area.

Curvature

The surface curvature represents the morphology of topography of the area and controls the hydrological condition of the soil. The curvature of the study area was generated from DEM data and classified as concave slope (negative value), convex slope (positive value) and flat surface Fig. 4B. The positive value indicates that the surface is upwardly convex (convex slope) at that cell and the negative value indicate that the surface is upwardly concave (concave slope) at that cell while the zero value indicates that the surface is linear (Gallant et al. 2000; Mersha & Meten 2020). Concave slopes have the ability to retain water and prospect for landslides occurrence than the convex slope (Mezughi et al. 2012; Meten et al. 2015b; Mersha & Meten 2020).

Land-use

Land use in the Dessie area has changed dramatically over the past five decades due to urbanization, deforestation, and infrastructure expansion (Fubelli et al. 2013; Vařilová et al. 2015). The construction process also causes the change in groundwater level and flow, modifying the slope, adding load to the slope, which initiate and reactivate landslides on hill slopes (Anbalagan 1992; Abay et al. 2019). Analysis shows that landslides occurred in the central Dessie town and its area were due to land use and construction processes, particularly in the lower slope.

Land use map of the study area is classified as built-up area, barren land and forest (Fig. 4C). The built-up area covers most of the centre of the area and along the road section. Forest area cover at toe of the Tossa and distributed overall the area and barren land covers the top of the Tossa, Azwa Gedal area and some areas of the northern part of the area. Google earth and Satellite image analysis show that most of the forest-covered and barren land areas were changing to infrastructures, which aggravate and reactivate landslides in the study area.

Distance from river

The river networks have a high probability of landslide occurrence as it erodes the slope and saturates the underwater section of the slope forming material (Dai et al. 2001; Moradi et al. 2012; Abay et al. 2019).

Most of the recorded landslides occurred in Dessie areas close to the vicinity of the rivers. That means if the area is closer to riverbanks, the area will be more susceptible to landslide occurrences (Guedjeo et al. 2013; Ali et al. 2019). The field analysis shows that high magnitude of landslide threat to built-up areas and properties were aligned along with E–W trending or following riverbanks. Consequently, as the distance to rivers increases, the susceptibility to landslide is decreasing shown in Fig. 4D.

Slope aspect

Aspect related parameters such as exposure to sunlight, prevailing winds and the amount of rainfall on a slope may vary depending on its orientation to control the occurrence of landslides (Dai & Lee 2002). The Aspect map of the study area was derived from the DEM of the study area and grouped into nine main directions (Fig. 4E). Slopes facing the N, NE, and E have more records of landslides.

Slope

The slope gradient is one the topographic factors that are directly related to terrain susceptibility of rockfall and landslide (Dai et al. 2001; Forbes & Broadhead 2013; Abay et al. 2019). In this study area, most of the landslides occurred in the slope range of 25°- 35° and 5°-15°. The slope map was extracted from DEM using ArcGIS tool (Fig. 4F).

Distance from Fault

The presence of main fault N-S northern and southern edges of the Dessie Basin and SW-NE trending transfer faults control the topography of the area, which contributes significantly to the occurrence of landslides in the area (Fubelli et al. 2008, 2013). However, most of the landslides in the area occurred in colluvial and alluvial deposits far away from the fault. Hence it needs detailed investigation whether the fault is causative landslide or not in Dessie area (Fig. 4G).

Distance from road

Distance to roads is one of the major human activities factors influencing landslide occurrences (El Jazouli et al. 2019). Road constructions change the nature of the topography and decrease the shear strength of the toe of the slope and cause tensile stress (Moradi et al. 2012). The slope may be stable under normal condition, but road construction can have an undesirable effect on slope and causes for infiltrating of water in slopes and enforces extra stresses due to traffic loads (e.g. Dessie to Kombolcha road). In the study area, landslides occurred closest alongside the road section (Fig. 4H).

Elevation

The intensive fieldwork shows that most of the landslide occurred relatively in the lower elevation (2516 m – 2635 m). This is due to the lower elevation covered by unconsolidated Quaternary deposits and easily affected by human activities and rainfall (Fig. 4I).

Figure 4 Landslide influencing factors maps: A) lithological map B) Curvature map, C) land use map, D) Distance from River map

3. Modeling

3.1 Frequency ratio

Frequency ratio model is one of the probability model based on spatial distribution of landslides and considered causative factor, to reveal the relationship between landslide location and its causative factors in the study area (Lee & Talib 2005; Lee & Pradhan 2007; Solaimani et al. 2013). To evaluate the contributions of causative factors towards landslide susceptibility, the training landslide data points were combined with data layers using the following procedure (Lee & Talib 2005). First, count the number of pixels of landslide in each factor class using tabulated ArcGIS tools. The number of pixels is calculated as the total area of the class divided by cell size area (30*30) using a field calculator in ArcGIS and Microsoft Excel. Second, change the landslide and non-landslide area into number of pixels, percentage and results were summarized in Table 2. Then, the Frequency Ratio of each factor class was calculated by dividing the percentage of landslide pixels in each causative factor class to the percentage of total area pixels of each causative factor class Eq. 1. Mathematically, FR calculation is expressed as;

$$FR = \frac{a}{b} = \frac{\frac{Nlpix}{Npix}}{\frac{Ntlpix}{Ntpix}} \quad \text{…… Eq. 1}$$

Where FR = Frequency Ration; Nlpix = number of pixels that contains the landslide in each factor class; Npix = number of pixels within the in each factor; Ntlpix = total number of pixels that contains the landslide in the whole study area; and Ntpix = total number of pixels of the whole study area.

The calculated FR value represents the level of correlation between landslide and certain class of the causative factor. If the value of FR is 1 an average value, but, if the value is greater than 1, it means a class has strong correlation with landslide which indicates high probability of landslide occurrence and values less than 1 means lower relation with landslide, which has low probability of landslide occurrence in a certain class of a landslide causative factor (Lee & Talib 2005; Solaimani et al. 2013; Mersha & Meten 2020). After, FR maps of each causative factor have been prepared by reclassifying using calculated FR values in ArcGIS tools. Finally, the landslide susceptibility index (LSI) map was prepared by adding the value of all landslide causative factors of FR maps using (Eq. 2) raster calculator of the spatial analysis tool in ArcGIS 10.3.

$$FR = FR1 + FR2 + FR3 + \cdots + FRn \quad \text{…… Eq. 2}$$

Where FR1, FR2, FR3…FRn are the frequency ratio raster maps of landslide causative factors, LSI represent landslide susceptibility index and n is the number of factors.

3.2 Logistic regression model

Logistic regression allows one to establish a multivariate regression relation between a dependent variable and several predictor variables. Logistic regression model is the most common multivariate analysis used to predict a result measured by a binary variable such as presence or absence of landslide based on the value of a set of independent variables (Meten *et al.*2015). Also, the independent variables may be either continuous or discrete or any combination of both types and do not necessarily have normal distributions (Lee & Talib 2005; Schicker & Moon 2012). In the case of landslide susceptibility mapping, the purpose of this model is to find the best fit to describe the relationship between dependent variable and set of independent parameters (Ayalew & Yamagishi 2005).

For this study, the dependent variable is a binary variable representing the presence (coded as 1) or absence (coded as 0) of landslide and the study area partitioned into equal numbers of landslide and non-landslide. Then, merge landslides and non-landslides points using ArcGIS tools and extract the values of each frequency ratio maps for each landslide causative factor and processes for coefficients of each independent factors in SPSS (Akgün & Bulut 2007; Meten et al. 2015c).

Mathematically, the relationship between landslide occurrence and its dependency on several variables can be expressed (Lee & Talib 2005; Lee & Sambath 2006; Lee & Pradhan 2007; Schicker & Moon 2012; Meten et al. 2015c; Shano et al. 2021) as:

$$P = \frac{1}{1+e^{-z}} \quad \text{..... Eq. 3}$$

Where p is the probability of landslide occurrence that varies from zero to one on the S-shape curve and Z is the linear combination. When probability (P) closer to 1 landslide is more probable to occur and closer to 0, landslide is less probable to occur. Logistic regression coefficient can be used to measure the contribution of the independent variables in the model.

Z can be defined as:

$$Z = B0 + B1X1 + B2X2 + B3X3 + \cdots + BnXn \quad \text{..... Eq. 4}$$

Where B_0 is the intercept of the model, the B1, B2, B3…Bn represent the coefficient of the model and X1, X2, X3…Xn represent independent variables and n is the number of factors.

In the case of a logistic regression model, all independent variables are not included in the model. If the observation is statistically significant (< 0.05 in this study), the predictor variable is used to develop the model.

3.3 Model validation

In landslide susceptibility zonation mapping, validation is very important, otherwise, the prediction models have no scientific significance (Chung & Fabbri 2003; Bui et al. 2012; Rasyid et al. 2016). For this study, to validate landslide susceptibility zonation maps, Area under Curve (AUC) of Receiver Operating Characteristic (ROC) has been applied (Mersha & Meten 2020). The AUC of predictive rate was obtained

by comparing known landslide location with LSZ mapping in SPSS (Chung & Fabbri 2003; Meten *et al.*2015). This is done by susceptibility index maps that are reclassified into 100 and rearranged in descending order. Then, the LSI maps values were extracted from ArcGIS in bdf format and the AUC (predictive rate) was generated in SPSS (Fig. 5). This resulted in 80.8% and 84.4% AUC values for FR and LR models, respectively. Rasyid et al. (2016) stated that the values range from 0.50–0.60 (model failed), 0.60–0.70 (model is weak), 0.70–0.80 (model is fair), 0.8–0.90 (model is good) and 0.90–1.0 (model is very good). Therefore, the models are good performance level, which means that the models are acceptable and will be used to predict future landslides and land-use management for urban expansion.

4. Result And Discussion

GIS-based landslide susceptibility zonation mapping was prepared using frequency ratio and logistic regression models. The analysis of FR values and LR coefficients determines the important influencing factors/classes of independent variables on landslides and the results have been summarized in Table 2 and Table 3.

Thus, the FR value analysis of slope shows high FR value (ratio > 1), from (5°-15°) and (25°-35°), which indicates that this class has high probability of landslide occurrences in the area. The other slope classes have a ratio < 1, which indicates the class has low correlation with landslides i.e. it has low probability of landslide occurrence in this slope class in normal conditions (without extensive human activities and earthquakes) (Table 2).

In the case of slope aspect classes, the high FR value (ratio > 1) detected in the direction of NE, SE and E slope faces, which have high probability of landslide occurrences. Similarly, the concave slope has a high probability of landslide occurrences (ratio value > 1), due to retaining water and saturating the slope forming material and initiation for landslide occurrences.

In the case of lithology classes, the FR Value analysis shows that the high probability of landslide occurrence is observed in colluvail deposit. The basalt and alluvial deposit are show low frequency ratio value (< 1), which indicates low probability of landslide occurrence in this area. In the case of distance from fault, river and road, the probability of landslide occurrence increases as close to river and road (Table 2). In the study area, most of the landslide observed closer to riverbank and road section.

Similarly, FR values analyses of land use shows that the built-up area has higher probability of landslide occurrence than the bare land and forest areas. This clearly shows that human activities aggravate and reactivate landslides in the area, particularly in the center of Dessie town. Hence, adequate geotechnical investigation must be conducted before construction or upgrade of buildings carried out. In the case of elevation, high probability of landslide occurrence is observed relatively in lower elevation classes (2516–2635 m). This is because settlements are located on the reclaimed land, which is often unstable due to human activity.

Table 2
frequency ration values of each landslide causative factors

Factor	Class	No. area pixels in each class	% area pixels [b]	No. landslide pixels	% Landslide pixels [a]	FR = a/b
Slope	0°-5°	15963	22.30	12	16.67	0.75
	5°-15°	37880	52.91	42	58.33	1.10
	15°-25°	10616	14.83	7	9.72	0.66
	25°-35°	4551	6.36	10	13.89	2.19
	> 35°	2587	3.61	1	1.39	0.38
Slope aspect	Flat	4672	6.50	3	4.17	0.64
	N	8139	11.33	15	20.83	1.77
	NE	11211	15.60	24	33.33	2.14
	E	7477	10.41	13	18.06	1.74
	SE	5400	7.52	5	6.94	0.92
	S	7813	10.87	1	1.39	0.13
	SW	9011	12.54	4	5.56	0.44
	W	9076	12.63	6	8.33	0.66
	NW	5404	7.52	1	1.39	0.18
Curvature	Convex slope	3177	4.42	1	1.39	0.31
	Flat	46476	64.68	39	54.17	0.84
	Concave slope	22200	30.90	32	44.44	1.44
Lithology	Colluvial	7325	10.22	37	51.39	5.03
	Alluvial	14606	20.37	8	11.11	0.55
	Basalt	49774	69.41	27	37.50	0.54
Distance from River(m)	0-100	17535	24.44	30	41.67	1.70
	100–200	15300	21.33	18	25.00	1.17
	200–300	12859	17.92	12	16.67	0.93

Factor	Class	No. area pixels in each class	% area pixels [b]	No. landslide pixels	% Landslide pixels [a]	FR = a/b
	300–400	9324	13.00	8	11.11	0.85
	> 400	16727	23.31	4	5.56	0.24
Distance form Road (m)	0-100	9298	12.96	35	48.61	3.75
	100–200	8232	11.47	16	22.22	1.94
	200–300	6488	9.04	6	8.33	0.92
	300–400	5517	7.69	5	6.94	0.90
	> 400	42210	58.83	10	13.89	0.24
Land use	Built-up	12877	17.94	24	33.33	1.86
	Bare-land	42226	58.84	37	51.39	0.87
	Forest	16659	17.94	11	15.28	0.85
Elevation(m)	2250–2516	4012	5.58	4	5.56	0.99
	2516–2635	13411	18.66	51	70.83	3.80
	2635–2731	31459	43.78	17	23.61	0.54
	2731–2824	16917	23.54	0	-	-
	2824–3064	6054	8.43	0	-	-
Fault	0-100	5139	7.16	14	19.44	2.71
	100–200	5408	7.54	13	18.06	2.40
	200–300	4400	6.13	10	13.89	2.26
	300–400	4128	5.75	11	15.28	2.66
	> 400	52670	73.41	24	33.33	0.45

Table 3
logistic regression of coefficients, multicolinearity test and model statistics

Landslide causative factors			Collinearity Statistics		Hosmer Lemesh		
	Coefficients	p-value	Tolerance	VI			
Fault	− .148	.790	.936	1.068			
Land cover	− .602	.048	.819	1.222	Nagelkerke R Square		.435
River	− .578	.339	.962	1.040	Hosmer-Lemeshow tes		
Slope	.970	.017	.936	1.068	Chi-square	df	sign
Curvature	.595	.046	.866	1.155	14.277	8	0.075
Slope aspect	.299	.523	.724	1.382	Cox & Snell R Square		.326
Road	− .296	.203	.548	1.823	**AUC curve**		
Lithology	.485	.002	.663	1.508	Frequency Ratio (FR)		78.6%
Elevation	.776	.001	.686	1.458	Logistic Regression (LR)		84.4%

Logistic regression model also used to prepare landslide susceptibility zonation maps. To analyze logistic regression equal proportions of landslide and non-landslide points were extracted using the ArcGIS tool and applied the principle of logistic regression and processed in SPSS software (Meten et al. 2015c; Rasyid et al. 2016). Then, the coefficient of independent factors has been calculated using SPSS. The coefficients of independent variables show negative sign and positive sign (Table 3). The positive coefficient indicates that as the independent variable increases, the probability of occurrence of landslide increases assuming that the other variables in the model are held constant (Chen & Wang 2007; Shano et al. 2021). The negative coefficient indicates that as the independent variable increases, the probability occurrence of landslide decreases whereas other variables are constant (Peng & So 2002; Chen & Wang 2007; Rasyid et al. 2016; Shano et al. 2021). It doesn’t mean that the independent variables with negative coefficients have no relationship with a landslide, but it is based on the researcher's data analysis perspective (Peng & So 2002; Chen & Wang 2007; Rasyid et al. 2016; Shano et al. 2021).

P-values or significance level is an important parameter in logistic regression analysis, which is used to determine whether the independent variable is statistically significant or not to include in the model analysis. The independent variable with low p-value (< 0.05) indicates changes in predictors are related to changes in the response variable, which is statistically significant. The low p-values of the independent variables included in the model analysis have high influence on landslide occurrence (Shano et al. 2021). Hence, in this study independent variables like slope aspects, fault, distance to river and road are

statistically insignificant and not included in model constructions according to the authors data. Thus, independent variables such as lithology slope, curvature and elevation are statistically significant and are used in model analysis. The other issue in the logistic regression model is about multicollinearity among the independent parameters. Multicollinearity has been detected by two important indexes of Tolerance (TOL) and its reciprocal, called variance inflation factor (VIF) (Menard 1995; Zhu & Huang 2006; Meten et al. 2015c). Menard (1995) states that if the Tolerance is smaller than 0.1, it shows that there is serious multicollinearity between independent parameters and the independent variables with VIF value > 5 are excluded from the LR analysis. In this study, the smallest tolerance is ≥ 0.557 (Table 3) showing that there is no multicollinearity problem among independent landslide factors. Then, the value of Z was calculated by substituting the coefficient of independent factors in Eq. 3. Lithology, slope, land use, curvature and elevation are statistically significant influences towards causing landslides. Using the calculated Z value, the probability of landslide occurrence of the study area was calculated by substituting the Z value in Eq. 3 and result shown on Fig. 6B.

Landslide susceptibility zonation mapping

The values of LSI were classified into five classes as very low susceptibility zone, low susceptibility zone, moderate susceptibility zone, high susceptibility zone and very high susceptibility zone area using the natural breaks method in Arc-Map (Fig. 6A and 6B). Natural breaks method has been used to determine the best arrangement of values into different classes and maximizes the variance between classes and reduces the variance within classes.

According to FR map result, the very low zone coverage is 25.9%. The low, medium and high zones cover 30%, 21.3% and 13.5% of the total area. The very high susceptibility zone covered 6.6% of the total study area. In the case of LSZ map prepared by LR method the very low zone covers 26.6% of the total study area. The low, moderate and high susceptibility zone consist of, respectively, 35.5%, 17.9% and 9.2% of the total study area coverage. The very high zone covers 10.7% of the total area.

The FR and LR methods result show that areas with very low and low susceptibility classes are dominantly covered in the northern part of the study area and on top of Tossa Chain Mountain. This area is covered by fewer infrastructures and grass land. The moderate susceptibility zone is widely distributed in the study areas and to some extent in the area of construction and deforestation. High susceptibility zones cover the area near roads and rivers that are rapidly being encroach into the urban. The very high susceptibility zone covers the centre of Dessie town, particularly in the Agere Gizat, Pepsi, Tekoam and some areas of Pissa and along the Dessie Kombolcha and Kutaber road segments (Fig. 6). In the study area, very high and high susceptibility zones are associated with relatively low elevation and colluvial deposits. This is due to the fact that the construction has been upgraded and the excavations have been carried out without detailed geotechnical investigation. However, landslides are less common on flat, steep slopes and grassland, where human activity is limited. Therefore, very high landslide susceptibility zones should be free of further infrastructure development and settlement planning without detailed geotechnical investigation and adequate landslide prevention measures taken.

6. Conclusion

Landslide susceptibility zonation maps in South Wello, Dessie area produced using GIS-based FR and LR approaches. These maps are classified into five classes: very low, low, medium, high and very high susceptibility zones. The study determines that factors and factor classes contribute more to the occurrence of landslides in the area. Hence, FR analysis shows that the probability of landslide occurrence is high in colluvial deposit, concave slope, built-up areas, elevation range from (2516–2635 m), slope from(5°-15°) and (25°-35°), slope facing to NE, E and SE and areas closer to road and river. Also, LR coefficient analysis shows that landslide causative factors such as lithology, slope, curvature and elevation have been revealed as high contribution to landslide occurrence in Dessie area of South Wello, Ethiopia. The maps have been validated by comparing LSZ maps with known landslide location using AUC of receiver operating characteristic curve (ROC). The analysis shows that the prediction rate of FR model achieved 80% while the LR model achieved 84.4%. The prediction accuracy shows that the logistic regression model (84.4%) is better in prediction than frequency ratio model (80.8%). These maps show areas prone to landslides that will be used for proper land use planning and further sustainable development activities in the area.

Declarations

ACKNOWLEDGEMENT

The authors were thankful to Bahir Dar University, who funded this research as well as School of Earth science and Department of Geology for all services during research activity.

References

Abay, A., Barbieri, G. & Woldearegay, K. 2019. GIS-based Landslide Susceptibility Evaluation Using Analytical Hierarchy Process (AHP) Approach: The Case of Tarmaber District, Ethiopia. *Momona Ethiopian Journal of Science*, **11**, 14–36.

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

Akgün, A. & Bulut, F. 2007. GIS-based landslide susceptibility for Arsin-Yomra (Trabzon, North Turkey) region. *Environmental geology*, **51**, 1377–1387.

Aleotti, P. & Chowdhury, R. 1999. Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the environment*, **58**, 21–44.

Alexander, D. 2005. Vulnerability to landslides. *Landslide hazard and risk*, 175–198.

Ali, S., Biermanns, P., Haider, R. & Reicherter, K. 2019. Landslide susceptibility mapping by using a geographic information system (GIS) along the China–Pakistan Economic Corridor (Karakoram Highway), Pakistan. *Natural Hazards & Earth System Sciences*, **19**.

Almond, D.C. 1986. Geological evolution of the Afro-Arabian dome. *Tectonophysics*, **131**, 301–332.

Anbalagan, R. 1992. Landslide hazard evaluation and zonation mapping in mountainous terrain. *Engineering geology*, **32**, 269–277.

Anbalagan, R., Kumar, R., Lakshmanan, K., Parida, S. & Neethu, S. 2015. Landslide hazard zonation mapping using frequency ratio and fuzzy logic approach, a case study of Lachung Valley, Sikkim. *Geoenvironmental Disasters*, **2**, 1–17.

Ayalew, L. 1999. The effect of seasonal rainfall on landslides in the highlands of Ethiopia. *Bull Eng Geol Env*, **58**, 9–19.

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

Ayenew, T. & Barbieri, G. 2005. Inventory of landslides and susceptibility mapping in the Dessie area , northern Ethiopia. **77**, 1–15, https://doi.org/10.1016/j.enggeo.2004.07.002.

Boccaletti, M., Bonini, M., Mazzuoli, R., Abebe, B., Piccardi, L. & Tortorici, L. 1998. Quaternary oblique extensional tectonics in the Ethiopian Rift (Horn of Africa). *Tectonophysics*, **287**, 97–116.

Bui, D.T., Pradhan, B., Lofman, O., Revhaug, I. & Dick, O.B. 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.

Chen, Z. & Wang, J. 2007. Landslide hazard mapping using logistic regression model in Mackenzie Valley, Canada. *Natural hazards*, **42**, 75–89.

Chorowicz, J., Collet, B., Bonavia, F. & Korme, T. 1999. Left-lateral strike-slip tectonics and gravity induced individualisation of wide continental blocks in the western Afar margin. *Eclogae Geologicae Helvetiae*, **92**, 149–158.

Chung, C.-J.F. & Fabbri, A.G. 2003. Validation of spatial prediction models for landslide hazard mapping. *Natural Hazards*, **30**, 451–472.

Dai, F.C. & Lee, C.F. 2002. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology*, **42**, 213–228.

Dai, F.C., Lee, C.F. & Zhang, X.H. 2001. GIS-based geo-environmental evaluation for urban land-use planning: a case study. *Engineering geology*, **61**, 257–271.

El Jazouli, A., Barakat, A. & Khellouk, R. 2019. GIS-multicriteria evaluation using AHP for landslide susceptibility mapping in Oum Er Rbia high basin (Morocco). *Geoenvironmental Disasters*, **6**, 3.

Forbes, K. & Broadhead, J. 2013. Forests and landslides: The role of trees and forests in the prevention of landslides and rehabilitation of landslide-affected areas in Asia. *Rap Publication*.

Fubelli, G., Abebe, B., Dramis, F. & Vinci, S. 2008. Geomorphological evolution and present-day processes in the Dessie Graben (Wollo, Ethiopia). *Catena*, **75**, 28–37.

Fubelli, G., Guida, D., Cestari, A. & Dramis, F. 2013. Landslide hazard and risk in the Dessie Town area (Ethiopia). *In*: *Landslide Science and Practice: Risk Assessment, Management and Mitigation*. 357–362., https://doi.org/10.1007/978-3-642-31319-6-48.

Gallant, J.C., Hutchinson, M.F. & Wilson, J.P. 2000. Future directions for terrain analysis. *Terrain Analysis: Principles and Applications*, 423–427.

Guedjeo, C., Kagou, D.A., Ngapgue, F., Nkouathio, D., Zangmo, T.G., Gountié, D.M. & Nono, A. 2013. Natural hazards along the Bamenda escarpment and its environs: The case of landslide, rock fall and flood risks (Cameroon volcanic line, North-West Region). *Global Advanced Research Journal of Geology and Mining Research*, **2**, 15–26.

Guzzetti, F. 2006. Landslide hazard and risk assessment.

Guzzetti, F., Carrara, A., Cardinali, M. & Reichenbach, P. 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*, **31**, 181–216.

Kanungo, D.P., Arora, M.K., Sarkar, S. & Gupta, R.P. 2012. Landslide Susceptibility Zonation (LSZ) Mapping–A Review.

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

Lee, S. & Sambath, T. 2006. Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environmental Geology*, **50**, 847–855.

Lee, S. & Talib, J.A. 2005. Probabilistic landslide susceptibility and factor effect analysis. *Environmental Geology*, **47**, 982–990.

Menard, S. 1995. An introduction to logistic regression diagnostics. *Applied logistic regression analysis*, 58–79.

Mersha, T. & Meten, M. 2020. GIS-based landslide susceptibility mapping and assessment using bivariate statistical methods in Simada area , northwestern.

Meten, M., PrakashBhandary, N. & Yatabe, R. 2015a. Effect of landslide factor combinations on the prediction accuracy of landslide susceptibility maps in the Blue Nile Gorge of Central Ethiopia. *Geoenvironmental Disasters*, **2**, 9.

Meten, M., Prakashbhandary, N. & Yatabe, R. 2015b. Effect of Landslide Factor Combinations on the Prediction Accuracy of Landslide Susceptibility Maps in the Blue Nile Gorge of Central Ethiopia, https://doi.org/10.1186/s40677-015-0016-7.

Meten, M., Bhandary, N.P. & Yatabe, R. 2015c. GIS-based frequency ratio and logistic regression modelling for landslide susceptibility mapping of Debre Sina area in central Ethiopia. *Journal of mountain science*, **12**, 1355–1372.

Mezughi, T.H., Akhir, J.M., Rafek, A.G. & Abdullah, I. 2012. Analytical hierarchy process method for mapping landslide susceptibility to an area along the EW highway (Gerik-Jeli), Malaysia. *Asian Journal of Earth Sciences*, **5**, 13.

Moradi, M., Bazyar, M.H. & Mohammadi, Z. 2012. GIS-based landslide susceptibility mapping by AHP method, a case study, Dena City, Iran. *Journal of Basic and Applied Scientific Research*, **2**, 6715–6723.

Pan, X., Nakamura, H., Nozaki, T. & Huang, X. 2008. A GIS-based landslide hazard assessment by multivariate analysis. *Journal of the Japan landslide society*, **45**, 187–195.

Parise, M. & Jibson, R.W. 2000. A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January, 1994 Northridge, California earthquake. *Engineering geology*, **58**, 251–270.

Peng, C.-Y.J. & So, T.-S.H. 2002. Logistic regression analysis and reporting: A primer. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences*, **1**, 31–70.

Rasyid, A.R., Bhandary, N.P. & Yatabe, R. 2016. Performance of frequency ratio and logistic regression model in creating GIS based landslides susceptibility map at Lompobattang Mountain, Indonesia. *Geoenvironmental Disasters*, **3**, 19.

Schicker, R. & Moon, V. 2012. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology*, **161**, 40–57.

Shano, L., Raghuvanshi, T.K. & Meten, M. 2020. Landslide susceptibility evaluation and hazard zonation techniques–a review. *Geoenvironmental Disasters*, **7**, 1–19.

Shano, L., Raghuvanshi, T.K. & Meten, M. 2021. Landslide Hazard Zonation using Logistic Regression Model: The Case of Shafe and Baso Catchments, Gamo Highland, Southern Ethiopia. *Geotechnical and Geological Engineering*, 1–19.

Solaimani, K., Mousavi, S.Z. & Kavian, A. 2013. Landslide susceptibility mapping based on frequency ratio and logistic regression models. *Arabian Journal of Geosciences*, **6**, 2557–2569.

Tefera, M., Chernet, T. & Haro, W. 1996. *Explanation of the Geological Map of Ethiopia*. Geological Survey of Ethiopia.

Vařilová, Z., Kropáček, J., Zvelebil, J., Šťastný, M. & Vilímek, V. 2015. Reactivation of mass movements in Dessie graben, the example of an active landslide area in the Ethiopian Highlands. *Landslides*, **12**, 985–996, https://doi.org/10.1007/s10346-015-0613-2.

Varnes, D.J. 1984. *Landslide Hazard Zonation: A Review of Principles and Practice*.

Woldearegay, K. 2013. Review of the occurrences and influencing factors of landslides in the highlands of Ethiopia : With implications for infrastructural development. **5**, 3–31.

Yalcin, A., Reis, S., Aydinoglu, A.C. & Yomralioglu, T. 2011. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena*, **85**, 274–287.

Yilmaz, I. 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). *Computers & Geosciences*, **35**, 1125–1138.

Zhu, L. & Huang, J. 2006. GIS-based logistic regression method for landslide susceptibility mapping in regional scale. *Journal of Zhejiang University-Science A*, **7**, 2007–2017.

Figures

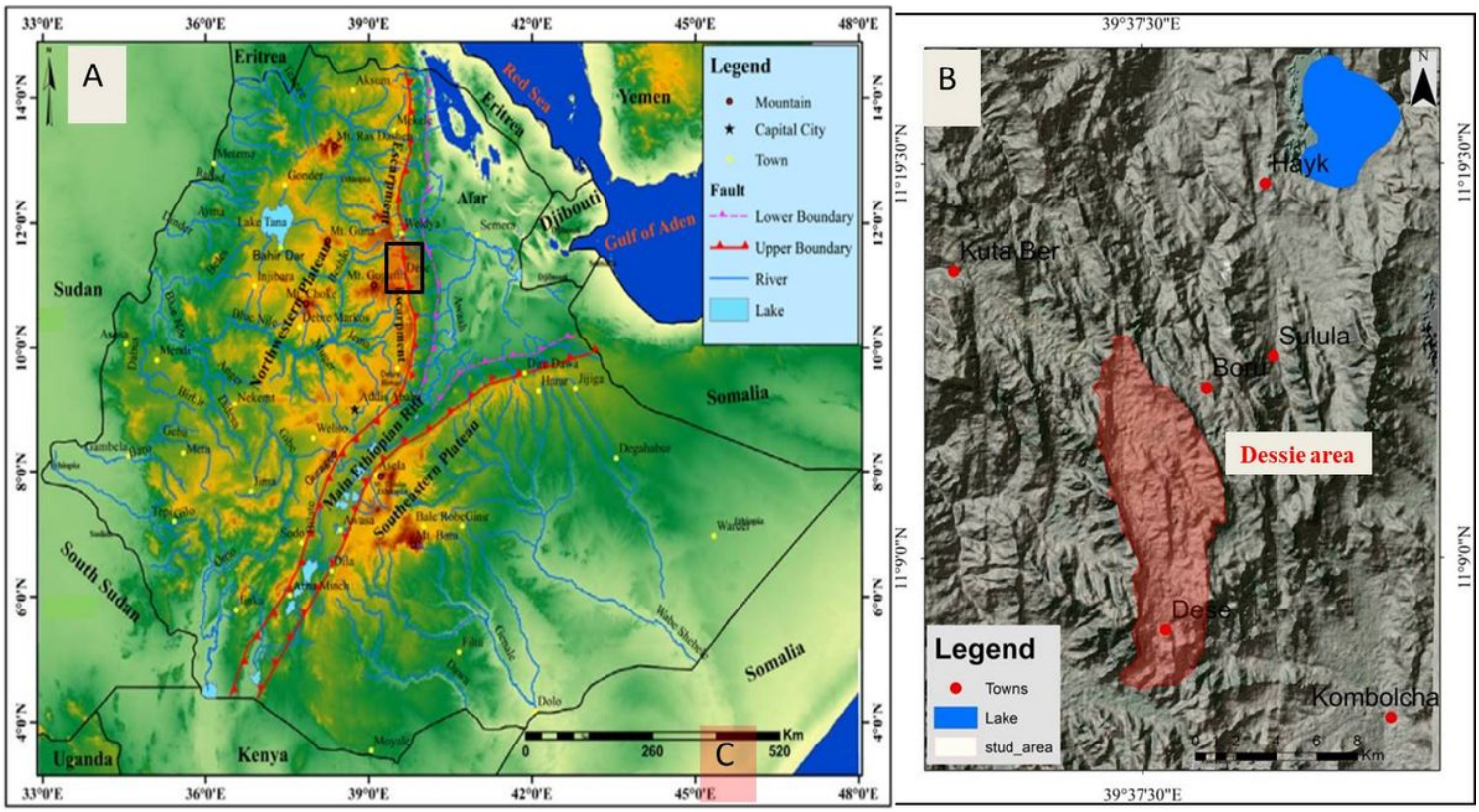

Figure 1

location map of the study area A) Geomorphological map of Ethiopia (the brown and yellow shades of the map show the Ethiopian highland while the green and light green shades indicate the lowland (Meten *et al.* 2015) B) indicate morphology of study area

Figure 2

Landslide damaged to infrastructures in the Dessie area A) damaged asphalt road and retaining wall B) deep incision develop by gully erosion C) damaged public building D) building floor cracked E) destructed of the individual home and F) cracked wall of public sectors

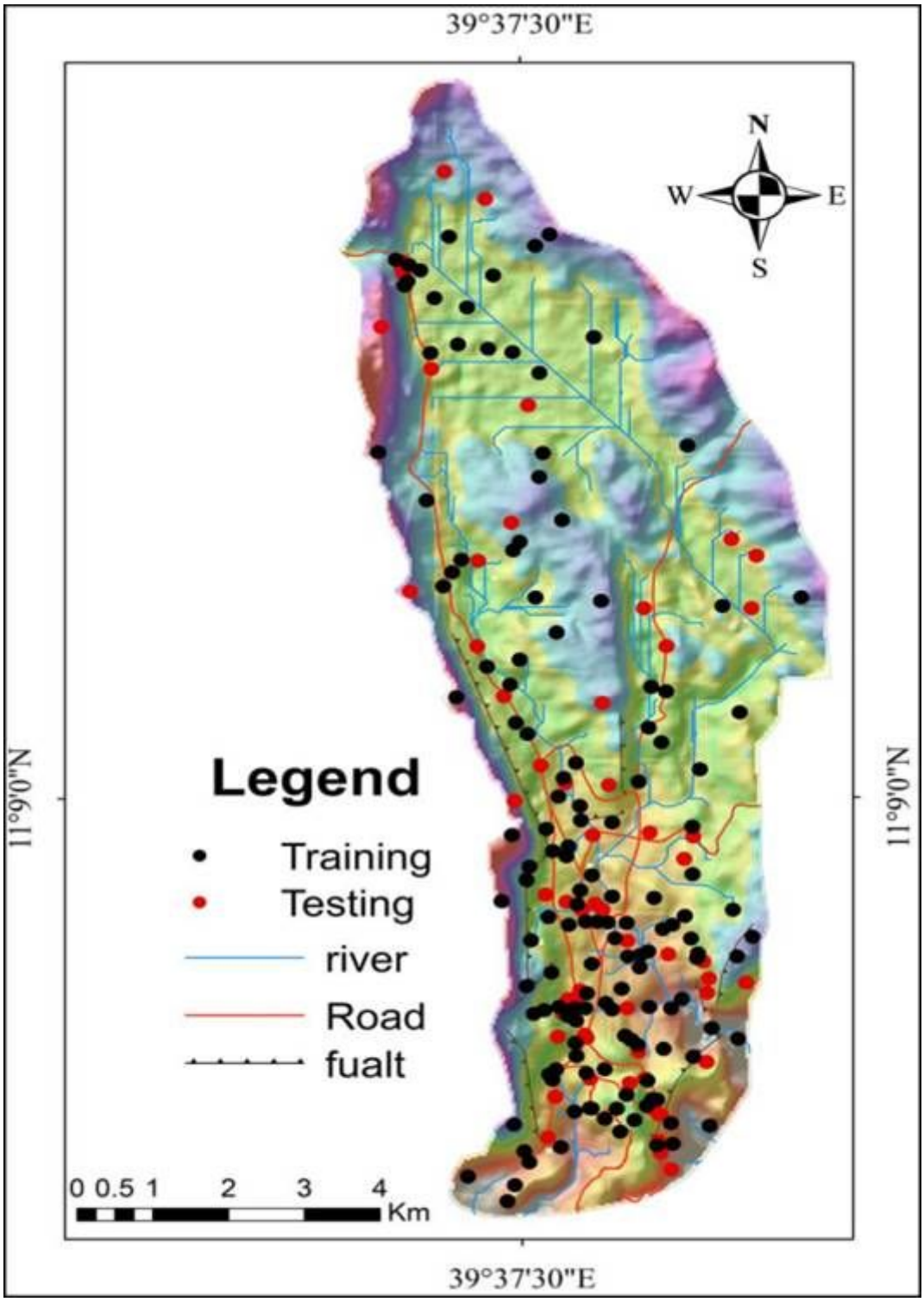

Figure 3

Landslide inventory map of the Dessie area

Figure 4

Landslide influencing factors maps: A) lithological map B) Curvature map, C) land use map, D) Distance from River map

landslide influencing factors maps: E) slope aspect map, F) slope map G) Distance to fault map, H) Distance to road map

landslide influencing factors map: I) Elevation

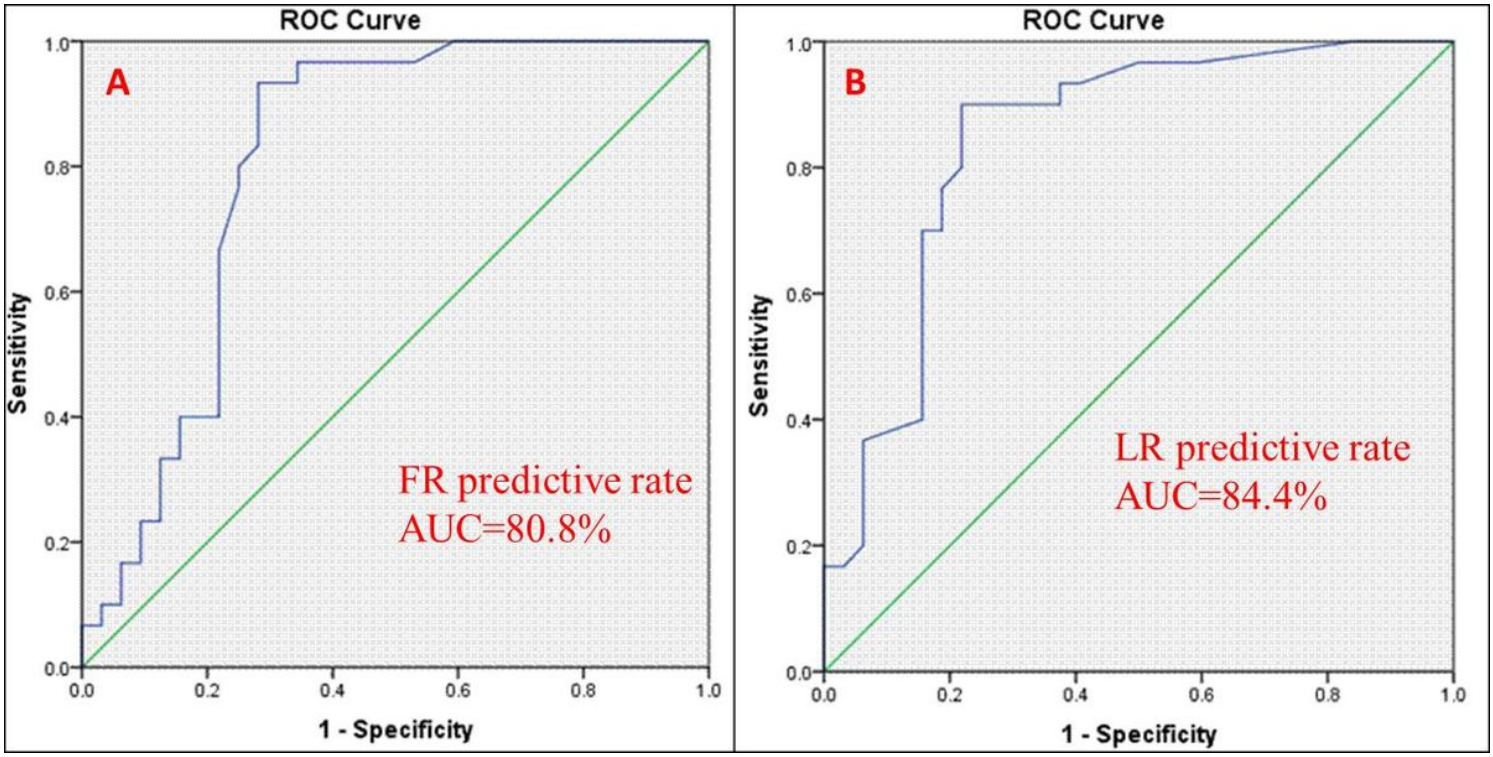

Figure 5

predictive rate curves from Frequency ratio (A) and Logistic Regression (B)

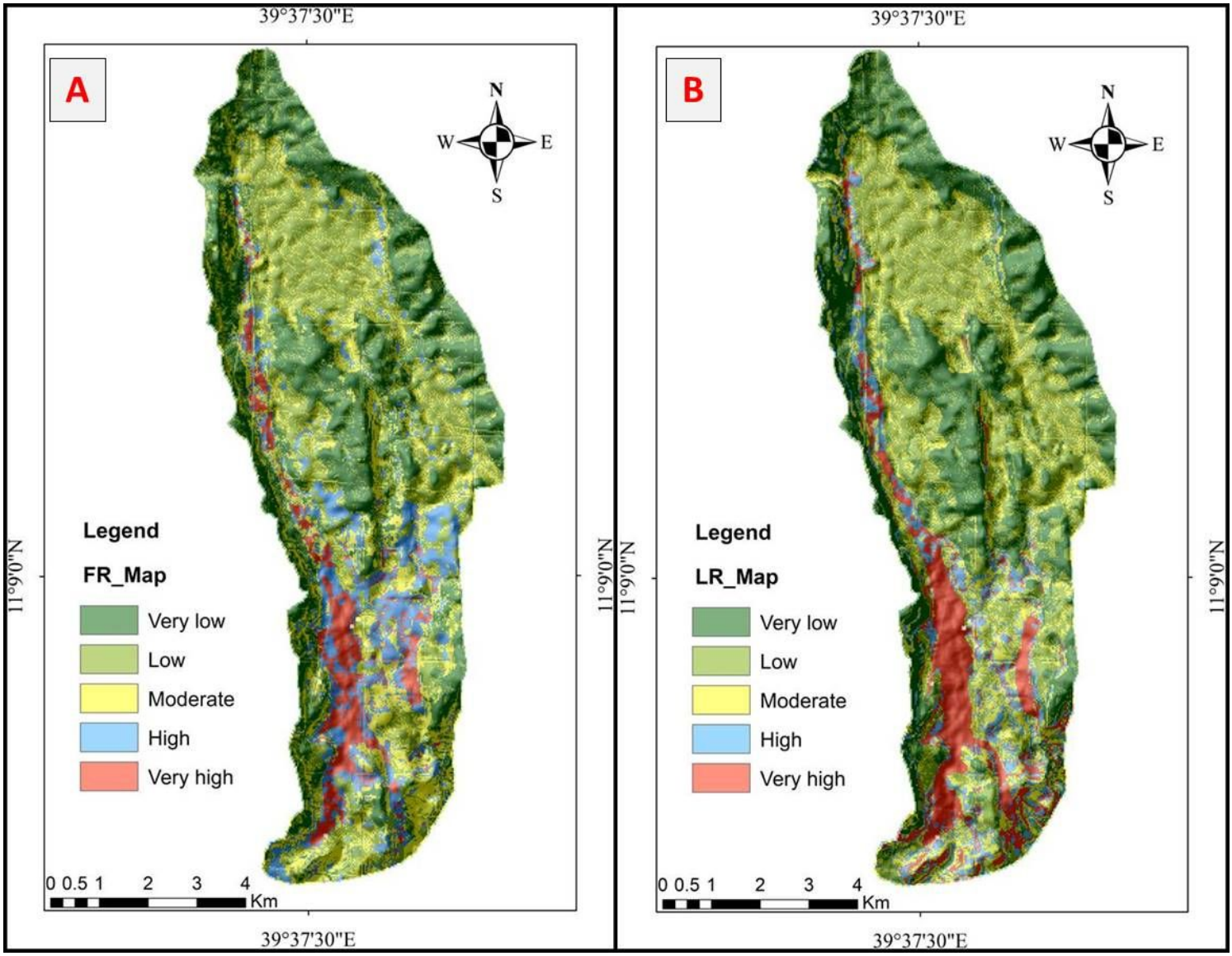

Figure 6

Figure show landslide susceptibility zonation maps A) frequency ratio map and B) logistic regression map