



Research article

GIS-based data-driven bivariate statistical models for landslide susceptibility prediction in Upper Tista Basin, India

Jayanta Das^a, Pritam Saha^b, Rajib Mitra^c, Asrafal Alam^a, Md Kamruzzaman^{d,*}^a Department of Geography, Rampurhat College, PO- Rampurhat, Dist- Birbhum, 731224, India^b Department of Geography, Cooch Behar Panchanan Barma University, P.O.- Cooch Behar, Dist- Cooch Behar, 736101, India^c Department of Geography and Applied Geography, University of North Bengal, PO- North Bengal University, Dist- Darjeeling, 734013, India^d Institute of Bangladesh Studies, University of Rajshahi, Bangladesh

ARTICLE INFO

Keywords:

Landslide susceptibility mapping
Evidence belief function
Frequency ratio
Index of entropy
ROC-AUC
Sikkim himalayan region

ABSTRACT

Predicting landslides is becoming a crucial global challenge for sustainable development in mountainous areas. This research compares the landslide susceptibility maps (LSMs) prepared from five GIS-based data-driven bivariate statistical models, namely, (a) Frequency Ratio (FR), (b) Index of Entropy (IOE), (c) Statistical Index (SI), (d) Modified Information Value Model (MIV) and (e) Evidential Belief Function (EBF). These five models were tested in the high landslides-prone humid sub-tropical type Upper Tista basin of the Darjeeling-Sikkim Himalaya by integrating the GIS and remote sensing. The landslide inventory map consisting of 477 landslide locations was prepared, and about 70% of all landslide data was utilized for training the model, and 30% was used to validate it after training. A total of fourteen landslide triggering parameters (elevation, slope, aspect, curvature, roughness, stream power index, TWI, distance to stream, distance to road, NDVI, LULC, rainfall, modified fourier index, and lithology) were taken into consideration for preparing the LSMs. The multicollinearity statistics revealed no collinearity problem among the fourteen causative factors used in this study. Based on the FR, MIV, IOE, SI, and EBF approaches, 12.00%, 21.46%, 28.53%, 31.42%, and 14.17% areas, respectively, identified in the high and very high landslide-prone zones. The research also revealed that the IOE model has the highest training accuracy of 95.80%, followed by SI (92.60%), MIV (92.20%), FR (91.50%), and EBF (89.90%) models. Consistent with the actual distribution of landslides, the very high, high, and medium hazardous zones stretch along the Tista River and major roads. The suggested landslide susceptibility models have enough accuracy for usage in landslide mitigation and long-term land use planning in the study area. Decision-makers and local planners may utilise the study's findings. The techniques for determining landslide susceptibility can also be employed in other Himalayan regions to manage and evaluate landslide hazards.

1. Introduction

Our natural environment faces severe natural hazards such as cyclones, drought, earthquakes, and landslides in different parts of the globe [1]. High rates of inherent risk are present in many parts of the world, including those prone to landslides [2]. Landslides are significant natural catastrophes that typically occur on natural slopes in hilly regions and are caused by various triggering factors,

* Corresponding author.

E-mail address: mkzaman@ru.ac.bd (M. Kamruzzaman).<https://doi.org/10.1016/j.heliyon.2023.e16186>

Received 13 February 2023; Received in revised form 23 March 2023; Accepted 9 May 2023

Available online 12 May 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>.

including rainfall and earthquakes [3]. The fundamental cause of landslides is slope instability, which occurs when the slope shifts from stable to unstable conditions due to gravitational force [4]. The hilly regions of the world are the most probable locations to be impacted by landslides, which are responsible for enormous damage to both property and human life [5]. As most of the hilly areas, such as the Himalayas, Rocky, Alps and Andes Mountains, are prone to landslides, countries such as India, the United States, China, and the Philippines have been impacted the hardest in recent years [6]. An estimated 100 million people were affected by the 346 disasters included in the 2015–2016 International Disaster Analysis (IDA) report, which led to the deaths of 22,773 persons across the globe [7]. According to the National Aeronautics and Space Administration (NASA) records of 10,084 landslide incidents from 2007 to 2017, the world experiences significant economic damage every year due to landslide hazards [8].

Landslides often cause the most significant amount of damage in Asia. India experiences most of its landslides during the monsoon season. There is a potential for landslides to occur over up to 15% of the total land area in India [9]. There are frequent landslides that occur mainly in the mountainous parts of the Himalaya and the Western Ghats of India. The Himalayan region accounts for around 80% of all landslides in India [10]. Approximately 43% of India's total landslide-prone area is located in the Darjeeling-Sikkim Himalayan region [11]. Landslides are the most severe geohazards at the top of rug cliffs in Sikkim Himalaya due to earthquakes, extreme rainfall, and anthropogenic activities [12]. In the landlocked state, due to the complicated geological setting and substantial rain in the terrain, the extent of damage caused by rapid earth movements in several parts of Sikkim each year is quite significant. In addition, unsustainable development practices, particularly the construction of haphazardly built settlements and roads, have further contributed to the frequent incidents of landslides in this area. Therefore, it is essential to identify locations vulnerable to landslides to prevent potential damage. Future landslide hazards must also be predicted in mountainous regions to ensure sustainable development. Forecasting future landslide occurrences is becoming essential to prevent the detrimental effects of landslides. It is necessary to identify the conditioning factors impacting slope stability to analyse the vulnerability of landslides and limit damages using suitable mitigation strategies.

Landslide investigations have recently begun to be conducted in a RS and GIS environment in order to acquire more accurate results [4,13,14]. Landslide studies can be classified as pre-disaster or post-disaster studies, with landslide prediction falling under pre-disaster studies [15,11]. In landslide prediction, a specific region may be classified into homogenous zones depending on the likelihood of landslide occurrences based on several conditional factors. The heuristic, deterministic, probabilistic, and machine learning approaches are the four main approaches used by many researchers to predict landslides [3,15,9,16]. The expert opinion is employed to provide weightage to various triggering factors in the heuristic method [17,18,19,20,2]; however, in the deterministic approach like safety factor calculation, the parameters may be figured out either on the field or in the laboratory [21]. The basic foundation of the probabilistic approach is the established association between landslide distribution and causative factors in a specific location. Some of the most used probabilistic models for studies of landslide prediction are the Weight of Evidence (WoE) [22,15,7], Frequency Ratio (FR) [23,11,24,25,26], Evidential Belief Function (EBF) [12], Statistical Index (SI) [27] Modified Information Value Model (MIV) [28], Index of Entropy (IOE) [29], Logistic Regression (LR) [30,26,31] and Multiple Regression Analysis (MRA) [32,33]. The EBF, one of the bivariate approaches, relies on the Dempster-Shafer theory of evidence [34,35], which has been demonstrated to be a reliable and efficient method in landslide prediction. EBF's ability to quantify the belief, disbelief, uncertainty, and plausibility values related to the model is its main advantage over other bivariate statistical models [15]. The machine learning approach is a nonlinear dynamic method that predicts the geographical distribution of landslides based on previous data. Artificial Neural Network (ANN) [36], Boosted Regression Tree (BRT) [10], Support Vector Machines (SVM) [37,38,39], Adaptive Neuro-Fuzzy Inference System (ANFIS) [40], Random Forest (RF) [41,42,43], Decision Tree [16], Multilayer Perceptron Neural Network (MLPNN) [36,44] and Convolutional Neural Network (CNN) [8,45] are globally used machine learning models that forecast the unknown association between the events of landslides and the causative factors responsible for such events. Any of the aforementioned methods may be used to create landslide prediction models; however, each method has drawbacks that limit its efficiency when employed individually. Although these different landslide prediction modelling studies included a variety of methodologies, relatively few studies compared the various statistical models they used in their research. Because of this, it is critical to examine multiple statistical methods to determine which ones are most appropriate for landslide prediction.

Several heuristic, probabilistic, and machine learning approaches have been used for landslide susceptibility mapping and analysis in the Sikkim Himalayan region in the last decade [11]. prepared a landslide susceptibility map (LSM) utilizing the frequency ratio method with ten important determining variables to identify the most dominant factors that cause frequent landslides in the Lachung River basin of North Sikkim. Recently [12], used an ensemble method of two statistical models (geographically weighted regression and evidential belief function) and one machine learning model (random forest) to investigate landslide prone zones in the Upper Rangit basin in West Sikkim district of Sikkim. According to the relative relevance of factor analysis, the landslide's occurrence was mostly influenced by the variables of distance to the river, land-use land-cover (LULC), road density, and drainage density [8]. prepare LSM in the Rorachu river basin of Sikkim Himalaya, using deep learning and advance machine learning algorithms. The East Sikkim Himalayan region's LSM was developed by Ref. [46] while taking into account a hybrid ensemble of machine learning techniques. According to their analysis rainfall emerges as the most crucial triggering factor for landslide occurrence among the factors used for modelling, followed by aspect, elevation, and LULC.

In the present study, landslide prediction mapping was carried out in the Upper Tista basin of Darjeeling-Sikkim Himalaya through field visits and GIS-based data-driven bivariate statistical models because few researchers had modelled their research on a specific watersheds or river basins in Sikkim Himalaya. Assigning weights for each input map by comparing an individual conditional factors map with a landslide inventory map is the fundamental advantage of bivariate statistical models [27]. In this context, the FR, IOE, SI,

MIV, and EBF models were used to determine which locations had a greater possibility for future landslide occurrences, and their modelling performances were evaluated using ROC curves. For this purpose, fourteen thematic maps that represented elevation, slope, aspect, curvature, roughness, stream power index (SPI), topographical wetness index (TWI), distance to stream (DtS), distance to road (DtR), normalized difference vegetation index (NDVI) land-use land-cover (LULC), rainfall, modified fournier index (MFI), and lithology were used for landslide prediction mapping of the area facing extreme landslide events in the past. The main novelty is that different bivariate statistical approaches have been utilized for the first time to identify which method is best suited for landslide prediction in the Sikkim Himalayan region. The reasons for landslides in Sikkim vary depending on location; however, heavy rainfall is the most prevalent trigger, resulting in a landslide on Sikkim's steep slope. As a result, various morphometric and climatic factors have received special attention in this work. The Upper Tista River basin is a crucial area in the state of Sikkim due to the fact that the state capital is in the centre of the basin, and the majority of Sikkim's national and state highways converge in this particular region. The destruction of homes, roads, and croplands caused by landslides in this region imposes significant stress on the local population. The landslides along the National Highway 10 (NH-10) caused great hardship for the people living there and for the state of Sikkim. The NH-10 is the primary route that links Sikkim to its neighbouring state of West Bengal, and most vehicles that transport goods to the state travel along this route. Hence, the prime objective of this present work is to determine a suitable statistical method to prepare a landslide prediction map of the Upper Tista basin. Policymakers and planners must be aware of the geographic distribution of landslide-prone regions to use natural resources while preventing landslides.

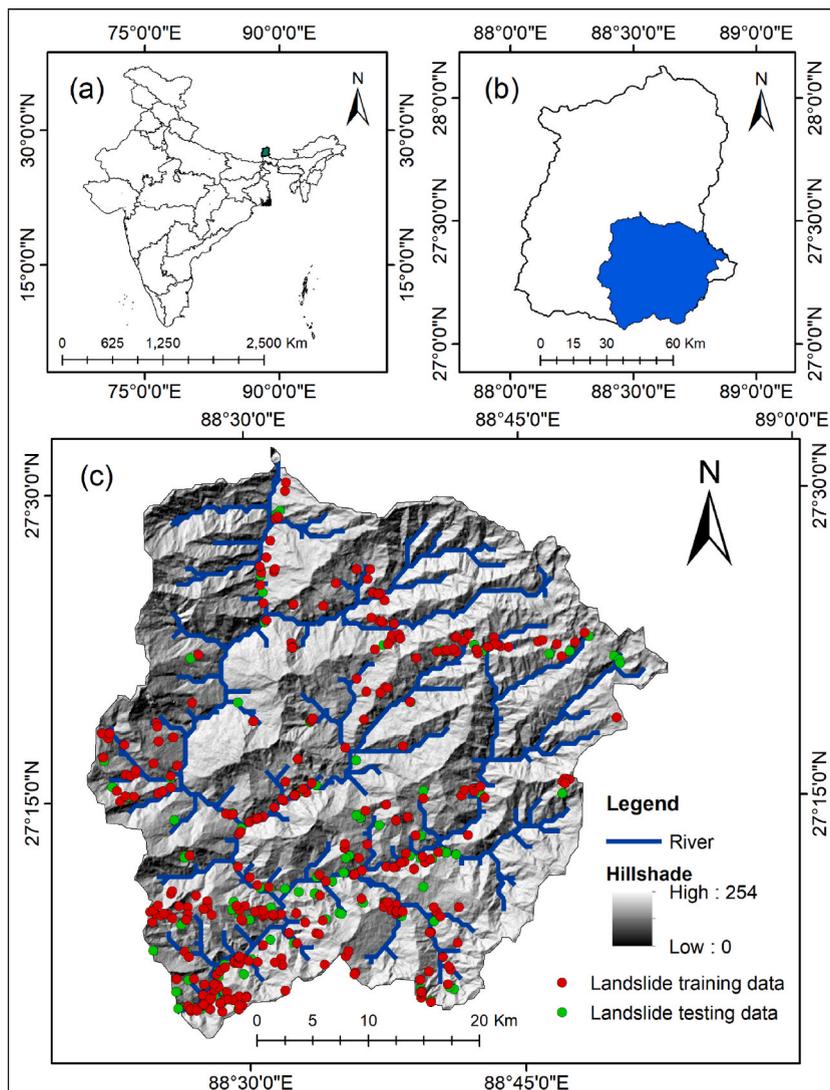


Fig. 1. Location map of the study area (a) India, (b) Sikkim, and (c) Upper Tista Basin with landslide training and testing data points.



Fig. 2. Field photographs of landslides within the study area: (a), (b), (c), (d), (e), (f), (g) Different landslides occurring places along NH-10 in Upper Tista Basin; (h), (i) Protection against landslide.

2. Database and methodology

2.1. Description of the study area

The south-eastern part of Sikkim and the northern part of the Indian state of West Bengal make up the study area, which is located in the Upper Tista Basin. It is physiographically part of the eastern Himalayan mountainous area. Elevations in the studied area generally range from 210 to 5124 m above sea level, with considerable topographic fluctuations. Rivers, gorges, and steep-sided deep valleys cut through the region. About 1759 km² area comprised the study area, and it geographically extended between the extension of approximately 27°03' to 27°32' eastern latitudes and 88°21' to 88°52' northern longitudes (Fig. 1). The climate of the study area is generally humid sub-tropical type, and due to the frequent intense rainfall activities throughout the monsoon season, the region is often perpetually saturated. The area is also tectonically active, and earthquakes have occurred with high magnitudes. Heavy rainfall associated with landslides in the hilly terrain and flooding in the plains is a frequent phenomenon throughout the Tista Basin [47]. Every year during the monsoon, landslide events happen, especially in the upper part of the Tista Basin, and affect the inhabitants' roads, houses, and livelihoods. The field photographs during 2021–2022 manifest the evidence of different types of landslides within the studied basin in Fig. 2.

2.2. Data used for thematic maps preparation

In the present study, different types of data have been utilized from authenticated sources, and therefore thematic maps are displayed by ESRI ArcGIS (version 10.4.1) software. The elevation, aspect, slope, curvature, TWI, SPI, and roughness thematic layers are generated using the USGS's SRTM DEM. From the SRTM DEM, the drainage network has been derived in the ArcGIS platform using different tools. On the other hand, NDVI and LULC maps are simultaneously made using Landsat-8 OLI/TIRS, Path/Row: 139/041. Each of the satellite images underwent radiometric correction before being used in subsequent calculations. The IMD's rainfall data (1986–2020) were gathered, and the long-term annual rainfall map was developed. The MFI was used to create the rainfall intensity map using these rainfall datasets. Using the 'Proximity' feature of 'Analysis Tools' in ArcGIS, Thiessen polygons were initially used to define the influencing region of a rain gauge station. The long-term annual rainfall and MFI were calculated in Microsoft Excel 2019, and themed layers were created in ArcGIS based on the provided values. Lithology layers were created using information obtained from Bhukosh GSI. Lastly, OpenStreetMap's dataset was used to construct the spatial layer of distance to road. The data sources of fourteen conditioning factors employed in the landslide prediction modelling are described in Table 1.

2.3. Landslide inventory map (LIM)

It is believed that using a LIM is necessary to understand the relationship between landslides and the factors that affect them [48,49,50,51]. The sites and patterns of historical landslides are shown on the LIM in Fig. 1c. The spatial pattern and frequency of the active and historical landslides must be well understood to produce a valid LSZ map that depicts the risk of landslides in a given research area [52,53]. LIM is the preliminary stage in LSZ mapping [54]. For the purpose of validating and training LSMs, a precise and comprehensive LIM is crucial [55,56,57]. In the present study, more than 500 historical landslide point data have been collected from the Geological Survey of India (<https://bhukosh.gsi.gov.in/Bhukosh/MapView.aspx>) and NASA Global Landslide Catalog Point (<https://gpm.nasa.gov/landslides/index.html>). The data was then validated and confirmed using Google Earth and a detailed field

Table 1
Source of different applied landslide conditioning factors (LCF) for the landslide prediction modelling.

Parameters	Descriptions	Source	Resolution/ scale	GIS data type
Elevation, aspect, slope, curvature, TWI, SPI, Distance to stream and roughness	Derived from ASTER DEM and prepared the thematic layer using ArcGIS	United States Geological Survey (USGS) Retrieved from: https://earthexplorer.usgs.gov	30 m	Raster
LULC and NDVI	Using Landsat 8 OLI/TIRS, Path/Row: 139/041, all the layers were prepared after mosaicing and atmospheric correction of the image	United States Geological Survey (USGS) Retrieved from: https://earthexplorer.usgs.gov	30 m	Raster
Long-term annual rainfall and MFI	Gridded rainfall for the period 1986–2020 has been used for calculation	India Meteorological Department (IMD) Retrieved from: https://www.imdpune.gov.in	0.25 × 0.25	NetCDF
Lithology	Digital lithological map of the study area has been extracted	Geological Survey of India (GSI) Retrieved from: https://bhukosh.gsi.gov.in/	–	Vector
Distance to road	Digital road networks of the study area have been extracted	OpenStreetMap Retrieved from: www.openstreetmap.org	–	Vector

survey conducted in 2021–2022. Finally, 477 landslide locations on the LIM were used to create the models, which was around 70% for training purposes of the total landslide data and the remaining 30% for validating purposes [28,58]. The methodological flowchart is depicted in Fig. 3.

2.4. Selection and construction of thematic maps of landslide conditioning factors (LCF) in the GIS environment

In this study, the LCF has been selected based on previous literature reviews [59,10,60,61] and expert opinion. The selected parameters are very much relevant to the landslide studies. Here for the construction of thematic maps, ArcGIS software is usually used, and 14 thematic layers of LCF have been prepared. A detailed description of the LCF has been given below.

2.4.1. Elevation

As a consequence of its gravitational potential force, elevation is crucial in LCF [62,63]. The geomorphic features, type of vegetation, and level of degradation are all impacted by elevational variability in any given region (Chen et al., 2017). Landslides are frequently influenced by topography indirectly or through other elements like aspect and slope gradient [64,65]. In the present study area, an elevation map was produced from the SRTM DEM, which varies from 210 to 5124 m (Fig. 4a), showing the region’s high altitudinal variation.

2.4.2. Slope

The slope is one of the most significant LCF employed globally by different scholars [9,66,67] for generating the LSMs. It can be utilized to determine how steep a terrain feature is, and it directly affects the frequency of landslides (Chen et al., 2019). The slope map was generated using the ‘spatial analyst’ tool in ArcGIS. The slope of the study region ranges between 0.06 to 73.57°, which manifests the very high steepness of the region. The slope map is shown in Fig. 4 (b).

2.4.3. Aspect

As a topographical key identifier, the aspect has an impact on the landslides [53,68]. It is known as the slope’s orientation and is connected to the flow’s azimuth and microclimatic characteristics in every given place [53]. Additionally, it influences the slope’s forest cover, humidity levels, and temperature fluctuations [69,67]. For this region, the aspect map was generated through the ‘3D analyst’ tool in ArcGIS, and here nine categories of aspect are found, as displayed in Fig. 4 (c).

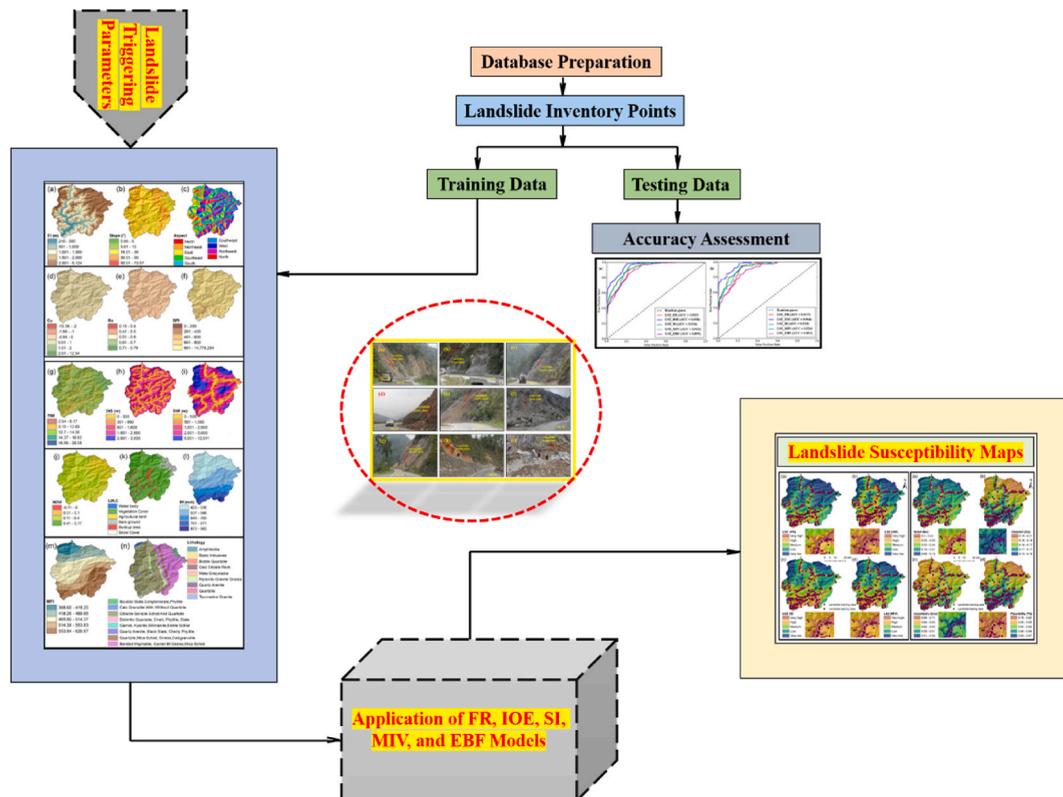


Fig. 3. Methodological flowchart of the study.

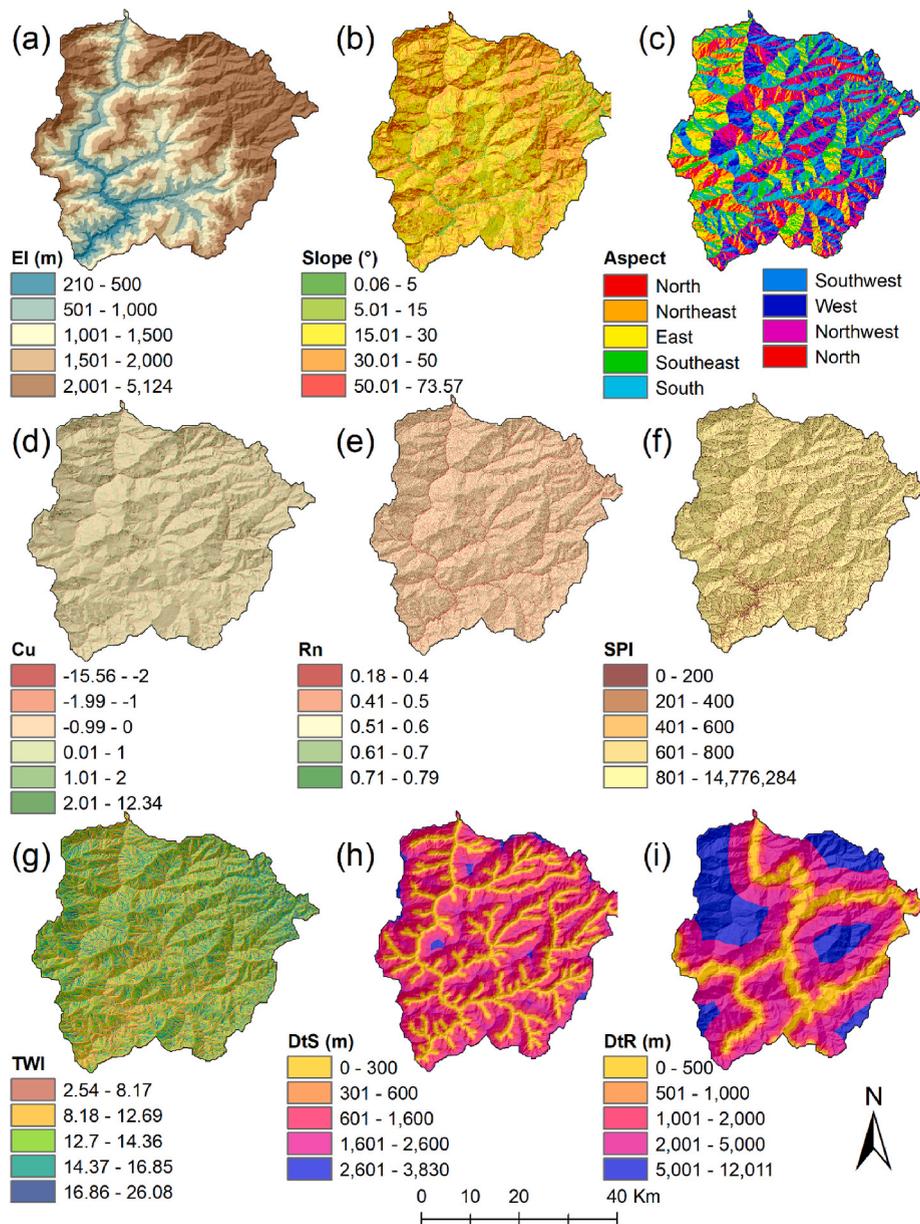


Fig. 4a. Landslide conditioning factors: (a) Elevation (EI), (b) Slope, (c) Aspect, (d) Curvature (Cu), (e) Roughness (Rn), (f) Stream power index (SPI), (g) Topographic wetness index (TWI), (h) Distance to stream (DtS), and (i) Distance to road (DtR).

2.4.4. Curvature

Several studies [70,71,66] assessed curvature as an important factor in LSMs. Concave, flat, and convex are the three elements that are represented by curvature. Positive values are related to convexity, whereas negative values are associated with concavity, and zero is represented by the flat region [72]. Curvature values that are both positive and negative implications that the terrain is highly susceptible to landslides. After a heavy downpour, a concave slope retains the water for a substantial time, causing the soil to get saturated and lose some of its mechanical qualities. In contrast, the process that causes landslides on a convex slope is described by rocks' decomposition and disintegration due to repeated expanding and contracting processes [64]. In the study area, using the 'spatial analyst' tool in ArcGIS curvature map was produced, and it varies from -15.56 to 12.34 (Fig. 4 d).

2.4.5. Roughness

It is another very important LCF used in different studies, viz., [73,74]. Using the "focal statistics" and "raster calculator" features on the ArcGIS software, it is determined from the filled SRTM DEM. This study calculated the LCF roughness based on Eq. (1) (Evans, 1972; [72]).

$$\text{Roughness} = \frac{(\text{FS}_{\text{mean}} - \text{FS}_{\text{min}})}{(\text{FS}_{\text{max}} - \text{FS}_{\text{min}})}, \tag{1}$$

where, FS_{mean} represent focal statistics mean, FS_{min} represent focal statistics minimum and FS_{max} represent the focal statistics maximum. The value of roughness for the Upper Tista Basin ranges from 0.18 to 0.79 (Fig. 4 e).

2.4.6. Stream power index (SPI)

The SPI illustrates the surface water flow’s erosional power, and it is a significant hydraulic factor for modelling LSMs [64,72,75]. SPI also describes how susceptible a terrain is to runoff-induced erosion and its indirect impact on slope stability. The SPI for the present study was computed by employing Eq. (2) (Moore et al., 1991):

$$\text{SPI} = A * \tan \beta, \tag{2}$$

where, A express the specific area, and $\tan \beta$ defines the gradient. The “Raster Calculator tool of “Spatial Analyst Tools” was used to produce the SPI map of the area. The SPI value varies from 0 to 14,776,284 (Fig. 4 f).

2.4.7. Topographic wetness index (TWI)

An important hydrological aspect in landslide research is the TWI, which depicts the saturation supply zone resulting from ground runoff with the effect of topographic factors [64]. The TWI map was made for the current study to locate possibly moist locations and classify those with high sensitivity to landslides because the presence of water leads soil to loosen and structurally degrade [65]. Eq. (3) was used to calculate the TWI (Beven and Kirkby, 1979 [72]):

$$\text{TWI} = \text{Ln}\left(\frac{a}{\tan B}\right), \tag{3}$$

where, a and B express specific catchment areas and slopes, respectively. The a was computed by Eq. (4).

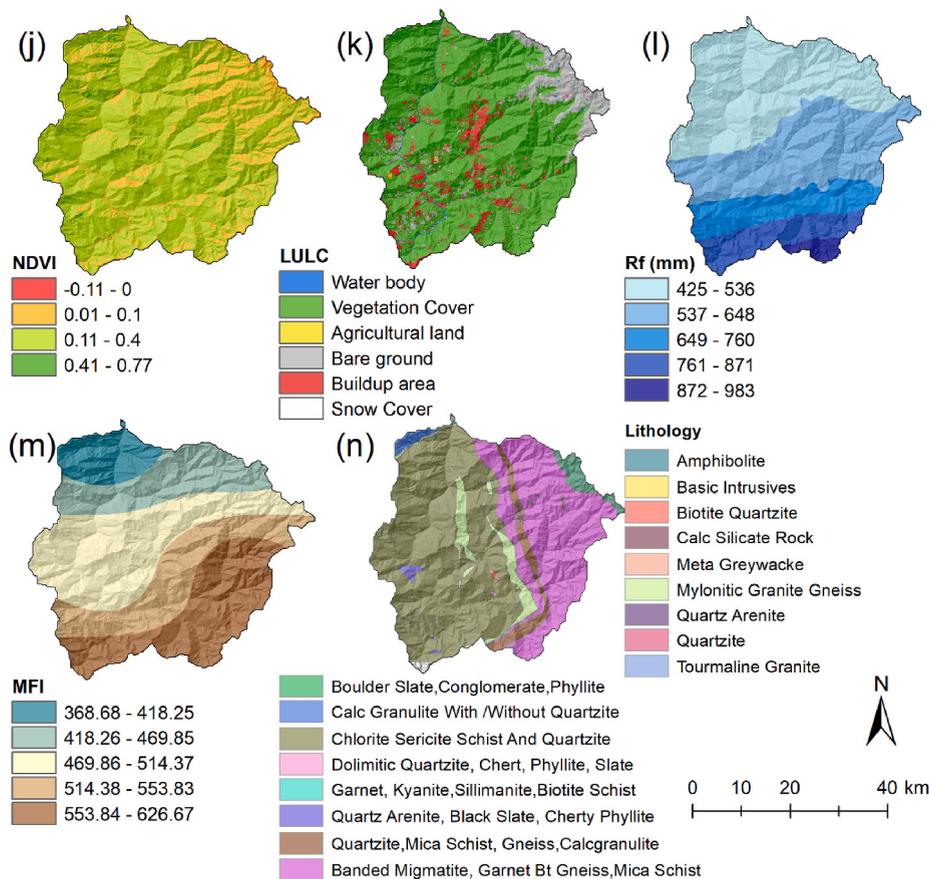


Fig. 4b. Landslide conditioning factors: (j) Normalized difference vegetation index (NDVI), (k) Land use land cover (LULC), (l) Annual rainfall (Rf), (m) Modified fournier index (MFI), and (n) Lithology.

$$a = \frac{A}{L}, \quad (4)$$

where, A represents the basin area, and L represents the length of the contour [75]. The TWI value ranges from 2.54 to 26.08 in the Upper Tista Basin (Fig. 4 g).

2.4.8. Distance to stream (DtS)

Throughout drainage lines, there are probable slope instability zones [76,77,78]. One of the important causes of landslides is the deterioration of soil's shear strength due to the high pore pressure in water [79,80,81]. The Upper Tista Basin is covered by an intensive drainage network. The DtS map manifests in Fig. 4 (h), where the value ranges from 0 to 3830 m.

2.4.9. Distance to road (DtR)

As anthropogenic elements, roads built in hilly terrain have a significant influence on the stabilization of slopes. During their development, the road network often sees substantial excavating, vegetation clearance, and the deployment of external pressures [64, 82]. They significantly influence decreasing the slope's resistance [65]. The value of DtR map of the Upper Tista Basin varied from 0 to 12011 m (Fig. 4 i).

2.4.10. Normalized difference vegetation index (NDVI)

In order to stabilize slopes, vegetation is crucial because roots help reinforce and consolidate soil layers. The slopes are less resistant to landslides the lighter the vegetation. NDVI is typically applied to distinguish between areas covered in vegetation and those that are not [72]. Numerous studies [83,84,85] demonstrate the connection between landslide occurrence and NDVI. It has a range of -1 to +1, and in the existing research region, it is between -0.11 and 0.48 (Fig. 4 j), using Eq. (5):

$$NDVI = (NIR - RED) / (NIR + RED), \quad (5)$$

where, NIR expresses the near-infrared band, and RED expresses the red band. The NDVI was estimated depending on Eq. (5) utilizing the "Raster Calculator" function of the "Spatial Analyst Tools" in ArcGIS.

2.4.11. Land use land cover (LULC)

Various land uses affect landslides differently, making them an important consideration for determining LSMs in several research [86,87]. Since plant roots are thought to protect the soil, landslides are less likely to happen in forested areas. They reduce the quantity of water in the soil, and evapotranspiration controls the slope's wetness. Human-induced changes in land use significantly impact the frequency of landslides [64]. The ERDAS imagine has implemented the supervised classification technique to create the LULC map for 2022 of the Upper Tista Basin with 85% accuracy [88]. Six major land use type is found in the region: water body, vegetation cover, agricultural land, bare ground, buildup area, and snow cover (Fig. 4 k).

2.4.12. Long-term annual rainfall

One of the main criteria for identifying landslide-prone regions is rainfall [89,90]. Landslides and rainfall are found to be positively correlated [91,92]. As a significant landslide conditioning component, an annual rainfall map was constructed. Fig. 4 (l) depicts the spatial pattern of the region's yearly long-term rainfall (1986–2020). Rainfall varies from 425 to 983 mm, and generally, in this region, rainfall tendency is high from north to south direction.

2.4.13. Modified fournier index (MFI)

The MFI was used to highlight the variability of rainfall intensity of the basin. Greater MFI values are associated with the areas more prone to landslides, which was regarded as a crucial landslide conditioning indicator [93,94]. MFI has been calculated based on Eq. (6) (Souissi et al., 2020 [72]);

$$MFI = \sum_{i=1}^{12} \frac{P_i^2}{P}, \quad (6)$$

here, P_i represents mean monthly precipitation, and P represents mean annual precipitation. In the Upper Tista Basin MFI value ranges from 368.68 to 626.67, and a higher MFI value is observed in the southern, south-western, south-eastern, and eastern portions of the basin (Fig. 4 m).

2.4.14. Lithology

The lithological characteristics of such a location are relevant to modelling landslide hazards [95,96]. The lithological structures affect the permeability and rock hardness of the earth's surface [53,97]. Since every lithological composition differs from the others in terms of its characteristics and organization, every lithology has a varied level of susceptibility to landslides [64]. Different lithological formations have been observed in the Upper Tista Basin, as depicted in Fig. 4 (n). The maximum area of the basin is covered with quartzite, banded migmatite, chlorite sericite schist, garnet Bt gneiss, calcgranulite and mica schist.

2.5. Multicollinearity assessment of the LCF

Two of the most common indicators used to assess the multicollinearity among the parameters utilized as input in a mathematical model are Variance Inflation (VIF) and Tolerance (TL). The VIF and TL are found in various landslide susceptibility studies, viz., [98, 99,100]. To detect collinearity throughout multiple regression, the VIF and TL constitute two highly connected statistical measures. These were dependent on the R-squared score of regressing a single predictor against every other predictor in the study [101]. The VIF and TL both are computed (Table 2) based on the following equations (Eq. 7 and 8):

$$TL = 1 - R_i^2 \tag{7}$$

where, R_i^2 depicts the coefficient of determination of the regression equation.

$$VIF = 1/T_i \tag{8}$$

2.6. Methods

2.6.1. FR model

A popular and reliable statistical technique that is widely employed is frequency ratio [102,103,104]. LSZ map was generated for the current investigation employing multi-class spatial data sources using this probabilistic approach. FR is described as the proportion of the research area to the area where the landslide happened. In order to describe the relationship or link between historical landslides and landslide-triggering variables, the FR framework is used. This characteristic still makes it a popular choice for LSMs investigations [28,54,105]. This study separated the influencing variables into subcategories, and the frequency ratios were subsequently calculated by counting the number of pixels containing landslides in each component’s subcategory (Table 3). The frequency ratio of the variable subcategories was calculated using Eq. (9) shown follows [28]:

$$Fr = \frac{Npix (L_{ij})/Npix (L)}{Npix (S_{ij})/Npix (S_a)} \tag{9}$$

where, $Npix (L_{ij})$ denotes the landslide pixel’s number in j th subclass of parameter i , $Npix (L)$ denotes the sum of landslide pixels in the study area, $Npix (S_{ij})$ denotes the pixel’s number in j th subclass of parameter i , and $Npix (S_a)$ denotes the sum of pixels in the study area. The FR value > 1 exhibits a substantial connection with landslides, the FR value 1 exhibits average results, and the FR value < 1 exhibits low connections with landslides. In order to calculate the LSZ, the aggregate of the FR for every subclass was computed, as shown in Eq. (10).

$$LSI = \sum_{i=1}^n Fr_i = Fr_1 + Fr_2 + \dots + Fr_n \tag{10}$$

Furthermore, the susceptibility map is constructed by spatially combining these data in GIS.

2.6.2. IOE model

The IOE model is focused on the assumption of bivariate evaluation, where the number of landslides inside a specific variable is calculated [106,107,108]. This method enables the weights for every input factor to be calculated [109,110,111]. The following procedures are employed to determine the weight value for the variable in its entirety or the information coefficient [112]:

Table 2
Collinearity statistics of landslide conditioning factors (LCF).

Sl. No	Parameters	Collinearity Statistics	
		Tolerance	Variance Inflation Factor (VIF)
1	Elevation	0.442	2.265
2	Slope	0.879	1.138
3	Aspect	0.754	1.327
4	Curvature	0.692	1.446
5	Roughness	0.841	1.189
6	SPI	0.831	1.203
7	TWI	0.716	1.396
8	DtS	0.904	1.106
9	DtR	0.825	1.211
10	NDVI	0.729	1.372
11	LULC	0.807	1.238
12	Rainfall	0.269	3.715
13	MFI	0.271	3.689
14	Lithology	0.835	1.198

Table 3
Spatial association of each conditioning factor and landslide by FR and IOE co-efficient.

Factors	Class	Reclass	Landslide pixels	Class pixels	FR	Pij	FFR	Hj	Hmax	Ij	Pj	Wj
Elevation	210–500	5	54,000	62,698	5.577	0.587	9.500	1.735	2.322	0.253	1.900	0.480
	501–1000	4	74,700	302,396	1.600	0.168						
	1001–1500	3	79,200	461,747	1.111	0.117						
	1501–2000	2	54,900	409,478	0.868	0.091						
	2001–5124	1	37,800	710,070	0.345	0.036						
Slope	0.06–5	1	0	17,965	0.000	0.000	2.338	1.218	2.322	0.475	0.468	0.222
	5.01–15	2	2700	164,985	0.106	0.045						
	15.01–30	3	162,900	1,004,685	1.050	0.449						
	30.01–50	4	135,000	739,577	1.182	0.506						
	50.01–73.57	1	0	19,177	0.000	0.000						
Aspect	North (0–22.5)	4	13,500	100,567	0.869	0.099	8.751	3.148	3.170	0.007	0.972	0.007
	Northeast (22.5–67.5)	1	22,500	191,004	0.763	0.087						
	East (67.5–112.5)	5	33,300	227,176	0.949	0.108						
	Southeast (112.5–157.5)	8	52,200	261,030	1.295	0.148						
	South (157.5–202.5)	6	45,000	257,421	1.132	0.129						
	Southwest (202.5–247.5)	7	44,100	246,452	1.159	0.132						
	West (247.5–292.5)	3	33,300	253,419	0.851	0.097						
	Northwest (292.5–337.5)	5	41,400	283,657	0.945	0.108						
	North (337.5–360)	2	15,300	125,664	0.788	0.090						
Curvature	–15.56––2	4	1800	11,595	1.005	0.184	5.471	2.523	2.585	0.024	0.912	0.022
	–1.99––1	6	18,900	94,556	1.294	0.237						
	–0.99 - 0	5	147,600	873,481	1.094	0.200						
	0.01–1	3	122,400	856,595	0.925	0.169						
	1.01–2	2	9000	99,911	0.583	0.107						
Roughness	2.01–12.34	1	900	10,250	0.569	0.104	3.390	1.963	2.322	0.155	0.678	0.105
	0.18–0.40	3	6300	59,502	0.682	0.201						
	0.41–0.50	5	164,700	919,742	1.154	0.340						
	0.51–0.60	4	125,100	913,355	0.883	0.260						
	0.61–0.70	2	4500	43,265	0.670	0.198						
SPI	0.71–0.79	1	0	1655	0.000	0.000	8.941	2.155	2.322	0.072	1.788	0.129
	0–200	1	40,500	290,614	0.902	0.101						
	201–400	5	14,400	29,938	3.114	0.348						
	401–600	3	11,700	46,870	1.616	0.181						
	601–800	4	18,900	51,064	2.397	0.268						
TWI	>800	2	215,100	1,527,904	0.912	0.102	4.389	2.179	2.322	0.061	0.878	0.054
	2.54–8.17	4	36,900	282,217	0.847	0.193						
	8.18–12.69	5	128,700	519,197	1.605	0.366						
	12.7–14.36	3	90,000	693,792	0.840	0.191						
	14.37–16.35	2	39,600	360,711	0.711	0.162						
DtS	16.86–26.08	1	5400	90,472	0.386	0.088	4.328	1.917	2.322	0.175	0.866	0.151
	0–300	5	97,200	374,555	1.682	0.389						
	301–600	4	57,600	332,168	1.124	0.260						
	601–1600	3	108,000	880,634	0.795	0.184						
	1601–2600	2	37,800	336,335	0.728	0.168						
DtR	2601–3830	1	0	24,079	0.000	0.000	5.923	1.832	2.322	0.211	1.185	0.250
	0–500	5	140,400	275,489	3.302	0.558						
	501–1000	4	27,900	206,396	0.876	0.148						
	1001–2000	3	43,200	351,246	0.797	0.135						
	2001–5000	2	70,200	690,694	0.659	0.111						

(continued on next page)

Table 3 (continued)

Factors	Class	Reclass	Landslide pixels	Class pixels	FR	Pij	FFR	Hj	Hmax	Ij	Pj	Wj
NDVI	>5000	1	18,900	423,946	0.289	0.049						
	-0.11 - 0	1	0	3920	0.000	0.000	2.034	1.000	2.000	0.500	0.508	0.254
	0.01-0.10	2	50,400	321,567	1.016	0.499						
	0.11-0.40	3	250,200	1,591,933	1.018	0.501						
LULC	0.41-0.77	1	0	30,348	0.000	0.000						
	Water body	6	2700	11,023	1.587	0.331	4.796	2.143	3.000	0.286	0.599	0.171
	Vegetation cover	4	227,700	1,571,000	0.939	0.196						
	Agricultural land	1	0	2050	0.000	0.000						
	Bare ground	2	900	18,955	0.308	0.064						
	Buildup area	5	68,400	335,390	1.321	0.276						
	Snow cover	3	900	9098	0.641	0.134						
Rainfall	425-536	1	64,800	791,940	0.530	0.070	7.544	2.107	2.322	0.093	1.509	0.140
	537-648	2	81,000	655,233	0.801	0.106						
	649-760	3	56,700	259,325	1.417	0.188						
	761-871	5	85,500	201,337	2.752	0.365						
	872-983	4	12,600	39,935	2.044	0.271						
MFI	368.68-418.25	1	11,700	174,688	0.434	0.096	4.516	2.170	2.322	0.065	0.903	0.059
	418.26-469.85	2	24,300	331,036	0.476	0.105						
	469.86-514.37	3	90,000	599,968	0.972	0.215						
	514.38-553.83	4	63,000	366,967	1.112	0.246						
	553.84-626.67	5	111,600	475,090	1.522	0.337						
Lithology	Amphibolite	1	0	23	0.000	0.000	7.362	2.340	4.087	0.428	0.433	0.185
	Banded Migmatite, Garnet Bt Gneiss, Mica Schist	2	72,000	622,835	0.749	0.102						
	Basic Intrusives	1	0	603	0.000	0.000						
	Biotite Quartzite	6	900	2432	2.398	0.326						
	Boulder Slate, Conglomerate, Phyllite	1	0	665	0.000	0.000						
	Calc Granulite With/Without Quartzite	1	0	20,703	0.000	0.000						
	Calc Silicate Rock	1	0	5101	0.000	0.000						
	Chlorite Sericite Schist And Quartzite	4	207,000	1,049,576	1.278	0.174						
	Dolimitic Quartzite, Chert, Phyllite, Slate	1	0	868	0.000	0.000						
	Garnet, Kyanite, Sillimanite, Biotite Schist	1	0	494	0.000	0.000						
	Meta Greywacke	1	0	1717	0.000	0.000						
	Mylonitic Granite Gneiss	3	13,500	106,187	0.824	0.112						
	Quartz Arenite	1	0	255	0.000	0.000						
	Quartz Arenite, Black Slate, Cherty Phyllite	5	3600	12,850	1.815	0.247						
	Quartzite	2	3600	78,222	0.298	0.041						
	Quartzite, Mica Schist, Gneiss, Calcgranulite	1	0	243	0.000	0.000						
	Tourmaline Granite	1	0	44,997	0.000	0.000						

$$P_{ij} = \frac{A_{sd}}{A_i} \tag{11}$$

$$(P_{ij}) = \frac{P_{ij}}{\sum_{j=1}^{S_j} P_{ij}} \tag{12}$$

The entropy values (H_j and H_{jmax}) are expressed as follows (Eqs. (13) and (14)):

$$H_j = \sum_{i=1}^{S_j} (P_{ij}) \log_2 (P_{ij}), j = 1, \dots, n \tag{13}$$

$$H_{jmax} = \log_2 S_j \tag{14}$$

where, S_j represents the number of classes.

$$I_j = \frac{H_{jmax} - H_j}{H_{jmax}} \tag{15}$$

$$W_j = I_j \times P_j \tag{16}$$

In the IOE approach, I_j represent the information coefficient, W_j is the weight value for the variable in its entirety and P_j represent the probability of slope failure, as described in Eqs. (15) and (16). The total of all parameter classes is utilized to generate the ultimate susceptibility value based on the estimated landslide density for each parameter class. The following equation (Eq. 17) was utilized to create the final LSZ applying the IOE model:

$$Y = \sum_{i=1}^n \frac{z}{mi} \times C \times W_j, \tag{17}$$

where, Y express the summation of all classes, i express the number of particular parametric map, z express the no. of classes inside the parametric map with the highest number of classes, mi express the number of classes inside the particular parametric map, C express the value of the class after secondary classification, and W_j express the weight of a factor.

2.6.3. SI model

The SI approach, put forth by Ref. [113]; is acknowledged as a bivariate statistical investigation. Bivariate statistical techniques are frequently employed to create correlations between landslide-controlling variables and the occurrence of landslides [114,115]. This method's quick and manageable process considers it appropriate for natural hazard prediction [116,117]. This framework specifies the weighting value for every category of conditioning parameter as the natural logarithm of the class's average landslide density divided by the average landslide density throughout the whole research area. This approach is centered on how often landslides occur for every variable class. The SI weights for every parameter are determined using the below method:

$$SI = \ln \left(\frac{\text{Densclass}}{\text{Densmap}} \right) = \ln \left[\frac{\left(\frac{N(L_i)}{A_i} \right)}{\left(\frac{\sum N(L_i)}{\sum A_i} \right)} \right], \tag{18}$$

where, SI represents the weight of i -th class, Densclass represents landslide density inside the factor class, Densmap represents landslide density inside the overall map, $N(L_i)$ represents landslide's number of the i -th factor class, A_i represents the i -th factor class area, $\sum N(L_i)$ represents pixels number of landslides of the research region, and $\sum A_i$ represents the area of the research region. Every LCF's SI weights were computed, and then, using the 'spatial analyst tool' in ArcGIS, each variable was classed using the derived SI values. Then 'raster calculator' was used to combine the reclassified variables to get the landslide probability index.

2.6.4. MIV model

A statistical modelling approach based on the principles of the information theory is the Information Value Model (IVM) [28,118]. According to Ref. [119]; there are certain issues with the IVM [119]. developed the modified information value model (MIV). Presently, the MIV is globally relevant in the LSM mapping [28,120,121]. With the following equation (Eq. 19) [119], described the MIV in order to prevent the issues of the IVM method and quantitatively represent the results:

$$I(H, xi) = \log_2 \left(\frac{N_{pix} (Si)/N_{pix} (Ni)}{\sum N_{pix} (Si)/\sum N_{pix} (Ni)} + 1 \right), \tag{19}$$

where, $I(H, xi)$ indicates the information value of subcategory i of each parameter, $N_{pix} (Si)$ represents the landslide pixel's number in subcategory i , $N_{pix} (Ni)$ represents the total pixels of subcategory I , $\sum N_{pix} (Si)$ represents the landslide pixel's number in the study

Table 4
Spatial relationship between each conditioning factor and landslide by SI, MIV and EBF models.

Factor	Class	Landslide pixels	Class pixels	SI	MIV	Bel	Dis	Unc	Pls
Elevation	210–500	54,000	62,698	1.719	2.717	0.893	0.161	0.252	0.839
	501–1000	74,700	302,396	0.470	1.378	0.047	0.167	0.664	0.833
	1001–1500	79,200	461,747	0.105	1.078	0.030	0.182	0.701	0.818
	1501–2000	54,900	409,478	−0.141	0.902	0.022	0.202	0.707	0.798
	2001–5124	37,800	710,070	−1.065	0.427	0.008	0.287	0.676	0.713
Slope	0.06–5.0	0	17,965	0.000	0.000	0.000	0.216	0.784	0.784
	5.01–15	2700	164,985	−2.245	0.145	0.038	0.235	0.720	0.765
	15.01–30	162,900	1,004,685	0.049	1.036	0.446	0.167	0.384	0.833
	30.01–50	135,000	739,577	0.167	1.126	0.515	0.165	0.329	0.835
	50.01–73.57	0	19,177	0.000	0.000	0.000	0.217	0.783	0.783
Aspect	North (0–22.5)	13,500	100,567	−0.140	0.902	0.097	0.114	0.787	0.886
	Northeast (22.5–67.5)	22,500	191,004	−0.271	0.818	0.083	0.115	0.797	0.885
	East (67.5–112.5)	33,300	227,176	−0.052	0.963	0.107	0.112	0.780	0.888
	Southeast (112.5–157.5)	52,200	261,030	0.258	1.198	0.156	0.104	0.748	0.896
	South (157.5–202.5)	45,000	257,421	0.124	1.092	0.132	0.108	0.763	0.892
	Southwest (202.5–247.5)	44,100	246,452	0.147	1.110	0.136	0.107	0.760	0.893
	West (247.5–292.5)	33,300	253,419	−0.162	0.888	0.094	0.114	0.789	0.886
	Northwest (292.5–337.5)	41,400	283,657	−0.057	0.960	0.107	0.112	0.780	0.888
	North (337.5–360)	15,300	125,664	−0.238	0.839	0.087	0.115	0.795	0.885
Curvature	−15.56–−2	1800	11,595	0.005	1.004	0.184	0.175	0.642	0.825
	−1.99–−1	18,900	94,556	0.258	1.198	0.250	0.170	0.594	0.830
	−0.99–0	147,600	873,481	0.090	1.066	0.204	0.137	0.663	0.863
	0.01–1	122,400	856,595	−0.078	0.945	0.167	0.165	0.666	0.835
	1.01–2	9000	99,911	−0.539	0.663	0.099	0.178	0.715	0.822
Roughness	2.01–12.34	900	10,250	−0.565	0.649	0.096	0.175	0.721	0.825
	0.18–0.40	6300	59,502	−0.382	0.751	0.194	0.214	0.584	0.786
	0.41–0.50	164,700	919,742	0.143	1.107	0.357	0.150	0.509	0.850
	0.51–0.60	125,100	913,355	−0.125	0.913	0.260	0.209	0.531	0.791
	0.61–0.70	4500	43,265	−0.400	0.740	0.190	0.214	0.588	0.786
SPI	0.71–0.79	0	1655	0.000	0.000	0.000	0.213	0.787	0.787
	0–200	40,500	290,614	−0.103	0.928	0.075	0.210	0.689	0.790
	201–400	14,400	29,938	1.136	2.041	0.427	0.202	0.450	0.798
	401–600	11,700	46,870	0.480	1.388	0.153	0.207	0.613	0.793
	601–800	18,900	51,064	0.874	1.764	0.270	0.200	0.532	0.800
TWI	>800	215,100	1,527,904	−0.093	0.935	0.075	0.181	0.717	0.819
	2.54–8.17	36,900	282,217	−0.167	0.885	0.184	0.210	0.597	0.790
	8.18–12.69	128,700	519,197	0.473	1.381	0.404	0.142	0.492	0.858
	12.7–14.36	90,000	693,792	−0.174	0.880	0.183	0.213	0.596	0.787
	14.37–16.35	39,600	360,711	−0.341	0.775	0.151	0.219	0.619	0.781
DtS	16.86–26.08	5400	90,472	−0.951	0.471	0.078	0.216	0.696	0.784
	0–300	97,200	374,555	0.520	1.423	0.424	0.158	0.454	0.842
	301–600	57,600	332,168	0.117	1.087	0.254	0.193	0.547	0.807
	601–1600	108,000	880,634	−0.230	0.844	0.169	0.223	0.593	0.777
	1601–2600	37,800	336,335	−0.317	0.789	0.153	0.216	0.616	0.784
DtR	2601–3830	0	24,079	0.000	0.000	0.000	0.211	0.789	0.789
	0–500	140,400	275,489	1.195	2.105	0.695	0.108	0.335	0.892
	501–1000	27,900	206,396	−0.133	0.908	0.105	0.203	0.649	0.797
	1001–2000	43,200	351,246	−0.227	0.846	0.094	0.207	0.658	0.793
	2001–5000	70,200	690,694	−0.418	0.730	0.076	0.234	0.655	0.766
NDVI	>5000	18,900	423,946	−1.242	0.366	0.031	0.249	0.703	0.751
	−0.11–0	0	3920	0.000	0.000	0.000	0.288	0.712	0.712
	0.01–0.10	50,400	321,567	0.015	1.011	0.499	0.276	0.224	0.724
	0.11–0.40	250,200	1,591,933	0.018	1.013	0.501	0.143	0.356	0.857
LULC	0.41–0.77	0	30,348	0.000	0.000	0.000	0.293	0.707	0.707
	Water body	2700	11,023	0.462	1.371	0.357	0.130	0.708	0.870
	Vegetation cover	227,700	1,571,000	−0.063	0.955	0.186	0.098	0.806	0.902
	Agricultural land	0	2050	0.000	0.000	0.000	0.131	0.869	0.869
	Bare ground	900	18,955	−1.179	0.387	0.055	0.132	0.837	0.868
	Built up area	68,400	335,390	0.300	1.215	0.282	0.126	0.736	0.874
Rainfall	Snow cover	900	9098	1.467	0.714	0.121	0.130	0.428	0.870
	425–536	64,800	791,940	−0.635	0.614	0.052	0.266	0.664	0.734
	537–648	81,000	655,233	−0.222	0.849	0.083	0.211	0.683	0.789
	649–760	56,700	259,325	0.348	1.273	0.164	0.180	0.632	0.820
	761–871	85,500	201,337	1.012	1.908	0.432	0.148	0.488	0.852
MFI	872–983	12,600	39,935	0.715	1.606	0.270	0.196	0.533	0.804
	368.68–418.25	11,700	174,688	−0.835	0.520	0.085	0.222	0.682	0.778
	418.26–469.85	24,300	331,036	−0.743	0.561	0.094	0.232	0.662	0.768
	469.86–514.37	90,000	599,968	−0.028	0.980	0.210	0.197	0.588	0.803

(continued on next page)

Table 4 (continued)

Factor	Class	Landslide pixels	Class pixels	SI	MIV	Bel	Dis	Unc	Pls
Lithology	514.38–553.83	63,000	366,967	0.107	1.079	0.246	0.194	0.560	0.806
	553.84–626.67	111,600	475,090	0.420	1.335	0.365	0.155	0.508	0.845
	Amphibolite	0	23	0.000	0.000	0.000	0.057	0.943	0.943
	Banded Migmatite, Garnet Bt Gneiss, Mica Schist	72,000	622,835	−0.289	0.807	0.084	0.061	0.888	0.939
	Basic Intrusives	0	603	0.000	0.000	0.000	0.057	0.943	0.943
	Biotite Quartzite	900	2432	0.875	1.765	0.380	0.057	0.780	0.943
	Boulder Slate, Conglomerate, Phyllite	0	665	0.000	0.000	0.000	0.057	0.943	0.943
	Calc Granulite With/Without Quartzite	0	20,703	0.000	0.000	0.000	0.057	0.943	0.943
	Calc Silicate Rock	0	5101	0.000	0.000	0.000	0.057	0.943	0.943
	Chlorite Sericite Schist And Quartzite	207,000	1,049,576	0.245	1.188	0.159	0.029	0.884	0.971
	Dolimitic Quartzite, Chert, Phyllite, Slate	0	868	0.000	0.000	0.000	0.057	0.943	0.943
	Garnet, Kyanite, Sillimanite, Biotite Schist	0	494	0.000	0.000	0.000	0.057	0.943	0.943
	Meta Greywacke	0	1717	0.000	0.000	0.000	0.057	0.943	0.943
	Mylonitic Granite Gneiss	13,500	106,187	−0.194	0.867	0.094	0.057	0.887	0.943
	Quartz Arenite	0	255	0.000	0.000	0.000	0.057	0.943	0.943
	Quartz Arenite, Black Slate, Cherty Phyllite	3600	12,850	0.370	1.493	0.252	0.056	0.845	0.944
	Quartzite	3600	78,222	2.036	0.377	0.031	0.056	0.421	0.944
	Quartzite, Mica Schist, Gneiss, Calcgranulite	0	243	−1.210	0.000	0.000	0.058	0.921	0.942
	Tourmaline Granite	0	44,997	0.000	0.000	0.000	0.057	0.943	0.943

area, and $\sum N_{pix} (N_i)$ represents the total pixels in the study area. With the help of the training landslide data, the MIV of each subcategory of 14 parameters was calculated (Table 4). The LSZ for each pixel is calculated from each conditioning parameter using Eq. (20).

$$I_{total} = LSI = \sum_{i=1}^n I(H, x_i) \tag{20}$$

2.6.5. EBF model

The bivariate and mathematical Dempster-Shafer approach is famous as the theory of EBF [58]. It is a common method widely utilized to predict the LSM in the studies of [122,123,124,125]. Generally, four fundamental evidential belief functions are employed. They are: Bel (degree of belief) and Pls (degree of plausibility), which specify the probability’s bottom and top boundaries, Dis (degree of disbelief), which express that a claim is false despite evidence to the contrary; and Unc (degree of uncertainty), which stands for the distinction between belief and plausibility [58]. The following steps have been used in the current study for the EBF model:

$$Bel_{C_{ij}} = \frac{W_{C_{ij}} (landslide)}{\sum_{j=1}^n W_{C_{ij}} (landslide)} \tag{21}$$

$$W_{C_{ij}} (landslide) = \frac{N(L \cap C_{ij})/N(L)}{[N(C_{ij}) - N(L \cap C_{ij})]/[N(C) - N(L)]}, \tag{22}$$

where, the denominator represents the proportion of factor class (Cij) pixels that are not from landslides, and the numerator represents the proportion of pixels that experience landslides in factor class (Cij).

$$Dis_{C_{ij}} = \frac{W_{C_{ij}} (non-landslide)}{\sum_{j=1}^n W_{C_{ij}} (non-landslide)} \tag{23}$$

$$W_{C_{ij}} (non-landslide) = \frac{[N(C_{ij}) - N(L \cap C_{ij})]/N(L)}{[N(C) - N(L) - N(C_{ij}) + N(L \cap C_{ij})]/[N(C) - N(L)]}, \tag{24}$$

where, the denominator represents the proportion of pixels in other attributes outside of the factor class (Cij) that are not from landslides, and the numerator represents the proportion of pixels from landslides that don’t occur in a factor class (Cij).

$$Unc_{C_{ij}} = 1 - Bel_{C_{ij}} - Dis_{C_{ij}} \tag{25}$$

$$Pls_{C_{ij}} = Bel_{C_{ij}} + Unc_{C_{ij}} \tag{26}$$

2.7. Validation of the models

After creating a landslide prediction map using several models, one of the essential steps is verifying the results’ accuracy [126,127, 128]. LSMs generated by five algorithms (IOE, SI, MIV, FR, and EBF) in the present investigation were validated by comparing the

susceptibility map with both the training and validating datasets [58,129]. To accomplish this, 477 landslide points were randomly divided into two categories, viz., 143 (30%) landslide points were used as validating data, while 334 (70%) landslide points were chosen as the training data. The accuracy of predicting LSMs created by the IOE, SI, MIV, FR, and EBF methods was validated using the receiver operating characteristic-area under curve (ROC-AUC) approach. ROC is a graph (two-dimensional) that represents 1-specificity (false positive rate) in the x-axis and sensitivity (true positive rate) in the y-axis as described below [72,130,75,53]:

$$x \text{ axis} = 1 - \text{specificity} = 1 - \left[\frac{TN}{(TN + FP)} \right] \tag{27}$$

$$y \text{ axis} = \text{sensitivity} = \left[\frac{TP}{(TP + FN)} \right], \tag{28}$$

where TN, FP, TP, and FN express true negative, false positive, true positive, and false negative, respectively (Swets, 1988). AUC was used to quantitatively assess the proposed maps' effectiveness based on five approaches in the study.

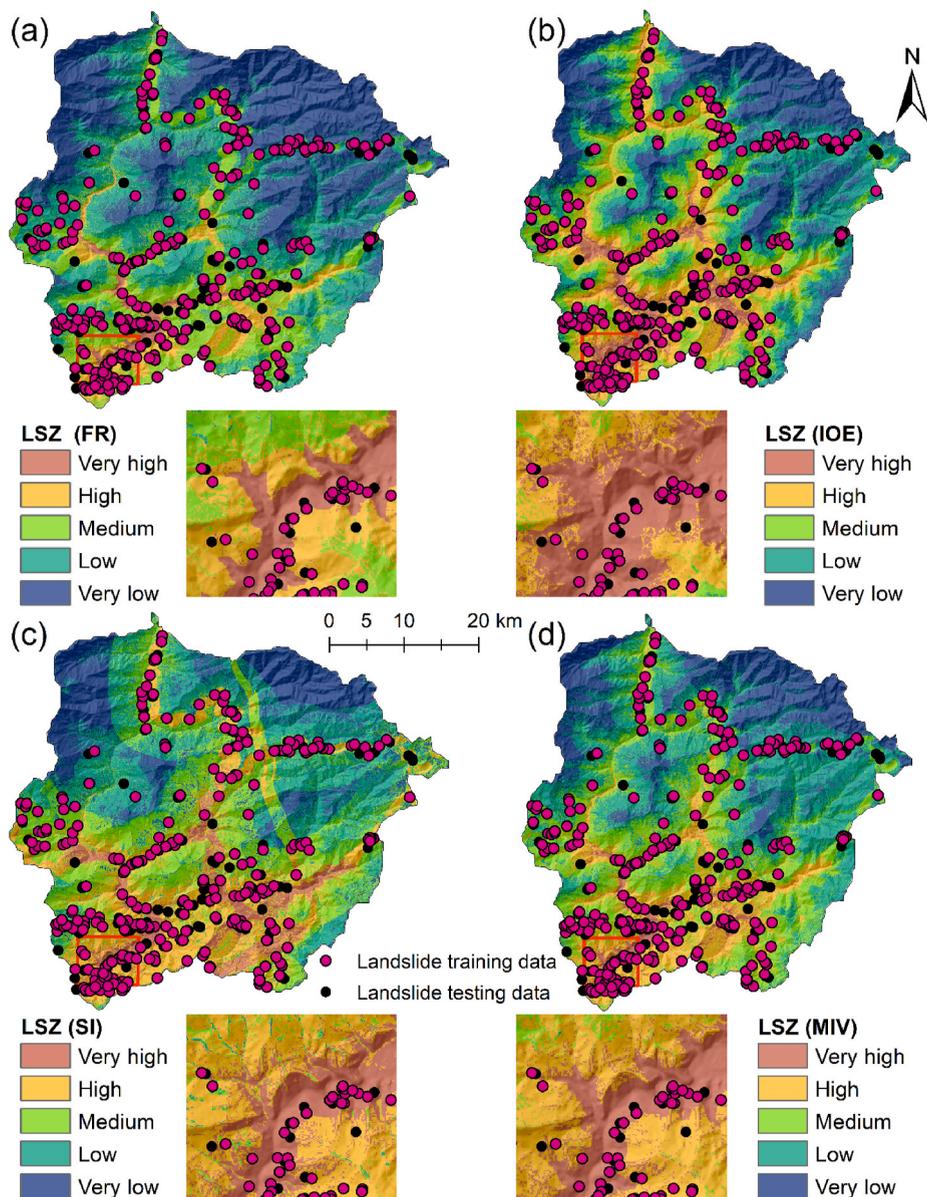


Fig. 5. Landslide Susceptibility maps using (a) FR model; (b) IOE model; (c) SI model; and (d) MIV model.

3. Results

3.1. Multicollinearity analysis

The findings of the multicollinearity test are shown in Table 2. The result showed no collinearity problem among the fourteen causative factors used in this study, as indicated by a tolerance value above 0.1 and a VIF value below 10 for each parameter at $p < 0.05$. As a result, all factors were chosen for the final modelling. It also verifies that no multicollinearity-related uncertainty penetrated the model output.

3.2. Application of FR model for LSZ mapping

The FR model is a straightforward probability model that may easily be implemented and can be used to generate accurate landslide prediction maps. The estimated value of FR indicates how closely landslides and the relevant class of causative factors are related. A number greater than 1 indicates a positive and strong association with higher chances of landslide phenomena, whereas less than 1 shows a negative link and a lower probability of landslides. The association between subclasses of causative parameters and landslides is shown by the FR values presented in Table 3. Elevation classes 210–500 m, 501–1000 m, and 1001–1500 m had higher FR values of 5.577, 1.600, and 1.111, respectively. Slope classes of 30.01–50 and 15.01–30 have higher FR values (1.182 and 1.050). Most landslides occur in Southeast, Southwest, and South facing directions regarding slope aspects. The FR values for these classes are 1.295, 1.159, and 1.132, respectively. FR values (1.294) for concave areas (−1.99 to −1) are higher than those for convex areas. According to the relationship among roughness and the likelihood of a landslide, class 0.41 to 0.50 seems to have the greatest value of FR (1.154). In the case of SPI, the FR value is higher (3.114, 2.397, and 1.616) in 201–400, 601–800, and 401–600 classes, respectively. The class between 8.18 and 12.69 has the greatest FR value in accordance with the relationship involving TWI and landslide probability (1.605). In the case of DtS, higher FR values 1.682 and 1.124 were found for distances between 0–300 m and 301–600 m. Evaluation of DtR revealed a strong association between the occurrence of landslides and a distance of 0–500 m. NDVI classes of 0.01–0.10 and 0.11–0.40 have higher FR values of 1.016 and 1.018, respectively. In the case of LULC, higher FR values were seen for water bodies (1.587) and built up (1.321) areas. For rainfall, the FR values manifest increasing rainfall has a significant trend in increasing landslides. The rainfall classes of 761–871 mm and 872–983 mm have higher FR values (2.752 and 2.044). The relationship connecting MFI and landslides demonstrated that when the MFI value rises, the likelihood of landslides happening likewise rises. In lithology, higher FR values have been noticed in Biotite Quartzite (2.398); Quartz Arenite, Black Slate, Cherty Phyllite (1.815); and Chlorite Sericite Schist and Quartzite (1.278) classes. This work accomplished the LSZ mapping in a GIS environment utilizing the FR model. Using the natural break method, the resulting LSZ map was divided into five categories: very high, high, medium, low, and very low. The study area is discovered to have maximum and lowest percentages of 33.95% and 3.05% for areas with low and very high landslide susceptibility, respectively (Fig. 5 a and Table 5). Regions with very low landslide susceptibility account for 30.96% of the total area, whereas areas with medium and high susceptibility account for 23.09% and 8.95%, respectively.

3.3. Application of IOE model for LSZ mapping

Entropy quantifies the degree of disorder, uncertainty, and instability in a given system. By using this model, it can be inferred that the entropy of a landslide event indicates the extent to which different conditioning factors impact the chances of landslide phenomena. The weight of several landslide conditioning factors for making predictions was determined using the IOE model. The specific class of each conditioning factor demonstrated a different likelihood, which was assessed using the P_{ij} value of the IOE model (Table 3). Since the FR value is taken into account by the probability density (P_{ij}) formula, it is predicted that there would be a correlation between the two categories of coefficients. However, the r-value of 0.79 (Fig. 6) shows a high correlation between FR and P_{ij} , and no superfluous data was used in the hybrid model training procedure. Table 3 displays the outcomes of the IOE model for the LSZ. The elevation, NDVI, and DtR landslide conditioning factors with weights of 0.480, 0.254, and 0.250, respectively, have reportedly had the most effects on the occurrence of landslides. Aspect, curvature, and TWI, with corresponding weights of 0.007, 0.022, and 0.054, had the least influence on landslide susceptibility. The following ranks are occupied by roughness, SPI, rainfall, DtS, LULC, and lithology. Eq. (17) were used to generate the final LSZ map using the IOE model. The natural break approach was used to divide the resulting LSZ map into five classes, spanning from very high to very low. The highest and lowest percentages for places with low and very high landslide susceptibility in the study area are found to be 25.60% and 9.03%, respectively (Fig. 5 b and Table 5). In the basin, 21.68% is located in regions with a very low landslide susceptibility, followed by 24.19% of land in medium susceptibility, and 19.49% in high

Table 5
The percentage of landslide susceptibility zones for FR, IOE, SI, MIV and EBF models.

Landslide Susceptibility Classes	FR	IOE	SI	MIV	EBF
Very low	30.96	21.68	16.19	17.24	17.33
Low	33.95	25.60	26.26	30.97	40.18
Medium	23.09	24.19	26.14	30.33	27.72
High	8.95	19.49	22.43	17.03	12.18
Very high	3.05	9.03	8.99	4.42	2.59

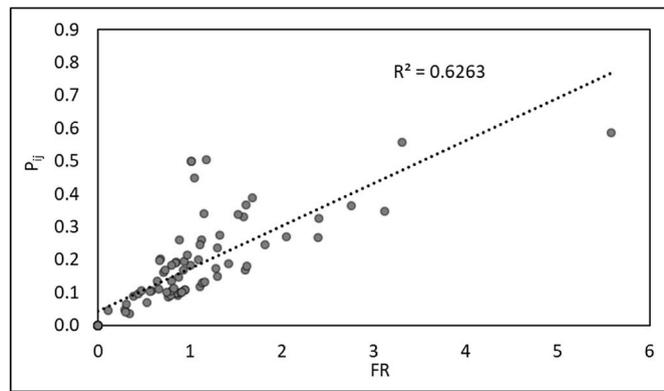


Fig. 6. Correlation between probability (P_{ij}) density and frequency ratio (FR).

susceptibility.

3.4. Application of SI model for LSZ mapping

The weights for each conditioning parameter in the SI model are obtained by dividing the natural logarithm of the landslide density class by the sum of the landslide densities across the entire map. When the number of landslides is higher than average, the weights of the values in natural algorithms will be positive, and when it is lower than average, they will be negative. Table 4 displays the association between each causative factor and landslide as determined by the SI model. The class with the highest positive SI value for elevation is 210–500 m, which is 1.719, followed by 501–1000 m (0.470), 1001–1500 m (0.105), 1501–2000 m (−0.141) and 2001–5124 m (−1.065). The slope angle 30.01–50 class has the highest SI value of 0.167, representing a higher potential for a landslide, while the highest SI value for the Southeast facing slope aspect is observed for the Southeast facing slope (0.258). The SI values of the other slope and aspect sub-classes are lower. The curvature −1.99 to −1 class has the greatest chance of landslide with the highest SI value of 0.258, and for the roughness, 0.41–0.50 class represents the maximum positive value (0.143), suggesting frequent chances of landslide. The SPI value between 201 and 400 has the maximum landslide probability with a positive SI value of 1.136, followed by 601–800 (0.874) and 401–600 (0.480). In the case of the TWI class of 8.18–12.69, the maximum SI value with a positive value (0.473) suggests a greater likelihood of landslide occurrence. The Dts 0–300 m class demonstrates the maximum positive value (0.520), followed by 301–600 m (0.117) and 601–1600 m (−0.230). The risk of a landslide generally increases with decreasing distance from a stream. If we compare Dtr's findings to those of Dts, they are almost identical. The Dtr 0–500 m class represents the maximum SI value (1.195), representing a higher potential for a landslide. In the case of NDVI, the 0.11–0.40 class has the highest SI value of 0.018, followed by 0.01–0.10 (0.015). The SI values of the other NDVI classes are lower. For LULC, the snow cover area has the maximum probability of landslide with an SI value of 1.467, followed by water body (0.462) and built-up area (0.300). According to the findings, landslides are more likely to occur in built-up regions and water bodies, whereas they are less likely to occur within vegetation cover. Regarding the rainfall parameter, the 761–871 mm class gets the maximum SI value of 1.012, followed by 872–983 mm (0.715) and 649–760 mm (0.348). In the case of MFI, the 553.84–626.67 class manifests the maximum SI value of 0.420. Biotite Quartzite lithological formation class has the maximum chances of landslide hazard with an SI value of 0.875. The LSZ mapping for this study was accomplished by utilizing the SI model using Eq. (18) within a GIS environment. Using the natural break approach, the studied basin was divided into five susceptibility zones: very high, high, medium, low, and very low. The maximum and minimum percentages for places with low and very high landslide susceptibility in the study area are found to be 26.26% and 8.99%, respectively (Fig. 5 c and Table 5). A quarter of the study area (26.14%) is situated in areas of medium landslide susceptibility, followed by 16.19% of land in very low susceptibility, and 22.43% in high susceptibility.

3.5. Application of MIV model for LSZ mapping

The Information value model represents an approach to statistical analysis that stems from information theory. As part of the model, landslide occurrence is defined by the information values of the various triggering parameters. The MIV weight values presented in Table 4 illustrate the relationship among subclasses of each triggering factor and the occurrences of landslides. The areas with the highest MIV values were found to be with an elevation of 210–500 m (2.717), a slope of 30.01–50 (1.126), Southeast facing slope aspect (1.198), a curvature of −1.99 to −1 class (1.198), roughness class between 0.41 and 0.50 (1.107), SPI value ranging from 201 to 400 (2.041), TWI of 8.18–12.69 (1.381), indicating the likelihood of landslide occurrence. The locations situated 0–300 m from drainage networks and 0–500 m from roads had the maximum MIV values of 1.423 and 2.105, respectively. The landslide inventory data supported this finding. Landslide risks are highest in the 0.11–0.40 NDVI class, where the MIV value is 1.013, and in the case of LULC maximum value is observed in water body areas, where it is 1.371. The class with the highest MIV value under rainfall is 761–871 mm, with a value of 1.908, and for MFI 553.84–626.67 class has the highest MIV value of 1.335. In Biotite Quartzite lithological formation class, the chances of landslide occurrence are maximum as it has the highest MIV value (1.765). To conduct the LSZ

mapping in the present study, the MIV method was used using Eq. (20). The produced LSZ map was categorized into five sub-classes using the natural break method: very high, high, medium, low, and very low. Areas with low and very high landslide susceptibility are observed to have the highest and lowest percentages of 30.97% and 4.42% of the study area, respectively (Fig. 5 d and Table 5). Areas with medium landslide susceptibility cover 30.33%, and areas with very low and high landslide susceptibility were distributed as 17.24% and 17.03%, respectively.

3.6. Application of EBF model for LSZ mapping

The bivariate EBF approach was utilized to assess geographical relationship between the subclasses of triggering parameters and the incidence of landslides. The estimated EBFs results for the four mass functions of Belief, Disbelief, Uncertainty, and Plausibility are shown in Table 4. A crucial restriction of EBF is that if Belief in a given class has no value, it means that there has never been a landslide in that class. Plausibility and Belief have values between 0 and 1 (Fig. 7). A relatively high value of Belief denotes a high probability of a landslide, whereas a low value of Belief denotes a low probability of a landslide. Based on the value of Belief, the classes with the greatest chance of landslide events were the same as those described for IOE, namely elevation (210–500 m), DtR (0–500 m), slope (30.01–50), NDVI (0.11–0.40), DtS (0–300 m), and rainfall (761–871 mm). The Belief map (Fig. 8a) and the Disbelief map (Fig. 8b) were compared, and the results indicated that Disbelief values were higher in places where belief values were low and vice versa. This was supported by Refs. [12,124]; who demonstrated that locations with high levels of Belief and low levels of disbelief for the occurrence were more likely to be affected by landslides. In regions with low Belief values, the degree of uncertainty associated with each factor type was high. The LSZ map by EBF has been developed by Eqs. (21)–(24) within a GIS environment and shown in Fig. 8. Using the natural break method, the resulting LSZ map was grouped into five categories: very high to very low. By integrating the four fundamental evidential belief functions (Bel, Dis, Unc and Pls) with GIS software maps were produced as illustrated in Fig. 8. The maximum and minimum percentages for places with low and very high landslide susceptibility in the studied river basin are found to be 40.18% and 2.59%, respectively (Fig. 9 and Table 5). Regions with medium landslide susceptibility account for 27.72% of the total area, whereas areas with very low and high susceptibility account for 17.33% and 12.18%, respectively.

3.7. Validation of LSZ maps

In the present work, landslide prediction maps generated by five models (FR, IOE, SI, MIV, EBF) were verified by comparing the prediction map to the training and validating data. About 70% of all landslide data was utilized for training the model, and 30% was used to validate it after training. The AUC values of each of the five models shown in Fig. 10 were used to determine the success and prediction rates.

AUC values less than 0.5 suggest that the model’s performance has no discriminating capacity. Whereas values between 0.5 and 0.7 demonstrate substantial discrimination ability. In contrast, values between 0.7 and 0.8 show acceptable discrimination capacity, and AUC values above 0.9 represent an excellent discrimination ability [131]. The findings showed that the IOE model has the highest AUC value of 0.958 for the success-rate curve, followed by SI (0.926), MIV (0.922), FR (0.915), and EBF (0.899) models (Fig. 10 a). The results indicated that the IOE model has the highest training accuracy of 95.80%, followed by SI (92.60%), MIV (92.20%), FR (91.50%), and EBF (89.90%) models. In the same way, the results of the prediction-rate curve showed that the IOE model has the highest AUC value of 0.964, followed by the SI (0.928), MIV (0.926), FR (0.917), and EBF (0.893) models (Fig. 10 b). This means that the IOE model has the highest prediction accuracy of 96.40%, followed by the SI (92.80%), MIV (92.60%), FR (91.70%), and EBF (89.30%) models. The AUC assessment indicated that all models had excellent accuracy in landslide prediction mapping for the present

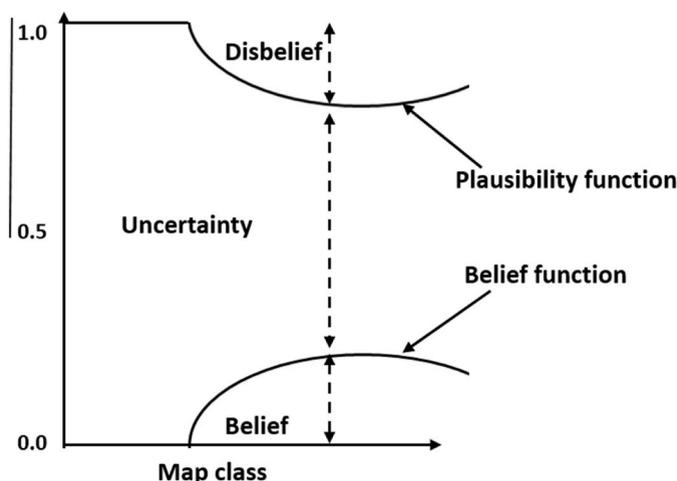


Fig. 7. Schematic relationships of evidential belief functions [124].

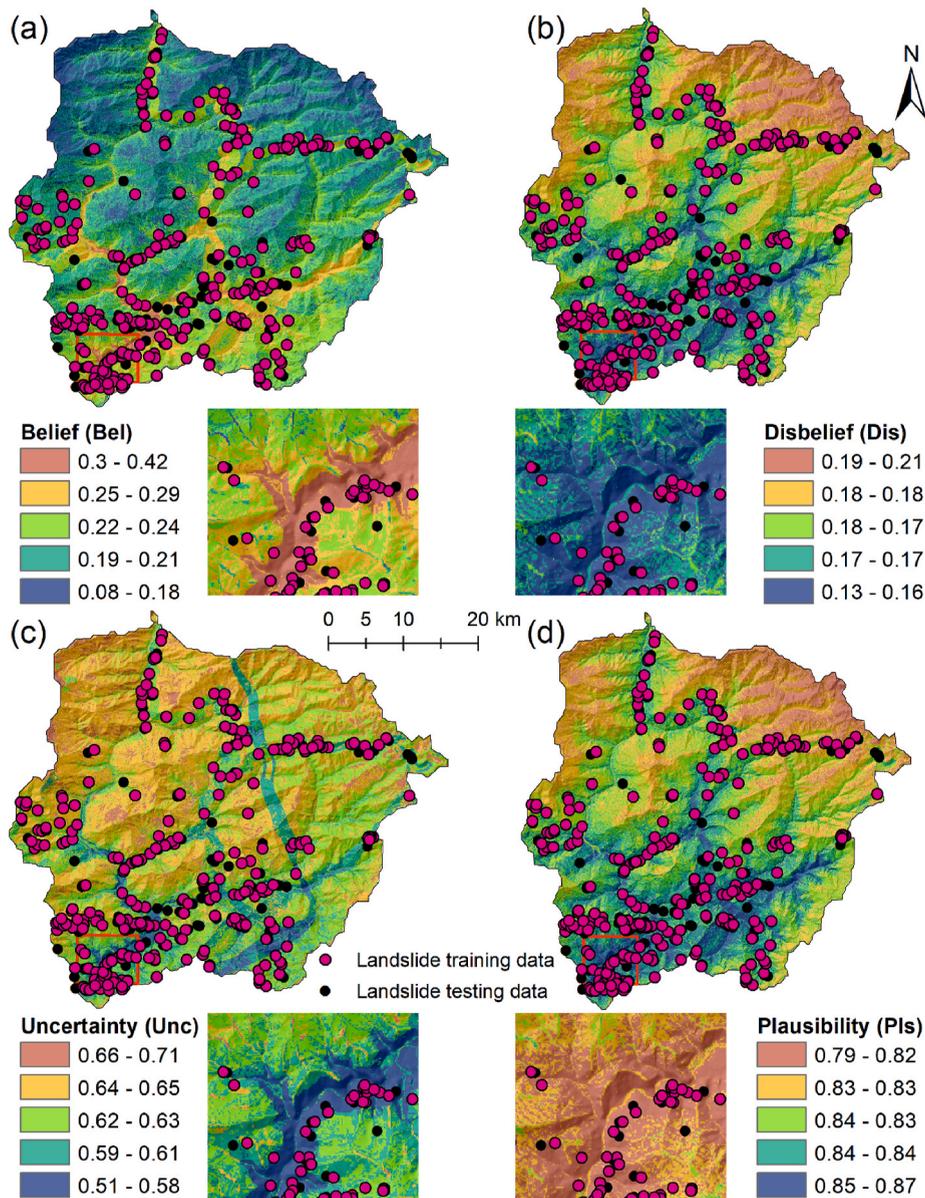


Fig. 8. Integrated results of Evidence belief function (a) Belief; (b) Disbelief; (c) Uncertainty; and (d) Plausibility.

study, supporting the findings of the scientific analysis that the success rate and prediction rate curves are almost equal. As a result, any of the models may be used to analyse potential future landslide occurrences and a risk assessment in the study area. For the current study, the landslide-susceptible map created by the IOE model performed the best in terms of landslide prediction mapping.

3.8. Discussion

Landslide prediction is becoming a critical worldwide issue for sustainable development in mountainous locations. In earlier studies, various methods that provide compelling findings have been used to predict or map the susceptibility of landslides [28,12,132, 131,19,133,6,46]. However, the prediction capability of landslide susceptibility is constrained by a single model. Sustainable development relies heavily on proper planning and efficient management, which have become more important as people understand landslide susceptibility zones better. Despite this, extensive studies are still required to fully comprehend the interrelationship between the conditioning factors and landslide distribution.

In the present study, landslide prediction mapping was carried out using data-driven bivariate statistical models (FR, IOE, SI, MIV, and EBF) to determine the most appropriate landslide prediction. Landslide susceptibility maps were generated utilizing the FR value, Pij value, SI value, MIV value, and Bel value from the five bivariate statistical models. Five different susceptibility maps were created

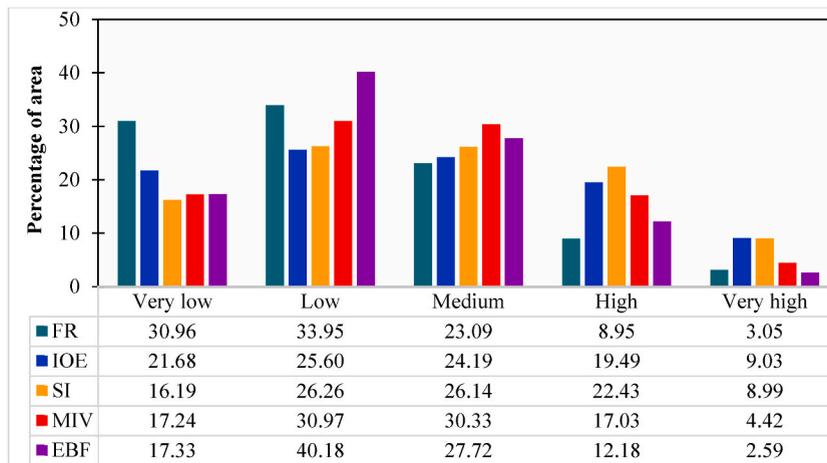


Fig. 9. Landslide susceptibility distribution within the five susceptibility levels.

using the same fourteen causative factors and compared and assessed. All five models provide comparable results when calculating the weight for each subclass of fourteen conditioning factors. The findings indicated that landslide risk dropped as the elevation rose. The thin layer of colluvial deposit covering the 210–500 m elevation class makes this zone prone to landslides. On the other hand, landslides are less common at higher altitudes due to the smaller areal extent of land [12]. found that locations at 519–1508 m elevation had the greatest landslide occurrence likelihood in their study of landslide susceptibility in the Sikkim Himalayan region. It is clear that landslides occur less often on gentle slopes owing to lower shear stress at slope angles, but they occur more frequently on steep slopes due to increased shear stress. According to a research conducted by Ref. [11] to analyse the landslide susceptibility in the Lachung River basin, North Sikkim, locations with a slope angle of 35.488°–43.580° had the greatest landslide occurrence probability. In the study area, most landslides happen on southeast-facing slopes. It is hypothesized that the slope aspect influences the soil strength because it affects the regional vegetation types. The curvature value significantly impacts the terrain’s morphology, substantially affecting debris flow’s acceleration, deceleration, convergence, and divergence. Positive curvature values signify convex upwards shape, whereas negative curvature values signify concave shape. The density of landslides is higher mostly in concave-curvature form because it enhances the soil moisture content and causes slope failure. In comparison to low convex slope, profile curvature classes with high concave slope had the greatest risk of landslip occurring, according to Ref. [11] findings. Another aspect influencing slope stability is the distance to the stream. It alters the hydrologic function of the slope and may impair its strength by weakening or wetting the slope’s material. The risk of landslides is greatest in the area nearest to the river. The results of a research on landslide susceptibility in the Artvin (Turkey) conducted by Akinci and Ozalp in 2021 showed that regions within 100 m of rivers were very susceptible to the appearance of landslides, but the susceptibility of landslides reduced as distance rose. The topography will inevitably change when highways are built across hilly areas. Consequently, the pressure on the rear of the slope rises, and tension fractures may form. The area closest to the roadways has the highest risk of landslides. The majority of landslides, according to Ref. [28]; occur near roads due to changes in slope stability, and they noted that the subclass within 200 m of roadways in the studied regions had the maximum values through all models used in the study. In regions with moderate NDVI values, shrubs and grassland are common, and shallow plant roots significantly impact severe landslides. According to Ref. [12]; locations with NDVI values ranging from 0.297 to 0.477 have the greatest risk of landslides. The prevalence of landslides in the study area has been linked to unsustainable development practices, particularly the construction of haphazardly built settlements and roads. According to Ref. [6] in addition to natural causes, human activities including excavation, deforestation, changes in land use, and slope profile can cause landslides to occur. In the case of rainfall, the result showed that increasing rainfall has a significant trend in increasing landslides. The relation between MFI and landslide showed that as the MFI value increases, the probability of landslide occurrences also increases. In lithology, the Biotite Quartzite class has the highest probability of landslide occurrences. This is because rainwater infiltrates through porous lithological structures and often triggers shallow landslides.

Elevation, DtR, Slope, NDVI, DtS, and Rainfall were discovered to be the most influential indicators for LSZ mapping in the study area. The final LSZ maps showed that the most landslide-prone areas were concentrated in the southwest and southeast part of the study area. The LSZ maps clearly show that roadside landslides are very susceptible to frequent occurrences. In most instances, barrier walls prevent the movement of the landslide, but in the unprotected region, debris, sometimes combined with huge boulders, slips onto the road and obstructs accessibility. The validation of the result through the ROC curve indicated the effectiveness of the five bivariate statistical models in meeting the study’s objectives. The IOE model has the highest AUC value for the success and prediction rate curves in the validation phase. In comparison to earlier studies carried out by various researchers in the Darjeeling-Sikkim Himalaya region [12,134,13,19,11,135], the LSZ maps created by the FR, IOE, SI, MIV, and EBF models have been delivering superior results. These statistical models may be enhanced further by ensemble them using machine learning algorithms. This sort of ensemble approach has been used in recent years by several researchers [12,42,44,43,7,10,46] to address a variety of real-world classification issues as well as to improve the accuracy of the LSZ maps.

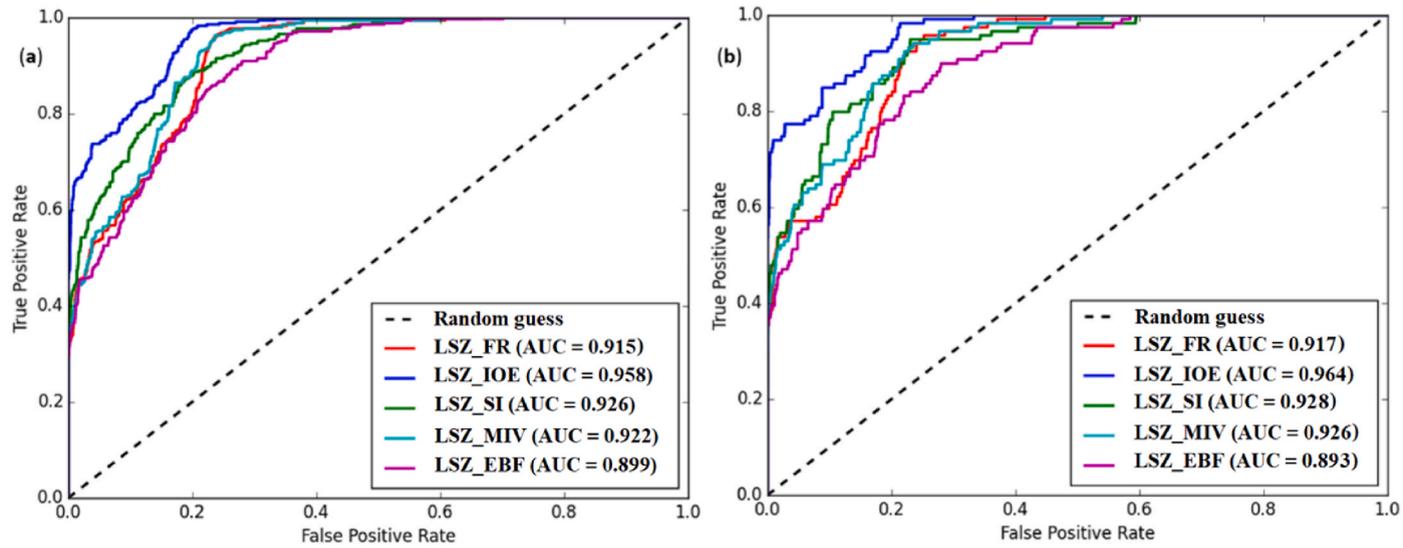


Fig. 10. Receiver operating characteristics ROC curve (a) Success rate; and (b) Prediction rate.

4. Conclusions

Landslide is one of the frequent calamities in the Darjeeling-Sikkim Himalayan region, which wreaks financial havoc, claiming human lives, causing ecological damages, and affecting the socio-economic life of this region. Evaluating landslide prediction has gained the attention of many researchers and policymakers worldwide. Risk reduction is necessary for landslide hazard evaluation and mitigation to encourage sustainable economic growth. Five GIS-based statistical landslide predicting models have been utilized in this research for the tectonically active Upper Tista Basin of the Darjeeling-Sikkim Himalayan region of India. All the models are widely used in contemporary periods due to their high performance, effective functions, and excellent reliability.

The current investigation divided the study area into five landslide susceptibility classes, varying from very high to very low, predicated based on the susceptibility indexed values of the respective models, i.e., FR, IOE, SI, MIV, and EBF. The FR statistical model demonstrates the most impacting LCF are elevation, SPI, aspect, MFI, and lithology; whereas in the IOE model, elevation, slope, DtR, NDVI, and lithology are most effective; in the MIV model, elevation, aspect, SPI, rainfall, and lithology are most effective; and in EBF model, elevation, slope, DtR, NDVI, and LULC are the significantly impacting factors. Parallely, the SI model shows seven factors are depicted with positive values, which are the most influencing factors (i.e., elevation, SPI, DtS, NDVI, LULC, rainfall, and lithology), and the rest of the seven factors show negative values, i.e., they are the less impacting factors in predicting the landslide. This study also portrayed anthropogenic interventions in the mountainous Upper Tista Basin that had a significant impact on slope instability based on selected LCF (LULC and DtR) and field survey. The comparative assessment of the five models reflects that the basin is situated in a high-risk-prone region. According to the FR, MIV, IOE, SI, and EBF models, the proportions of high and very high landslide-prone areas are 12.00%, 21.46%, 28.53%, 31.42%, and 14.17%, respectively. The positive aspect of this research manifests that the success rate and prediction rate curves are almost equivalent and were supported by the AUC assessment, which showed that all algorithms had high accuracy in predicting landslide zones for the current research.

This kind of statistical model-based study demonstrated the spatial distribution of landslide-prone areas and identified the factors that have the greatest impact on the occurrence of such vulnerability. But, proper and scientific field-based research is needed to better understand the vast impacts of such calamities. Database paucity is one of the largest issues in developing nations like India; consequently, landslide prediction studies will be greatly advanced by improving the intake layers. Future research should focus on in-depth field-based investigations to detect the other driving factors of slope instability. The integration of machine learning approaches with GIS should be employed for better improvement in this field. This research indicates the employed models are appropriate and acceptable for the study area, and these data-driven bivariate statistical models can also be utilized in other frequently landslide-susceptible areas of the world.

Author contribution statement

Jayanta Das: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Pritam Saha; Rajib Mitra: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Asraful Alam; Md. Kamruzzaman: Performed the experiments; Contributed reagents, materials, analysis tools or data.

Data availability statement

Data will be made available on request.

Additional information

Supplementary content related to this article has been published online at [URL].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] F. Huang, Z. Cao, J. Guo, S.H. Jiang, S. Li, Z. Guo, Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping, *Catena* 191 (2020), 104580, <https://doi.org/10.1016/j.catena.2020.104580>.
- [2] T. Stanley, D.B. Kirschbaum, A heuristic approach to global landslide susceptibility mapping, *Nat. Hazards* 87 (2017) 145–164, <https://doi.org/10.1007/s11069-017-2757-y>.
- [3] G. Balamurugan, V. Ramesh, M. Touthang, Landslide susceptibility zonation mapping using frequency ratio and fuzzy gamma operator models in part of NH-39, Manipur, India, *Nat. Hazards* 84 (2016) 465–488, <https://doi.org/10.1007/s11069-016-2434-6>.
- [4] A. Chawla, S. Pasupuleti, S. Chawla, A.C.S. Rao, K. Sarkar, R. Dwivedi, Landslide susceptibility zonation mapping: a case study from darjeeling district, eastern himalayas, India, *J. Indian Soc. Remote Sens.* 47 (2019) 497–511, <https://doi.org/10.1007/s12524-018-0916-6>.
- [5] J. Mathew, V.K. Jha, G.S. Rawat, Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India, *Int. J. Rem. Sens.* 28 (2007) 2257–2275, <https://doi.org/10.1080/01431160600928583>.

- [6] A. Saha, S. Saha, Application of statistical probabilistic methods in landslide susceptibility assessment in Kurseong and its surrounding area of Darjeeling Himalayan, India: RS-GIS approach, *Environ. Dev. Sustain.* 23 (2021) 4453–4483, <https://doi.org/10.1007/s10668-020-00783-1>. Springer Netherlands.
- [7] A. Saha, S. Saha, Comparing the efficiency of weight of evidence, support vector machine and their ensemble approaches in landslide susceptibility modelling: a study on Kurseong region of Darjeeling Himalaya, India, *Remote Sens. Appl.: Society and Environ.* 19 (2020), 100323, <https://doi.org/10.1016/j.rsase.2020.100323>.
- [8] K. Mandal, S. Saha, S. Mandal, Applying deep learning and benchmark machine learning algorithms for landslide susceptibility modelling in Rorachu river basin of Sikkim Himalaya, India, *Geosci. Front.* 12 (5) (2021), 101203, <https://doi.org/10.1016/j.gsf.2021.101203>.
- [9] S.K. Gupta, D.P. Shukla, M. Thakur, Selection of weightages for causative factors used in preparation of landslide susceptibility zonation (LSZ), *Geomatics, Nat. Hazards Risk* 9 (2018) 471–487, <https://doi.org/10.1080/19475705.2018.1447027>.
- [10] S. Saha, A. Arabameri, A. Saha, T. Blaschke, P.T.T. Ngo, V.H. Nhu, S.S. Band, Prediction of landslide susceptibility in Rudraprayag, India using novel ensemble of conditional probability and boosted regression tree-based on cross-validation method, *Sci. Total Environ.* 764 (2021), 142928, <https://doi.org/10.1016/j.scitotenv.2020.142928>.
- [11] S.C. Pal, I. Chowdhuri, GIS-based spatial prediction of landslide susceptibility using frequency ratio model of Lachung River basin, North Sikkim, India, *SN Appl. Sci.* 1 (2019) 1–25, <https://doi.org/10.1007/s42452-019-0422-7>.
- [12] I. Chowdhuri, S.C. Pal, A. Arabameri, P.T.T. Ngo, R. Chakraborty, S. Malik, B. Das, P. Roy, Ensemble approach to develop landslide susceptibility map in landslide dominated Sikkim Himalayan region, India, *Environ. Earth Sci.* 79 (2020) 1–28, <https://doi.org/10.1007/s12665-020-09227-5>.
- [13] S. Ghosh, R. Das, B. Goswami, Developing GIS-based techniques for application of knowledge and data-driven methods of landslide susceptibility mapping, *Indian Journal of Geosciences* 67 (2013) 249–272, <https://doi.org/10.1016/j.geomorph.2011.04.019>.
- [14] Y. Hong, R. Adler, G. Huffman, Use of satellite remote sensing data in the mapping of global landslide susceptibility, *Nat. Hazards* 43 (2007) 245–256, <https://doi.org/10.1007/s11069-006-9104-z>.
- [15] B. Feby, A.L. Achu, K. Jimnisha, V.A. Ayisha, R. Reghunath, Landslide susceptibility modelling using integrated evidential belief function based logistic regression method: a study from Southern Western Ghats, India, *Remote Sens. Appl.: Society and Environment* 20 (2020), 100411, <https://doi.org/10.1016/j.rsase.2020.100411>.
- [16] H. Saito, D. Nakayama, H. Matsuyama, Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: the Akaishi Mountains, Japan, *Geomorphology* 109 (2009) 108–121, <https://doi.org/10.1016/j.geomorph.2009.02.026>.
- [17] T. Basu, S. Pal, RS-GIS based morphometrical and geological multi-criteria approach to the landslide susceptibility mapping in Gish River Basin, West Bengal, India, *Adv. Space Res.* 63 (2019) 1253–1269, <https://doi.org/10.1016/j.asr.2018.10.033>.
- [18] S. Bera, B. Guru, V. Ramesh, Evaluation of landslide susceptibility models: a comparative study on the part of Western Ghat Region, India, *Remote Sens. Appl.: Society and Environment* 13 (2019) 39–52, <https://doi.org/10.1016/j.rsase.2018.10.010>.
- [19] B. Mandal, Analytical hierarchy process (AHP) based landslide susceptibility mapping of Lish river basin of eastern Darjeeling Himalaya, India, *Adv. Space Res.* 62 (2018) 3114–3132, <https://doi.org/10.1016/j.asr.2018.08.008>.
- [20] H.R. Pourghasemi, M. Mohammady, B. Pradhan, Landslide susceptibility mapping using index of entropy and conditional probability models in GIS: safarood Basin, Iran, *Catena* 97 (2012) 71–84, <https://doi.org/10.1016/j.catena.2012.05.005>.
- [21] P. Aleotti, R. Chowdhury, Landslide hazard assessment: summary review and new perspectives, *Bull. Eng. Geol. Environ.* 58 (1999) 21–44, <https://doi.org/10.1007/s100640050066>.
- [22] A.K. Batar, T. Watanabe, Landslide susceptibility mapping and assessment using geospatial platforms and weights of evidence (WoE) method in the Indian Himalayan region: recent developments, gaps, and future directions, *ISPRS Int. J. Geo-Inf.* 10 (2021) 114, <https://doi.org/10.3390/ijgi10030114>.
- [23] S. Lee, J.A. Talib, Probabilistic landslide susceptibility and factor effect analysis, *Environ. Geol.* 47 (2005) 982–990, <https://doi.org/10.1007/s00254-005-1228-z>.
- [24] B.T. Pham, D. Tien Bui, P. Indra, M. Dholakia, Landslide susceptibility assessment at a part of Uttarakhand Himalaya, India using GIS-based statistical approach of frequency ratio method, *Int. J. Eng. Res. Technol.* 4 (2015) 338–344, <https://doi.org/10.17577/ijertv4is110285>.
- [25] V. Ramesh, S. Anbazhagan, Landslide susceptibility mapping along Koli hills Ghat road section (India) using frequency ratio, relative effect and fuzzy logic models, *Environ. Earth Sci.* 73 (2015) 8009–8021, <https://doi.org/10.1007/s12665-014-3954-6>.
- [26] M. Sahana, H. Sajjad, Evaluating effectiveness of frequency ratio, fuzzy logic and logistic regression models in assessing landslide susceptibility: a case from Rudraprayag district, India, *J. Mt. Sci.* 14 (2017) 2150–2167, <https://doi.org/10.1007/s11629-017-4404-1>.
- [27] D.T. Bui, O. Lofman, I. Revhaug, O. Dick, Landslide susceptibility analysis in the Hoa Binh province of Vietnam using statistical index and logistic regression, *Nat. Hazards* 59 (2011) 1413–1444, <https://doi.org/10.1007/s11069-011-9844-2>.
- [28] H. Akinci, A. Yavuz Ozalp, Landslide susceptibility mapping and hazard assessment in Artvin (Turkey) using frequency ratio and modified information value model, *Acta Geophys.* 69 (2021) 725–745, <https://doi.org/10.1007/s11600-021-00577-7>.
- [29] Y. Wang, Z. Fang, H. Hong, R. Costache, X. Tang, Flood susceptibility mapping by integrating frequency ratio and index of entropy with multilayer perceptron and classification and regression tree, *J. Environ. Manag.* 289 (2021), 112449, <https://doi.org/10.1016/j.jenvman.2021.112449>.
- [30] T. Basu, S. Pal, Identification of landslide susceptibility zones in Gish River basin, West Bengal, India, *Georisk* 12 (2018) 14–28, <https://doi.org/10.1080/17499518.2017.1343482>.
- [31] E.R. Sujatha, V. Sridhar, Landslide susceptibility analysis: a logistic regression model case study in Coonoor, India, *Hydrology* 8 (2021) 41, <https://doi.org/10.3390/hydrology8010041>.
- [32] M. Komac, A landslide susceptibility model using the Analytical Hierarchy Process method and multivariate statistics in perialpine Slovenia, *Geomorphology* 74 (2006) 17–28, <https://doi.org/10.1016/j.geomorph.2005.07.005>.
- [33] M. Onagh, V.K. Kumra, P.K. Rai, Landslide susceptibility mapping in a part of Uttarakashi district (India) by multiple linear regression method, *Int. J. Geol. Earth and Environ. Sci.* 2 (2012) 102–120.
- [34] A.P. Dempster, Upper and lower probabilities induced by a multivalued mapping, in: *Classic Works of the Dempster-Shafer Theory of Belief Functions*, Springer, Berlin, Heidelberg, 2008, pp. 57–72, https://doi.org/10.1007/978-3-540-44792-4_3.
- [35] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976, p. 42.
- [36] B.T. Pham, D. Tien Bui, H.R. Pourghasemi, P. Indra, M.B. Dholakia, Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: a comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods, *Theor. Appl. Climatol.* 128 (2017) 255–273, <https://doi.org/10.1007/s00704-015-1702-9>.
- [37] Y. Huang, L. Zhao, Review on landslide susceptibility mapping using support vector machines, *Catena* 165 (2018) 520–529, <https://doi.org/10.1016/j.catena.2018.03.003>.
- [38] D. Kumar, M. Thakur, C.S. Dubey, D.P. Shukla, Landslide susceptibility mapping & prediction using support vector machine for mandakini River Basin, Garhwal Himalaya, India, *Geomorphology* 295 (2017) 115–125, <https://doi.org/10.1016/j.geomorph.2017.06.013>.
- [39] J. Roy, S. Saha, A. Arabameri, T. Blaschke, D.T. Bui, A novel ensemble approach for landslide susceptibility mapping (LSM) in darjeeling and, *Rem. Sens.* 11 (2019) 2886, <https://doi.org/10.3390/rs11232866>.
- [40] B. Peethambaran, R. Anbalagan, D.P. Kanungo, A. Goswami, K.V. Shihabudheen, A comparative evaluation of supervised machine learning algorithms for township level landslide susceptibility zonation in parts of Indian Himalayas, *Catena* 195 (2020), 104751, <https://doi.org/10.1016/j.catena.2020.104751>.
- [41] J.N. Goetz, A. Brenning, H. Petschko, P. Leopold, Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling, *Comput. Geosci.* 81 (2015) 1–11, <https://doi.org/10.1016/j.cageo.2015.04.007>.
- [42] A. Merghadi, A.P. Yunus, J. Dou, J. Whiteley, B. ThaiPham, D.T. Bui, R. Avtar, B. Abderrahmane, Machine learning methods for landslide susceptibility studies: a comparative overview of algorithm performance, *Earth Sci. Rev.* 207 (2020), 103225, <https://doi.org/10.1016/j.earscirev.2020.103225>.
- [43] H.R. Pourghasemi, O. Rahmati, Prediction of the landslide susceptibility: which algorithm, which precision? *Catena* 162 (2018) 177–192, <https://doi.org/10.1016/j.catena.2017.11.022>.

- [44] B.T. Pham, D. Tien Bui, I. Prakash, M.B. Dholakia, Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS, *Catena* 149 (2017) 52–63, <https://doi.org/10.1016/j.catena.2016.09.007>.
- [45] Y. Wang, Z. Fang, H. Hong, Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China, *Sci. Total Environ.* 666 (2019) 975–993, <https://doi.org/10.1016/j.scitotenv.2019.02.263>.
- [46] S. Saha, J. Roy, B. Pradhan, T.K. Hembram, Hybrid ensemble machine learning approaches for landslide susceptibility mapping using different sampling ratios at East Sikkim Himalayan, India, *Adv. Space Res.* 68 (2021) 2819–2840, <https://doi.org/10.1016/j.asr.2021.05.018>.
- [47] R. Mitra, D. Kumar Mandal, Assessment of livelihood vulnerability in the riparian region of the Tista River, West Bengal, India, *Geojournal* 1–29 (2022), <https://doi.org/10.1007/s10708-022-10645-0>.
- [48] J.J. Jennifer, Feature elimination and comparison of machine learning algorithms in landslide susceptibility mapping, *Environ. Earth Sci.* 81 (2022) 1–23, <https://doi.org/10.1007/s12665-022-10620-5>.
- [49] R. Mahalingam, B. Kim, Factors affecting occurrence of landslides induced by the M7. 8 April 2015, Nepal earthquake, *KSCSE J. Civ. Eng.* 25 (2021) 78–91, <https://doi.org/10.1007/s12205-020-0508-1>.
- [50] B. Pokharel, M. Alvioli, S. Lim, Assessment of earthquake-induced landslide inventories and susceptibility maps using slope unit-based logistic regression and geospatial statistics, *Sci. Rep.* 11 (2021) 1–15, <https://doi.org/10.1038/s41598-021-00780-y>.
- [51] I. Ullah, B. Aslam, S.H.I.A. Shah, A. Tariq, S. Qin, M. Majeed, H.B. Havenith, An integrated approach of machine learning, remote sensing, and GIS data for the landslide susceptibility mapping, *Land* 11 (2022) 1265, <https://doi.org/10.3390/land11081265>.
- [52] Z. Mansour, T. Yanick, S. Aissa, R. Soraya, H. Abderahmane, A. Abdelkader, K. Abdelkader, S. Zahera, M. Souad, M. Kacem, The susceptibility analysis of landslide using bivariate and multivariate modeling techniques in western Algeria: case of Fergoug watershed (Beni-Chougrane Mountains), *Arabian J. Geosci.* 14 (19) (2021) 1–19, <https://doi.org/10.1007/s12517-021-07919-1>.
- [53] D. Roy, S. Das, R. Mitra, An application of geospatial-based multi-criteria decision-making technique to identify landslide susceptibility zones in the Ragnu Khola River Basin of Darjeeling Himalayan region, India, *Applied Geomatics* (2022) 1–19, <https://doi.org/10.1007/s12518-022-00468-6>.
- [54] G. Berhane, M. Kebede, N. Alfarah, E. Hagos, B. Grum, A. Giday, T. Abera, Landslide susceptibility zonation mapping using GIS-based frequency ratio model with multi-class spatial data-sets in the Adwa-Adigrat mountain chains, northern Ethiopia, *J. Afr. Earth Sci.* 164 (2020), 103795, <https://doi.org/10.1016/j.jafrearsci.2020.103795>.
- [55] M.T. Abraham, N. Satyam, R. Lokesh, B. Pradhan, A. Alamri, Factors affecting landslide susceptibility mapping: assessing the influence of different machine learning approaches, sampling strategies and data splitting, *Land* 10 (2021) 989, <https://doi.org/10.3390/land10090989>.
- [56] M.A. Hussain, Z. Chen, Y. Zheng, M. Shoaib, S.U. Shah, N. Ali, Z. Afzal, Landslide susceptibility mapping using machine learning algorithm validated by persistent scatterer In-SAR technique, *Sensors* 22 (2022) 3119, <https://doi.org/10.3390/s22093119>.
- [57] A. Trigila, C. Iadanza, C. Esposito, G. Scarascia-Mugnozza, Comparison of logistic regression and random forests techniques for shallow landslide susceptibility assessment in giampilieri (NE sicily, Italy), *Geomorphology* 249 (2015) 119–136, <https://doi.org/10.1016/j.geomorph.2015.06.001>.
- [58] K. Cui, D. Lu, W. Li, Comparison of landslide susceptibility mapping based on statistical index, certainty factors, weights of evidence and evidential belief function models, *Geocarto Int.* 32 (2017) 935–955, <https://doi.org/10.1080/10106049.2016.1195886>.
- [59] V. Gupta, N. Chauhan, J. Penna, R. Hermans, J. Dehls, A. Sengupta, R.K. Bhasin, Geomorphic evaluation of landslides along the Teesta river valley, Sikkim Himalaya, India, *Geol. J.* 57 (2022) 611–621, <https://doi.org/10.1002/gj.4377>.
- [60] T. Sivasankar, S. Ghosh, M. Joshi, Exploitation of optical and SAR amplitude imagery for landslide identification: a case study from Sikkim, Northeast India, *Environ. Monit. Assess.* 193 (2021) 1–17, <https://doi.org/10.1007/s10661-021-09119-6>.
- [61] I. Sonker, J.N. Tripathi, A.K. Singh, Landslide susceptibility zonation using geospatial technique and analytical hierarchy process in Sikkim Himalaya, *Quaternary Science Advances* 4 (2021), 100039, <https://doi.org/10.1016/j.qsa.2021.100039>.
- [62] W. Feng, H. Bai, B. Lan, Y. Wu, Z. Wu, L. Yan, X. Ma, Spatial-temporal distribution and failure mechanism of group-occurring landslides in Mibe village, Longchuan County, Guangdong, China, *Landslides* (2022) 1–14, <https://doi.org/10.1007/s10346-022-01904-9>.
- [63] W.S. Udin, N.N. Yahaya, S.I.M. Shariffuddin, Landslide susceptibility assessment using geographic information system in Aring, Gua Musang, Kelantan, in: *IOP Conference Series: Earth and Environmental Science* vol. 842, 2021, 012008.
- [64] A. Abdi, A. Bouamrane, T. Karech, N. Dahri, A. Kaouachi, Landslide susceptibility mapping using GIS-based fuzzy logic and the analytical hierarchical processes approach: a case study in constantine (North-East Algeria), *Geotech. Geol. Eng.* 39 (2021) 5675–5691, <https://doi.org/10.1007/s10706-021-01855-3>.
- [65] K. Pawluszek, A. Borkowski, Impact of DEM-derived factors and analytical hierarchy process on landslide susceptibility mapping in the region of Rożnów Lake, *Poland, Nat. Hazards* 86 (2017) 919–952, <https://doi.org/10.1007/s11069-016-2725-y>.
- [66] M.B. Hassangavyar, H.E. Damaneh, Q.B. Pham, N.T.T. Linh, J. Tiefenbacher, Q.V. Bach, Evaluation of re-sampling methods on performance of machine learning models to predict landslide susceptibility, *Geocarto Int.* 37 (2022) 2772–2794, <https://doi.org/10.1080/10106049.2020.1837257>.
- [67] S. Jaksic, J. Ninkov, S. Milic, J. Vasin, M. Zivanov, D. Jaksic, V. Komlen, Influence of slope gradient and aspect on soil organic carbon content in the region of Niš, Serbia, *Sustainability* 13 (2021) 8332, <https://doi.org/10.3390/su13158332>.
- [68] D. Yang, H. Qiu, S. Hu, Y. Pei, X. Wang, C. Du, Y. Long, M. Cao, Influence of successive landslides on topographic changes revealed by multitemporal high-resolution UAS-based DEM, *Catena* 202 (2021), 105229, <https://doi.org/10.1016/j.catena.2021.105229>.
- [69] T.E. Bilir, I. Fung, T.E. Dawson, Slope-aspect induced climate differences influence how water is exchanged between the land and atmosphere, *J. Geophys. Res.: Biogeosciences* 126 (2021), 2020JG006027, <https://doi.org/10.1029/2020JG006027>.
- [70] T.X. Bien, P.T. Truyen, T. Van Phong, D.D. Nguyen, M. Amiri, R. Costache, D.M. Duc, H. Van Le, H.B.T. Nguyen, I. Prakash, B.T. Pham, Landslide susceptibility mapping at sin Ho, lai chau province, vietnam using ensemble models based on fuzzy unordered rules induction algorithm, *Geocarto Int.* (2022) 1–21, <https://doi.org/10.1080/10106049.2022.2136253>.
- [71] Y.S. Cheng, T.T. Yu, N.T. Son, Random forests for landslide prediction in tsengwen river watershed, central taiwan, *Rem. Sens.* 13 (2021) 199, <https://doi.org/10.3390/rs13020199>.
- [72] R. Mitra, J. Das, A comparative assessment of flood susceptibility modelling of GIS-based TOPSIS, VIKOR, and EDAS techniques in the Sub-Himalayan foothills region of Eastern India, *Environ. Sci. Pollut. Res.* (2022), <https://doi.org/10.1007/s11356-022-23168-5>.
- [73] I. Bostjancic, M. Filipović, V. Gulam, D. Pollak, Regional-scale landslide susceptibility mapping using limited LiDAR-based landslide inventories for Sisak-Moslavina County, Croatia, *Sustainability* 13 (2021) 4543, <https://doi.org/10.3390/su13084543>.
- [74] P. Prasad, V.J. Loveson, S. Das, P. Chandra, Artificial intelligence approaches for spatial prediction of landslides in mountainous regions of western India, *Environ. Earth Sci.* 80 (2021) 1–20, <https://doi.org/10.1007/s12665-021-10033-w>.
- [75] R. Mitra, P. Saha, J. Das, Assessment of the performance of GIS-based analytical hierarchical process (AHP) approach for flood modelling in Uttar Dinajpur district of West Bengal, India, *Geomatics, Nat. Hazards Risk* 13 (2022) 2183–2226, <https://doi.org/10.1080/19475705.2022.2112094>.
- [76] A. Chueasamat, T. Hori, H. Saito, T. Sato, Y. Kohgo, Experimental tests of slope failure due to rainfalls using 1g physical slope models, *Soils Found.* 58 (2018) 290–305, <https://doi.org/10.1016/j.sandf.2018.02.003>.
- [77] J. Rodriguez, R. Macciotta, M.T. Hendry, M. Roustaei, C. Grapel, R. Skirrow, UAVs for monitoring, investigation, and mitigation design of a rock slope with multiple failure mechanisms—a case study, *Landslides* 17 (9) (2020) 2027–2040, <https://doi.org/10.1007/s10346-020-01416-4>.
- [78] X. Xiong, Z. Shi, Y. Xiong, M. Peng, X. Ma, F. Zhang, Unsaturated slope stability around the Three Gorges Reservoir under various combinations of rainfall and water level fluctuation, *Eng. Geol.* 261 (2019), 105231, <https://doi.org/10.1016/j.enggeo.2019.105231>.
- [79] E.M. Amah, V.Y. Katte, R.T. Ghogomu, V.K. Kamgano, An assessment of landslides along mountain forest roads: bamenda ring road segment through bafut and befang forests on the Cameroon volcanic line, *Am. J. Environ. Stud.* 5 (2022) 60–86, <https://doi.org/10.47672/ajes.1160>.
- [80] J. Liu, Q. Xu, S. Wang, S. Siva Subramanian, L. Wang, X. Qi, Formation and chemo-mechanical characteristics of weak clay interlayers between alternative mudstone and sandstone sequence of gently inclined landslides in Nanjiang, SW China, *Bull. Eng. Geol. Environ.* 79 (2020) 4701–4715, <https://doi.org/10.1007/s10064-020-01859-y>.

- [81] J. Zhan, Q. Wang, W. Zhang, Y. Shangguan, S. Song, J. Chen, Soil-engineering properties and failure mechanisms of shallow landslides in soft-rock materials, *Catena* 181 (2019), 104093, <https://doi.org/10.1016/j.catena.2019.104093>.
- [82] H. Bourenane, Y. Bouhadad, M.S. Guettouche, M. Braham, GIS-based landslide susceptibility zonation using bivariate statistical and expert approaches in the city of Constantine (Northeast Algeria), *Bull. Eng. Geol. Environ.* 74 (2015) 337–355, <https://doi.org/10.1007/s10064-014-0616-6>.
- [83] H.A. Al-Najjar, B. Pradhan, G. Beydoun, R. Sarkar, H.J. Park, A. Alamri, A novel method using explainable artificial intelligence (XAI)-based Shapley Additive Explanations for spatial landslide prediction using Time-Series SAR dataset, *Gondwana Res.* (2022), <https://doi.org/10.1016/j.gr.2022.08.004>.
- [84] N. Manchar, C. Benabbas, R. Hadji, F. Bouaicha, F. Grecu, Landslide susceptibility assessment in Constantine region (NE Algeria) by means of statistical models, *Studia Geotechnica Mech.* 40 (2018) 208–219, <https://doi.org/10.2478/sgem-2018-0024>.
- [85] D.K. Rai, D. Xiong, W. Zhao, D. Zhao, B. Zhang, N.M. Dahal, Y. Wu, M.A. Baig, An investigation of landslide susceptibility using logistic regression and statistical index methods in dailekh district, Nepal, *Chin. Geogr. Sci.* 32 (2022) 834–851, <https://doi.org/10.1007/s11769-022-1304-2>.
- [86] M. Liao, H. Wen, L. Yang, Identifying the essential conditioning factors of landslide susceptibility models under different grid resolutions using hybrid machine learning: a case of Wushan and Wuxi counties, China, *Catena* 217 (2022), 106428, <https://doi.org/10.1016/j.catena.2022.106428>.
- [87] W. Zhang, S. Liu, L. Wang, P. Samui, M. Chwala, Y. He, Landslide susceptibility research combining qualitative analysis and quantitative evaluation: a case study of yunyang county in chongqing, China, *Forests* 13 (2022) 1055, <https://doi.org/10.3390/f13071055>.
- [88] P. Saha, R. Mitra, K. Chakraborty, M. Roy, Application of multi layer perceptron neural network Markov chain model for LULC change detection in the Sub-Himalayan North Bengal, *Remote Sens. Appl.: Society and Environment* 26 (2022), 100730, <https://doi.org/10.1016/j.rsase.2022.100730>.
- [89] A.A. Apostolov, L.A. Yelistratova, I.F. Romanciu, I.U.L. Zakharchuk, Identifying potential landslide areas by employing the erosion relief index and meteorological criteria in Ukraine, *Rev. Roum. Géogr./Rom. Journ. Geogr.* 65 (2021) 125–141.
- [90] N. Eitvandi, R. Sarikhani, S. Derikvand, Landslide susceptibility mapping by integrating analytical hierarchy process, frequency ratio, and fuzzy gamma operator models, case study: north of Lorestan Province, Iran, *Environ. Monit. Assess.* 194 (2022) 1–26, <https://doi.org/10.1007/s10661-022-10206-5>.
- [91] A. Ávila, F. Justino, A. Wilson, D. Bromwich, M. Amorim, Recent precipitation trends, flash floods and landslides in southern Brazil, *Environ. Res. Lett.* 11 (2016), 114029.
- [92] M. Xia, G.M. Ren, X.L. Ma, Deformation and mechanism of landslide influenced by the effects of reservoir water and rainfall, Three Gorges, China, *Nat. Hazards* 68 (2013) 467–482, <https://doi.org/10.1007/s11069-013-0634-x>.
- [93] E. Gregori, M.C. Andrenelli, G. Zorn, Assessment and classification of climatic aggressiveness with regard to slope instability phenomena connected to hydrological and morphological processes, *J. Hydrol.* 329 (2006) 489–499, <https://doi.org/10.1016/j.jhydrol.2006.03.001>.
- [94] T. Lukić, D. Bjelajac, K.E. Fitzsimmons, S.B. Markovic, B. Basarin, D. Mladan, T. Micić, R.J. Schaeztl, M.B. Gavrilov, M. Milanović, G. Sipos, Factors triggering landslide occurrence on the Zemun loess plateau, Belgrade area, Serbia, *Environ. Earth Sci.* 77 (2018) 1–15, <https://doi.org/10.1007/s12665-018-7712-z>.
- [95] S. Neamat, H. Karimi, A systematic review of GIS-based landslide hazard mapping on determinant factors from international databases, in: 2020 International Conference on Advanced Science and Engineering (ICOASE), 2020. December, pp. 180–183, <https://doi.org/10.1109/ICOASE51841.2020.9436611>.
- [96] T. Xiao, S. Segoni, L. Chen, K. Yin, N. Casagli, A step beyond landslide susceptibility maps: a simple method to investigate and explain the different outcomes obtained by different approaches, *Landslides* 17 (2020) 627–640, <https://doi.org/10.1007/s10346-019-01299-0>.
- [97] T. Watakabe, Y. Matsushi, Lithological controls on hydrological processes that trigger shallow landslides: observations from granite and hornfels hillslopes in Hiroshima, Japan, *Catena* 180 (2019) 55–68, <https://doi.org/10.1016/j.catena.2019.04.010>.
- [98] B. Aslam, A. Zafar, U. Khalil, Comparative analysis of multiple conventional neural networks for landslide susceptibility mapping, *Nat. Hazards* (2022) 1–35, <https://doi.org/10.1007/s11069-022-05570-x>.
- [99] L.L. Liu, Y.L. Zhang, S.H. Zhang, B. Shu, T. Xiao, Machine learning with a susceptibility index-based sampling strategy for landslide susceptibility assessment, *Geocarto Int.* 1–31 (2022), <https://doi.org/10.1080/10106049.2022.2102221>.
- [100] G. Yin, J. Luo, F. Niu, Z. Lin, M. Liu, Machine learning-based thermokarst landslide susceptibility modeling across the permafrost region on the Qinghai-Tibet Plateau, *Landslides* 18 (2021) 2639–2649, <https://doi.org/10.1007/s10346-021-01669-7>.
- [101] J. Miles, *Tolerance and Variance Inflation Factor*, Wiley statsref: statistics reference online, 2014.
- [102] A. Aditiana, T. Kubota, Y. Shinohara, Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia, *Geomorphology* 318 (2018) 101–111, <https://doi.org/10.1016/j.geomorph.2018.06.006>.
- [103] A. Arabameri, K. Rezaei, A. Cerda, C. Conoscenti, Z. Kalantari, A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran, *Sci. Total Environ.* 660 (2019) 443–458, <https://doi.org/10.1016/j.scitotenv.2019.01.021>.
- [104] Z. Umar, B. Pradhan, A. Ahmad, M.N. Jebur, M.S. Tehrani, Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia, *Catena* 118 (2014) 124–135, <https://doi.org/10.1016/j.catena.2014.02.005>.
- [105] I. Yilmaz, Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey), *Comput. Geosci.* 35 (2009) 1125–1138, <https://doi.org/10.1016/j.cageo.2008.08.007>.
- [106] H.R. Pourghasemi, B. Pradhan, C. Gokceoglu, Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran, *Nat. Hazards* 63 (2012) 965–996, <https://doi.org/10.1007/s11069-012-0217-2>.
- [107] A. Wubalem, B. Getahun, Y. Hailemariam, A. Mesele, G. Tesfaw, Z. Dawit, E. Goshe, Landslide susceptibility modeling using the index of entropy and frequency ratio method from nefas-mewcha to weldiya road corridor, northwestern Ethiopia, *Geotech. Geol. Eng.* 40 (2022) 5249–5278, <https://doi.org/10.1007/s10706-022-02214-6>.
- [108] A.M. Youssef, M. Al-Kathery, B. Pradhan, Landslide susceptibility mapping at Al-Hasher area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models, *Geosci. J.* 19 (2015) 113–134, <https://doi.org/10.1007/s12303-014-0032-8>.
- [109] K.C. Devkota, A.D. Regmi, H.R. Pourghasemi, K. Yoshida, B. Pradhan, I.C. Ryu, M.R. Dhital, O.F. Althuwaynee, 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, *Nat. Hazards* 65 (2013) 135–165.
- [110] Y. Liu, A. Yuan, Z. Bai, J. Zhu, GIS-based landslide susceptibility mapping using frequency ratio and index of entropy models for She County of Anhui Province, China, *Appl. Rheol.* 32 (2022) 22–33, <https://doi.org/10.1515/arih-2022-0122>.
- [111] C.J. Van Westen, Geo-information tools for landslide risk assessment: an overview of recent developments, *Landslides: evaluation and stabilization* 1 (2004) 39–56.
- [112] J. Vicko, P. Wagner, Z. Rychlikova, Evaluation of regional slope stability, *Miner. Slovaca* 12 (3) (1980) 275–283.
- [113] C.J. Van Westen, *Statistical Landslide Hazards Analysis, ILWIS 2.1 for Windows Application Guide*, ITC Publication, 1997.
- [114] S. Akter, S.A. Javed, GIS-based assessment of landslide susceptibility and inventory mapping using different bivariate models, *Geocarto Int.* (2022) 1–31, <https://doi.org/10.1080/10106049.2022.2076907> (just-accepted).
- [115] H. Bourenane, A.A. Meziani, D.A. Benamar, Application of GIS-based statistical modeling for landslide susceptibility mapping in the city of Azazga, Northern Algeria, *Bull. Eng. Geol. Environ.* 80 (2021) 7333–7359, <https://doi.org/10.1007/s10064-021-02386-0>.
- [116] M. Shafapour Tehrani, L. Kumar, M. Neamah Jebur, F. Shabani, Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods, *Geomatics, Nat. Hazards Risk* 10 (2019) 79–101, <https://doi.org/10.1080/19475705.2018.1506509>.
- [117] G. Zhang, Y. Cai, Z. Zheng, J. Zhen, Y. Liu, K. Huang, Integration of the statistical index method and the analytic hierarchy process technique for the assessment of landslide susceptibility in Huizhou, China, *Catena* 142 (2016) 233–244, <https://doi.org/10.1016/j.catena.2016.03.028>.
- [118] S. Farooq, M.S. Akram, Landslide susceptibility mapping using information value method in Jhelum Valley of the Himalayas, *Arabian J. Geosci.* 14 (2021) 1–16, <https://doi.org/10.1007/s12517-021-07147-7>.
- [119] Q. Wang, D. Wang, Y. Huang, Z. Wang, L. Zhang, Q. Guo, W. Chen, W. Chen, M. Sang, Landslide susceptibility mapping based on selected optimal combination of landslide predisposing factors in a large catchment, *Sustainability* 7 (2015) 16653–16669, <https://doi.org/10.3390/su71215839>.

- [120] Q. Ba, Y. Chen, S. Deng, J. Yang, H. Li, A comparison of slope units and grid cells as mapping units for landslide susceptibility assessment, *Earth Science Informatics* 11 (2018) 373–388, <https://doi.org/10.1007/s12145-018-0335-9>.
- [121] M. Klose, D. Gruber, B. Damm, G. Gerold, Spatial databases and GIS as tools for regional landslide susceptibility modeling, *Z. Geomorphol.* 58 (2014) 1–36.
- [122] O.F. Althuwaynee, B. Pradhan, H.J. Park, J.H. Lee, A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping, *Catena* 114 (2014) 21–36, <https://doi.org/10.1016/j.catena.2013.10.011>.
- [123] M. Mohammady, H.R. Pourghasemi, B. Pradhan, Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster–Shafer, and weights-of-evidence models, *J. Asian Earth Sci.* 61 (2012) 221–236, <https://doi.org/10.1016/j.jseae.2012.10.005>.
- [124] H. Nampak, B. Pradhan, M. Abd Manap, Application of GIS based data driven evidential belief function model to predict groundwater potential zonation, *J. Hydrol.* 513 (2014) 283–300, <https://doi.org/10.1016/j.jhydrol.2014.02.053>.
- [125] N.W. Park, Application of Dempster-Shafer theory of evidence to GIS-based landslide susceptibility analysis, *Environ. Earth Sci.* 62 (2011) 367–376, <https://doi.org/10.1007/s12665-010-0531-5>.
- [126] M. Daviran, M. Shamekhi, R. Ghezelbash, A. Maghsoudi, Landslide susceptibility prediction using artificial neural networks, SVMs and random forest: hyperparameters tuning by genetic optimization algorithm, *Int. J. Environ. Sci. Technol.* (2022) 1–18, <https://doi.org/10.1007/s13762-022-04491-3>.
- [127] U. Khalil, I. Imtiaz, B. Aslam, I. Ullah, A. Tariq, S. Qin, Comparative analysis of machine learning and multi-criteria decision making techniques for landslide susceptibility mapping of Muzaffarabad district, *Front. Environ. Sci.* 10 (2022) 1–19, <https://doi.org/10.3389/fenvs.2022.1028373>.
- [128] S. Nikoobakht, M. Azarafza, H. Akgun, R. Derakhshani, Landslide susceptibility assessment by using convolutional neural network, *Appl. Sci.* 12 (2022) 5992, <https://doi.org/10.3390/app12125992>.
- [129] H. Hong, J. Liu, A.X. Zhu, Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble, *Sci. Total Environ.* 718 (2020), 137231, <https://doi.org/10.1016/j.scitotenv.2020.137231>.
- [130] R. Mitra, D. Roy, Delineation of groundwater potential zones through the integration of remote sensing, geographic information system, and multi-criteria decision-making technique in the sub-Himalayan foothills region, India, *International Journal of Energy and Water Resources* (2022) 1–21, <https://doi.org/10.1007/s42108-022-00181-5>.
- [131] T. Kavzoglu, E. Kutlug Sahin, I. Colkesen, An assessment of multivariate and bivariate approaches in landslide susceptibility mapping: a case study of Duzkoy district, *Nat. Hazards* 76 (2015) 471–496, <https://doi.org/10.1007/s11069-014-1506-8>.
- [132] J. Dou, A.P. Yunus, Y. Xu, Z. Zhu, C.W. Chen, M. Sahana, K. Khosravi, Y. Yang, B.T. Pham, Torrential rainfall-triggered shallow landslide characteristics and susceptibility assessment using ensemble data-driven models in the Dongjiang Reservoir Watershed, China, *Nat. Hazards* 97 (2019) 579–609, <https://doi.org/10.1007/s11069-019-03659-4>.
- [133] S. Mondal, S. Mandal, Landslide susceptibility mapping of Darjeeling Himalaya, India using index of entropy (IOE) model, *Applied Geomatics* 11 (2019) 129–146, <https://doi.org/10.1007/s12518-018-0248-9>.
- [134] S. Ghosh, E.J.M. Carranza, C.J. Van Westen, V.G. Jetten, D.N. Bhattacharya, Selecting and weighting spatial predictors for empirical modeling of landslide susceptibility in the darjeeling himalayas (India), *Geomorphology* 131 (2011) 35–56, <https://doi.org/10.1016/j.geomorph.2011.04.019>.
- [135] J. Roy, S. Saha, Landslide susceptibility mapping using knowledge driven statistical models in Darjeeling District, West Bengal, India, *Geoenvironmental Disasters* 6 (2019) 1–18, <https://doi.org/10.1186/s40677-019-0126-8>.