Landslide Susceptibility Mapping using Statistical Methods in Uatzau Catchment Area, Northwestern Ethiopia

Azemeraw Wubalem, alubelw@gmail.com

Engineering Geologist, College of Natural and Computational Sciences, Department of Geology, University of Gondar, Gondar, Ethiopia

Abstract: landslide susceptibility mapping is important to hazard management and to have planning development activities in the mountainous country like Ethiopia. In the present study, the landslide susceptibility mapping of the Uatzau basin is made using certainty factor, information value and logistic regression methods. Preparation of landslide inventory map from detailed fieldwork and Google Earth image interpretation was the first activity. Thus, 514 landslides were mapped and out of which 490 (70 %) of landslides were randomly selected keeping their spatial distribution to build landslide susceptibility models while the remaining 155 (30 %) of the landslides were used to models validation. It is clear that the effectiveness of the landslide susceptibility model using GIS and statistical methods is depending on the selection of the causative factors, which have a prevailing effect on landslide occurrence. In this study, six factors including lithology, land use/cover, distance to stream, slope gradient, slope aspect, and slope curvature were the landslide factors that were evaluated. After preparation of these factor maps, the effects of them on slope instability was determined by comparing with landslide inventory raster map using GIS. Finally, the landslide susceptibility model for the Uatzau area was developed and validated using the receiver operating characteristics curve (ROC). The results of ROC showed that for landslide susceptibility map using frequency ratio model (FRM) with an AUC value of 0.8883 has the highest prediction accuracy of 88.83%. The landslide susceptibility map, which is produced using Certainty factor and information value methods also showed that 87.03 % and 84.83% of prediction accuracy respectively. Besides the prediction accuracy of the model, the success rate curve for all models was applied and the result showed that more than 80 % accuracy (i.e, 80.83% for the information value model, 87.19 % for the certainty factor model and 83.27 % for frequency ratio model). The present research finds out that all methods/ models, which have employed in this study showed that reasonably very good accuracy in predicting landslide susceptibility of the Uatzau area. Therefore, these landslide susceptibility maps can be used for regional land use planning and landslide hazard mitigation purposes.

Keywords: landslide; susceptibility; Geographic Information System (GIS); certainty factor; frequency ratio; information value; Ethiopia.
1. Background

Our planets are under the change throughout time. These dynamic natures of the Earth are due to the existence of the internal and external processes that resulted in a change in the Earth system. Landslide is one of the processes that bring change to Earth. This change is the result of the combined factors including natural and anthropogenic activities. It is caused thousands of victims and deaths, hundreds of billion dollars of damages and environmental losses every year (Aleotti and Chowdhury 1999; Gutiérrez et al. 2015; El Jazouli et al. 2019). It has occurred when the driving force exceeds the resistance force due to destabilization of natural soil or rock slopes because of natural and anthropogenic factors like improper land use practice, the presence of loss sediment, heavy and prolonged rainfall, highly weathered and fractured rocks, gully and riverbank erosion, earthquake, due to superficial soil-rock interfere and unplanned urban explanation.

Ethiopia is characterized by steep fragile topography; deep river incises valley, intensive gully and riverbank erosion, intercalation of massive rock unit with loss rock unit, loss sediments, heavy rainfall, shallow groundwater depth, improper land use practice, frequent fault escarpment, and active geodynamic process. These conditions are favorable for the occurrence of various natural hazards including flooding, earthquake and landslide hazard.

As a summary of (Ayalew 1999; Temesgen et al. 2001; Woldearegay, 2008; Ibrahim 2011), landslides in Ethiopia have resulted in a loss of human and animal lives, damages in infrastructures and properties in the last five decades. From 1960 – 2010, alone, about 388 people died, 24 people injured, a wide area of cultivated and non-cultivated land, environment, infrastructure, and houses was affected.

From 2018 – 2019, rainfall triggered landslides also caused 60 people to died, 30 people were injured, 5,091 households were displaced, houses were damaged, and a widely cultivated and non-cultivated land was destructed in different parts of the country (Wubalem and Meten 2019). Although the landslide problem is critical in Ethiopia, still there is no adequate slope stability assessment in the different parts of the country (Wubalem and Meten 2019). The study area is one of the areas that was recently affected by the landslide incidences and so far, the area has not been studied. This contribution provides the originality of this study. Landslide in the study area has resulted in the devastation of houses, gravel roads, farmlands and loss of animal and human lives.

Therefore, landslide susceptibility mapping and assessment in this area can be provided useful information that helps us to disaster loss reduction and serve as a guideline for sustainable land use planning.

Landslide susceptibility is the likelihood of a landslide occurrence in an area depend on the terrain condition (Brabb 1984). It is an estimate of where landslides will be occurred. Landslide susceptibility mapping is not only to ascertain the factors that are most influential to the landslides occurred in the region but also to estimate the relative contribution of each factor for slope failures (Chen and Wang 2007). It also helps to establish a relationship between the factors and landslides to predict the landslide hazard in the future (Chen and Wang 2007). Before nowadays, because of the lacks of remotely sensing data and advancements of GIS tool, landslide susceptibility mapping has been difficult tasks, however at the present day, preparation of landslide susceptibility map is easy due to the advancement of computers, remote sensing and GIS (Jia et al. 2010; Karimi et al. 2010; Wang et al. 2011; Pradhan et al. 2011; Bednarik et al., 2012). Although several approaches are developed for landslide susceptibility mapping, generally they can be categorized into (1) deterministic (or engineering or geotechnical), (2) the heuristic (or index) (3) the statistical methods (Varnes 1984, Regmi et al. 2014) and
(4) machine learning methods or data mining methods. In the statistical approach, multivariate and bivariate statistical techniques are most widely used throughout the world and provides reliable results (Dai and Lee 2002; Donati and Turrini 2002; Ayalew and Yamagishi 2005; Duman et al. 2006; Sakar et al. 2013; Meten et al. 2015; Chandak et al. 2016; Zhang et al. 2017; Kouhpeima et al. 2017; Wubalem and Meten 2020). Certainty factor is one of the probability bivariate statistical methods, which can be provided reliable results and help to determine the correlation between landslide factor and landslide occurrence (Kanungo et al. 2011; Pourghasemi et al. 2012a; Sujatha et al. 2012; Pourghasemi et al. 2013c; Liu et al. 2014). Frequency ratio model is one of the bivariate statistical methods which is easy and provide reliable models (Chung and Fabbri 2003, 2005; Lee and Pradhan 2006, 2007; Akgu’n et al. 2008; Pradhan et al. 2010c, 2011, 2012; Meten et al. 2015). Another commonly practiced method in landslide susceptibility mapping is an information value method which is easily operated and provided reliable results (Saha et al. 2005; Sarkar et al. 2006; Kanungo et al. 2009; Wubalem and Meten 2019).

The Uatzau basin is one of the areas characterized by populated settlements and intensive farming. Thus, to maintain the smooth operation of agricultural activity as well as to minimize the losses of life and property, landslide susceptibility assessment is essential for this basin. In this case, frequency ratio (FR), certainty factor (CF) and information value model have been used to develop the landslide susceptibility map of the study area. These methods are easy to apply and it gives a very well-meaning result. In literature, various bivariate approaches for landslide susceptibility mapping are available, however, a comparison among CF, FR, and IV models yet have not been encountered. Therefore, a comparison of the results of these methods is discussed in this paper.

2. Study Area

The study area is located 5 km far from Debre Markos town in northwestern highlands of Ethiopia. It lies within the latitude, 117, 215 N to 1, 138, 231 N and the longitude 349, 253 E to 364, 786 E. The study area covers an area of about 138 km². The minimum and maximum altitudes of the area are 1332 m at the river gorge and 2, 498 m in hills and plateau lands (Fig. 1). Many tributaries are available in the entire study area and joined the Uatzau River, which drains into the Abay River. The various streams in the area are caused to remove soil through stream bank erosion. The study area is characterized by variable topographic conditions including ridge, cliff, hill, plateau, deep River gorge, and gentle slope. These fragile nature of topography have been facilities the rate of soil erosion. 54 % of this region is covered by agricultural lands and rocky lands/bar lands, Residential, and Grazing lands cover remaining lands. Tropical to subtropical climatic condition prevails in the study area. The main characteristic of the climates in the study area are the monsoon rainfall, which occurs between June and September and delivers an average of 90% of the total rainfall of the year.

The study area comprised mainly two lithological units besides recent soil sediments at the slope toe of the study area, which is grouped in early Mesozoic and Cenozoic Era of sedimentary (sandstone) and volcanic rocks (basalt). The sedimentary rock is red sandstone, which is characterized by highly wreathed, closely spaced vertical and horizontal joints. These joints used as a conduit for the rainfall water to infiltrate into the ground that helps to raise the groundwater level and cause exposure to the surface at the hillside of the sandstone lithology. The vertical joints are characterized by wide aperture, not infilled with secondary material and have rusty color along the discontinuity. The horizontal joints are run as perpendicular to the vertical joints. The volcanic rock is underlined the highly fractured sandstone and covered by thick dark color soil deposit. This unit is characterized by a high degree of weathering and fracturing. At the contacts of the basalt rock and
sandstone as well as the interfere of the weathered and fractured basalt with dark color soil, spring waters are emanating. At the toe of cliff basaltic rock slope, very loss /unconsolidated soil deposit is formed due to slope failure and gravity effects. In this soil deposit, unplanned intensive agricultural activities are common.

3. Methodology

In the present research work, various activities and steps are employed in order to achieve the goal of the present research work. These are data collection, landslide inventory mapping, landslide factor evaluation and mapping, landslide susceptibility modeling and model validation.

3.1 Data Collection

The reliable data for the generation of the landslide susceptibility map/model were collected and extracted from various data sources. Slope angle, slope aspect, slope curvature and distance to stream are derived from DEM of SRTM. The geological map of the study area is prepared and digitized from 1:50,000 Geological maps and reports of the Geological Survey of Ethiopia. Land use/land cover map is digitized from Google Earth.

3.2 Landslide Inventory Mapping

In landslide susceptibility mapping, landslide inventory mapping is one of the key element, which can be prepared using various techniques like the aerial photograph or Google Earth image interpretation, field investigation, and evaluation of archived data coupled with GIS tool. In the present research work, the landslide location is gathered and evaluated using
detail and extensive Google Earth image interpretation and fieldwork. In the study area, 514 landslides were determined and digitized into polygons using GIS tool with the help of Google Earth, finally, landslide inventory map is produced (Fig. 2).

3.3 Evaluation of Landslide Factors

In landslide susceptibility mapping, the selection of landslide factors is one of the most important elements. However, there is no well-defined standard to select the most significant landslide factors. The factors that initiate the landslide incidence in the study area are selected based on literature review, local people interview and field evaluation. These are slope angle,
slope aspect, slope curvature, land use, lithology and distance to stream/river have been taken into account to examine the spatial relationship between them and landslide occurrence in the study area. Distance to stream (five classes), slope angle (five classes), slope aspect (ten classes), and slope curvature (three classes) maps have been constructed from SRTM DEM (Fig 3a, 3b, 3c, and 3d). The lithological map of the study area is prepared through digitization from 1:50,000 geological map of Geological Survey of Ethiopia, which has three classes (weathered basalt, sandstone, and unconsolidated/colluvial sediments). Land use map of the study area is digitized using high-resolution Google Earth image interpretation that has five classes such as grazing land, cultivated land, bare land, residential, and scatters bush (Fig 3e). Even though rainfall is one of the factors that can be triggered landslide incidence, it is not included in this landslide susceptibility modeling because of the lack of rain gage station in the target area. In order to determine the effects of each landslide factor class on landslide occurrence, weight rating through landslide factor raster combined with landslide raster map is important. For this purpose, all landslide factor maps are converted into raster and reclassified with the same pixel size (30m x 30m) and the same projection using GIS tool under the Arc toolbox in conversion as well as a spatial analysis tool. Then, the landslide inventory raster map is overlaid through combine in spatial analysis tool under local toolbox with landslide factor raster class to extracted landslide pixels for each landslide factor class. Then the effects of each factor class were determined using the equation of frequency ratio, certainty factor and information value methods and the results are summarized in Table 1.

Figure 3 landslide Factors a) Slope b) Aspect c) Curvature d) Distance to stream e) Land use f) Lithology
3.5 Modeling Approaches

3.5.1 Landslide Susceptibility mapping using Frequency Ratio Model

It is one of the bivariate probability methods, which is applicable to determine the correlation between landslide occurrence and landslide causative factor classes. The frequency ratio is the ratio of areas where the landslide occurred to the areas of the landslide factor class. When the ratio value is greater than one, it indicates the higher correlation between factor class and landslide occurrence in a given terrain, however, the ratio value less than one is indicated that low coloration between landslide occurrence and landslide factors, which means low probability of landslide occurrence (Bonham-Carter, 1994; Lee and Talib 2005). It can be calculated using Eq. 1.

$\text{FR} = \frac{a}{b}$

Where FR is frequency ratio, a is area where landslide affects and b the area where the landslide occurred in landslide factor class. In the present research work, the frequency ratio for each causative factor class was calculated using equation one and the results are summarized in Table 1.

After calculation of the frequency ratio for each landslide factor class using Microsoft excel and GIS, the frequency ratio value for each factor class is assigned through the join in the ArcGIS tool. Then the weighted landslide factors are rasterized using the lookup tool in spatial analysis. Landslide susceptibility index is indicated the degree of susceptibility of the area for landslide occurrence. The landslide susceptibility index (LSI) of the study area is calculated by carefully sum up the weighted rasterized factor raster maps using equation 2 by rater calculator in Map Algebra of the spatial analysis tool. In order to get the landslide susceptibility index, the frequency ratio of each factor type or class were summed as in Eq. 2

$I = \sum \text{FR}$

Where LSI is the landslide susceptibility index and FR is the frequency ratio of each landslide factor type or classes. After landslide susceptibility index calculation, the index values were classified into a different level of landslide susceptibility zones using natural breaks in the GIS tool. The higher the value of landslide susceptibility index, the higher the probability of landslide occurrence but the lower the LSI is indicated, the lower the probability of landslide occurrence.

Based on natural break classification, the landslide susceptibility map of the study area has five classes such as very low, low moderate, high and very high landslide susceptibility class (Fig 5a).

3.5.2 Landslide Susceptibility mapping using Information Value Model

The information value method is one of the probability methods of a bivariate statistical method, which is used to envisage the correlation between landslides and landslide factor classes (Sarkar et al. 2006).

The information values for each factor class have been determined through the combination of reclassified landslide raster to reclassified landslide factor raster based on the presence of landslide in a given map unit (Fig 5c). These values are important to define the role of each causative factor in classes for landslide occurrence (Kanungo et al., 2009). This can be calculated as in Eq.3.

$\text{IV} = \ln \left( \frac{\text{Conditional probability}}{\text{Prior probability}} \right)$

Where IV is the information value.
Where Conditional probability is the ratio of the pixel of a landslide in class to the pixel of a class and prior probability is the ratio of the total number of pixels of landslide to the total number of pixels of the study area. When the IV > 0.1, the landslide occurrence with the factor classes have a high correlation, means it will have a high probability of landslide occurrence however when the IV < 0.1 or IV < 0, it is low coloration between landslide factors and landslide occurrence which indicated a low probability of landslide occurrence.

After calculation of the information value for each landslide factor class using Microsoft excel and GIS, the information value for each factor class is assigned through the join in the ArcGIS tool. Then, the weighted landslide factors are rasterized using the lookup tool in spatial analysis and the landslide susceptibility index (LSI) of the study area is calculated as in Eq. 4.

\[
LSI = \sum IV
\]

\[
LSI = IV \times \text{Slope raster} + IV \times \text{Slope aspect raster} + IV \times \text{Slope curvature raster} + IV \times \text{Lithology raster} + IV \times \text{Land use raster} + IV \times \text{Distance to stream raster}
\]

Where LSI is landslide susceptibility index and IV is the information value of each factor class. The higher value of LSI has indicated the higher probability of landslide occurrence.

3.5.3 Landslide Susceptibility mapping using Certainty Factor Model

The certainty factor is one of the probability methods that widely used to landslide susceptibility mapping for different data (Kanungo et al. 2011; Sujatha et al. 2012; Pourghasemi et al. 2013c; Liu et al. 2014). Shortliffe and Buchanan (1975) were proposed the certainty factor (the probability function) for landslide susceptibility mapping later Heckeman (1986) has been improved and it expresses mathematically as:

\[
CF = \frac{(PP_a - PP_b)}{(PP_a - PP_b)} \text{ if } PP_a \geq PP_b
\]

\[
CF = \frac{(PP_a - PP_b)}{(1 - PP_b)} \text{ if } PP_a < PP_b
\]

Where \(PP_a\) is the conditional probability of landslide in the defined area \(a\) and \(PP_b\) is the prior probability of landslide in the defined entire study area \(b\). The CF value ranges from -1 to 1, a positive value indicates increasing certainty of landslide occurrence and a negative value indicates decreasing of certainty of landslide occurrence. If the certainty value is closing to zero, it means there is no adequate information about the relation between landslide factor classes and landslide occurrence, therefore, it is difficult to give any certainty of landslide occurrence (Sujantha et al., 2012; Dou et al., 2014). After the calculation of CF for each landslide factor class, the landslide susceptibility index (LSI) is determined as in Eq. 6.

\[
LSI = \sum CF
\]

\[
LSI = CF \times \text{Slope raster} + CF \times \text{Slope aspect raster} + CF \times \text{Slope curvature raster} + CF \times \text{Lithology raster} + CF \times \text{Land use raster} + CF \times \text{Distance to stream raster}
\]

Where LSI is the landslide susceptibility index, and CF is the certainty factor.

3.6 Model Validation

Landslide susceptibility map without validation has no sense in the scientific world (Wubalem and Meten 2019). Therefore, validation of the landslide susceptibility model is very important to evaluate the degree of accuracy of modeling using different validation techniques (Gorsevski et al 200; Chung and Fabbri 2003). For this purpose, the landslide area has been
classified based on time, space and random partition (Chung and Fabri, 2003, Lee and Pradhan, 2007, and Meten et al., 2015).

As stated by Yesilnacar and Topal (2005), the area under the curve (AUC) value is used to evaluate the performance of the model and its value range from 0.5 – 1. When the AUC value is in between the range of 0.9 – 1, the model has excellent performance; if AUC value is in between the range of 0.8 – 0.9, the model has very good performance. If the AUC value is in between the range of 0.7 – 0.8, the model has good performance. If AUC value is between the range of 0.6 – 0.7, the model has an average performance. However, if AUC value is between the range of 0.5 – 0.6 and equal to 0.5 or less than 0.5, the model has poor performance (Yesilnacar and Topal 2005).

In the present work, the landslide area was randomly classified as 70 % landslides for training and 30 % landslides for model validation by keeping their spatial distribution into the account using the random partition technique (Chung and Fabri, 2003, Meten et al., 2015). After model development, the models were validated by Receiver Operating Characteristics (ROC) curves.
4. Results and Discussion

4.1 Frequency Ratio (FR)

The frequency ratio for all landslide factor classes was rating and show important effects of each factor class on slope instability. As it can be observed from Table 1, the lithology class colluvial deposit and weathered basalt have a high value of FR (1.3 and 1.1 respectively) which is > 1, indicating high landslide probability but sandstone class has low FR value (0.6) which is < 1, indicating a low probability of landslide occurrence.

Because the colluvial deposit is a recent deposit in the study area which is characterized by loose/ unconsolidated, low shear strength nature and a series of spring water. The presence of spring water has been reducing the normal force in slope material when the pore space in the soil grain filled with water, it will be generated pore water pressure.

Besides this, a series of the stream is passed through the slope toe of this loose soil deposit which caused the removal of the slope toe by the stream bank erosion. This resulted in the reduction of resisting force in slope material when the slope toe is eroded. As we know that landslide may have occurred when driving force exceeds the resisting force in the slope material. This is happening due to various constraints. In this research case, slope toe erosion by a stream is the key element to driven landslide incidence in the study area.

Basalt rock has a high positive relation to landslide occurrence than sandstone due to the effects of weathering because basalt rock in the study area is highly affected by weathering but sandstone has a low degree of weathering because of the presence of quartz cement.

As designed in Table 1, the slope class 0° - 7°, and 7° – 14° have low FR value (0.89 & 0.76 respectively) and high value of FR (1.04, 1.3, & 2.09) for slope classes 14° – 21°, 21° – 28°, 28° – 68°, respectively. This correlation indicated that landslide probability increase as the slope gradient increases (Sun, 2009), however, it may not be always true when the steep slope comprised of massive and strong slope material. Landslide may have occurred in a gentle slope when the slope material is loose and the slope subjected for modification due to anthropogenic and natural activity (Wubalem & Meten, 2019). However, in the present study area, the result of the FR value indicated in Table 1 and Fig 6 as the slope angle increased, landslide probability is increased. This is because of the presence of shallow loose soil deposit, highly weathered rock, active soil erosion, and improper land use practice.

In the case of slope aspect factor class, the FR value is > 1 for south-facing (1.33), southwest facing (1.68) and west-facing (1.41), indicating high landslide probability. However, the remaining slope aspect classes have FR value < 1, indicating a low probability of landslide occurrence.

The FR value for slope curvature class of -26 - -2 (1.42) & 2 – 23 (1.32) is > 1, indicating high landslide probability. This is because of the effects of slope shape for rainwater impounding and gravity effect. However, the slope curvature class -2 – 2 has FR value (0.85) is < 1, indicating a low probability of landslide occurrence.

In the case of distance to stream, as designated in Table 1, as a distance to stream increase, the probability of landslide occurrence decreased. At a distance of 0 – 50m, 50 – 100m, and 100 – 150m, the value of FR (1.2) is > 1, indicating high landslide probability, however, at a distance > 150m, the value of FR is < 1, indicating low landslide probability. This is because of the effects of slope modification, gully erosion, riverbank erosion and river undercutting.
As noticed in Table 1, the value of FR for land use/cover class of agriculture land (1.1) and bar land (10.7) is > 1, indicating high landslide probability. This is because the cultivated land has increased soil moisture. Whenever the soil moisture increased in the slope, the weight of slope material and the pore water pressure in the slope material are increased in parallel. This could have resulted in a reduction in the normal force in the soil mass. This leads to slope failure when the driving force exceeds a resisting force.

In the case of barland class, FR value has shown a higher correlation to the probability of landslide occurrence. Because barland in the study area is highly affected by a gully soil erosion, which caused a reduction of shear strength of soil material. The remain classes including settlement and grazing land have FR value < 1, indicating a low probability of landslide occurrence. Because settlement and grazing land in the study area has been practicing in gentle slope gradient parts of the study area.

4.2 Certainty Factor (CF)

The certainty factor rating for different landslide factor classes has been calculated by overlay landslide raster with landslide factor raster layer and it shows the important effects of each factor class on slope instability. As designated in Table 1, the lithology class such as colluvial deposit and weathered basalt have a positive and high value of CF (0.24 and 0.11 respectively), indicating high landslide probability but sandstone class has negative CF value (-0.4) which indicating a low probability of landslide occurrence.

As observed in Table 1, the slope class 0 - 7°, and 7° – 14° have negative CF value (-0.11 & -0.25 respectively) is indicating low landslide probability and positive value of CF (0.04, 0.24 & 0.54) for slope classes, 14° – 21°, 21° – 28°, and 28° – 68°, respectively is indicating high landslide probability.

In the case of slope aspect factor class, the CF value is positive for south-facing (0.25), southwest facing (0.42) and west-facing (0.3), indicating high landslide probability. However, the remaining slope aspect classes have negative CF value, indicating a low probability of landslide occurrence.

The CF value for slope curvature class of -26 - -2 (0.31) & 2 – 23 (0.25) is positive, indicating high landslide probability. However, the slope curvature class -2 – 2 has a negative CF value (-0.16), indicating a low probability of landslide occurrence.

At a distance of 0 – 50m and 100 – 150m, the value of CF (0.19 and 0.25) is positive, indicating high landslide probability, however, at a distance 50 – 100m and > 150m, have negative value, indicating low landslide probability.

As noticed in Table 1, the value of CF for land use/cover class of agriculture land (0.07) and bar land (0.91) is positive, indicating high landslide probability. The remaining factor classes as settlement, scatter bush and grazing land have negative CF value, indicating a low probability of landslide occurrence.

4.3 Information Value (IV)

The information value rating for different landslide factor classes has been calculated by overlay landslide raster with landslide factor raster layer and it shows the important effects of each factor class on slope instability. When the IV value is > 0.1, the given factor class will have a positive correlation for landslide occurrence but the IV < 0.1 indicates a low probability of landslide occurrence. As designated in Table 1, the IV is > 0.1 for lithology class such as colluvial deposit.
and weathered basalt (0.27 and 0.12 respectively), indicating high landslide probability but the IV < 0.1 for sandstone class (-0.5) which indicates a low probability of landslide occurrence.

As observed in Table 1, the IV < 0.1, for slope class 0° - 7°, 7° – 14° and 14° - 21° (IV = -0.12, -0.28 & 0.04 respectively) is indicating low landslide probability and IV > 0.1 for slope classes, 21° – 28° and 28° – 68°, respectively (IV = 0.27 & 0.74) is indicating high landslide probability.

In the case of slope aspect factor class, the IV > 0.1 is for south facing (IV = 0.28), southwest facing (IV = 0.34), indicating high landslide probability. However, IV < 0.1 for the remaining slope aspect classes are indicating a low probability of landslide occurrence. The IV > 0.1 for slope curvature class of -26 - -2 (IV = 0.35) & 2 – 23 (IV = 0.28) is indicating high landslide probability. However, the IV < 0.1 for the slope curvature class -2 – 2 (IV = -0.17) is indicating low probability of landslide occurrence.

At a distance of 0 – 50m and 100 – 150m, the value of IV is > 0.1 which is 0.2 and 0.16, indicating high landslide probability, however, at a distance 50 – 100m and > 150m, the IV < 0.1, indicating low landslide probability. As noticed in Table 1, the value of IV for land use/cover class of agriculture land (0.07) and bar land (0.91) is > 0.1, indicating high landslide probability. The IV for the remaining factor classes like settlement, scatter bush and grazing land is < 0.1, indicating a low probability of landslide occurrence.

Table 1 Weight rating of landslide factor class

| Factors                  | Class                  | Class pixel | % Class Area (CA) | Landslide pixel | % landslide in class (LA) | FR | CP | PP | CF | IV  |
|--------------------------|------------------------|-------------|-------------------|-----------------|--------------------------|----|----|----|----|-----|
| Lithology                | Colluvial Deposit      | 39011       | 25.2              | 524             | 32.9                     | 1.3| 0.01| 0.01| 0.24| 0.27|
|                          | Sandstone              | 51186       | 33.1              | 320             | 20.1                     | 0.6| 0.01| 0.01| -0.40| -0.50|
|                          | Weathered Basalt       | 64149       | 41.7              | 748             | 47.0                     | 1.1| 0.01| 0.11| 0.12| 0.12|
| Land use/cover           | Agriculture            | 83991       | 54.3              | 375             | 8.4                      | 1.1| 0.004| 0.04| 0.07| 0.1  |
|                          | Settlement             | 43115       | 27.9              | 133             | 4.3                      | 0.7| 0.003| 0.04| -0.26| -0.3 |
|                          | Grazing Land           | 25490       | 16.5              | 76              | 2.6                      | 0.7| 0.003| 0.04| -0.29| -0.3 |
|                          | Bar Land               | 1364        | 0.9               | 61              | 0.1                      | 10.7| 0.045| 0.04| 0.91| 2.4  |
|                          | Scatter Bush           | 616         | 0.4               | 0               | 0.1                      | 0.0| 0.000| 0.04| -1.00| -1.00|
| Distance to stream       | 0 – 50                 | 47432       | 30.7              | 2115            | 37.5                     | 1.2| 0.04| 0.19| 0.20| 0.20 |
|                          | 50 – 100               | 42442       | 27.4              | 1487            | 26.4                     | 1.0| 0.04| 0.04| -0.03| -0.04|
|                          | 100 – 150              | 33253       | 21.5              | 1416            | 25.1                     | 1.2| 0.04| 0.15| 0.16| 0.16 |
|                          | 150 – 200              | 22225       | 14.4              | 469             | 8.3                      | 0.6| 0.02| 0.43| -0.55| -0.55|
|                          | 200 – 2, 355           | 9270        | 6.0               | 149             | 2.6                      | 0.4| 0.02| 0.57| -0.82| -0.82|
| Slope angle              | 0° – 7°                | 48562       | 31.4              | 1577            | 28.0                     | 0.89| 0.03| 0.04| -0.11| -0.12|
|                          | 7° – 14°               | 48082       | 31.1              | 1329            | 23.6                     | 0.76| 0.03| 0.04| -0.25| -0.28|
|                          | 14° – 21°              | 29704       | 19.2              | 1124            | 19.9                     | 1.04| 0.04| 0.04| 0.04| 0.04 |
|                          | 21° – 28°              | 19003       | 12.3              | 899             | 16.0                     | 1.30| 0.05| 0.24| 0.26| 0.26 |
|                          | 28° – 68°              | 9271        | 6.0               | 707             | 12.5                     | 2.09| 0.08| 0.54| 0.74| 0.74 |
| Flat (-1)                |                        | 37796       | 24.4              | 1143            | 20.3                     | 0.83| 0.03| 0.04| -0.18| -0.19|
| North (0 - 22.5)         |                        | 2094        | 1.4               | 23              | 0.4                      | 0.30| 0.01| 0.04| -0.71| -1.20|
| NE (22.5 - 67.5)         |                        | 17652       | 11.4              | 388             | 6.9                      | 0.60| 0.02| 0.04| -0.41| -0.51|
| E (67.5 - 112.5)         |                        | 6855        | 4.4               | 146             | 2.6                      | 0.58| 0.02| 0.04| -0.42| -0.54|
| SE(112.5 - 157.5)        |                        | 6756        | 4.4               | 187             | 3.3                      | 0.76| 0.03| 0.04| -0.25| -0.28|
| Slope Aspect             | S (157.5 - 202.5)      | 26889       | 17.4              | 1299            | 23.0                     | 1.33| 0.05| 0.25| 0.28| 0.28 |
4.4 Landslide Susceptibility Mapping

After the calculation of the landslide susceptibility index, it is important to classify the LSI into different susceptibility classes based on the LSI value. The landslide susceptibility index map of the study area of the information value method, certainty factor method and frequency ratio method is classified into five levels of susceptibility classes using the natural break method in ArcGIS 10.1. Using natural breaks method in ArcGIS 10.1, the landslide susceptibility map generated with the information value model was reclassified into five classes like very low, low, moderate, high and very high landslide susceptibility classes (Fig 5c). From the results of the analysis (Table 2), it was found that 15.5 % and 24.3 % of the study area falls in very low and low susceptibility classes. Moderate, high and very high landslide susceptibility classes have comprised 31.5 %, 21.1 % and 7.6 % of the study area, respectively. As designated in table 2, 6.3 % and 11.1 % of landslides fall in very low and low susceptibility classes of the study area, respectively. The remaining 23.8 %, 31.8 % and 26.3 % of landslides fall in moderate, high and very high landslide susceptibility classes.

Landslide susceptibility map produced using certainty factor model (Table 2), very low and low susceptibility classes cover 17.8 % and 31.0 % of the total study area, however, 28.8 %, 19.0 % and 3.4 % of the total area fall in moderate, high and very high landslide susceptibility classes respectively. As indicated in Table 2, 4.7 % and 12.3 % of landslide fall in very low and low susceptibility classes of the study area, respectively. The remaining 17.5 %, 34.8 % and 30.7 % of landslides fall in moderate, high and very high landslide susceptibility classes, respectively.

As it can be observed from Table 2, the landslide susceptibility map produced using frequency ratio model, very low and low landslide susceptibility classes cover 22.7 % and 30.8 of the total area, however, 22.4 %, 19.3 % and 4.8 % of the total area fall in moderate, high and very high landslide susceptibility classes respectively. As designated in Table 2, 5.4 % and 14.7 % of landslide fall in very low and low susceptibility classes of the study area, respectively. The remain 19.5 %, 43.7 % and 16.7 % of landslides fall in moderate, high and very high landslide susceptibility classes, respectively.
Table 2 Landslide Susceptibility percent area, percent of both validation and training landslide in each susceptibility class and AUC value

| Information value method | LSI | LSM  | % Class area | %VLSA | % TLSA | VLS_AUC  | TLS_AUC  | TLS_AUC |
|--------------------------|-----|------|--------------|-------|--------|----------|----------|---------|
|                          | -0.5 - 0.9 | VLS  | 15.5         | 3.9   | 6.3    |          |          | 0.808265|
|                          | 0.9 - 1.5   | LS   | 24.3         | 7.9   | 11.8   |          |          |         |
|                          | 1.5 – 2     | MS   | 31.5         | 20.1  | 23.8   | 0.848323 |          |         |
|                          | 2.0 - 2.6   | HS   | 21.1         | 40.8  | 31.8   |          |          |         |
|                          | 2.6 - 4.1   | VHS  | 7.6          | 27.3  | 26.3   |          |          |         |

| Certainty Factor (CF)   | -2.2 - -0.97| VLS  | 17.8         | 4.7   | 6.0    |          |          | 0.871933|
|                        | -0.97 - -0.47| LS   | 31.0         | 12.3  | 16.4   |          |          |         |
|                        | -0.47 - 0.04| MS   | 28.8         | 17.5  | 24.8   | 0.870348 |          |         |
|                        | 0.04 - 0.74 | HS   | 19.0         | 34.8  | 33.0   |          |          |         |
|                        | 0.74 - 2.61 | VHS  | 3.4          | 30.7  | 19.7   |          |          |         |

| Frequency Ratio (FR)    | 3.1 - 4.3   | VLS  | 22.7         | 5.4   | 9.3    |          |          | 0.832718|
|                        | 4.3 - 4.8   | LS   | 30.8         | 14.7  | 17.8   |          |          |         |
|                        | 4.8 - 5.3   | MS   | 22.4         | 19.5  | 20.0   | 0.888337 |          |         |
|                        | 5.3 – 6     | HS   | 19.3         | 43.7  | 35.0   |          |          |         |
|                        | 6 - 7.7     | VHS  | 4.8          | 16.7  | 17.8   |          |          |         |

LSI is landslide susceptibility index, LSM is landslide susceptibility map, VLSA is validation landslide area, VLS is validation landslide, TLSA is training landslide area, VLS is very low susceptibility, LS is low susceptibility, MS is moderate susceptibility, HS is high susceptibility, VHS is very high susceptibility and AUC is area under the curve.

4.5 Model Validation

Model validation is the last step in landslide susceptibility mapping, which helps to evaluate the accuracy of the model, generated using different statistical methods. Various model validation techniques are available like success and predictive rate curve, landslide relative density index (\( R – index \)), receiver operating characteristics curve (ROC), and area under the curve (AUC). However, in the present research work, receiver operating characteristics curve and area under the curve are used to evaluate the accuracy of landslide susceptibility model generated by frequency ratio, certainty factor, and information value methods. In this case, the landslide in the study area was classified into a 70 % (499) training data set and 30 % (155)
validation data set randomly keeping their spatial distribution. Then, the ROC and AUC are calculated by comparing the training and validation landslide rater with landslide susceptibility classes. Then, the pixel of landslide for each landslide susceptibility class was extracted. Finally, the ROC and AUC were generated using Real statistics software in Excel. As the results of the analysis shown in Fig 6 and Table 2, the closer the ROC curve to the left of the top of the curve, indicating the higher the accuracy of the model. The accuracy of the models also evaluated using AUC which ranges from 0.5 – 1. The closer the AUC value to one, the higher the model accuracy. As indicated in Table 2, the AUC value is closer to one, indicating the higher accuracy of the model. As shown in Table 2, the AUC value for CF is 0.870 and 0.872 of validation and training data set respectively. This means more or less the AUC value for two data set is indicated closer to each other. In the case of FR, the AUC value is 0.888 and 0.833 for the predictive rate curve of validation landslide data sets and the success rate curve for training landslide data sets. The AUC value for IV is 0.848 and 0.808 for the predictive rate curve of validation landslide data sets and the success rate curve for training landslide data sets. These results indicated that the FR, CF and IV models have successfully estimated the landslide susceptibility classes of the region and these models have been employed in this study to have reasonable accuracy in predicting the landslide susceptibility classes of the study area. However, based on AUC value CF and FR models revealed that a better result than the IV model for landslide susceptibility mapping in the study area (Fig 6).

![Receiver operating characteristics curve (ROC)](image-url)

Fig. 6 Receiver operating characteristics curve (ROC)
5. Conclusion
In summary, the study area (Uatzau) is characterized by recent unconsolidated soil deposits, rugged topography, active gulley and riverbank erosion and improper land use practice which makes it very prone to different landslides including earth/debris flow, soil slide, and soil creep.

Frequency ratio (FR), Certainty factor (CF) and information value (IV) models were used to prepare landslide susceptibility maps of the Uatzau area with very low, low, moderate, high and very high susceptibility classes. From Frequency ratio (FR), Certainty factor (CF) and information value model, the weight/rating of each landslide factor classes were calculated for each factor classes that tell about the contribution of each factor classes on landslide occurrence when its value is greater than 1, positive and greater than 0.1 respectively.

Based on the results of information value, Certainty factor and frequency ratio model, those values which greater than 0.1, positive, and > 1, the landslide factor classes which cause slope instability problem include slope (> 14°), land use (bare land, and agriculture), lithology (colluvial soil and highly weathered basalt), distance to stream (0 – 50m and 100 – 150 m), Curvature (concave and convex), and slope aspect (south, southwest and west slope facing) were determined.

The accuracy of the landslide susceptibility models, which generated using IV, FR and CF, have been evaluated using the receiver operating characteristics (ROC) curve through comparison of training and validation landslide raster with the models. Then, the success rate curve and predictive rate curve of receiver operating characteristics (ROC) curves were developed as well as the area under the curve (AUC) for three models have been determined. The value of AUC for three models are closed to 1, indicating the very good accuracy of the models. Based on the AUC values, the frequency and certainty factor methods have better performance than the information value model. Therefore, the frequency ratio, certainty factor, and information value methods are important for landslide susceptibility mapping for areas, which have fragile topography. Geologists, engineers, and planners for land use planning and landslide hazard management can use these models.

Acknowledgments
First, I would like to thanks the almighty God who allowed me to accomplish this research work. Next, I would like to give my gratitude to my lovely wife Ms. Seble Asmare and my brother Mr. Zemene Wubalem for their valuable support and advice during this research work. I would like to give my special thanks to Dr. Veera and my friends for their continuous support during the research work. Finally, I would like to thank the University of Gondar for geological equipment; National Metrological Agency staffs, Geological Survey of Ethiopia, disaster head office, natural resource management head office and the rural community, for their valuable data.

Authors’ Contributions
All activities starting from conception and design of the work, the developments of the models as well as the statistical analysis and interpretations of the results were done by me.
Funding
It is not applicable in this case.

Available of data and material
All the datasets that have been used and analyzed during the current study is available from the corresponding author on reasonable request.

Competing interests
I have declare that there is no any competing interests.

Author details’
Department of Geology, University of Gondar, Ethiopia

Reference
1. Akgu’n A, Dag S, Bulut F (2008) Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. Environ Geol 54:1127–1143
2. Bednarik M, Yilmaz I, Marschalko M (2012) Landslide hazard and risk assessment: a case study from the Hlohovec–Sered’ landslide area in south-west Slovakia. Nat Hazards, doi:10.1007/s11069-012-0257-7
3. Bonham-Carter GF (1994) Geographic information systems for geoscientists. Modeling with GIS. Pergamon, Oxford, p 398
4. Brabb EE (1984) Innovative approaches to landslide hazard mapping. In: Proceedings 4th international symposium on landslides, Toronto 1: 307–324
5. Lee S, Talib JA (2005) Probabilistic landslide susceptibility and factor effect analysis. J Environ Geol 47:982–990
6. Dai F. C. and Lee C. F. (2002) Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. Geomorphology 42: 213-228
7. Donati, L. and Turrini, M. C. (2002) An objective method to rank the importance of the factors predisposing to landslides with the GIS methodology: application to an area of the Apennines (Valnerina; Perugia, Italy). Engg. Geol. 63: 277-289
8. Duman, T.Y., Can, T., Gokceoglu, C., Nefesliogocu, H. A., and Sonmez, H. (2006) Application of logistic regression for landslide susceptibility zoning of Cekmee area, Istanbul, Turkey. Verlag, 242 – 256
9. Chandak, P.G., Sayyed, S.S., Kulkami, Y.U., Devtale, M.K. (2016) Landslide hazard zonation mapping using information value method near Parphi village in Garhwal Himalaya. Ljemas, 4: 228 – 236
10. Chung CJF., Fabbri AG (2003) Validation of Spatial Prediction Models for Landslide Hazard Mapping. Natural Hazards 30(3): 451-472
11. Chung CJ, Fabbri AG (2005) Systematic procedures of landslide hazard mapping for risk assessment using spatial prediction models. In: Glade T, Anderson MG, Crozier MJ (eds) Landslide hazard and risk. Wiley, New York, 139–177
12. Chen, Z. and Wang, J. (2007) Landslide hazard mapping using a logistic regression model in Mackenzie Valley, Canada. Nat. Hazard 42(1), 75-89.
13. Dou J, Oguchi T, Hayakawa Y S, Uchiyama S, Saito H and Paudel U (2014) GIS-based landslide susceptibility mapping using a certainty factor model and its validation in the Chuetsu area, central Japan; In Landslide Sci- ence for a Safer Geoenvironment (Springer International Publishing), 419–424
14. Gorsevski P V, Gessler P and Foltz R B (2000) Spatial prediction of landslide hazard using discriminant analysis and GIS; In: GIS in the Rockies 2000 Conference and Workshop.
15. Heckeman (1986) Probabilistic interpretation of MYCIN’s certainty factors; In: Uncertainty in artificial intelligence (eds) Kanal L N and Lemmer J F (New York: Elsevier), 298–311
16. 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
17. Jia N, Xie M, Mitani Y, Ikemi H, Djamaluddin I (2010) A GIS-based spatial data processing system for slope monitoring. Int Geoinf Res Dev J 1(4)
18. Kanungo D.P., Arora M.K., Sarkar S, Gupta R.P (2009) Landslide susceptibility zonation mapping a review. J South Asia Disaster stud. 2: 81 – 105
19. Kanungo D P, Sarkar S and Sharma S 2011 Combining neural network with fuzzy, certainty factor south-facing and likelihood ratio concepts for spatial prediction of landslides; Nat. Hazards 59(3) 1491–1512
20. Karimi Nasab S, Ranjbar H, Akbar S (2010) Susceptibility assessment of the terrain for slope failure using remote sensing and GIS, a case study of Maskoon area, Iran. Int Geoinf Res Dev J 1(3)
21. Kifle Woldearegay (2008) Characteristics of a large-scale landslide triggered by heavy rainfall in Tarmaber area, central highlands of Ethiopia Geophysical Research Abstracts 10
22. Kouropeima S. Feiznia H. Ahmadi and Moghadamnia A.R. (2017) Landslide susceptibility mapping using logistic regression analysis in Latyan catchment. Desert, 85 - 95
23. Liu M, Chen X and Yang S (2014) Collapse landslide and mudslide hazard zonation; In Landslide science for a safer geoenvironment (Springer International Publishing), 457–462
24. Luelseged Ayalew and 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
25. Luelseged, Ayalew (1999) The effects of the seasonal rainfall on a landslide in the highland of Ethiopia. Bull Engineering Geology and Environment, 58: 9-19
26. Lee S, Pradhan B (2006) Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia. J Earth Sys Sci 115(6):661–667
27. Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models, Landslides 4:33–41
28. Meten, M., Bhandary, N.P., and Yatabe, R. (2015) GIS-based frequency ratio and logistic regression modeling for landslide susceptibility mapping of Debre Sina area in central Ethiopia. J.Mt.Sci., 12(6): 1355 – 1372
29. Pourghasemi HR, Pradhan B, Gokceoglu C, Mohammadi M, Moradi HR (2012a) Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. Arab J Geosci. doi:10.1007/s12517-012-0532-7
30. Pourghasemi H R, Pradhan B, Gokceoglu C, Mohammadi M Moradi H R 2013c Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz Watershed, Iran; Arabian J. Geosci., 6(7) 2351–2365
31. Pradhan B, Lee S, Buchroithner MF (2010c) Remote sensing and GIS based landslide susceptibility analysis and its cross-validation in three test areas using a frequency ratio model. Photogramm Fernerkun, 1:17–32. doi:10.1127/14328364/2010/0037
32. Pradhan B, Mansor S, Pirasteh S, Buchroithner M (2011) Landslide hazard and risk analyses at a landslide prone catchment area using statistical-based geospatial model. Int J Remote Sens, 32(14): 4075–4087. doi:10.1080/01431161.2010.484433
33. Pradhan B, Chaudhari A, Adinarayana J, Buchroithner MF (2012) Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: a case study at Penang Island, Malaysia. Environ Monit Assess, 184(2):715–727
34. Regmi A. D., Yoshida K., Pourghasemi H. R., Dhital M. R., and Pradhan B. (2014) Landslide susceptibility mapping along Bhalubang-Shiwapur area of mid-western Nepal using frequency ratio and conditional probability models. Jour. Mountain Sci, 11(5): 1266-1285
35. Varnes D. J. (1984) Landslide hazard zonation, a review of principles and practice, International Association of Engineering Geology Commission on Landslides and Other Mass Movements on Slopes, UNESCO, Paris, 63 p.
36. Temesgen B., Mohammed U, Korme T (2001) Natural hazard assessment using GIS and remote sensing methods, with particular reference to the Landslides in the Wondogenet Area, Ethiopia. Physics and Chemistry of the Earth Part C: Solar Terrestrials & Planetary Science (C), 26: 665-615
37. Sarkar S., Rjan Martha T., and Roy A. (2013) Landslide susceptibility Assessment using information value method in parts of the Darjeeling Himalayas. Geological Society of India, 82: 351 – 362
38. Sun W F (2009) Study of landslide hazard assessment on typical loess area in Qianhe valley, Qianyang County (Ph.D dissertation: Chinese Academy of Geological Science,
39. Kifle Woldearegay (2013) Review of the occurrences and influencing factors of landslides in the highlands of Ethiopia with implications for infrastructural development.
40. El Jazouli, Ahmed Barakat* and Rida Khelouk (2019) GIS-multicriteria evaluation using AHP for landslide susceptibility mapping in Oum Er Rbia high basin (Morocco).
41. Aleotti P., and R. Chowdhury (1999) Landslide hazard assessment: summary review and new perspectives. Bulletin of Engineering Geology and the Environment, 58: 21–44
42. Gutiérrez, F., R. Linares, C. Roqué, M. Zarroca, D. Carbonel J. Rosell, and M. Gutiérrez (2015) Large landslides associated with a diapiric fold in Canelles reservoir (Spanish Pyrenees): Detailed geological–geomorphological mapping, trenching and electrical resistivity imaging. Geomorphology 241: 224–242
43. Saha AK, Gupta RP, Sarkar I, Arora KM, Csaplovics E (2005) An approach for GIS-based statistical landslide susceptibility zonation with a case study in the Himalayas. Landslides 2(1):61–69
44. Sarkar S., Kanungo D, Ptra A., Kumar P (2006) Disaster mitigation of debris flow, slope failure, and landslides. GIS-based landslide susceptibility case study in Indian Himalaya. Universal Academy Press, Tokyo, Japan, 617 – 624
45. Shortliffe EH and Buchanan B G (1975) A model of inexact reasoning in medicine; Math. Biosci. 23(3) 351–379
46. Sujatha E R, Rajamanickam G V and Kumaravel P (2012) Landslide susceptibility analysis using probabilistic certainty factor approach: A case study on Tevankarai stream watershed, India; J. Earth Syst. Sci. 121(5) 1337–1350
47. Wang HB, Wu SR, Shi JS, Li B (2011) Qualitative hazard and risk assessment of landslides: a practical framework for a case study in China. Nat Hazards. doi:10.1007/s11069-011-0008-1
48. Wubalem A.*, Meten M (2019) Landslide susceptibility mapping using information value and logistic regression methods in Goncha Siso Eneses, northwestern Ethiopia.
49. Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: A comparison of logistic regression and neural networks method in a medium scale study, Hendek region (Turkey). Engineering Geology 79: 251 – 266