Comparing Probabilistic and Statistical Methods in Landslide Susceptibility Modeling in Rwanda /Centre-Eastern Africa

Jean Baptiste Nsengiyumva[a,b,c,d], Geping Luo[a,b,*], Amobichukwu Chukwudi Amanambu[a,b,e], Richard Mind'je[a,b,d], Gabriel Habiyaremye[d,h], Fidele Karamage[b,f], Friday Uchenna Ochege[a,b,g] and Christophe Mupenzi [d]

[a.] State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, No. 818, South Beijing Road, Urumqi 830011, China.
[b.] University of Chinese Academy of Sciences, Beijing 100049, China.
[c.] Ministry in Charge of Emergency Management, P.O. Box 4386, Kigali, Rwanda.
[d.] Faculty of Environmental Studies, University of Lay Adventists of Kigali (UNILAK), P.O.Box 6392, Kigali, Rwanda.
[e.] Water, Engineering and Development Centre, School of Civil and Building Engineering, Loughborough University, Leicestershire, UK.
[f.] Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China.
[g.] Department of Geography and Environmental Management, University of Port Harcourt, PMB 5323 Choba, East-West, Port Harcourt, Nigeria.
[h.] Lancaster University, Lancaster Environment Centre, Faculty of Science and Technology. LA 1 4 YQ, UK.

* Corresponding Author: luogp@ms.xjb.ac.cn; Tel.: +86-991-7823127; Fax: +86-991-7885320

Abstract:

Application of suitable methods to generate landslide susceptibility maps (LSM) can play a key role in risk management. Rwanda, located in centre-eastern Africa experiences frequent and intense landslides which cause substantial impacts. The main aim of the current study was to effectively generate susceptibility maps through exploring and comparing different statistical and probabilistic models. These included weights of evidence (WoE), logistic regression (LR), frequency ratio (FR) and statistical index (SI). Experiments were conducted in Rwanda as a study area. Past landslide locations have been identified through extensive field surveys and historical records. Totally, 692 landslide points were collected and prepared to produce the inventory map. This was applied to calibrate and validate the models. Fourteen maps of conditioning factors were produced for landslide susceptibility modelling, namely: elevation, slope degree, topographic wetness index (TWI), curvature, aspect, distance from rivers and streams, distance to main roads, lithology, soil texture, soil depth, topographic factor (LS), land use/land cover (LULC), precipitation and normalized difference vegetation index (NDVI). Thus, the produced susceptibility maps were validated using the receiver operating characteristic curves (ROC/AUC). The findings from this study disclosed that prediction rates were 92.7%, 86.9%, 81.2% and 79.5% respectively for WoE, FR, LR and SI models. The WoE achieved the highest AUC value (92.7%) while the SI produced a lowest AUC value (79.5%). Additionally, 20.42% of Rwanda (5,048.07km$^{2)}$ was modelled as high susceptible to landslides with the western part the highly susceptible comparing to other parts of the country. Conclusively, the comparison of produced maps revealed that all applied models are promising approaches for landslide susceptibility studying in Rwanda. The results of the present study may be useful for landslide risk mitigation in the study area and in other areas with similar terrain and geomorphological conditions. More studies should be performed to include other important conditioning factors that exacerbate increases in susceptibility especially anthropogenic factors.

Keywords: Landslide; Susceptibility; Rwanda; Frequency ratio; Statistical index; Logistic regression.

---

1. Introduction

Landslide is one the of most devastating natural disasters that causes loss of human lives, properties and infrastructure in many parts around the globe (Chen et al. 2018a; Chen et al. 2018b). Many countries in the world are susceptible to landslide hazards with unacceptable levels of natural underlying risks (Pisano et al. 2017). Their fatalities are recurrently recorded, especially in mountainous prone zones. The impacts of landslide hazards are therefore still numerous in most parts of the globe (Pisano et al. 2017; Zêzere et al. 2017; Chen et al. 2018a). According to the international disaster analysis in 2015, it was disclosed that 346 disaster cases were reported, with 22,773 deaths and 98.6 million people affected. Additionally, 66.5 billion USD were lost due to these natural disasters with five countries being the most hit, including China; USA; India; Philippines and Indonesia (Pisano et al. 2017; Ahmed and Dewan 2017). Furthermore, 174 landslides cases were recorded worldwide in 2014, leading to major devastating effects and impacts. Previous studies reported that landslides are categorized as the third cause of the most global serious and deadly natural disasters (Ramani et al. 2011; Ahmed and Dewan 2017; Nsengiyumva et al. 2018).

The Management of landslide risks requires a lot concerted efforts, but landslide susceptibility mapping becomes the most significant tool to minimize their impacts through resilience building (EAC 2012; Nsengiyumva 2012; Zschau and Küppers 2013; Chen et al. 2018a). Therefore, susceptibility maps reveal the spatial distribution of probabilities of landslide occurrences in a given area based on certain conditioning factors. Generally, landslide susceptibility is controlled by a number of parameters including conditioning factors, types of landslides, failure mechanisms, and coverage of affected areas, frequency and intensity among others.

In the previous decades, the study of landslide susceptibility attracted the attention of many researchers worldwide, but still, landslides constitute a major threat to human life. The literature on landslide studies avails various susceptibility mapping techniques and approaches, ranging from very simple to more complex. These include inventory based (Nichol and Wong 2005; Van Westen et al. 2006; Yalcin et al. 2011; Akgun 2012; Van Den Eeckhaut et al. 2005), data-driven methods composed by bivariate and multivariate statistics (weights of evidence, frequency ration, logistic regression, cluster analysis, artificial neural networks (Dahal et al. 2008b; Neuhäuser and Terhorst 2007; Dahal et al. 2008a; Mohammady et al. 2012; Ayalew and Yamagishi 2005; Yilmaz 2009; Ramani et al. 2011; Sujatha et al. 2012; Zêzere et al. 2017; Chen et al. 2018a); and the knowledge-driven methods for landslide susceptibility studies (fuzzy logic, analytical hierarchy process, spatial multi-criteria evaluation, multi-class overlay and Boolean logic) (Gorsevski et al. 2006; Pradhan 2010b, 2010a; Neaupane and Piantanakulchai 2006). Additionally, some landslide susceptibility studies use probabilistic methods composed of both parameter uncertainty and temporal prediction (Refice and Capolongo 2002; Zhou et al. 2003; Mazzanti et al. 2015; Brenning 2005; Zêzere et al. 2004) and physically-based and deterministic methods (Cervi et al. 2010; Gökceoglu and Aksoy 1996; Godt et al. 2008; Yalcin 2008; Baum et al. 2008; Terlien et al. 1995; Kuriakose et al. 2009; McDougall and Hungr 2005; Wu et al. 2009; Turner et al. 2015; Schilirò et al. 2016; Sinarta et al. 2017).

From the above literature review, it was revealed that several data driven models exist such as statistical and probabilistic methods but have never been compared for landslide susceptibility modeling in Africa. In addition, the application of data driven methods may play a big role in accurately predicting landslide susceptibility for African prone regions (Nsengiyumva et al. 2018; Monsieurs et al. 2018; Bizimana and Sönmez 2015; MIDIMAR 2015a). Thus, it is therefore important to compare probabilistic and statistical methods to achieve suitable and accurate outputs for landslide susceptibility mapping (Youssef et al. 2016). The current study aims therefore to make a comparative analysis of four models, including statistical index (SI), frequency ratio (FR), logistical regression (LR), and weights-of-evidence (WoE) models to predict landslide susceptibility in Rwanda.

2. General description of the study area

The present study covered the entire territory of Rwanda, a country located in the great lakes region of the central-east Africa (Nahayo et al. 2018; Karamage et al. 2016). Rwanda is a land-locked country occupying a total surface area of 26,338 square kilometres with a total population of 12, 601,482 in 2018. Rwanda is one of the most densely-populated countries in Africa (Nsengiyumva et al. 2018).

Rwanda extends over the eastern shoulder of the Kivu-Tanganyika rift in Africa (Fig.1). Despite its proximity to the equator, Rwanda enjoys a tropical climate moderated by hilly topography varying between 920 and 4495 m above sea level, stretching from east to west (Ndayisaba et al. 2016). The country has four climatic seasons in which long rainy (late February to late May) and short rainy seasons (end September to early December) alternate with long dry (June–September) and short dry (mid-December–mid-February) seasons. The two rainy seasons correspond to agricultural seasons, season B and season A, respectively. The annual precipitation ranges between 700 and 1400 mm, with the eastern part having short rains.

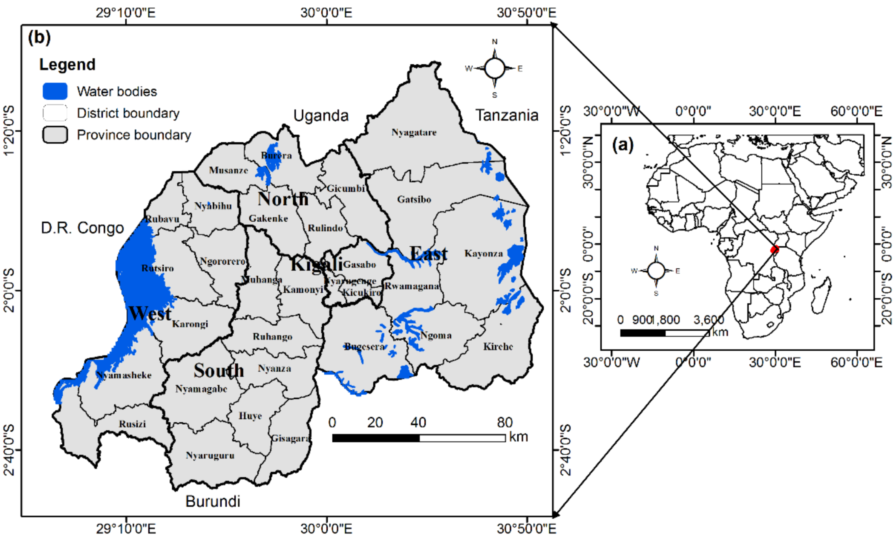


Fig.1. Location map of the study area

Rwanda is characterized by a strongly heterogeneous landscape, with very different terrain features over the 30 districts. High mountains up to 4km above sea level are found in the west and northwest parts of the country, including the Congo Nile Ridge, the Volcanic Range and the Buberuka highlands. The plains are found in the eastern part of the country, including the eastern savanna, the eastern plateau, the central plateau and the Bugesera-Mayaga. A large wetland reservation in the Akagera National Park is found in the northeast of the country (Ndayisaba et al. 2016). Based on the statistical analysis from the geospatial lithological and soil types, it is confirmed that Rwandan soils are mainly composed by fragile soils (a physico-chemical alteration of basic schistose, quartzite, granite, basic igneous rock, and volcanic rocks). The underlying geology consists of the Acrisols (47.5%), Ferralsols (17.5%), Regosols (13%), Andosols (5%), Histosols (4.1%) and Cambisols (3.2%), Vertisols (1.8%), Greysols (1.5%), Nitosols (0.4%), and water bodies cover the remaining 6%. Agriculture which occupies 58.31% of the land remains entirely rainfed and is mainly practiced on hill slopes. (Karamage et al. 2016).

Due to its topographic nature with steep slopes, Rwanda is prone to natural hazards including mass wasting especially landslides which are the most recurrent hazards in the sub-region. Landslide hazards are very common natural phenomena in the centre-eastern Africa. The geomorphology in the study area presents, therefore, a favorable uniqueness to explore and test the four methods (FR, SI, WoE and LR). Moreover, as confirmed by the Ministry of Disaster Management and Refugees, 124 people were killed by landslides, 141 injured and 897 houses destroyed in the study area from 2011 to May 2017 (Nsengiyumva et al. 2018). In addition, from January to October 2018, natural disasters including landslides killed 234 people, injured 218, destroyed 15,264 houses and 9,412 hectares of crops, 31 roads and 52 bridges damaged, 86 classrooms completely destroyed as well as 797 livestock killed (MIDIMAR 2018). Therefore, this testifies how much the study area is a landslide prone zone.

3. Data and methods

3.1. The landslide inventory

For any susceptibility mapping activity, it is essential to detect and understand the relationship between conditioning factors and past landslide distribution (Nsengiyumva et al. 2018; Youssef et al. 2016; Chen et al. 2018a; Chen et al. 2018b). This information is therefore obtained from the past landslide locations inventory and landslides are mostly represented as points (Regmi et al. 2014). Thus, for this study, the inventory map was generated with 692 past landslide points. These points represent areas in Rwanda where landslides had occurred in the past. These landslides locations were mapped using GPS or global positioning system through field surveys from March 2015 to December

2017. Interviews were conducted with local residents in the hazard-prone zones. Historical records and disaster reports from the MIDIMAR and district offices were also consulted. Additionally, some landslide data were extracted from existing provincial topographic maps (1: 500,000-scale).

Therefore, the inventory map and the maps of the landslide predictors were produced at the national scale of 1:1,000,000 since the current study covered the entire Rwanda. Thus, 75% of collected past landslides (519 points) were used to simulate the models while 25% (equivalent to 173 landslide points) were used for validation (Fig. 2). To extract and split the total points into training and testing points, authors used the geostatistical analyst extension of ArcMap 10.3, through the subset feature that divides the original dataset into two parts: one was used for modeling the spatial structure and to produce a surface, while the other was used for comparing and validating the output surface by sub-setting the data (Youssef et al. 2016; Zêzere et al. 2017).

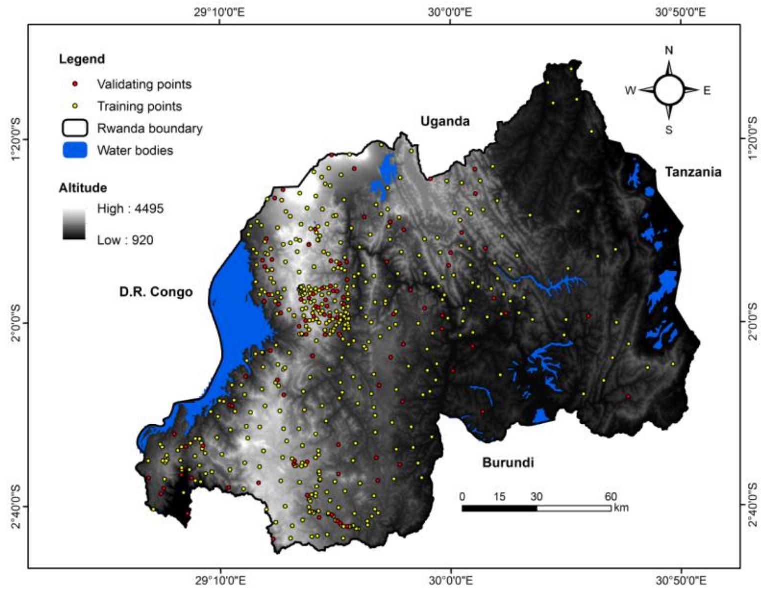


Fig.2. Landslide inventory map

3.2. The landslide conditioning factors

To carry out the landslide susceptibility mapping, various datasets have to be used including landslide conditioning factors and inventory maps (Chen et al. 2017b). For the current study, fourteen landslide conditioning factors have been used including precipitation, distance from main roads, desistance from rivers/streams, slope, elevation, NDVI, lithology, TWI, topographic factor (LS), soil depth, curvature, LULC, soil texture and aspect. The selection of these factors based on available datasets, historical records, fieldworks in the study area, study area context as well as objectives of the study. Therefore, conditioning factors play a critical role in modeling landslide susceptibility.

To investigate landslide susceptibility in the study area, the digital elevation model (DEM) of 30m resolution was used. This was obtained from the Global Digital Elevation Model-GDEM (Maes et al. 2018). The DEM was used to derive six landslide causal factors including slope, aspect, elevation, curvature, topographic factor and topographic wetness index (Fig.3). The spatial extension toolset of ArcMap 10.3 was used to deduce these factors.

Furthermore, land cover/land use (LULC) is also considered as an important landslide conditioning factor (Ramani et al. 2011). The latest LULC map of 2017 (Fig. 3k), has been produced from Landsat-8 OLI images. These images were obtained from the United States Geological Survey (USGS) through global visualization tool (Maes et al. 2018; USGS). This was accomplished by using the maximum likelihood classification technique in Envi 5.3 software. Subsequent to radiometric corrections, masking of cloud shadows and gaps filling; the LULC map has been classified. The classification was done following the previous classification by the regional centre for mapping of resources for development (RCMRD) for East-Africa region. Similarly, the current study applied type one of USGS classification techniques (Nsengiyumva et al. 2018; Karamage et al. 2017). Thus, the study area was then classified into six classes (forestland: 15.38%, grassland: 14.31%, cropland: 58.31%, built up land: 1.86%, wetland: 4.02% and water bodies: 6.12%). In addition, for the accuracy assessment, authors randomly composed sixty points for each land use/land cover type, which were overlaid to a classified image in Google Earth to make verification. An overall satisfactory accuracy of 92 % was therefore achieved.

Existing geological maps with good scale (1:100,000) were obtained from Rwanda Natural Resources. These datasets were used to deduce lithology factor for landslide susceptibility modeling in Rwanda. Additionally, soil datasets were obtained from the Ministry of Agriculture and Rwanda agriculture board. They were originated from extensive soils mapping and surveys nationwide in 1995. Soil is a very important conditioning factor of landslide susceptibility (Dou et al. 2018; Chen et al. 2017a; Coppola 2006). Therefore, three factors were generated namely soil depth, lithology and texture (Fig. 3i, 3j and 3l).

For the precipitation factor, this study applied monthly mean rainfall for 21 years (1996-2017). Authors utilized rainfall datasets from climate hazards group infraRed precipitation with station data (CHIRPS) as described by Funk. et al., (2017). These datasets were coupled with rainfall data from meteorological stations in the study area as provided by Rwanda Meteorological Agency. In most landslide studies, rainfall is considered as a severe trigger of landslide hazards especially in mountainous areas (NASA; Schilirò et al. 2016).

The topographic factor represents the product of slope length (*L*) and steepness (*S*) factor. LS illustrates the influence of topography on landslide/soil erosion occurrence (Ramani et al. 2011) and has a high value if the length and slope of terrain are high. If the length and steepness of slope are more, the landslide will be high and vice versa. Thus, it can be estimated through field measurements or can be derived from digital elevation model (DEM). LS equation has been developed to generate the topographic factor map based on DEM (Moore and Wilson 1992). For this study, LS factor was estimated from the Shuttle Radar Topography Mission (STRM), 30m resolution provided by the National Aeronautics and Space Administration (NASA). The spatial analyst extension tool of ArcMap 10.3 was used to derive the L and S values of each pixel using equation 1 for L factor developed by Desmet and Govers (1996) and equation 2 for S factor applying McCool et al. (1987) method.

$$L_{i.j} = \frac{(A_{i.j-in} + D^2)^{m+1} - A_{i.j-in}^{m+1}}{D^{m+2} \cdot x_{i.j}^{m} \cdot (22.13)^{m}} \qquad (1)$$

$$m = \frac{\beta}{1+\beta} \qquad (1a)$$

$$\beta = \frac{\mathrm{Sin}\ \theta/0.0896}{3(\mathrm{Sin}\ \theta)^{0.8} + 0.56} \qquad (1b)$$

$$S_{i.j} = \begin{cases} 10.8\sin\theta_{i.j} + 0.03, \tan\theta_{i.j} < 9\% \\ 16.8\sin\theta_{i.j} - 0.50, \tan\theta_{i.j} \geq 9\% \end{cases} \qquad (2)$$

where $L_{i.j}$ = slope length factor for the grid cell with coordinates (i.j); D = the grid-cell size (m); $X_{i.j}$ = (Sin $a_{i.j}$ + Cos $a_{i.j}$); $a_{i.j}$ = aspect direction for the grid-cell with coordinates (i.j); $A_{i.j-in}$ is the flow accumulation or contributing area at the inlet (m^2) of a grid-cell with coordinates (i.j). Besides, the slope-length exponent m is related to the ratio β of rill erosion (caused by flow) to interrill erosion (principally caused by raindrop impact); β is the ratio of rill to interrill erosion for conditions when the soil is moderately susceptible to both rill and interrill erosion; θ is the slope angle in degrees (Karamage et al. 2017).

For the present study, NVDI has been considered as one of the landslide conditioning factors. The Normalized difference vegetation index highlights the vegetation stability in a given area (Akgun 2012).

NDVI was extracted from Landsat-8 for 2016 with 30m resolution. This was achieved using spatial analyst tool in ArcMap 10.3 based on equation 3 as follows:

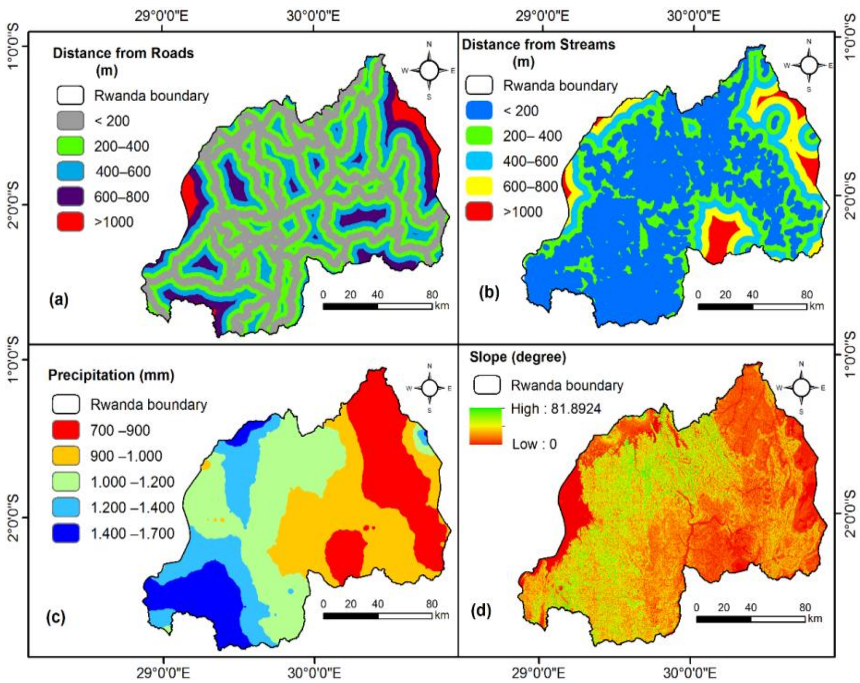
$$\text{NDVI} = \frac{\text{Band 5} - \text{Band 4}}{\text{Band 5} + \text{Band 4}} \tag{3}$$

For landslide susceptibility modeling, topographic wetness index (TWI) is considered as a significant conditioning factor (Ahmed and Dewan 2017; Youssef et al. 2016). For the present study, TWI computation was achieved by using the flow accumulation obtained from the flow direction. All these were given by DEM using hydrological tool from spatial analyst tools of ArcMap 10.3. Equation 4 was therefore applied to calculate TWI and it ranges from 1.92 to 27.28 (Fig. 2h).

$$\text{TWI} = \ln\left(\frac{A_s}{\tan\alpha}\right) \tag{4}$$

Where A_s is the catchment area and α is the slope gradient (in degrees). The curvature can be influenced by the slope erosion processes as convergence or divergence of water during downhill flow and it constitutes one of the landslide conditioning factors.

Fieldworks in the study area revealed that some landslide cases were caused by proximity to roads and streams/rivers. Therefore, authors decided to consider both distance from roads and distance from rivers as landslides conditioning factors in Rwanda. Road datasets were freely obtained from the Rwanda Transport Development Agency while geospatial datasets on rivers and streams networks, were obtained from the Ministry of Lands and Forests. The distances were calculated using Euclidian distance of spatial analyst tool of ArcMap10.3. All the 14 conditioning factors were applied to models (FR, SI, WoE and LR) in generating landslide susceptibility maps for Rwanda.



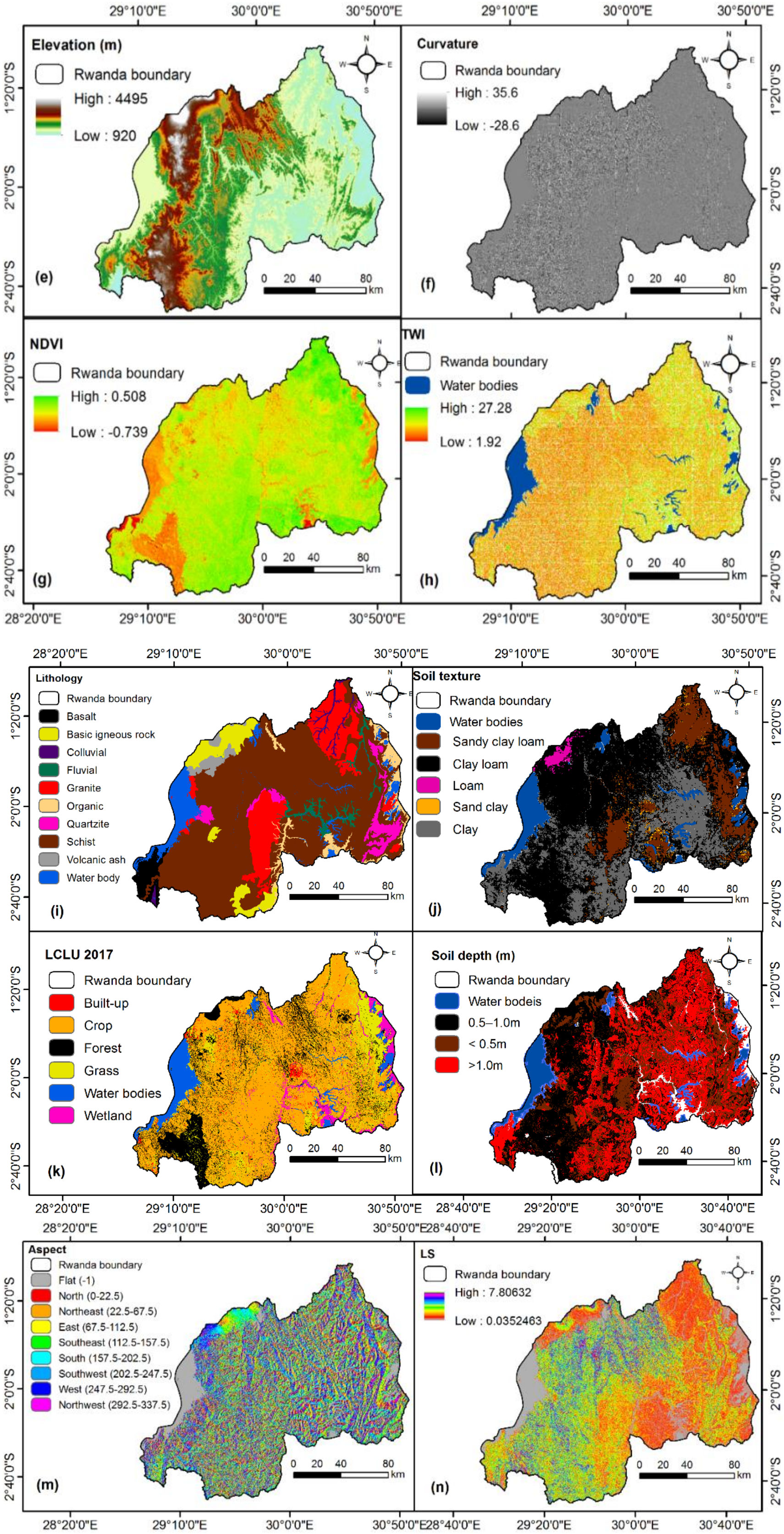


Fig.3. Landslide conditioning factors: (a) Distance from roads; (b) Distance from streams/rivers; (c) Precipitation; (d) Slope; (e) Elevation; (f) Curvature; (g) Normalized Difference Vegetation Index (NDVI); (h) Topographic Wetness Index (TWI); (i) Lithology; (j) Soil texture; (k) Land use/cover 2017 (LULC); (l) Soil depth; (m) Aspect; (n) Topographic factor (LS).

For this study, each of the conditioning factors was correlated with the landslide occurrences.

3.3. Landslide susceptibility modeling

3.3.1. Study design

Fig.4 presents the study framework for landslide susceptibility assessment in the study area.

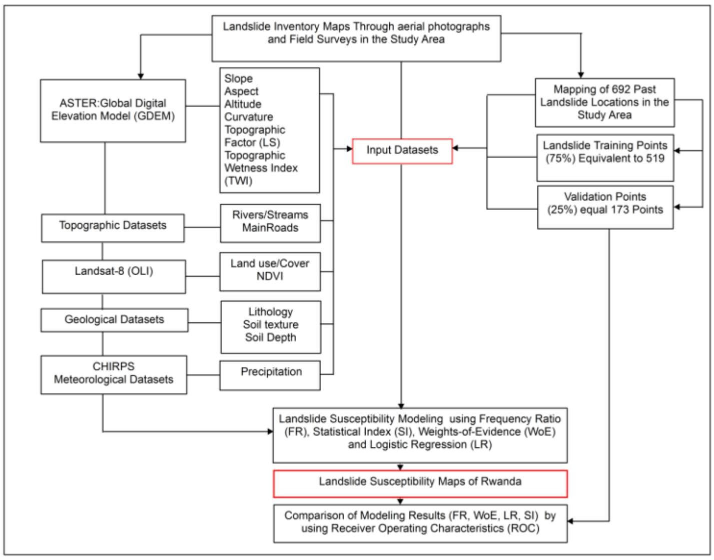


Fig.4. Flowchart of the study

In susceptibility modeling, the accuracy and quality of produced maps largely depend on input datasets, study area complexity and applied methodologies as well. This section presents a detailed description of applied methods. Thus, the authors explored and compared probabilistic and statistical models to predict landslide susceptibility in the study area.

3.3.2. The Statistical Index Model (SI)

In the literature related to landslide hazards and disasters, statistical index model has widely been applied by different researchers (Bui et al. 2011; Ahmed and Dewan 2017). SI is a statistical, bivariate approach to study susceptibility in landslide prone zones. The weighting value is obtained by dividing the density of landslides in the class by the landslide in the entire map. Therefore, the SI can be modeled using equation 5.

$$W_{SI} = \ln\left(\frac{E_{xy}}{E}\right) = \ln\left(\frac{L_{xy}/L_T}{P_{xy}/P_L}\right) \quad (5)$$

where, W_{SI}, is the statistical index weight assigned to a given landslide class x of the y factor; E_{xy}, stands for the density of landslides in x class of y factor; E, entire density of landslide in the total map; L_{xy}, amount of landslides in a given x class of y the parameter; P_{xy}, amount of pixels of the x class for the y factor; LT, total landslide in the whole map; PL, number of pixels of the whole mapped area. The landslide susceptibility map through SI is therefore obtained with equation 6.

$$LSM_{SI} = W_{SI\ 1} + W_{SI\ 2} + W_{SI\ 3} + \cdots W_{SIn} \quad (6)$$

whereby LSM_{SI} represents the landslide susceptibility for statistical index model, and W_{SI} = the assigned weight of a given landslide conditioning factor or a range of conditioning factors.

3.3.3. Frequency Ratio Model (FR)

Generally, present and past landslide occurrences are assumed to be useful in predicting future potential landslide events, and it is commonly believed that landslides would occur from similar circumstances (Pradhan and Lee 2010; Regmi et al. 2014). Based on this principle, it is required to determine the relationship between conditioning factors and past landslides while modeling susceptibility. As confirmed by previous studies (Regmi et al. 2014), frequency ratio model discloses the correlation between observed landslides and conditioning factors.

Landslide prediction is performed through the relation between the causal factor and landslides event inventory (Regmi et al. 2014). The estimation of the FR is given by the ratio of area of the landslide points to area that has not been affected by landslide. This is then computed for each class factors.

For the production of susceptibility index (LSI), the values for each factor's frequency ratio are to be summed up using appropriate equations. As previously indicated by studies (Mohammady et al. 2012), FR ratios are calculated and summed for each considered factor to generate hazard susceptibility. Therefore, the higher the value, the greater the probability of the landslide to occur and inversely, the lower values represent the lower occurrences of landslide hazards. Frequency ratio is modeled using equation 7 below:

$$\mathrm{LSI} = \mathrm{Fr}_1 + \mathrm{Fr}_2 + \mathrm{Fr}_3 + \cdots \mathrm{Fr}_n \quad (7)$$

Hence, as described by Regmi et al.(2014), landslide susceptibility by FR, is yielded by equation 8:

$$\mathrm{LSI} = \textstyle\sum_1^n \mathrm{FR} \quad (8)$$

from equation 8, LSI = landslide susceptibility index and Fr stands for each factor's rating. As previously stated by Ahmed (2011), the frequency ratio method is expressed in more details with equations 9 and 10.

$$\mathrm{FR}_{ij} = \frac{\mathrm{FrX}_{ij}}{\mathrm{FrY}_{ij}} \quad (9)$$

from equation 9, FR_{ij} is the landslide occurrence proportion for the i class of the j factor; FrX_{ij} denotes the frequency of observed landslides within class i of factor j; also FrY_{ij} represents the frequency of the percentage for non-observed landslides in class i of factor j. For this equation, it is deduced that a greater ratio shows stronger correlation between occurrence and the factor's class and the lower ratio shows less relationship between occurrences and considered factors (Regmi et al. 2014).

From this assumption, landslide susceptibility is determined by applying equation 10.

$$\mathrm{LSI}_{\mathrm{FR}} = \sum_{j=1}^{n} \mathrm{W}_{ij} \quad (10)$$

where $\mathrm{LSI}_{\mathrm{FR}}$ is the landslide susceptibility index by frequency ration model; W_{ij}= the weight for class i within j conditioning factor and n = the total number of all considered conditioning factors.

3.3.4. Weights of Evidence Model (WoE)

The mapping of susceptibility has attracted attention of more researchers through application of statistical approaches (Mohammady et al. 2012; Regmi et al. 2014). WOE is one of the fundamental models in studying landslide susceptibility worldwide. The WoE application to susceptibility modeling has also been widely recognized in the literature related to landslide studies (Regmi et al. 2014). The performance of WoE model requires the calculation of positive and negative parameters (W^+ and W^- Weights).

Generally, past landslide datasets are the key factors for weights determination (Monsieurs et al. 2018). It has also been confirmed that landslide susceptibility modeling has to rely on the theory that landslide events can be caused by similar conditions to those which triggered past landslides (Regmi et al. 2014). Thus, the WoE is modelled using equations 11 and 12.

$$\mathrm{W}^+ = \ln\left(\frac{\mathrm{P}\left(\frac{\mathrm{X}}{\mathrm{Y}}\right)}{\mathrm{P}\left(\frac{\mathrm{X}}{\mathrm{y}}\right)}\right) \quad (11)$$

$$W^{-} = \ln\left(\frac{P\left(\frac{x}{y}\right)}{P\left(\frac{x}{y}\right)}\right) \quad (12$$

whereby P = the probability of occurrence, ln = the natural logarithm. Besides, X and x represent the presence and absence of landslide conditioning factors, whilst Y and y stand for the presence and absence of landslides events. When modeling landslide susceptibility, WoE computes each landslide conditioning factor's weight x according to the existence or absence of landslide hazards in the study area as per equations 13 and 14.

$$W^{+} = \ln\left\{\frac{\left(\frac{NLP_{xy}}{TNLP_{y}}\right)}{\left(\frac{NSP_{xy}}{TNSP_{y}}\right)}\right\} \quad (13)$$

where W^{+} is the positive weighting for the x class of the y factor, NLP_{xy} equals to the total points of landslide hazards within x class of y factor, $TNLP_{y}$ is the sum of landslide points for each y factor, NSP_{xy} represents number of pixels in stable condition for x class of y factor. Thus, $TNSP_{y}$ = the number of total pixels of y factor in stable condition. The negative weight is therefore determined using equation 14.

$$w^{-} = \ln\left\{\frac{\left(\frac{NLP_{ny}}{TNLP_{y}}\right)}{\left(\frac{NSP_{ny}}{TNSP_{y}}\right)}\right\} \quad (14)$$

from equation 14, W^{-} is the negative weight to be assigned when the class x of the factor y is absent, NLP_{ny} is the amount of landslide points in further n classes of y factor, $TNLP_{y}$ equals total points of factor y, NSP_{ny} denotes stable pixels in extra n classes of y factor, $TNSP_{y}$ is the total amount of pixels with stability of y factor. Furthermore, positive weighting (W^{+}) determines the presence of the landslide conditioning factors in place and this testifies the strong link between the presence of hazard events and conditioning factors. In case of the negative weighting (W^{-}), it is confirmed the non-existence of conditioning factors which shows lack of correlation (Regmi et al. 2014).

Accordingly, the weight of contrast (W_{f}) is obtained from the difference between W^{+} and W^{-}. The LSI is therefore obtained from weight of contrast (Youssef et al. 2016). Thus, the W_{f} is calculated using equation 15.

$$W_{f} = W^{+} - W^{-} \quad (15$$

$$LSI_{WofE} = \sum_{j=1}^{n} W_{fxy} \quad (16)$$

where W_{f} = the weight of contrast, LSI =landslide susceptibility index, WofE = weights of evidence, Wf_{xy} = final weight of x class in y conditioning factor, and n the total number of conditioning factors.

3.3.5. Logistic regression (LR) model

Logistic Regression which is a multivariate model has extensively been applied in many different studies related to landslide susceptibility studies (Chen et al. 2017a; Chen et al. 2017b). This model produces results basing on one or more independent variables, and the result is measured through dichotomous variables such as true and false or 0 and 1. In this study, the application of logistic regression in susceptibility modeling, served to define the linkage between the presence and lack of landslide events with related conditioning factors. LR generates coefficients that predict landslides in a given area. Logistic regression model was therefore applied using equations 17 and 18.

$$z = a_{0} + a_{1}x_{1} + a_{2}x_{2} + a_{3}x_{3} + \cdots a_{n}x_{n} \quad (17)$$

where z = the linear combination of the dependent variables representing the absence (0) or the presence (1) of landslide, and variable values from $-\infty$ to $+\infty$, a_{0} stands for the intercept of the model, a_{1}, a_{2}, …a_{n} represent the coefficients of logistic regression model, and x_{1}, x_{2}, …x_{n} denote the conditioning factors or independent variables (Devkota et al. 2013).

In a simplified way, logistic regression model is expressed using equation 18:

$$P = \ln\left(\frac{P}{1-P}\right) = \frac{1}{1+e^{-z}} \tag{18}$$

with P = the probability (varying between 0 and 1) of a landslide to occur, and z = the linear model for considered variables. At this stage, with logistic regression model, the landslide susceptibility index is calculated using equation 19.

$$SI = \exp(z) \;/\; (1 + \exp(z)) \tag{19}$$

where, SI = the landslide susceptibility index.

3.4. Model performance validation

Studies on landslide have ascertained that susceptibility maps are not useful unless they are validated (Chen et al. 2018b). For susceptibility mapping, it is required to assess the validity of the models applied since they have no scientific significance without validation (Chen et al. 2018a) Appropriate methods are therefore essential to validate landslide susceptibility maps generated using models. To validate the LSMs in this study, authors applied the receiver operating characteristic (ROC). Thus, ROC presents the percentages of true positive rating of past landslides against the false positive rating percentage of susceptibility index in a cumulative decreasing order. This helps to get the ROC curve of the rate of success (Ahmed and Dewan 2017; Chen et al. 2018a). The area under the ROC curve (AUC) is useful to detect which of the applied models is the best predictor of landslide susceptibility for the area under investigation.

In case of the poor prediction or non-improvement, the AUC value becomes less or equal to 0.5 while the best and ideal susceptibility modeling is obtained when AUC value is higher or equal to 0.7 (Devkota et al. 2013; Ahmed and Dewan 2017). The literature confirms that AUROC curves are one of the most common tools used to validate and compare landslide susceptibility modeling methods (Zêzere et al. 2017). The AUC was therefore calculated using equation 20 (Chen et al. 2018b).

$$AUC = \frac{(\sum TP + \sum TN)}{(P + N)} \tag{20}$$

where P is the total number of landslides and N is the total number of non-landslides; TP = the true positive and TN =the true negative.

4. Results and discussion

4.1. Relationship between conditioning factors and landslide locations

The spatial relationship between each landslide conditioning factor and landslide locations was calculated using the four models FR, SI, W_f and LR, and the results are shown in Tables 1 and 2.

As shown in Table 1, each factor has been assigned different values depending on the previously observed landslides including percentage number of pixels with the applied models. It was observed that some factor classes gained high values for all the four models. These factors include distance from roads (classes of <200m and 200-400m), precipitation (classes of 1400-1700mm and 1200-1400mm), slope degrees (>26.11^o and 18.30-26.11^o), elevation (2,196-2,813m and 2,813-4,495m), lithology (schist, basic igneous rock and volcanic rocks), soil texture (clay loam, clay), land use/land cover (cropland, built up and forestland), soil depth (<0.5m and 0.5-1.0m), slope aspect (south and southwest), curvature (0.64-35.68), and LS (2.50-7.81m ad 1.77-2.50m). In contrast, the models presented some differences for class factor relationship between TWI, NDVI and distance from streams.

Table 1 Spatial relationship between each landslide conditioning factor and landslide locations using FR and WoE, SI and LR

Conditioning Factor	Factor class	% Pixels	% OL	FR	SI	W_f	LR
TWI	1.92– 4.71	38.24	34.33	0.814	0.248	-0.304	0.810
	4.71– 6.50	30.10	45.17	2.127	0.342	0.275	0.726
	6.50– 8.89	20.36	13.26	0. 6791	0.297	-0.185	0.825
	8.89– 12.57	8.27	6.02	0.853	0.364	-0.044	0.901
	12.57–27.28	3.03	1.22	0.960	0.527	0.795	1.000
NDVI	(-0.73) – (0.31)	6.74	5.10	0.873	0.314	-0.174	0.832
	(-0.31) – (-0.19)	33.89	11.20	0.597	0.213	-0.162	0.684
	(-0.19) – (-0.09)	18.96	38.30	3.471	0.737	1.245	0.851
	(-0.09) – (0.10)	15.84	31.00	2.034	0.417	1.085	0.906
	(0.10) – (0.50)	24.57	14.40	0.695	0.289	-0.239	0.985
Distance from roads (m)	< 200 m	47.67	49.61	1.806	0.627	1.868	1.000
	200 – 400 m	22.94	26.40	2.302	0.412	1.074	0.941
	400 – 600 m	15.76	12.04	0.667	0. 122	-0.457	0.890
	600 –800 m	13.63	9.52	0.562	0.056	-0.807	0.673
	>1000 m	3,89	2.43	0.970	0.030	-0.545	0.456
Distance from streams (m)	< 200 m	52.86	31.8	0.415	0.594	-0.565	0.937
	200 – 400 m	19.57	23.18	1.073	0.428	0.787	0.852
	400 – 600 m	13.84	17.4	1.163	0.205	0.908	0.778
	600 –800 m	7.01	18.6	2.728	0.164	1.095	0.508
	>1000 m	6.72	9.02	1.987	0.097	0.819	0.493
Precipitation (mm)	700 – 900 mm	20.16	4.74	0.344	0.152	-0.545	0.523
	900 – 1000 mm	33.32	12.62	0.672	0.203	-0.720	0.642
	1000 – 1200 mm	24.53	30.24	2.527	0.224	0.978	0.738
	1200 –1400 mm	18.25	29.60	2.928	0.492	1.127	0.997
	1400–1700 mm	3.74	22.80	3.573	0.618	1.572	1.000
Slope (degrees)	0^o – 4.81^o	31.81	0.02	0.089	0.012	-1.041	0.238
	4.81^o – 11.24^o	23.74	13.66	0.118	0.249	-1.087	0.591
	11.24^o – 18.30^o	15.04	16.85	1.210	0.441	-0.499	0.657
	18.30^o – 26.11^o	5.33	27.32	4.397	0.713	1.319	0.994
	> 26.11^o	24.08	42.15	4.861	0.966	1.678	1.000
Elevation (m)	920 – 1,537 m	26.9	1.30	0.026	0.122	-0.154	0.051
	1,537 – 1,832 m	22.50	5.11	0.166	0.301	-0.436	0.886
	1,832 – 2,196 m	19.12	14.09	0.351	0.186	-1.240	0.725
	2,196 – 2,813 m	24.07	61.48	4.607	0.911	3.687	0.994
	2,813 – 4,495 m	7.41	18.02	3.010	0.673	2.030	1.000
Curvature	(-28.60) – (-0.87)	7.25	15.39	1.341	0.752	0.532	0.284
	(-0.87) – (-0.37)	18.54	11.64	0.620	0.601	-0.914	0.661
	(-0.37) – (0.13)	29.93	20.57	0.403	0.493	-0.979	0.535
	(0.13) – (0.64)	23.86	18.12	0.767	0.602	-0.759	0.099
	(0.64) – (35.68)	20.42	34.28	1.720	0.815	0.748	0.967
Lithology	Basalt	2.27	1.12	0.968	0.422	0.565	0. 676
	Basic igneous rock	5.24	9.73	2.589	0.635	1.279	0.989
	Colluvial	1.04	0.31	1.063	0.224	0.067	0.804
	Fluvial	2.24	0.16	0.958	0.318	-0.119	0.695
	Granite	13.19	1.18	0.412	0.172	0.361	0.426
	Organic	4.14	1.12	0.879	0.187	0.067	0.263
	Quartzite	3.87	2.02	0.998	0.235	0.085	0.038
	Schist	60.21	80.06	6.987	0.931	4.026	1.000
	Volcanic rocks	1.70	4.30	0.576	0.460	0.765	0.985
	Water bodies	6.12	0.00	0.000	0.000	0.000	0.000
Soil texture	Water bodies	6.12	0.00	0.000	0.000	0.000	0.000
	Sand clay loam	13.57	11.77	0.853	0.225	0.519	0.762
	Clay loam	54.44	62.51	2.127	0.753	1.270	1.000
	Loam	1.05	2.13	1.543	0.242	0.862	0.719
	Sand clay	1.12	0.95	0.801	0.395	0.425	0.065
	Clay	23.72	22.64	0.985	0.520	0.619	0.885

Land use/cover	Built-up land	1.86	4.12	1.234	0.341	0.658	0.983
	Crop land	58.30	78.29	6.7921	0.913	3.271	1.000
	Forest land	15.38	12.35	0.786	0.567	0.590	0.754
	Grass land	14.31	5.14	0.582	0.231	-0.157	0.432
	Wet land	4.05	0.10	0.187	0.167	-0.008	0.029
	Water bodies	6.12	0.00	0.000	0.000	0.000	0.000
Soil depth (m)	< 0.5 m	17.63	10.20	0.708	0.383	0.423	0.897
	0.5 –1.0 m	29.48	82.22	5.227	0.997	2.604	1.000
	> 1.0 m	52.89	7.58	0.014	0.261	-0.979	0.553
Slope aspect	Flat (-1)	9.42	0.00	0.000	0.122	0.000	0.000
	North (0 – 22.5)	4.14	11.83	1.032	0.233	0.028	0.640
	Northeast (22.5 – 67.5)	10.47	8.16	0.899	0.229	-0.220	0.829
	East (67.5 – 112.5)	8.57	7.91	0.967	0.175	-0.689	0.516
	Southeast (112.5–157.5)	11.26	8.43	0.726	0.224	-0.465	0.758
	South (157.5 – 202.5)	10.14	19.08	1.859	0.294	0.658	0.994
	Southwest (202.5 – 247.5)	13.21	26.91	1.934	0.332	0.780	0.960
	West (247.5 – 292.5)	16.95	8.03	0.764	0.193	-0.574	0.536
	Northwest (292.5 – 337.5)	15.84	9.65	0.896	0.242	-0.098	0.958
LS (m)	0.03 –0.49	27.16	6.02	0.188	0.177	-1.371	0.207
	0.49 –1.10	25.98	13.26	0.511	0.242	-0.882	0.560
	1.10 –1.77	18.35	24.09	1.296	0.406	0.485	0.803
	1.77 –2.50	20.19	36.33	5.624	0.992	2.213	0.972
	2.50 –7.81	8.32	20.30	3.0157	0.683	1.364	1.000

With FR, SI, WoE and LR models, the results of spatial relationship between conditioning factors and landslide locations revealed that for the distance to roads, the classes of <200m has the highest values 1.806, 0.627, 1.868 and 1.000 for all models respectively. It was observed during fieldworks that many landslides in Rwanda occur alongside the roads due to slope stability modification. Results of the study also disclosed that spatial relationship values for precipitation classes have indications with increasing of precipitation in the study area. Precipitation class of 1400-1700mm has the highest value for FR, SI, WoE and LR models (3.573, 0.618, 1.572 and 1.000 respectively) followed by the class of 1200-1400 mm (2.928, 0.492,1.127 and 0.997 values for FR, SI, WoE and LR respectively). This is entirely the western part of the country where most landslides events are frequently recorded. Additionally, the relationship between slope degrees and landslide probability showed that the class of > 26.11^{o} has the highest FR, SI, WoE and LR values (4.861, 0.966, 1.678 and 1.000 respectively), whereas the class of 0^{o}-4.81^{o} gives the lowest values (0.089, 0.012, -1.041 and 0.238 respectively for FR, SI, WoE and LR models). Basically, as the slope increases, the shear stress increases, and gentle slope angles are normally expected to have lower weights values since they are associated with lower shear stresses (Pourghasemi et al. 2012).

Regarding elevation factor, the classes of 2,196-2,813m and 2,813-4,495m have the highest values (4.607; 0.911; 3.687; 0.994 and 3.010; 0.673; 2.030 and 1.000 respectively for FR, SI, WoE and LR models) while the class of 920-1,537m has the lowest values for all the models. The results of this study revealed that the spatial relationship values increased with increasing elevation. Besides, the study showed that areas with schist lithology are highly susceptible to landslides in Rwanda. Also, clay loam soils and cropland areas were found highly susceptible in the study area. Soil depth class of 0.5-1.0m was proven to have the highest spatial relationship with 5.227; 0.997; 2.604 and 1.00 values for FR, SI, WoE and LR models respectively.

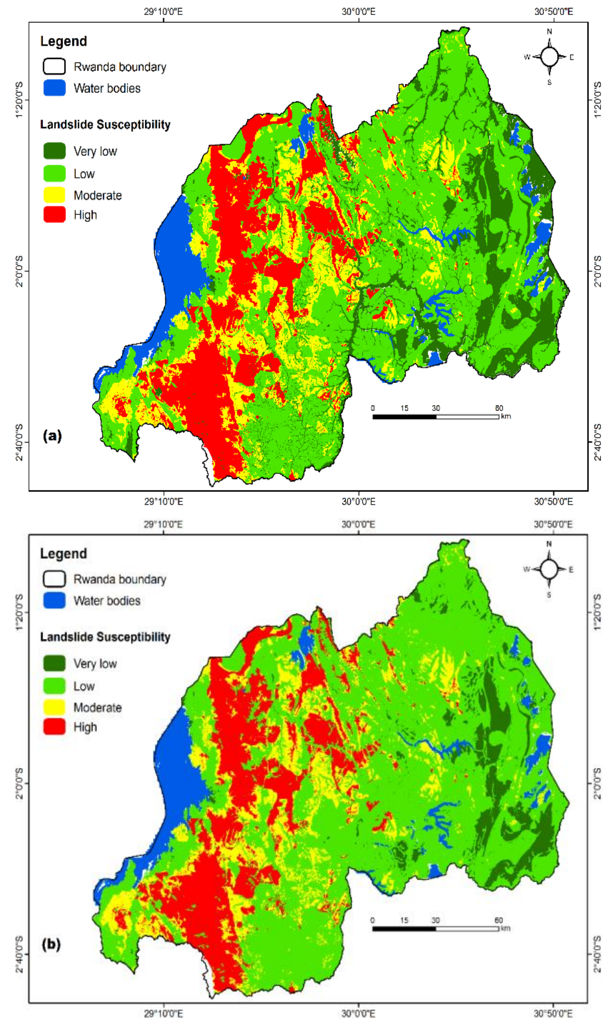
The overall analysis of conditioning factors revealed that 10 factors are more influential than others. They include slope degree, precipitation, elevation, curvature, aspect, soil depth, land use/land cover, soil texture, distance to roads and topographic factors (Table 2).

Table 2 Statistical coefficients generated by R Software

Predicting factors	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	-2.664e-08	6.479e+02	0	1
Elevation	4.344e-16	2.431e-04	0	1
TWI	-4.147e-14	2.544e-02	0	1
Soiltexture	2.081e-14	4.092e-02	0	1
Slope	7.173e-08	1.744e+03	0	1
Rainfall	8.725e-06	4.740e-04	0	1
NDVI	-2.740e-14	5.781e-01	0	1
LS	7.559e-07	1.838e+04	0	1
Lithology	-8.647e-15	3.695e-02	0	1
LANDUSE	1.556e-13	8.501e-02	0	1
Distance_river	-3.287e-17	3.104e-05	0	1
Distance_road	1.570e-17	1.995e-05	0	1
Curvature	1.031e-13	1.237e-01	0	1
Aspect	1.770e-16	6.344e-04	0	1
Soil depth	2.1851e-11	4.032e-02	0	1

4.2. Generation of landslide susceptibility maps

The present study explored and compared different probabilistic and statistical methods to produce landslide susceptibility maps for the area under investigation (Fig. 5). Application of FR, SI, LR and WoE helped authors to generate LSM for Rwanda using 14 factors. Four landslide susceptibility maps were therefore produced (Fig. 5).



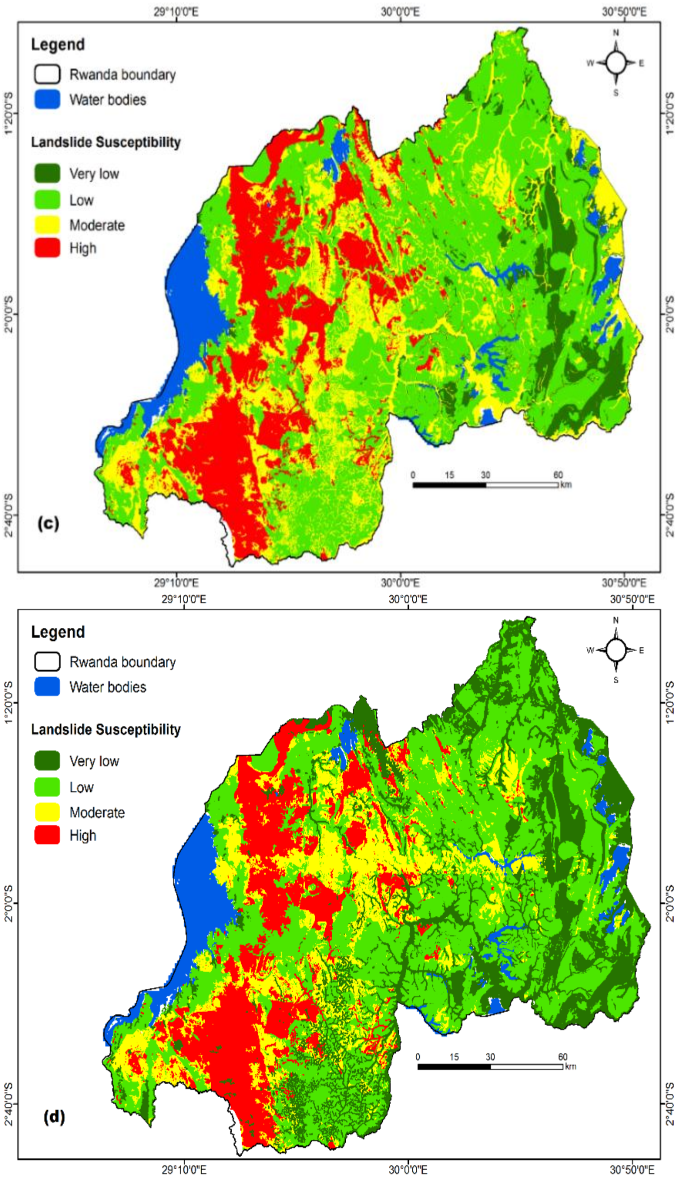


Fig.5. Landslide susceptibility maps using: (a) FR model; (b) LR model; (c) SI model and (d) WoE model.

The FR, SI, WoE and LR models were constructed using the training points and layers of landslide conditioning factors (Fig. 2). Previous studies recommended that the combination of more factors play a big role in generating accurate landslide susceptibility maps (Piller 2016; Nsengiyumva et al. 2018; Pradhan and Lee 2010). Thus, the calculated landslide susceptibility index (LSI) values for the entire study area using FR, SI, WoE and LR models were between 0.00 and 1.00. Finally, all landslide susceptibility index values of the area under investigation were divided into four classes using the equal interval method to generate landslide susceptibility maps of Rwanda. Areas were respectively classified as high susceptibility (0.75−1), moderate susceptibility (0.50−0.75, low susceptibility (0.25−0.50) and very low susceptibility (0.00–0.25), (Fig.5.)

The produced susceptibility maps (Fig.5) reflect what was observed from field-work in the study area (Fig. 2). Moreover, the disaster losses database available in the Ministry in Charge of Emergency Management in Rwanda, and interviews with local experts in the study area, disclose that landslide susceptibility is spatially dispersed across the study area. However, steep slope zones become the highly susceptible. Additionally, higher precipitations were found the major triggers of landslide events in the study area. Landslide hazards affect people, livestock, crops, family houses and other different important infrastructure including roads and bridges. Additionally, fieldwork confirmed that majority of past landslide events occurred in crop land, built-up land and forest land (Fig. 3k and Fig.2) and this can also undermine the agriculture sector.

4.3. Validation and comparison of models

In landslide susceptibility modeling, most scientists agree that appropriate methods need to be applied to evaluate the performance of landslide susceptibility models. However, there is no clear agreement concerning which methods are the best or must be used given regional variability. For this study, the results of the four landslide susceptibility models were validated using validation datasets obtained during fieldwork a stated earlier. Additionally, we used the AUC to evaluate the model results (Fig.6). Results confirmed that all four models have good susceptibility prediction capacity. Therefore, the AUC values of 92.7%, 86.9%, 81.2% and 79.5% respectively for WoE, FR, LR and SI models showed reasonable prediction for all the models. However, the results indicated that the WoE model performed the best (92.7%) in mapping landslide susceptibility in the study area whereas SI model showed the smallest AUC value (79.5%).The weight of evidence proved the best model capable of combining expert knowledge with field datasets in susceptibility modeling.

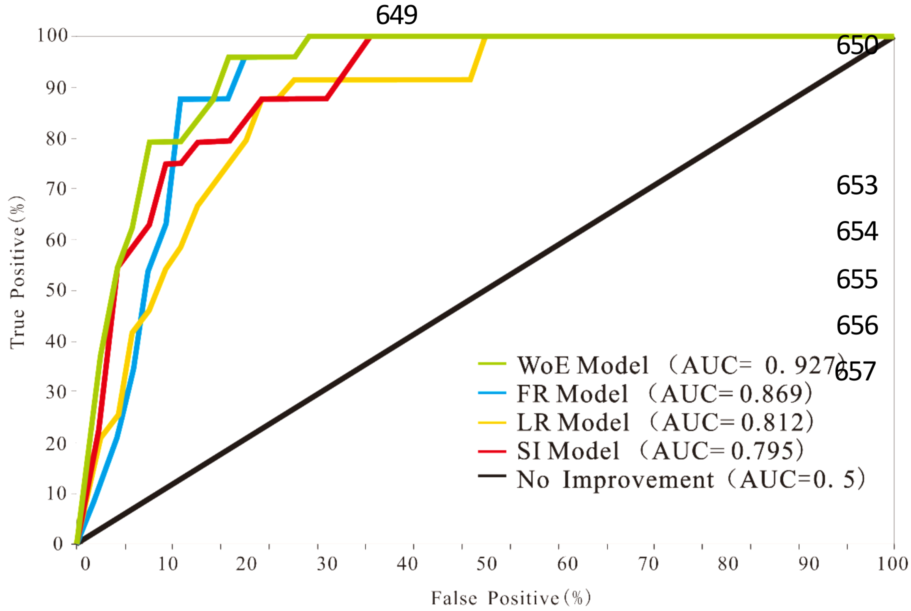


Fig.6. Prediction rates with AUC for model performance.

The landslide susceptibility maps are commonly considered as a fundamental stage in managing landslide risks (Chen et al. 2018a). They play a big role in identifying critical risk zones; inform relocation of families from hazard- prone zones as well as development of landslide mitigation infrastructure. This stage of risk management cycle can also help to identify significant triggers of landslides. Significantly, it was noted from results of this study that most of past landslide events occurred between March and May of the previous years and rain has been confirmed a major trigger of landslides.

Analytically, landslide risks and exposure in the study area was found to be very high. Majority of the affected population are rural people living in extreme poverty. Most of them live in poorly constructed houses that are located in highly vulnerable landslide zones, and are unable to cope at the event of any landslide disaster. Districts were predicted to be highly prone to landslide hazards including Ngororero, Nyabihu, Karongi, Nyamasheke, Gakenke, Muhanga, Rusizi, Nyamagabe, Rulindo, Musanze and Nyaruguru (Fig. 1, 2 and 5). Rwanda is tenderly referred to as a country of thousand hills (Das et al. 2010), as depicted by its topography with volcanoes and dominant Congo-Nile ridge hills. All prone Districts are located within high elevation zones. Inversely, nine districts were found stable to landslide hazards namely Kirehe, Rwamagana, Ngoma, Kayonza, Gatsibo, Nyagatare, Bugesera, Gisagara and Nyanza. These have mostly been classified as low and very low susceptible (Fig. 1 and 5). The population growth increases the pressure on land even in steep slope areas through informal and illegal settlements by the local community members. Inappropriate land use has continued exacerbate the impacts of landslide hazards in the study area. This situation requires landslide resilience building from national to local level.

The spatial distribution of different landslide susceptibility classes is illustrated in Table 3. The study has revealed that landslide susceptibility is spatially dispersed across the entire Rwanda. The areas covered under different susceptibility categories (high, moderate, low and very low) were calculated using ArcMap 10.3, as shown in Table 3. In the case of WoE model, it can be observed that the high susceptibility class accounts for 18.03% of the study area. The moderate, low and very low susceptibility classes account for 27.4%, 44.27% and 10.3% of the study area, respectively. For the landslide susceptibility map generated by the FR model, 8.46% of the study area belongs to very low susceptibility class. The low susceptibility class covers 45.3% of the study area and the moderate susceptibility class accounts for 24.54% of the study area, while the high susceptibility class accounts for 21.7% of the study area. Regarding landslide susceptibility map produced using LR model, the very low and low susceptibility classes account 6.16% and 47.52 of the study area, respectively. 24.63% of the study area falls into the moderate susceptibility class and 21.71% falls into the high susceptibility class. This comparative study confirms that the four models (FR, SI, WoE and LR) are promising approaches to map landslide in Rwanda since they all produced reasonable susceptibility maps. Additionally based on the landslide susceptibility map produced using SI model, it is revealed that 6.11% of the study area falls into the very low susceptibility class, while 44% of the study area falls into the low susceptibility class. Both moderate and high susceptibility classes account for 29.64% and 20.25% of the study area, respectively.

Table 3 Landslide susceptible areas for WOE, FR, LR and SI models

Susceptibility class	**WOE**		**FR**		**LR**		**SI**	
	Zone under category (%)	Area (Km^2)	Zone under category (%)	Area (Km^2)	Zone under category (%)	Area (Km^2)	Zone under category (%)	Area (Km^2)
Very low susceptible	10.3	2546.79	8.46	2091.83	6.15	1520.66	6.11	1510.77
Low susceptible	44.27	10946.25	45.3	11200.93	47.51	11747.38	44	10879.5
Moderate susceptible	27.4	6774.96	24.54	6067.79	24.63	6090.04	29.64	7328.82
High susceptible	18.03	4458.12	21.7	5365.57	21.71	5368.04	20.25	5007.03
Total	**100**	**24726.12**	**100**	**24726.12**	**100**	**24726.12**	**100**	**24726.12**

In general, it is shown that 20.42% (5,048.07 km^2) of Rwanda falls into high susceptibility class whereas 7.75% (1,917.5 km^2) falls into very low susceptibility class (Table 3).

The Results of the current susceptibility modeling study is in conformity with the previous studies in the study area (Nsengiyumva 2012; Nsengiyumva et al. 2018; Piller 2016; Nduwayezu 2017; MIDIMAR 2012, 2015b). This agreement confirms the western and the northern parts as landslide prone zones while the eastern part is the least susceptible. Objectives of the study have been reached since the applied models (FR, SI, WoE and LR) produced reasonable susceptibility maps, and AUC/ROC was used to assess the model performance. The landslide susceptibility mapping should be followed by detailed risk assessment and vulnerability analysis to improve risk reduction practices. Disaster risk is normally a product of hazard, risk, vulnerability and exposure. Further quantitative studies should therefore be conducted to address landslide management uncertainties. However, the findings from the current study confirmed that the comparison of FR, SI, LR and WoE is a very reasonable and a promising approach to generate landslide susceptibility maps within prone areas of Rwanda as well as the centre-eastern-Africa region. Landslide control practices such as contouring, strip-cropping and terracing should be adopted especially for areas falling into moderate and high susceptibility classes to lessen the impacts.

5. Conclusion

Landslide hazards are very recurrent in the study area. In the current study, WOE, FR, LR, and SI models were applied to map landslide susceptibility in Rwanda. The four models have never been compared before in the entire literature related to landslide susceptibility studies for the Africa region.

A landslide inventory map and 14 maps of conditioning factors were applied to simulate the models. Thus, ROC Curves were used to evaluate the performance of the models. For this case study, it was disclosed that the WOE model achieved the highest AUC value (92.7%) while the SI model produced a lowest AUC value (79.5%). However, all the four models employed in this study are promising approaches for landslide susceptibility studying in Rwanda. Generally, the western part of Rwanda was modeled as highly susceptible to landslides comparing to other parts of the country. Therefore, further detailed studies should be conducted to compare quantitative and process-driven models using different conditioning factors. Conclusively, the results of the current study may be useful for landslide risk mitigation and land use planning in the study area, and in other areas with similar terrain conditions as well as environmental settings. More studies should be performed to include other important conditioning factors that exacerbate increases in susceptibility especially anthropogenic factors.

Acknowledgments

The authors are thankful to Professor Wei Ouyang, Associate Editor of the journal of Science of the Total Environment and the anonymous reviewers for their valuable contribution to this Manuscript. With their comments, inputs, advice and suggestions, we were able to significantly improve the quality of this paper. Special thanks are also addressed to Professor Geping Luo, for his scientific inspiration and mentorship. Therefore, the authors express thanks and appreciation to the Rwanda Ministry in Charge of Emergency Management for the provision of data and information. This research has been supported by the China–Africa Joint Research Centre Project of the Chinese Academy of Sciences (SAJC201610). The authors declare no conflicts of interest.

References

Ahmed, B., and A. Dewan. 2017. Application of Bivariate and Multivariate Statistical Techniques in Landslide Susceptibility Modeling in Chittagong City Corporation, Bangladesh. *Remote Sensing* 9 (4):304.

Akgun, A. 2012. A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey. *Landslides* 9 (1):93-106.

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

Baum, R. L., W. Z. Savage, and J. W. Godt. 2008. TRIGRS-A Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis, version 2.0: US Geological Survey.

Bizimana, H., and O. Sönmez. 2015. Landslide Occurrences in The Hilly Areas of Rwanda, Their Causes and Protection Measures. *Disaster Science and Engineering* 1 (1):1-7.

Brenning, A. 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Natural Hazards and Earth System Science* 5 (6):853-862.

Bui, D. T., O. Lofman, I. Revhaug, and O. Dick. 2011. Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression. *Natural hazards* 59 (3):1413.

Cervi, F., M. Berti, L. Borgatti, F. Ronchetti, F. Manenti, and A. Corsini. 2010. Comparing predictive capability of statistical and deterministic methods for landslide susceptibility mapping: a case study in the northern Apennines (Reggio Emilia Province, Italy). *Landslides* 7 (4):433-444.

Chen, W., J. Peng, H. Hong, H. Shahabi, B. Pradhan, J. Liu, A.-X. Zhu, X. Pei, and Z. Duan. 2018a. Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. *Science of The Total Environment* 626:1121-1135.

Chen, W., H. R. Pourghasemi, and Z. Zhao. 2017a. A GIS-based comparative study of Dempster-Shafer, logistic regression and artificial neural network models for landslide susceptibility mapping. *Geocarto international* 32 (4):367-385.

Chen, W., X. Xie, J. Peng, J. Wang, Z. Duan, and H. Hong. 2017b. GIS-based landslide susceptibility modelling: a comparative assessment of kernel logistic regression, Naïve-Bayes tree, and alternating decision tree models. *Geomatics, Natural Hazards and Risk* 8 (2):950-973.

Chen, W., S. Zhang, R. Li, and H. Shahabi. 2018b. Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Science of The Total Environment* 644:1006-1018.

Coppola, D. P. 2006. *Introduction to international disaster management*: Elsevier.

Dahal, R. K., S. Hasegawa, A. Nonomura, M. Yamanaka, S. Dhakal, and P. Paudyal. 2008a. Predictive modelling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of-evidence. *Geomorphology* 102 (3):496-510.

Dahal, R. K., S. Hasegawa, A. Nonomura, M. Yamanaka, T. Masuda, and K. Nishino. 2008b. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. *Environmental Geology* 54 (2):311-324.

Das, I., S. Sahoo, C. van Westen, A. Stein, and R. Hack. 2010. Landslide susceptibility assessment using logistic regression and its comparison with a rock mass classification system, along a road section in the northern Himalayas (India). *Geomorphology* 114 (4):627-637.

Desmet, P., and G. Govers. 1996. A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. *Journal of soil and water conservation* 51 (5):427-433.

Devkota, K. C., A. D. Regmi, H. R. Pourghasemi, K. Yoshida, B. Pradhan, I. C. Ryu, M. R. Dhital, and O. F. Althuwaynee. 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. *Natural Hazards* 65 (1):135-165.

Dou, J., H. Yamagishi, Z. Zhu, A. P. Yunus, and C. W. Chen. 2018. TXT-tool 1.081-6.1 A Comparative Study of the Binary Logistic Regression (BLR) and Artificial Neural Network (ANN) Models for GIS-Based Spatial Predicting Landslides at a Regional Scale. In *Landslide Dynamics: ISDR-ICL Landslide Interactive Teaching Tools*: Springer, 139-151.

EAC. 2012. Disaster Risk Reduction and Management Strategy. EAC-Arusha Tanzania: East African Community, 63.

Godt, J., R. Baum, W. Savage, D. Salciarini, W. Schulz, and E. Harp. 2008. Transient deterministic shallow landslide modeling: requirements for susceptibility and hazard assessments in a GIS framework. *Engineering Geology* 102 (3):214-226.

Gökceoglu, C., and H. Aksoy. 1996. Landslide susceptibility mapping of the slopes in the residual soils of the Mengen region (Turkey) by deterministic stability analyses and image processing techniques. *Engineering Geology* 44 (1-4):147-161.

Gorsevski, P. V., P. Jankowski, and P. E. Gessler. 2006. An heuristic approach for mapping landslide hazard by integrating fuzzy logic with analytic hierarchy process. *Control and Cybernetics* 35:121-146.

Karamage, F., C. Zhang, A. Kayiranga, H. Shao, X. Fang, F. Ndayisaba, L. Nahayo, C. Mupenzi, and G. Tian. 2016. Usle-based assessment of soil erosion by water in the Nyabarongo River Catchment, Rwanda. *International journal of environmental research and public health* 13 (8):835.

Karamage, F., C. Zhang, T. Liu, A. Maganda, and A. Isabwe. 2017. Soil Erosion Risk Assessment in Uganda. *Forests* 8 (2):52.

Kuriakose, S. L., L. Van Beek, and C. Van Westen. 2009. Parameterizing a physically based shallow landslide model in a data poor region. *Earth Surface Processes and Landforms* 34 (6):867-881.

Maes, J., C. Parra, K. Mertens, B. Bwambale, L. Jacobs, J. Poesen, O. Dewitte, L. Vranken, A. de Hontheim, and C. Kabaseke. 2018. Questioning network governance for disaster risk management: Lessons learnt from landslide risk management in Uganda. *Environmental Science & Policy* 85:163-171.

Mazzanti, P., F. Bozzano, I. Cipriani, and A. Prestininzi. 2015. New insights into the temporal prediction of landslides by a terrestrial SAR interferometry monitoring case study. *Landslides* 12 (1):55-68.

McCool, D., L. Brown, G. Foster, C. Mutchler, and L. Meyer. 1987. Revised slope steepness factor for the Universal Soil Loss Equation. *Transactions of the ASAE* 30 (5):1387-1396.

McDougall, S., and O. Hungr. 2005. Dynamic modelling of entrainment in rapid landslides. *Canadian Geotechnical Journal* 42 (5):1437-1448.

MIDIMAR. 2012. Disaster high risk zones on floods and landslides, edited by R. R. a. Preparedness. Prevention Web: UNISDR, 33.

———. 2018. *The National Risk Atlas* 2015a [cited 30 November 2018].

———. *National Risk Atlas of Rwanda. Available online: http://midimar.gov.rw/index.php?id=76&tx_pagebrowse_pi1%5Bpage%5D=3&cHash=f1359c48518cebbd859ec0e04d7c02f3 (Accessed on 16 January 2018).* 2015b [cited.

———. 2018. Rwanda Rapid Post Disaster Needs Assessment (PDNA). Kigali-Rwanda: Ministry in Charge of Emergency Management, 232.

Mohammady, M., H. R. Pourghasemi, and B. Pradhan. 2012. Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster–Shafer, and weights-of-evidence models. *Journal of Asian Earth Sciences* 61:221-236.

Monsieurs, E., L. Jacobs, C. Michellier, J. B. Tchangaboba, G. B. Ganza, F. Kervyn, J.-C. M. Mateso, T. M. Bibentyo, C. K. Buzera, and L. Nahimana. 2018. Landslide inventory for hazard assessment in a data-poor context: a regional-scale approach in a tropical African environment. *Landslides*:1-15.

Nahayo, L., L. Li, G. Habiyaremye, M. Richard, V. Mukanyandwi, E. Hakorimana, and C. Mupenzi. 2018. Extent of disaster courses delivery for the risk reduction in Rwanda. *International journal of disaster risk reduction* 27:127-132.

NASA. 2018. *DEM, http://www.dwtkns.com/strm30m/, Accessed,* [cited 30 July 2018].

Ndayisaba, F., H. Guo, A. Bao, H. Guo, F. Karamage, and A. Kayiranga. 2016. Understanding the spatial temporal vegetation dynamics in Rwanda. *Remote Sensing* 8 (2):129.

Nduwayezu, E. N., Jean-Baptiste; BUgnon, Pierre-Charles; Jaboyedoff, Michel; Derron, Marc-Henri. 2017. Debris flows susceptibility mapping under tropical rain conditions in Rwanda. In *19th EGU General Assembly, EGU2017, proceedings from the conference held 23-28 April, 2017, p.16352*. Vienna, Austria: COPERNICUS.

Neaupane, K. M., and M. Piantanakulchai. 2006. Analytic network process model for landslide hazard zonation. *Engineering Geology* 85 (3):281-294.

Neuhäuser, B., and B. Terhorst. 2007. Landslide susceptibility assessment using "weights-of-evidence" applied to a study area at the Jurassic escarpment (SW-Germany). *Geomorphology* 86 (1):12-24.

Nichol, J., and M. Wong. 2005. Satellite remote sensing for detailed landslide inventories using change detection and image fusion. *International journal of remote sensing* 26 (9):1913-1926.

Nsengiyumva, J. B. 2012. Disaster High Risk Zones on Floods and Landslides. *Kigali: MIDIMAR.*

Nsengiyumva, J. B., G. Luo, L. Nahayo, X. Huang, and P. Cai. 2018. Landslide susceptibility assessment using spatial multi-criteria evaluation model in Rwanda. *International journal of environmental research and public health* 15 (2):243.

Piller, A. N. 2016. Precipitation Intensity Required for Landslide Initiation in Rwanda.

Pisano, L., V. Zumpano, Ž. Malek, C. M. Rosskopf, and M. Parise. 2017. Variations in the susceptibility to landslides, as a consequence of land cover changes: A look to the past, and another towards the future. *Science of The Total Environment* 601:1147-1159.

Pourghasemi, H., B. Pradhan, C. Gokceoglu, and K. D. Moezzi. 2012. Landslide susceptibility mapping using a spatial multi criteria evaluation model at Haraz Watershed, Iran. In *Terrigenous mass movements*: Springer, 23-49.

Pradhan, B. 2010a. Application of an advanced fuzzy logic model for landslide susceptibility analysis. *International Journal of Computational Intelligence Systems* 3 (3):370-381.

———. 2010b. Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches. *Journal of the Indian Society of Remote Sensing* 38 (2):301-320.

Pradhan, B., and S. Lee. 2010. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software* 25 (6):747-759.

Ramani, S. E., K. Pitchaimani, and V. R. Gnanamanickam. 2011. GIS based landslide susceptibility mapping of Tevankarai Ar sub-watershed, Kodaikkanal, India using binary logistic regression analysis. *Journal of Mountain Science* 8 (4):505-517.

Refice, A., and D. Capolongo. 2002. Probabilistic modeling of uncertainties in earthquake-induced landslide hazard assessment. *Computers & Geosciences* 28 (6):735-749.

Regmi, A. D., K. C. Devkota, K. Yoshida, B. Pradhan, H. R. Pourghasemi, T. Kumamoto, and A. Akgun. 2014. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arabian Journal of Geosciences* 7 (2):725-742.

Schilirò, L., L. Montrasio, and G. S. Mugnozza. 2016. Prediction of shallow landslide occurrence: Validation of a physically-based approach through a real case study. *Science of The Total Environment* 569:134-144.

Sinarta, I. N., A. Rifa'i, T. Faisal Fathani, and W. Wilopo. 2017. Slope Stability Assessment Using Trigger Parameters and SINMAP Methods on Tamblingan-Buyan Ancient Mountain Area in Buleleng Regency, Bali. *Geosciences* 7 (4):110.

Stanley, T., and D. B. Kirschbaum. 2017. A heuristic approach to global landslide susceptibility mapping. *Natural Hazards* 87 (1):145-164.

Sujatha, E. R., P. Kumaravel, and V. Rajamanickam. 2012. Landslide susceptibility mapping using remotely sensed data through conditional probability analysis using seed cell and point sampling techniques. *Journal of the Indian Society of Remote Sensing* 40 (4):669-678.

Terlien, M. T., C. J. Van Westen, and T. W. van Asch. 1995. Deterministic modelling in GIS-based landslide hazard assessment. In *Geographical information systems in assessing natural hazards*: Springer, 57-77.

Turner, D., A. Lucieer, and S. M. De Jong. 2015. Time series analysis of landslide dynamics using an unmanned aerial vehicle (UAV). *Remote Sensing* 7 (2):1736-1757.

USGS. Usgs global visualization viewer: Earth resources observation and science center (eros). http://glovis.usgs.gov/index.shtml (20 December 2016). .

Van Den Eeckhaut, M., J. Poesen, G. Verstraeten, V. Vanacker, J. Moeyersons, J. Nyssen, and L. Van Beek. 2005. The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides. *Geomorphology* 67 (3):351-363.

Van Westen, C., T. W. Van Asch, and R. Soeters. 2006. Landslide hazard and risk zonation—why is it still so difficult? *Bulletin of Engineering Geology and the Environment* 65 (2):167-184.

Wu, J.-H., J.-S. Lin, and C.-S. Chen. 2009. Dynamic discrete analysis of an earthquake-induced large-scale landslide. *International Journal of Rock Mechanics and Mining Sciences* 46 (2):397-407.

Yalcin, A. 2008. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. *Catena* 72 (1):1-12.

Yalcin, A., S. Reis, A. Aydinoglu, and T. Yomralioglu. 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 (3):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 (6):1125-1138.

Youssef, A. M., H. R. Pourghasemi, B. A. El-Haddad, and B. K. Dhahry. 2016. Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir Region, Saudi Arabia. *Bulletin of Engineering Geology and the Environment* 75 (1):63-87.

Zêzere, J., S. Pereira, R. Melo, S. Oliveira, and R. Garcia. 2017. Mapping landslide susceptibility using data-driven methods. *Science of The Total Environment* 589:250-267.

Zêzere, J., E. Reis, R. Garcia, S. Oliveira, M. Rodrigues, G. Vieira, and A. Ferreira. 2004. Integration of spatial and temporal data for the definition of different landslide hazard scenarios in the area north of Lisbon (Portugal). *Natural Hazards and Earth System Science* 4 (1):133-146.

Zhou, G., T. Esaki, Y. Mitani, M. Xie, and J. Mori. 2003. Spatial probabilistic modeling of slope failure using an integrated GIS Monte Carlo simulation approach. *Engineering Geology* 68 (3):373-386.

Zschau, J., and A. N. Küppers. 2013. *Early warning systems for natural disaster reduction*: Springer Science & Business Media.