Geoscience Frontiers 12 (2021) 857–876

HOSTED BY

Contents lists available at ScienceDirect

Geoscience Frontiers

journal homepage: www.elsevier.com/locate/gsf

Research Paper

GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms

Sk Ajim Ali [a], Farhana Parvin [a], Jana Vojteková [b], Romulus Costache [c,d], Nguyen Thi Thuy Linh [e], Quoc Bao Pham [f,g,*], Matej Vojtek [b], Ljubomir Gigović [h], Ateeque Ahmad [a], Mohammad Ali Ghorbani [i]

[a] Department of Geography, Faculty of Science, Aligarh Muslim University (AMU), Aligarh, UP 202002, India
[b] Department of Geography and Regional Development, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, Trieda A. Hlinku 1, 94901 Nitra, Slovakia
[c] Research Institute of the University of Bucharest, 90-92 Sos. Panduri, 5th District, Bucharest 050663, Romania
[d] National Institute of Hydrology and Water Management, Bucureşti-Ploieşti Road, 97E, 1st District, Bucharest 013686, Romania
[e] Thuyloi University, 175 Tay Son, Dong Da, Hanoi, Viet Nam
[f] Institute of Research and Development, Duy Tan University, Danang 550000, Viet Nam
[g] Faculty of Environmental and Chemical Engineering, Duy Tan University, Danang 550000, Viet Nam
[h] Department of Geography, University of Defence, 11000 Belgrade, Serbia
[i] Sustainable Management of Natural Resources and Environment Research Group, Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Viet Nam

ARTICLE INFO

Article history:
Received 6 May 2020
Received in revised form 30 June 2020
Accepted 1 September 2020
Available online 10 October 2020

Handling Editor: E. Shaji

Keywords:
Landslide susceptibility modeling
Geographic information system
Fuzzy DEMATEL
Analytic network process
Naïve Bayes classifier
Random forest classifier

ABSTRACT

Hazards and disasters have always negative impacts on the way of life. Landslide is an overwhelming natural as well as man-made disaster that causes loss of natural resources and human properties throughout the world. The present study aimed to assess and compare the prediction efficiency of different models in landslide susceptibility in the Kysuca river basin, Slovakia. In this regard, the fuzzy decision-making trial and evaluation laboratory combining with the analytic network process (FDEMATEL-ANP), Naïve Bayes (NB) classifier, and random forest (RF) classifier were considered. Initially, a landslide inventory map was produced with 2000 landslide and non-landslide points by randomly divided with a ratio of 70%:30% for training and testing, respectively. The geospatial database for assessing the landslide susceptibility was generated with the help of 16 landslide conditioning factors by allowing for topographical, hydrological, lithological, and land cover factors. The ReliefF method was considered for determining the significance of selected conditioning factors and inclusion in the model building. Consequently, the landslide susceptibility maps (LSMs) were generated using the FDEMATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier models. Finally, the area under curve (AUC) and different arithmetic evaluation were used for validating and comparing the results and models. The results revealed that random forest (RF) classifier is a promising and optimum model for landslide susceptibility in the study area with a very high value of area under curve (AUC = 0.954), lower value of mean absolute error (MAE = 0.1238) and root mean square error (RMSE = 0.2555), and higher value of Kappa index (K = 0.8435) and overall accuracy (OAC = 92.2%).

1. Introduction

Landslide is a very frequent geological disaster, which especially occurs in the hilly region that causes severe damage to the landscape and human lives (Vakhshoori et al., 2019; Xiao et al., 2019). Landslide is a general term, which is used to define the downslope movement of rock, soil, or accumulated debris (Dang et al., 2020). It occurs as a result of different associated natural hazards or anthropogenic activities like an earthquake, flash flood, road construction, deforestation, and mineral exploration (Sarker and Rashid, 2013; Dang et al., 2018; Saito et al., 2018; Hussain et al., 2019; Shao, 2019). The global statistics on landslide damage revealed that there were more than 3876 landslides from 1995 to 2014, in which 11,689 injuries and 163,658 deaths were reported (Haque et al., 2019). Research findings recorded that only in the year 2014, about 174 landslides occurred worldwide resulting in devastating consequences on human as well as natural resources (Nsengiyumva et al., 2019). Another research also revealed that due to landslides, more than 55,000 lives lost worldwide from 2004 to 2016 (Froude and Petley, 2018).

As a result, landslides have become a major threat to human lives as well as natural properties despite the fact that a number of studies on landslide modeling were published since last decades (Ramani et al., 2011; Sujatha et al., 2012; Sarker and Rashid, 2013; Pham et al., 2016a; Nsengiyumva et al., 2018; Tien Bui et al., 2019). Although the

* Corresponding author at: Institute of Research and Development, Duy Tan University, Danang 550000, Viet Nam.
E-mail address: phambaoquoc@duytan.edu.vn (Q.B. Pham).

https://doi.org/10.1016/j.gsf.2020.09.004

research on landslide susceptibility is unable to totally stop the hazardous phenomena, it can at least help in reducing the extent of impacts by incorporating the susceptibility maps into policy-making and planning the development of a region (Iovine and Sheridan, 2009; Choi and Cheung, 2013; Iovine and Cohen, 2014; Cieslik et al., 2019).

The study area, which is represented by the Kysuca river basin, belongs to highly affected regions by natural hazards and disasters, especially by floods and landslides. Slope deformations and landslides are common geodynamic events in Slovakia (Barančoková and Kenderessy, 2014). 21,190 slope deformations were recorded in Slovakia until 2008 (Kopecký et al., 2008). In 2010, more than 550 new and numerous reactivated landslides were reported in Slovakia (Liščák and Káčer, 2013). Overall, more than 15,000 potential landslides were registered in Slovakia accounting for 63% of the total landslides (Barančoková and Barančok, 2019).

Therefore, landslides cannot be completely prevented, but accurate prediction of the landslide susceptible areas can save natural resources as well as human lives. Landslide susceptibility analysis reveals the spatial distribution of potential landslide areas based on different conditioning factors (Nsengiyumva et al., 2018). Landslides are occurred and controlled by one or more conditioning factors, like topography, slope, failure mechanism, intensity of rainfall, land cover, geological formation, strength of rocks, and many more (Schilirò et al., 2016; Pisano et al., 2017; Zêzere et al., 2017).

Based on the literature, various approaches and methods have been applied in landslide susceptibility mapping including inventory-based approaches (Van Den Eeckhaut et al., 2005; Akgun, 2012), bivariate and multivariate statistics like logistic regression (LR), frequency ratio (FR); statistical index, cluster analysis, artificial neural networks (ANNs) (Dahal et al., 2008; Yilmaz, 2009; Ramani et al., 2011; Mohammady et al., 2012; Sujatha et al., 2012; Buša et al., 2019), expert's knowledge-based methods like analytic hierarchy process (AHP), analytic network process (ANP), fuzzy logic, Boolean logic (Neaupane and Piantanakulchai, 2006; Othman and Naim, 2012; Kayastha et al., 2013; Leonardi et al., 2016; Vojteková and Vojtek, 2020), probabilistic methods (Zêzere et al., 2004; Brenning, 2005), data mining and machine learning (ML) methods (Micheletti et al., 2014; Lee et al., 2017; Zêzere et al., 2017; Kavzoglu et al., 2019).

Recently, the application of machine learning (ML) algorithms in landside modeling get much more popularity and attract the attention of researchers because of their capability to handle multi-governing factors (Goetz et al., 2015a; Colkesen et al., 2016; Chen et al., 2018a). To predict and model landslides, various machine learning technique, like Bayesian tree (BT), Decision tree (DT), Alternating decision tree (ADT), Naïve Bayes (NB), random forest (RF) can be applied and compared for choosing the most accurate method (Pham et al., 2016b; Ada and San, 2018). Along with machine learning algorithms, various statistical techniques, such as weights-of-evidence or generalized additive models and different methods of multi-criteria decision analysis (MCDA), such as, Vlse Kriterijumska Optimizacija I Kompromisno Resenje in Serbian (VIKOR), Simple additive weighting (SAW) or fuzzy set, can also be applied and compared for accurate susceptibility mapping. However, recent studies showed higher accuracy and power of machine learning algorithms in comparison to conventional statistical methods or MCDM approaches for susceptibility modeling (Goetz et al., 2015b; Khosravi et al., 2019). Moreover, ensemble machine learning methods, such as Alternating Decision tree (ADT), Bagging, MultiBoost, Rotation subspace or Rotation Forest, have got more preferences in comparison to single machine learning technique because ensemble learning helps in generating hybrid models which are more efficient to handle complex input data and generally offer more accurate result (Dang et al., 2018; Hong et al., 2018; Pham et al., 2018a, 2018b, 2018c).

The above-mentioned literatures disclosed that quite a lot of studies on landslide susceptibility modeling (LSM) were performed worldwide using different MCDM approaches, single and ensemble machine learning algorithms. However, a comparative assessment of MCDM and machine learning has not been carried out for LSM in Slovakia. Therefore, the current study aimed to make a comparative analysis between fuzzy multi-criteria methods and machine learning algorithms. The current study was planned in the following sections: a geo-physical narrative of the study area and landslide inventory map (LIM) is provided in the second section, the conditioning factors of landslide are reviewed in the third section, an insight of the applied modeling approaches, i.e. fuzzy decision-making trial and evaluation laboratory (FDEMATEL), analytic network process (ANP), Naïve Bayes (NB) classifier, and random forest (RF) classifier, are described in the next section. The fifth section gives the performance evaluation of the applied models. The sixth section highlights the comparative results of these models, which is followed by the discussion section and conclusions.

2. Materials and methods

2.1. Study area

The present study area is represented by the Slovak part of the Kysuca river basin. The total area of the Kysuca river basin, which also extends into Czechia and Poland, has an area of 1053 km^2. The studied Slovak part of the basin covers 989 km^2. The Kysuca river has a length of 65.5 km and it creates a right-sided tributary of the Váh river. The spring of the Kysuca river is located on the northern slopes of the Hričovec peak (1062 m).

The geographical coordinates of the study area are the following: north most point is 49°30′35″ N and 18°35′42″ E; south most point is 49°14′11″ N and 18°52′17″ E; west most point is 49°21′24″ N and 18°23′51″ E; east most point is 49°19′40″ N and 19°9′41″ E (Fig. 1).

From the viewpoint of the geomorphological division, the following geomorphological units are located in the Kysuca river basin (Mazúr and Lukniš, 1986): Žilinská kotlina (basin), Jablunkovské medzihorie (foothills), Kysucká vrchovina (highland), Podbeskydská vrchovina (highland), Turzovská vrchovina (highland), Moravsko-sliezske Beskydy (mountain), Javorníky (mountain), Kysucké Beskydy (mountain), and Oravské Beskydy (mountain). The highest elevation point which is represented by the Rycierova hora peak (1226 m) is situated in the eastern part. The lowest elevation point (326 m) is situated at the confluence of the rivers Kysuca and Váh.

The average annual rainfall varies between 800 mm, referring to lower elevations of the basin, and 1500 mm, which regards the highest mountainous areas (Bochníček et al., 2015). The average annual discharge at the Kysucké Nové Mesto gauging station, which is located at the 8.0 river kilometer, is 16.6 m^3/s (Ministry of Environment of the Slovak Republic, 2018). The left-side tributaries are Bystrica river, Predmieranka stream or Čierňanka stream. The right-sided tributaries include, for example, Raková stream or Rudinský potok stream.

Based on the administrative and NUTS division, the study area belongs to NUTS II Central Slovakia and NUTS III Žilina Region.

2.2. Landslide inventory map (LIM)

The preparation of inventory map is a significant aspect for landslide susceptibility assessment. The landslide inventory map helps to understand the relationship between the distributions of historical landslides and selected conditioning factors (Youssef et al., 2016; Chen et al., 2018a). The location of past landslides can be obtained through different sources like GNSS-based field survey, high-resolution satellite images, Google Earth images, previous projects and reports, atlases, and the like (Galli et al., 2008; Borrelli et al., 2015; Shirzadi et al., 2017; Tien Bui et al., 2018). Therefore, the evidence which is obtained from past landslides is known as an inventory map and these are mainly location of existing landslides. This is a fundamental part of landslide susceptibility assessment to acquire information from past landslides.

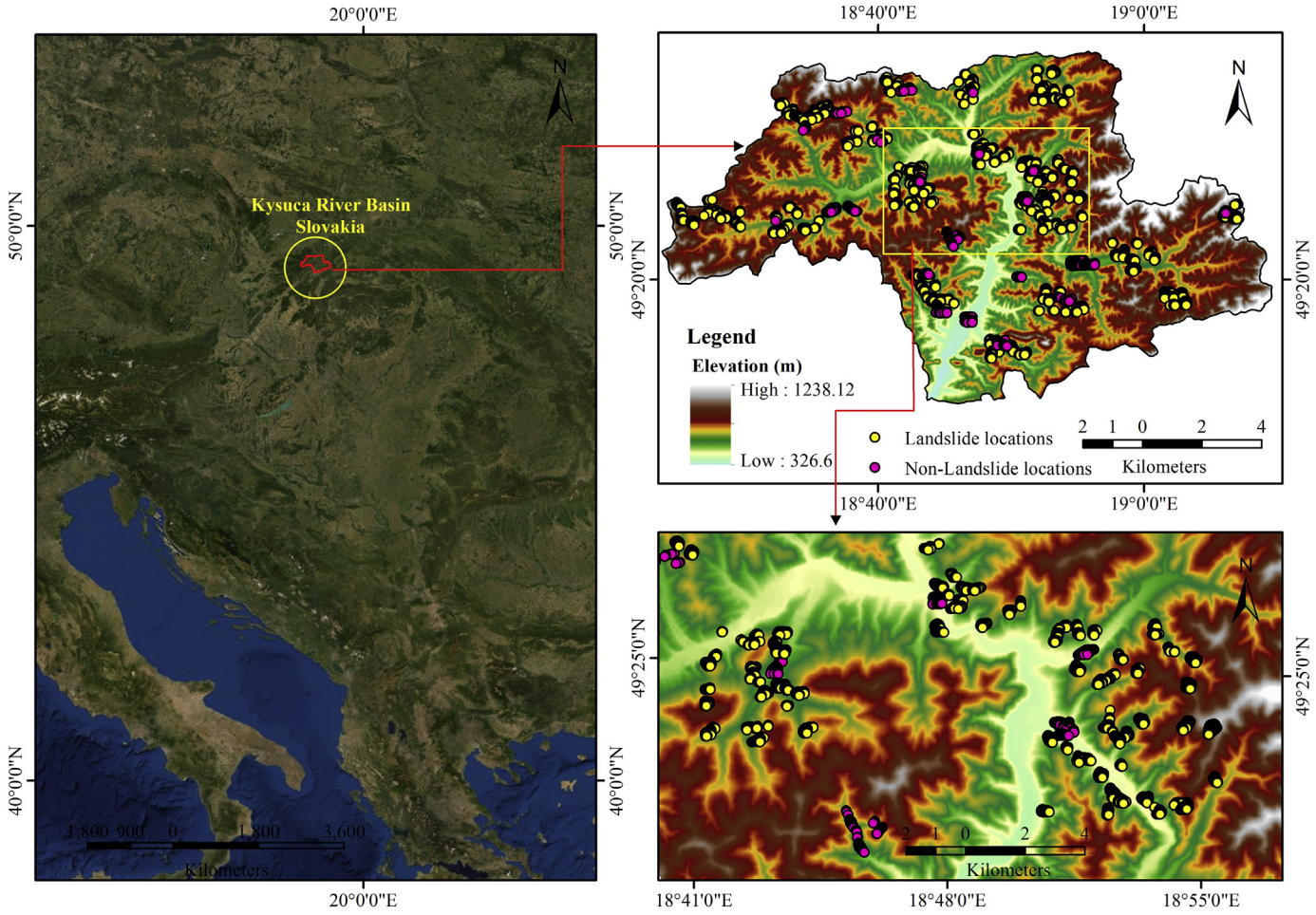

Fig. 1. Location map of the study area.

In a first stage, in order to create the landslide inventory map, many polygons representing the areal extension of the landslides, were extracted from the Atlas of Slope Stability Maps of the Slovak Republic at a scale of 1:50,000 (Šimeková et al., 2006). Further, in order to be included in the GIS workflow, the collected polygons were converted in raster with the same cell size as the Digital Elevation Model, and then the raster was converted into points. Thus, each cell of the raster corresponds to a point. Finally, a total of 1000 points, which represent the past landslide areas locations in the Kysuca Basin, were used in the present analysis. In this study, the training dataset included both landslide and non-landslide locations. The non-landslide locations were randomly selected points based on the condition to have a slope angle lower than 3° and its number is the same as landslide locations. Thus, a total of 2000 georeferenced points containing both landslides and non-landslides locations were used to train and validate the landslide susceptibility map. The input training dataset was spilt into two groups, particularly, 70% for training and 30% for validation of the result (Fig. 2). Many recent studies evidenced the use of landslide and non-landslide points to prepare the input dataset for landslide susceptibility assessment (Razavizadeh et al., 2017; Nsengiyumva et al., 2019; Vakhshoori et al., 2019; Xiao et al., 2019).

2.3. Landslide conditioning factors

There are not any exact or accepted principles for selecting the conditioning factors because these factors depend on the availability of source of data, geo-environmental settings of the study region, as well as the nature of landslides occurred in that area (Vakhshoori et al., 2019). Hence, looking towards the nature of the present study, 16 landslide conditioning factors together with slope (°), aspect, elevation (m), general curvature, plan curvature, profile curvature, sediment transport index (STI), rainfall (average annual precipitation total in mm), annual solar radiation (WH/m^2), stream power index (SPI), topographic wetness index (TWI), distance to rivers (m), lithology, distance to faults (m), land cover, distance to roads (m) were taken into consideration. All the thematic layers were processed in GIS using the WGS_1984_UTM_Zone_34N projection system with a 20 m spatial resolution.

Slope, aspect, elevation, and curvature were extracted from digital elevation model (DEM) of the study area (Fig. 3a–f). Slope is an important landslide conditioning factors which were used in different studies (Akgun and Türk, 2010; Chen et al., 2017b). Elevation is another important factor that is commonly used in landslide susceptibility studies. It is the general altitude of land above the mean sea level. It is believed that higher elevation is more prone to the landslides in comparison to lower elevation (Devkota et al., 2013). The rocks have more strength at higher elevations than in lower or foothills (Chen et al., 2018b). The organic matters, namely carbon and nitrogen in soil, at high elevation are causes of dehydrogenases activity, which has positive effect on landside (Błońska et al., 2018). Aspect is the direction of slope (Moosavi and Niazi, 2016). The correlation between the slope aspect and landslide occurrences was found in different studies (Akgun et al., 2012; Colkesen et al., 2016). The degree of weathering and moisture content differs with different slope aspects (Shirzadi et al., 2018). Curvature is also responsible for landslide occurrence which was evidenced in different studies (Nefeslioglu et al., 2008a; Pham et al., 2015). General curvature indicates the shape of the ground surface which has a great impact on

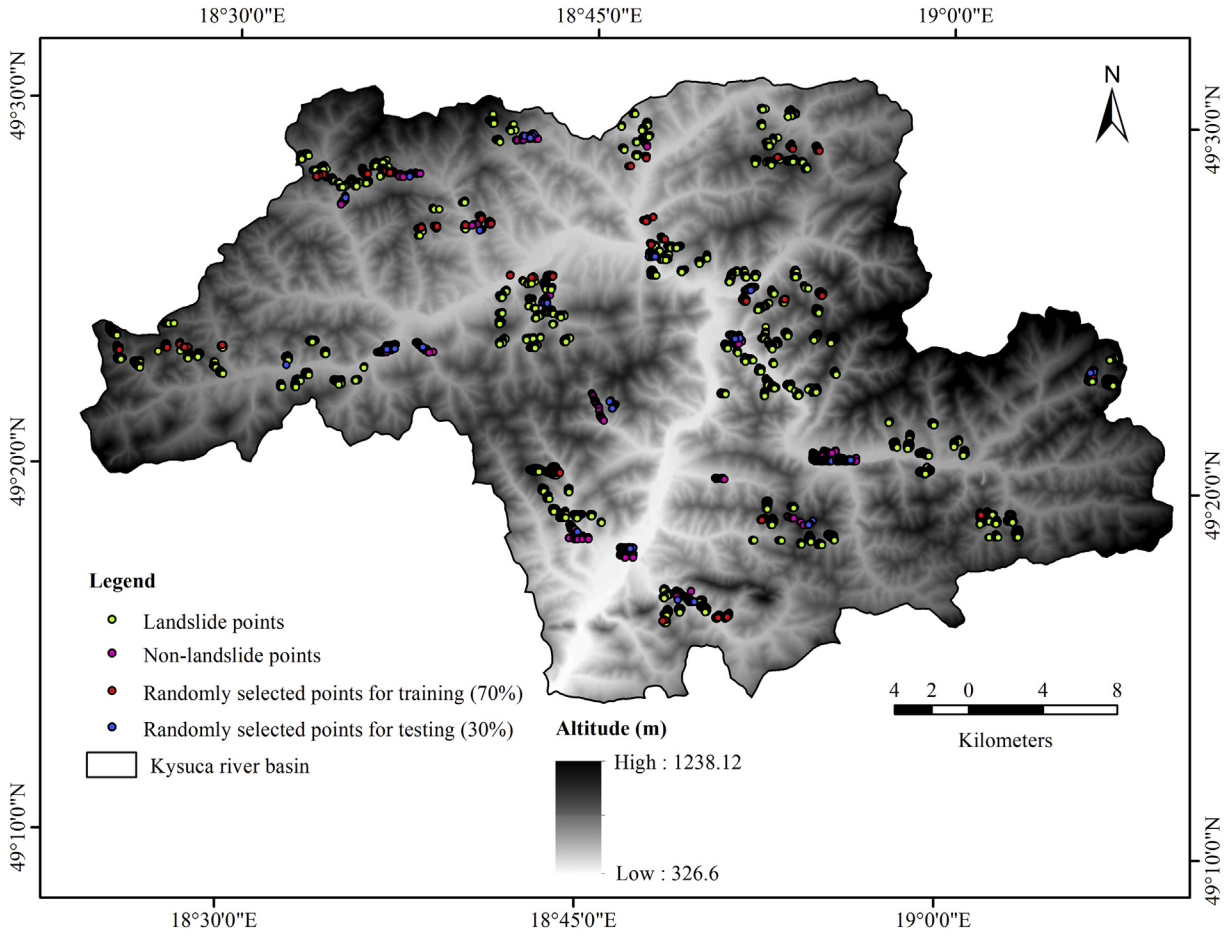

Fig. 2. Landslide inventory map.

landslides. Plan curvature indicates the steepness of slope which is formed due to the intersection of the horizontal plane and surface landforms (Chen et al., 2018c). Consequently, the profile curvature denotes the vertical plane parallel to slope direction (Yilmaz et al., 2012; Ding et al., 2017). These three curvature factors interlink with each other controlling the acceleration and deceleration of surface flow and thus affecting the process of the landsliding (Moosavi and Niazi, 2016). The stream power index (SPI) and sediment transport index (STI) are also important factors that indicate the erosion power of overland flow and sediment transported by overland flow (Moore et al., 1991; Moore and Wilson, 1992; Pradhan and Kim, 2014; Nefeslioglu et al., 2008b). To determine the SPI and STI, the slope gradient (β) and slope contributing area (A_s) were considered using the DEM (Gritzner et al., 2001; Yilmaz, 2009; Pourghasemi et al., 2012). These were calculated using Eqs. (1) and (2) (Figs. 3g and 4b).

$$SPI = A_s \times \tan\beta \tag{1}$$

$$STI = \left(\frac{A_s}{22.13}\right)^{0.6}\left(\frac{\sin\beta}{0.0896}\right)^{1.3} \tag{2}$$

Rainfall plays a vital role in landslide susceptibility mapping. Thus, the annual rainfall in the study area was considered for susceptibility assessment of landslide, and the study area was divided into five categories ranging from 800 to 1500 mm (Fig. 3h). The high rainfall areas generally depict high susceptibility to the landslide.

Annual solar radiation (WH/m^2) is another important landslide conditioning factor which was also considered in the present study. Annual solar radiation is expressed in terms of mean solar radiation of a year converged at a particular pixel (Brown, 2012). Higher annual solar radiation is the indication of a lower probability of landsliding because of more available pore space of rock and soil (Fig. 4a). The 'Area Solar Radiation' function in ArcGIS was used to calculate the annual solar radiation based on aspect and slope. Topographic wetness index (TWI) is another vital landslide conditioning factor (Fig. 4c). TWI was computed using Eq. (3).

$$TWI = In\left(\frac{A_s}{\tan\beta}\right) \tag{3}$$

where A_s and β are the slope contributing area and slope gradient respectively.

The landslide susceptibility also depends on the lithological formation which determines the geotechnical characteristics of rocks. The rocks having low permeability properties have high potential of landsliding (Shirzadi et al., 2018; Pham et al., 2018d). According to the lithological structure, the study area was classified into seven categories, namely claystone-limestone rocks, conglomerate-sandstone rocks, deluvial deposits, floodplain deposits, flyschoid rocks, limestone-dolomite rocks, and river terrace deposits (Fig. 5a). Different land cover types have different effect on landsliding (Moosavi and Niazi, 2016; Ding et al., 2017; Shirzadi et al., 2018). The land cover map was represented by the vector layer of CORINE Land Cover from 2018. According to the state-of-the-art literature related to the landslide susceptibility assessment across medium or large size study areas/river basins (with surfaces higher than 500 km^2), it was concluded that the CORINE Land Cover database is almost unanimously used as data source for land use/land cover conditioning factor (Myronidis et al., 2016; Di Traglia

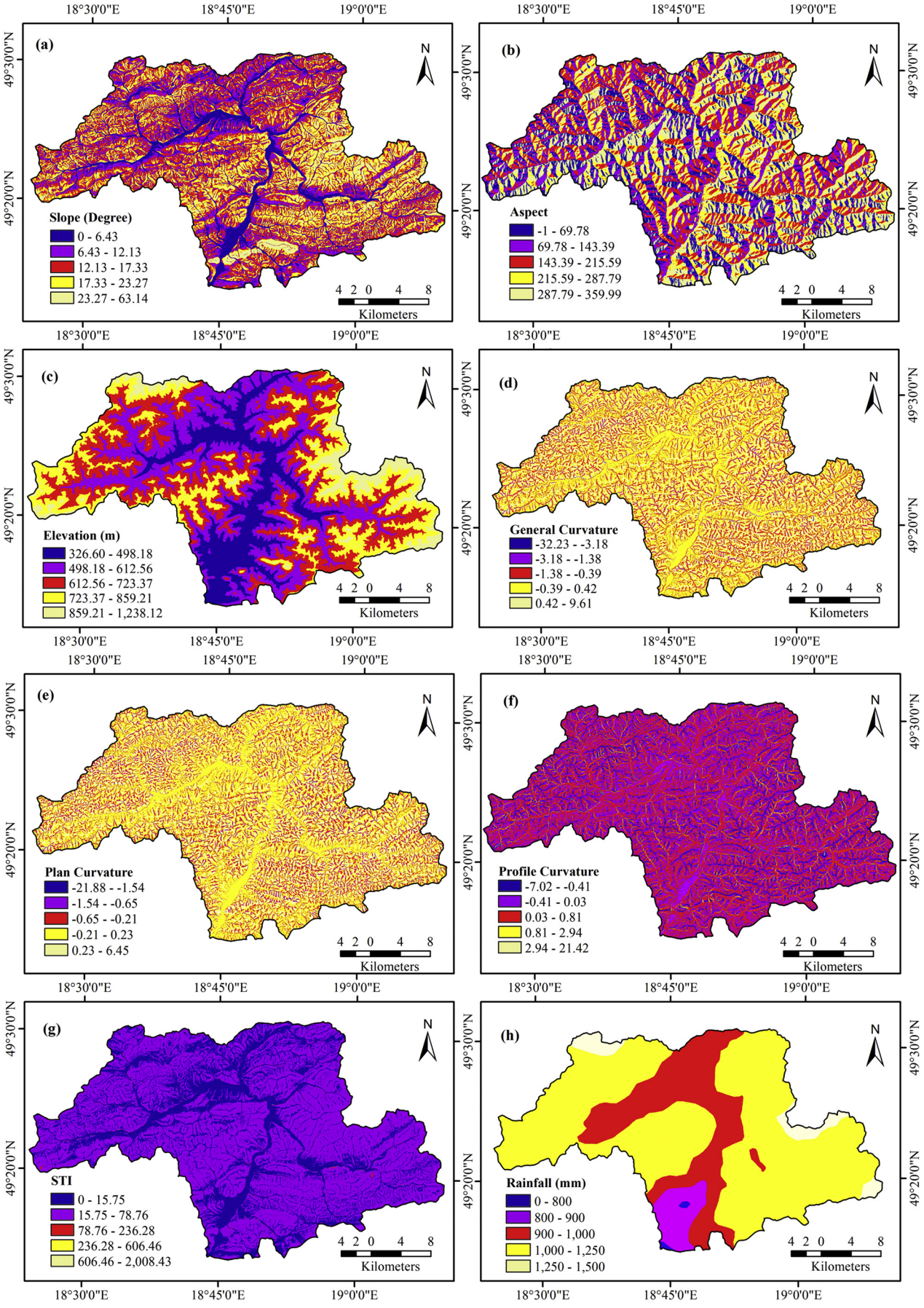

Fig. 3. Landslide conditioning factors: (a) slope; (b) aspect; (c) elevation; (d) general curvature; (e) plan curvature; (f) profile curvature; (g) sediment transport index (STI); (h) Rainfall (average annual precipitation total in mm).

et al., 2018; Meneses et al., 2019; Tevi and Stoica, 2019; Kavoura and Sabatakakis, 2020; Segoni et al., 2020). The land cover predictor across the study area includes the following 15 classes: discontinuous urban fabric, industrial or commercial units, road and rail networks and associated land, sport and leisure facilities, non-irrigated arable land, pastures, annual crops associated with the permanent crops, complex cultivation patterns, land principally occupied by agriculture, broad-leaved forest, coniferous forest, mixed forest, natural grasslands, transitional woodland-shrub, water bodies (Fig. 5c).

The distance to specific features like a fault, river or road also plays an important role in landslide susceptibility (Althuwaynee et al., 2016). In this study, distance to faults, river, and roads were also considered (Figs. 4d, 5b and d). Several studies proved that landslides were observed close to faults, roads, and rivers (Poudyal et al., 2010; Dehnavi et al., 2015; Shirzadi et al., 2017).

2.4. Methods

Recently, several data driven approaches and spatial analysis techniques were applied to landslide susceptibility modeling (Colkesen et al., 2016; Chen et al., 2017a, 2018d). In the present study, two main techniques were considered for comparative assessment of landslide susceptibility. The first technique is represented by the fuzzy DEMATEL and ANP methods. The second technique focused on machine learning (ML) algorithms namely random forest (RF) classifier and Naïve Bayes (NB) classifier.

Here, the study design was planned in two stages. The first stage includes preparing the LIM of the study area and determination of landslide conditioning factors, normalizing of these factors, extracting weights using fuzzy DEMATEL-ANP approaches, and preparing the landslide susceptibility model. The second stage includes feature selection for accepting important landslide conditioning factors and rejecting unimportant factors for model building, generating landslide susceptibility maps based on two ML algorithms mentioned above, evaluating the validation of the models, and comparing the performance of the two applied approaches (Fig. 6).

2.4.1. Fuzzy decision-making trial and evaluation laboratory (FDEMATEL)

In multi-criteria decision analysis, one criterion directly or indirectly depended on other criteria. The DEMATEL judges the interrelationship and dependence of decision criteria (Kijewska et al., 2018). It was first developed by the Science and Human Affairs from 1972 to 1979 (Sumrit and Anuntavoranich, 2013). The DEMATEL expertly evaluates the causal relationship among the criteria and the weight of each relationship in a decision-making process (Arabsheibani et al., 2016).

In DEMATEL, the direct relation matrix is prepared to structure the complex causal relationships which represent the dependencies and intensity of direct relation of one criterion with other (Tseng and Lin, 2009; Dou et al., 2014). For making a comparison, DEMATEL uses a scale range from 0 to 1 which indicates the influences on criteria against each other. For example, if criterion$_1$ is ranked with 1 and criterion$_6$ with 4, it means that criterion$_6$ has a very high influence upon criterion$_1$. The DEMATEL is an accepted decision-making tool to resolve the interrelationship among decision criteria (Yang et al., 2008; Tzeng et al., 2007). But the personal bias and inconsistency in the variables is the common issue of multi-criteria approaches. Hence, for a consistent decision making, the vagueness in the linguistic variables should be removed (Sangaiah et al., 2017). Fuzzy is an operational method for removing such vagueness (Zadeh, 1965; Vinodh et al., 2016). Table 1 shows the linguistic variables and triangular fuzzy numbers correspondingly for lower, middle and upper (l_{ij}^k, m_{ij}^k, u_{ij}^k). DEMATEL by combining with fuzzy logic consists of the following steps (Lin et al., 2018).

Step 1: Calculation of the fuzzy direct-relation matrix $(\widetilde{A})$.

To calculate the fuzzy direct-relation matrix $\widetilde{A}$, experts give several pairwise comparison ranks to each distinct criterion. $Ex_{ij}^n = (l_{ij}^n, m_{ij}^n, u_{ij}^n)$ is a triangular number given by 'n' number of decision makers. For $\widetilde{A}$, the average of 'n' number of experts is considered which is calculated using Eq. (4).

$$(\widetilde{A}) = \frac{\left(Ex^1 + Ex^2 + Ex^3 + \ldots Ex^N\right)}{N} \tag{4}$$

where N is the total number of experts involved in the assessment and decision making. $Ex^1, Ex^2, Ex^3 \ldots\ldots\ldots Ex^N$ are the expert group. Hence, $(\widetilde{A})$ is computed using the matrix as expressed in Eq. (5).

$$(\tilde{A}) = \begin{bmatrix} 0 & Ex_{21} & \cdots & Ex_{n1} \\ Ex_{12} & 0 & \cdots & Ex_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ Ex_{1n} & Ex_{2n} & \cdots & 0 \end{bmatrix} \tag{5}$$

Step 2: Calculation of the normalized fuzzy direct-relation matrix $(\widetilde{X})$

$\widetilde{X}$ is acquired by the matrix $\widetilde{A}$ using Eqs. (6) and (7).

$$\tilde{X} = \begin{bmatrix} 0 & \tilde{x}_{21} & \cdots & \tilde{x}_{n1} \\ \tilde{x}_{12} & 0 & \cdots & \tilde{x}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{1n} & \tilde{x}_{2n} & \cdots & 0 \end{bmatrix} \tag{6}$$

$$\widetilde{x}_{ij} = \frac{Ex_{ij}}{s}\left(\frac{l_{ij}}{s}, \frac{m_{ij}}{s}, \frac{u_{ij}}{s}\right), s = \max_{\leq i \leq n}\left(\sum_{j=1}^{n} uij\right) \tag{7}$$

Step 3: Obtaining the fuzzy total-relation matrix $\widetilde{T}$

The direct and indirect relations among the factors is shown by the fuzzy total-relation matrix, $\widetilde{T}$, which is acquired using Eqs. (8) and (9).

$$\widetilde{T} = \begin{bmatrix} \widetilde{t}_{11} & \widetilde{t}_{21} & \cdots & \widetilde{t}_{n1} \\ \widetilde{t}_{12} & \widetilde{t}_{22} & \cdots & \widetilde{t}_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{t}_{1n} & \widetilde{t}_{2n} & \cdots & \widetilde{t}_{nn} \end{bmatrix}, \widetilde{t} = \left(l_{ij}^n, m_{ij}^n, u_{ij}^n\right), i \text{ and } j = 1, 2, 3 \ldots, n \tag{8}$$

$$\begin{aligned} \widetilde{T} &= E \times (I - E)^{-1}\left(l_{ij}^n\right) = E \times (I - E)^{-1}, \left(l_{ij}^n\right) = E \times (I - E)^{-1}, \left(l_{ij}^n\right) \\ &= E \times (I - E)^{-1} \end{aligned} \tag{9}$$

where I represents the identity matrix.

For comparing the inter-effect between the criteria, it is suggested to convert $\widetilde{T}$ to defuzzified numbers because it gives a well understanding of the interrelation among the criteria. Hence, elements of $\widetilde{T}$ should be defuzzified using Eq. (10). Then, the α-cut total-relation matrix T_a is estimated, if $t_{ij} < a$, the $t_{ij}^n = 0$, other it will carry the same value as calculated for T_{ij}^{def}.

$$T_{ij}^{def} = \frac{l_{ij}^n + 4\left(m_{ij}^n\right) + u_{ij}^n}{6} \tag{10}$$

Step 4: Estimating the fuzzy value of $\widetilde{D}$ and $\widetilde{R}$

In total-relation matrix $\widetilde{T}$, the summation of columns and rows are represented by $\widetilde{D}$ and $\widetilde{R}$ correspondingly. These are calculated using Eqs. (11) and (12).

$$\widetilde{D} = \left[\widetilde{d}_i\right]_{n \times 1} = \left(\sum_{j=1}^{n} t_{ij}\right) \tag{11}$$

$$\widetilde{R} = \left[\widetilde{r}_j\right]_{n \times 1} = \left(\sum_{j=1}^{n} t_{ij}\right) \tag{12}$$

where $\widetilde{d}_i$ and $\widetilde{r}_j$ are the sum of i^{th} column and row in a matrix $\widetilde{T}$ respectively.

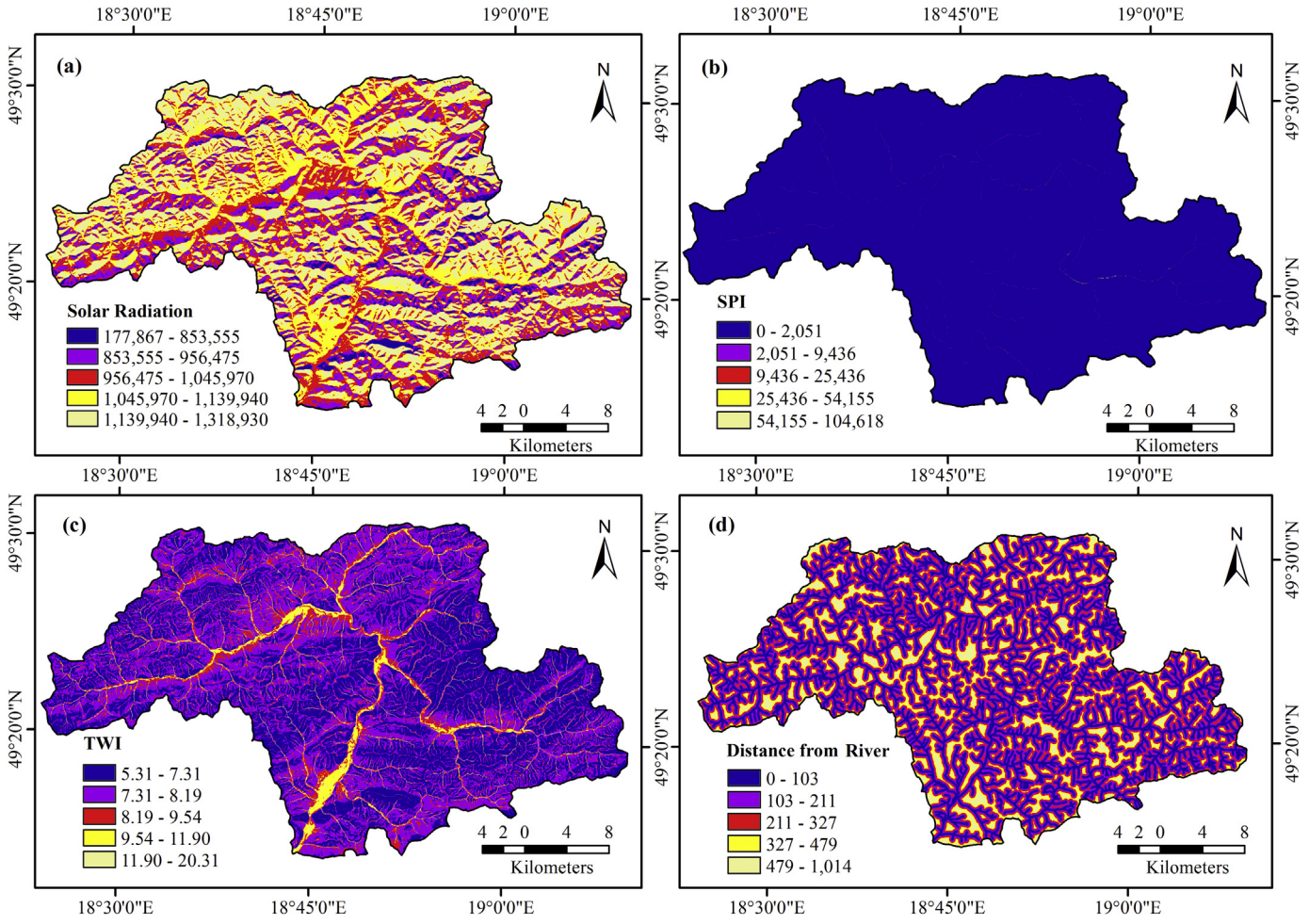

Fig. 4. Landslide conditioning factors: (a) annual solar radiation (WH/m^2); (b) stream power index (SPI); (c) topographic wetness index (TWI); (d) distance to rivers (m).

Step 5: Derivation the causal relationship by the fuzzy value of $\widetilde{D}$ *and* $\widetilde{R}$

To understand the influence of the criteria, the causal relationship between the criteria has be constructed by the sets of $\widetilde{R}+\widetilde{D}$ *and* $\widetilde{R}-\widetilde{D}$ which are defuzzified using Eq. (13).

$$\widetilde{R}\pm\widetilde{D}=\frac{\left(\widetilde{R}\pm\widetilde{D}\right)_l^{\text{fuzzy}}+4\left(\widetilde{R}\pm\widetilde{D}\right)_m^{\text{fuzzy}}+\left(\widetilde{R}\pm\widetilde{D}\right)_u^{\text{fuzzy}}}{6} \tag{13}$$

On calculating $\widetilde{R}\pm\widetilde{D}$, if the value of i-th criterion is negative, it will fall under the effects group which gets influence from other criteria. On the other hand, if the value of i-th criterion is positive, it will fall under the cause group which influences other criteria. A higher value of $\widetilde{R}\pm\widetilde{D}$ means that it has greater priority for a particular instance.

For landslide susceptibility mapping, the influence of conditioning factors was judged through fuzzy direct-relation matrix as shown in Eq. (5), which was further normalized using Eqs. (6) and (7) for calculating the fuzzy total-relation matrix (Eqs. (8) and (9)). The main objective of the application of fuzzy DEMATEL was to identify the interdependence and inter-relation among the conditioning factors which is very crucial because landslide susceptibility is defined as the aggregation of conditioning factors. Thus, Eqs. (11) and (12) were used to calculate the dependences and understand the cause and effect relationship among the conditioning factors. The causal factors of landslide imply its greater role in landslide incidence. The defuzzified values of each factor, as estimated by Eq. (13), was used to determine the scale of importance for preparing the super-matrix of analytic network process (ANP) which is shown in the following section.

2.4.2. Analytic network process (ANP)

ANP is considered more accurate and comprehensive than analytic hierarchy process (AHP) (Saaty, 1996). In AHP, it is supposed that there is no direct relationship between the criteria and each criterion is independent as a one-way hierarchy, but in real circumstances, there might be dependencies between the criteria (Papaioannou et al., 2015). To overcome such limitations of AHP, the ANP model organizes the decision criteria into a network of clusters and nodes (Ghorbanzadeh et al., 2018). ANP is relatively a simple technique that can estimate individual criteria in the decision-making model. In this study, the ANP was applied for calculating the final weights using the following steps (Kanani-Sadat et al., 2019).

Step 1: Formulation of the super-matrix by comparison of the decision criteria W

The formulation of W includes a pair-wise comparison between the criteria. The relative importance of the conditioning factors in pair-wise matrix ranges from 1 to 9. The pairwise comparison of super-matrix is constructed as shown in Eq. (14).

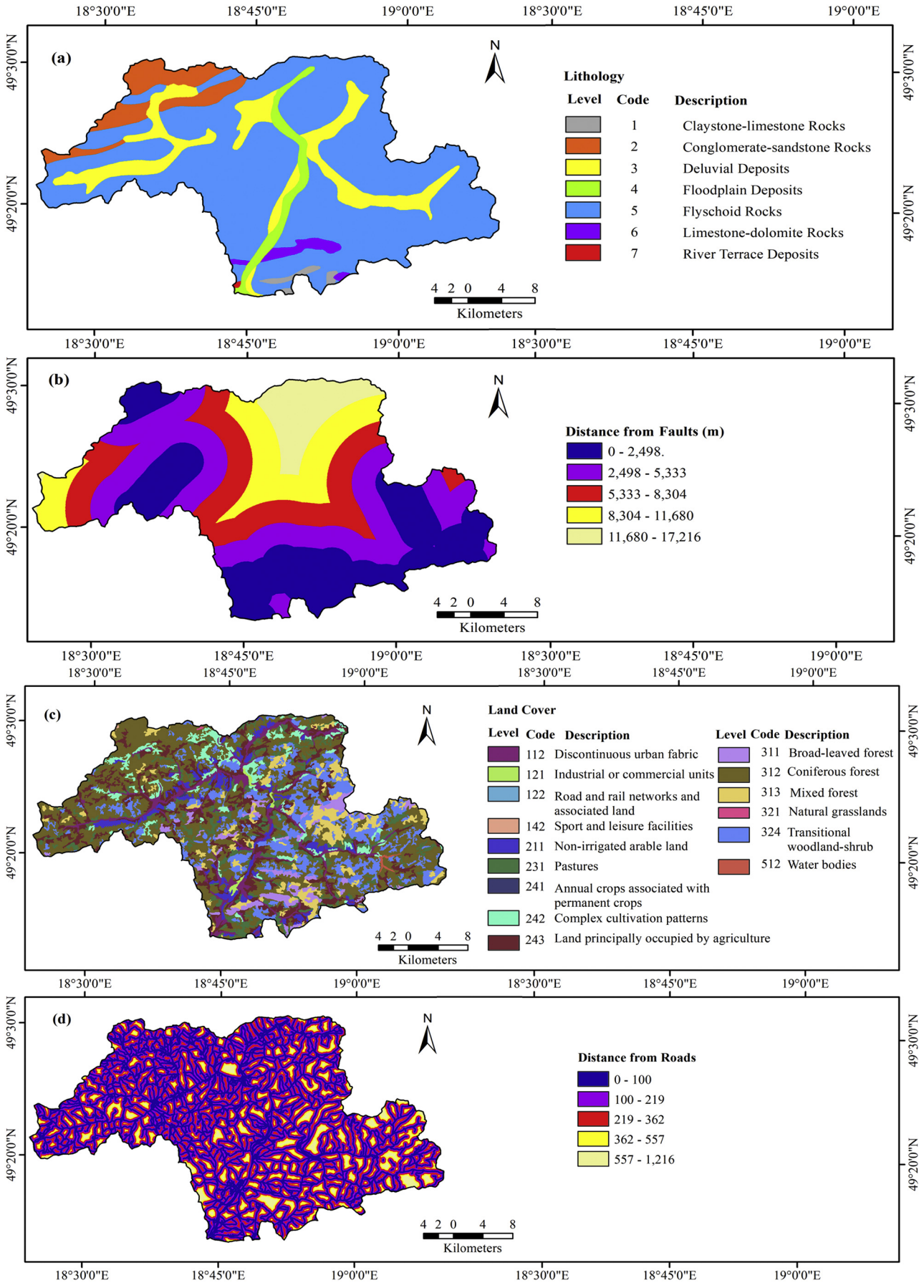

Fig. 5. Landslide conditioning factors: (a) lithology; (b) distance to faults (m); (c) land cover; (d) distance to roads (m).

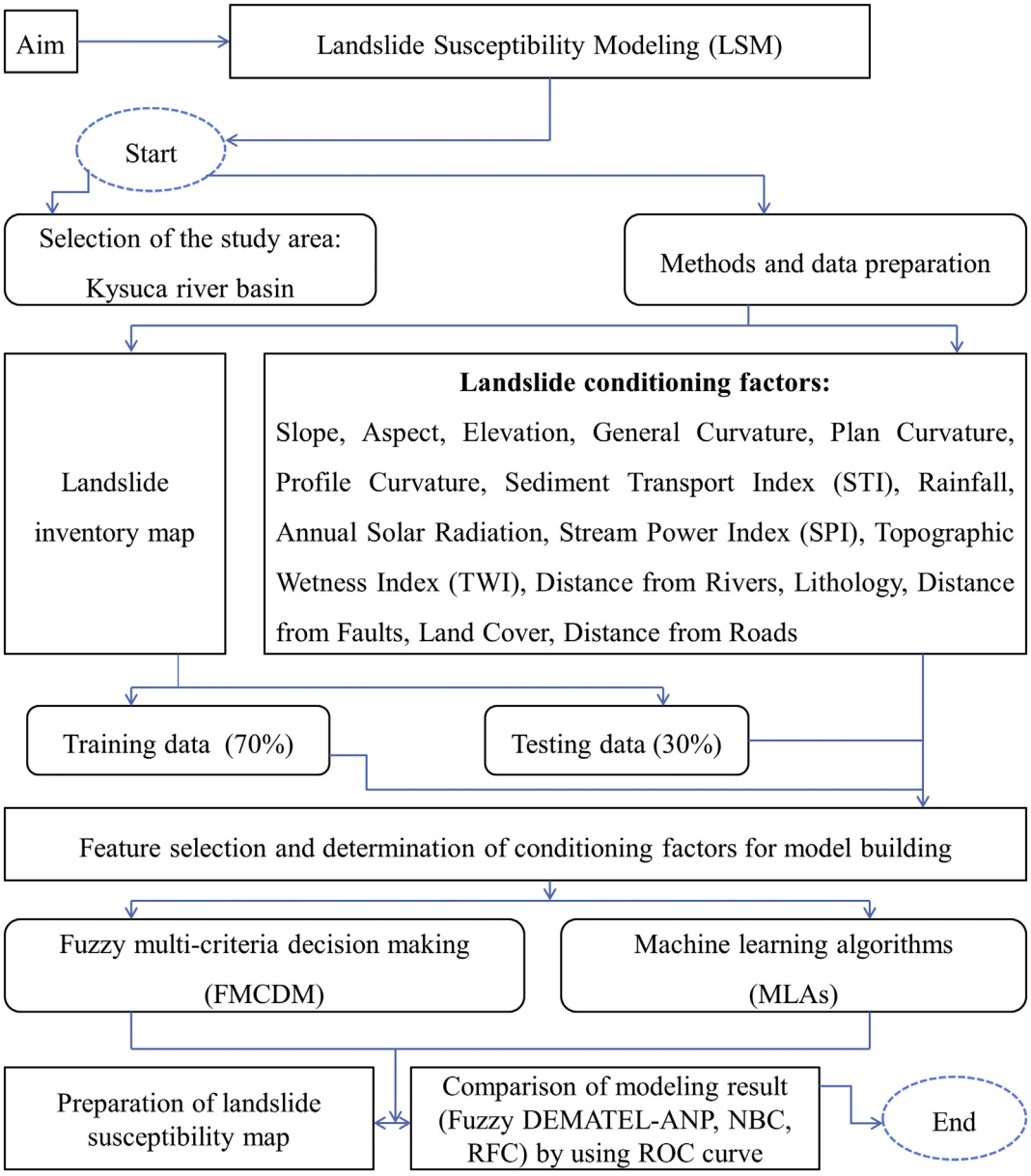

Fig. 6. Flowchart of the methodology adopted in this study.

$$W = \begin{array}{c} \\ C_1 \\ C_2 \\ \vdots \\ C_n \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \cdots & C_n \end{array} \\ \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{n1} \\ W_{21} & W_{22} & \cdots & W_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ W_{1n} & W_{2n} & \cdots & W_{nn} \end{bmatrix} \end{array} \tag{14}$$

Table 1
Linguistic variables and triangular fuzzy numbers used for calculating the fuzzy direct-relation matrix.

Linguistic term	Value for DEMATEL	Triangular fuzzy numbers
No. influence	0	[0, 0, 0.25]
Very low influence	1	[0, 0.25, 0.5]
Low influence	2	[0.25, 0.5, 0.75]
High influence	3	[0.5, 0.75, 1.0]
Very high influence	4	[0.75, 1.0, 1.0]

where, C_n defines the total number of criteria, and W_{ik} refers to importance of k-th criterion on i-th criteria in the pairwise comparison.

Step 2: Acquiring weighted super-matrix D_i

The summation value of each row of the threshold value α, which is calculated through the fuzzy DEMATEL, is expressed using Eq. (15):

$$D_i = \sum_{j=1}^{n} t_{ij}^{\alpha} \tag{15}$$

where the threshold value α of total decision matrix T_s is computed using Eq. (16):

$$T_s = \begin{bmatrix} t_{11}^s & \cdots & t_{1j}^s & \cdots & t_{1n}^s \\ \vdots & \vdots & \vdots & & \vdots \\ t_{i1}^s & \cdots & t_{ij}^s & \cdots & t_{in}^s \\ \vdots & \vdots & \vdots & & \vdots \\ t_{n1}^s & \cdots & t_{nj}^s & \cdots & t_{nn}^s \end{bmatrix} \tag{16}$$

The weighted super-matrix W_w is achieved using the following equation to unweighted super-matrix through Eq. (17):

$$W_w = \begin{bmatrix} t^s_{11} \times W_{11} & \cdots & t^s_{1j} \times W_{j1} & \cdots & t^s_{1n} \times W_{n1} \\ \vdots & \vdots & \vdots & & \vdots \\ t^s_{i1} \times W_{1i} & \cdots & t^s_{ij} \times W_{ji} & \cdots & t^s_{in} \times W_{ni} \\ \vdots & \vdots & \vdots & & \vdots \\ t^s_{n1} \times W_{1n} & \cdots & t^s_{nj} \times W_{jn} & \cdots & t^s_{nn} \times W_{nn} \end{bmatrix} \quad (17)$$

Step 3: Estimation of the weighted super-matrix to a suitably large power kW_l

As given in Eq. (18), the super-matrix becomes a stable matrix by limiting W_w to a large power k. To finish, W_l provides the final weights of each selected criterion.

$$W_l = \lim_{n\to\infty} (W_w)^k \quad (18)$$

The ANP was used for deriving the final weight. The weighted super-matrix, as shown in Eq. (17), was the final decision matrix used in fuzzy DEMATEL-ANP. After stabilizing the matrix, the final weight of each conditioning factor was computed. The landslide susceptibility index (LSI) using the fuzzy DEMATEL-ANP was calculated by multiplying the weight with respective factors, which is shown in Eq. (22).

2.4.3. Naïve Bayes (NB) classifier

Naïve Bayes (NB) classifier is a Bayes-based classifier method which is used to maximize posterior possibility in a Bayesian setting. Naïve Bayes (NB) classifier helps in improving classification probability in which no dependency subsists between variables and conditioning factors exist (Soni et al., 2011; Pham et al., 2017a). The core function of Naïve Bayes (NB) classifier comprises (1) selecting relevant training dataset, (2) determining the posterior probability of each class (landslide and non-landslide location), (3) estimating class level and calculating co-variance matrix, and (4) finally construct a discriminant function of each class level (Khosravi et al., 2019).

In the present study, '*f*' is the conditioning factors of landslide ($f = f_1, f_2, f_3, \ldots, f_n$) and '*l*' is the classifier variables ($l = l_1, l_2$ corresponds to landslide and non-landslide location respectively). The classification probability is then obtained for all classes through the Bayes (NB) classifier. The prediction model to maximize the posterior probability can be obtained using Eq. (19) (Khosravi et al., 2019).

$$l_{NB} = argmax_l P(l_i) \prod_i^n P\left(x_{i/l_i}\right) l_i \in (landslide, non-landslide\ location) \quad (19)$$

where $P(l_i)$ is the posterior probability and $P\left(x_{i/l_i}\right)$ is the conditional probability which is defined by Eq. (20).

$$P\left(x_{i/l_i}\right) = \frac{1}{\sqrt{2\pi}\lambda} e^{-(aj-\mu)^2/2\lambda^2} \quad (20)$$

where μ and λ correspond to mean deviation and standard deviation of the training datasets (x_i), respectively.

The spatial prediction of landslide susceptibility was done based on the largest posterior probability, as shown in Eq. (19), while the prior probability was made based on the output class in the training dataset which helps in increasing the accuracy of prediction. For applying the Naïve Bayes (NB) classifier, the training dataset, containing the landslide and non-landslide locations, was transported into Weka 3.8.4 software. Additionally, a 10-fold cross validation process was applied to run the classification and to derive the weights of landslide conditioning factors. The derived weights were used to generate the landslide susceptibility index (LSI), which is shown in Eq. (23).

2.4.4. Random forest (RF) classifier

Random forest (RF) classifier is a Tree-based classifier. Random forest (RF) classifier is the improved form of bagged tree classifier which is taken to resolve the complications of prediction and multi-classification (Quiroz et al., 2018). It combines the bagging ensemble learning and random subspace which make it a powerful tool to create a classification tree for pattern recognition of large-scale and multivariate data (Breiman, 2001). The core function of Random forest (RF) classifier take account of (1) random selection of the sample features '*k*' from a total sample '*m*' where '*k*' < '*m*', (2) calculation of the node tree '*d*' by using suitable split point among the selected features '*k*', (3) again using the best spilt point, split '*d*' into daughter nodes tree '*dn*', and (4) repetition of the above steps until '*I*' the number of nodes tree is achieved.

The random forest (RF) classifier decides maximum nodes to its output class (Bonissone et al., 2010). Hence, for an input data '*x*' is computed its output '*y*' from maximum ensemble which is expressed in Eq. (21).

$$y = (x) = max\left[\sum_k I(t)\right] \quad (21)$$

where $I(t)$ is an indicator function defined as:

$$I(t)\left\{ \begin{array}{l} 1, t =' YES' \\ 0, t =' NO' \end{array} \right\}$$

For the training dataset in this study, *'YES'* is the landslide location and *'NO'* is the non-landslide location. Random forest (RF) classifier was also previously applied in different studies on landslide modeling (Chen et al., 2017b; Dang et al., 2020).

Similar to the Naïve Bayes (NB) classifier, the weights of conditioning factors were derived using random forest (RF) classifier in Weka 3.8.4 software. The weight values were calculated based on bagging with 100 iterations and base learner. Finally, the weights were multiplied with respective conditioning factors using map algebra tool of ArcGIS 10.5 software. In order to generate the landslide susceptibility index (LSI), Eq. (24) was applied.

$$LSI_{FDEMATEL-ANP} = \sum_{i=1}^{n} W_i * f_i \quad (22)$$

$$LSI_{NBC} = \sum_{i=1}^{n} W_i * f_i \quad (23)$$

$$LSI_{RFC} = \sum_{i=1}^{n} W_i * f_i \quad (24)$$

where n is the number of landslide predictors, W_i is the weight of landslide conditioning factors determined by the fuzzy DEMATEL-ANP, Naïve Bayes (NB) classifier and random forest (RF) classifier, respectively, and f_i is the respective factor.

2.4.5. Feature selection

Regarding the machine learning algorithms, feature selection is an important method used in preliminary phase of model building (Dang et al., 2020). Before the model training phase, it is essential to judge the relevance of the selected factors used for landslide susceptibility mapping. In this study, the ReliefF technique was considered to estimate the significance of conditioning factors based on attribute rank. The ReliefF is a probabilistic method used for data classification which inspects conditional dependencies and discriminative power of selected factors (Robnik-Šikonja and Kononenko, 2003). The ReliefF method estimates the weight of each factor to quantify its importance and relevance. Higher weight depicts more importance and essentiality of the

factor, whereas '0' depicts irrelevant factor. Hence, it is required to choose a threshold value of the weights and below this threshold value the factor is considered irrelevant for the present analysis. In the present study, '0' was selected as a threshold value; thus, the normalized weight of any factor belonged to '0' or '<0', were removed from further analysis. Since all of the factors weights were not '0' or null, there was not an insignificant factor and all of them were taken for modeling and further analysis. Other recent studies also showed that the factors with zero weight are not important to landslide susceptibility and should be removed from modeling, but greater than zero can be included while feature selection is performed for random forest and other machine learning algorithms (Dang et al., 2020; Hong et al., 2018).

In the present study, the training dataset was prepared using landslide and non-landslide locations based on multi-values to point extraction in which each point covers a specific pixel value of each conditioning factor. In the training dataset, the 'type' was numbered with 1 and 0, indicating landslide and non-landslide, respectively. The final dataset was exported into the Weka 3.8.4 software and ReliefF was selected for evaluating the relevance of the selected factors.

2.4.6. Performance evaluation of the models

2.4.6.1. Receiver operating characteristic (ROC) curve. The ROC curve is a widely used and accepted technique used in the geospatial analysis for determining the validation of the model (Tehrany et al., 2013; Chen et al., 2018a). ROC uses the area under the curve (AUC) for numerical evaluation of the predicted model, in which 1-specificity and sensitivity are plotted on 'X' and 'Y' axis respectively. The ROC curve ranges from 0.5 to 1.0, where ≤0.5 and 1 indicate inaccurate and highly accurate of the prediction model, respectively (Chen et al., 2020).

2.4.6.2. Arithmetic measures. Along with ROC curves, the arithmetic measures were also performed to evaluate the prediction accuracy of the models. A set of quantitative analysis including mean absolute error (MAE), root mean square error (RMSE), Kappa index (K), sensitivity, specificity, and overall accuracy (OAC) were estimated to measure the accuracy of the landslide susceptibility models. The following formulas are accepted for these statistical measures (Khosravi et al., 2019):

$$MAE = \frac{1}{n}\sum_{i=1}^{i=n}|X_{ei} - X_{oi}| \tag{25}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{i=n}(X_{ei} - X_{oi})^2} \tag{26}$$

$$K = \left(\frac{L_c - L_{exp}}{1 - L_{exp}}\right) \tag{27}$$

$$Sensitivity = \left(\frac{TP}{TP + FN}\right) \tag{28}$$

$$Specificity = \left(\frac{TN}{TN + FP}\right) \tag{29}$$

$$OAC = \left(\frac{TP + TN}{TP + TN + FP + FN}\right) \tag{30}$$

where *TP* and *TN* are the true positive and true negative, representing numbers of pixel correctly classified; *FP* and *FN* are the false positive and false negative, representing numbers of pixel incorrectly classified; L_c is the numbers of a pixel that are correctly classified as landslide and non-landslide, L_{exp} are the expected agreements, X_{ei} and X_{oi} are the model predicted and observed values respectively, and n is the number of data points.

After preparing the landslide susceptibility map using fuzzy DEATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier, their performance and accuracy were measured. In order to extract the true positive, true negative, false positive, and false negative, the susceptibility maps were reclassified into five classes using natural breaks classification method and the multi-values to point extraction technique was used to extract the pixel values of susceptibility maps into landslide and non-landslide points. The final data was exported in Excel sheet and the MAE, RMSE, Kappa index, sensitivity, specificity, and overall accuracy was calculated using Eqs. (25)–(30), respectively.

3. Results

3.1. Weights of the conditioning factors

The ranked weight and normalized weight of all selected conditioning factors resulted in the value of <1 and > 0, which indicates that all of the selected factors are more or less important because there are no multicollinearity problems (Table 2). The application the ReliefF technique, used to assess the weight-based importance of all landslide conditioning factors, revealed that slope (NW = 0.1516) had the highest normalized weight (NW) and thus the highest influence on landslide occurrence (Fig. 7). The next influential factor is lithology (NW = 0.1182) followed by slope aspect (NW = 0.0966), land cover (NW = 0.0900), distance to faults (NW = 0.0874), TWI (NW = 0.0797), annual solar radiation (NW = 0.0670), rainfall (NW = 0.664), elevation (NW = 0.0620), STI (NW = 0.0530), distance to roads (NW = 0.0499), distance to river (NW = 0.0452), plan curvature (NW = 0.0106), general curvature (NW = 0.0096), profile curvature (NW = 0.0088), and SPI (NW = 0.0040). The results showed that among the selected 16 conditioning factors, three types of curvature and SPI had lower effect on landslide occurrence, but these were not removed from the model building because they have NW > 0.

3.2. Landslide susceptibility maps (LSMs)

The present study highlighted different techniques and compared their results for proposing an optimal model for LSMs in the study area. The statistical measures were used to perform the model evaluation by using historical landslide records. Primarily, 16 relevant conditioning factors of landslide were considered to integrate and prepare the susceptibility maps using two different approaches i.e. fuzzy multi-criteria and machine learning methods and then the results were compared between the FDEMATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier models. The detailed data sets for running these three models can be found in the Supplementary Files S1 – S8.

Table 2
Feature selection for determining the importance of landslide conditioning factors.

Landslide conditioning factors	Rank weight (out of 1)	Normalized weight (%)
Slope	0.3490	15.1573
Aspect	0.2224	9.6591
Elevation	0.1428	6.2036
General Curvature	0.0222	0.9642
Plan Curvature	0.0244	1.0598
Profile Curvature	0.0203	0.8800
Sediment transport index (STI)	0.1219	5.2958
Rainfall	0.1530	6.6431
Annual solar radiation	0.1543	6.7000
Stream Power Index (SPI)	0.0091	0.3952
Topographic Wetness Index (TWI)	0.1835	7.9708
Distance to rivers	0.1042	4.5236
Lithology	0.2720	11.8151
Distance to faults	0.2013	8.7440
Land cover	0.2071	8.9959
Distance to roads	0.1150	4.9926

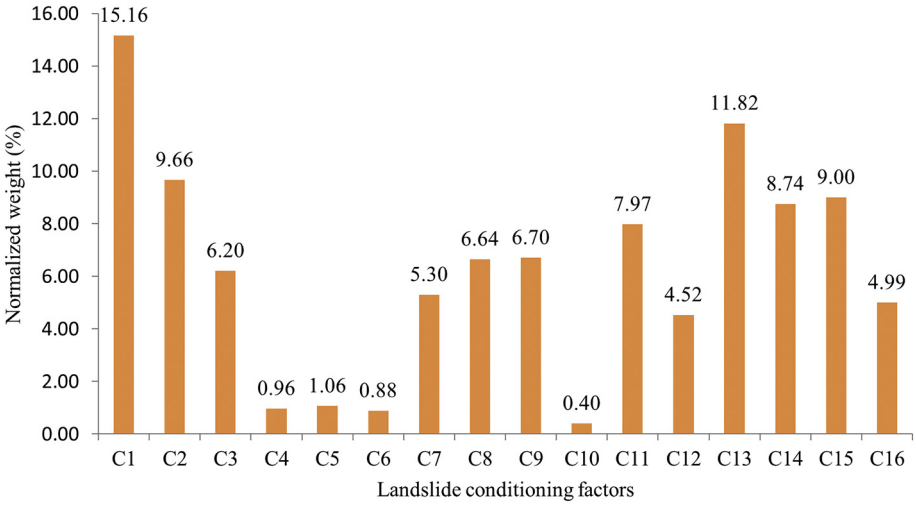

Fig. 7. Importance of landslide conditioning factors using the RelieffF method. C1-Slope; C2-Aspect; C3-Elevation; C4-General Curvature; C5-Plan Curvature; C6-Profile Curvature; C7-Sediment Transport Index (STI); C8-Rainfall; C9-Annual solar radiation; C10-Stream Power Index (SPI); C11-Topographic Wetness Index (TWI); C12-Distance to rivers; C13-Lithology; C14-Distance to faults; C15-Land cover; C16-Distance to roads.

Landslide susceptibility maps were prepared separately for the FDEMATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier models by using the respective weights of selected conditioning factors. In the case of FDEMATEL-ANP, it evaluates the interdependences of each selected criterion concerning others. The *D* and *R* (sum of the columns and rows respectively) value of the defuzzified total-relation matrix validates the role of all conditioning factors in LSM (Table 3). The higher defuzzified value of *Ri* + *Di* means that it has more relation with other criteria whereas defuzzified value of *Ri*–*Di* denotes the kind of relationship or dependencies of conditioning factors with others. If the value of *Ri*–*Di* is plus (+), it means it comes under cause group which affects the other factors. Subsequently, if the value of *Ri*–*Di* is minus (−), it refers it comes to effect group which get influenced by other factors. For the present study area, the aspect, STI, TWI, distance from rivers, lithology, distance from faults, and land cover are causal factors and the rest are effect factors for landslide susceptibility modeling. For FDEMATEL-ANP, the normalized weights were used, which were derived from the limit super-matrix (Table 4). The weighted overlay was applied by putting the normalized weight to respective conditioning factors and the final susceptibility map was generated (Fig. 8). In case of the Naïve Bayes (NB) classifier model, the attribute ranks were multiplied with respective conditioning factors and the landslide susceptibility map was generated (Fig. 9). As for the random forest (RF) classifier model, the Bagging with 100 iterations and base learners were considered and similarly as in Naïve Bayes (NB) classifier, the ranks were multiplied to produce the final landslide susceptibility map for the Kysuca river basin (Fig. 10).

After generating the LSMs by using fuzzy multi-criteria and machine learning algorithms, the study area was reclassified into 5 susceptibility classes, i.e. VLS for very low susceptibility, LS for low susceptibility, MS for moderate susceptibility, HS for high susceptibility, and VHS for very high susceptibility, using 'Natural Breaks' classification schemes. The histogram of all applied models corresponding to pixel coverage by each landslide susceptibility class is shown in Fig. 11. The result reveals that each model covers a different number of pixels and dissimilarity between each model is also displayed. For more clarification, the spatial distribution of landslide susceptibility classes and recorded landslides in each class are also shown in Fig. 12. The prepared histograms of all three models reveal that nearly about 80% of the recorded landslides are accurately found in very high to moderate classes, except for the Naïve Bayes (NB) classifier model in which more than 35% of the recorded landslides are found in both low and very low susceptibility classes. From that finding, the random forest (RF) classifier and fuzzy DEMATEL-ANP model are more accurate for landslide susceptibility

Table 3
Crisp and fuzzy values of selected landslide conditioning factors.

Criteria	Fuzzy values		Defuzzified values	
	Ri + *Di*	*Ri* − *Di*	*Ri* + *Di*	*Ri* − *Di*
Slope	[2.7,3.58,7.5]	[−0.5, −0.4, −0.3]	4.081	−0.38
Aspect	[3.7,4.44,8.2]	[0.04,0.05, −0.1]	4.947	0.031
Elevation	[2.9,4.39,8.4]	[−0.4,0,0.1]	4.819	−0.06
General Curvature	[4.8,5.44,9.3]	[−0.3, −0.2, −0.1]	5.971	−0.2
Plan Curvature	[4.6,5.73,9.7]	[−0.9, −0.5, −0.1]	6.201	−0.48
Profile Curvature	[3,4.13,8.2]	[−0.6, −0.5, −0.6]	4.62	−0.56
Sediment transport index (STI)	[5.5,6.12,9.9]	[0.18,0.16,0]	6.646	0.143
Rainfall	[3.1,4.32,8.4]	[−1.5, −1.3, −1.4]	4.786	−1.36
Annual solar radiation	[4.8,6,10]	[0.77,0.52,0.2]	6.483	0.51
Stream Power Index (SPI)	[5.3,6.29,10]	[−0.2, −0.4, −0.5]	6.822	−0.36
Topographic Wetness Index (TWI)	[3.2,4.88,9]	[0.85,0.79,0.8]	5.289	0.79
Distance to rivers	[1,2.5,6.5]	[0.61,0.52,0.5]	2.902	0.53
Lithology	[1.6,3.36,7.4]	[0.81,0.59,0.7]	3.737	0.642
Distance to faults	[1.2,2.7,6.7]	[0.89,0.76,0.9]	3.126	0.798
Land cover	[0.4,1.81,5.7]	[0.31,0.21,0.2]	2.234	0.226
Distance to roads	[0.1,1.4,5.2]	[−0.1, −0.3, −0.4]	1.822	−0.28

Table 4
Final weights for landslide susceptibility modeling.

Criteria	Weight (w_i)	
	Idealized	Normalized[a]
C1-Slope	0.7848	0.1707
C2-Aspect	0.0855	0.0186
C3-Elevation	0.2409	0.0524
C4-General Curvature	0.0897	0.0195
C5-Plan Curvature	0.2006	0.0436
C6-Profile Curvature	0.0675	0.0147
C7-Sediment transport index	0.4508	0.0981
C8-Rainfall	0.2129	0.0463
C9-Annual solar radiation	0.0890	0.0194
C10-Stream Power Index	0.1062	0.0231
C11-Topographic Wetness Index	1.0000	0.2176
C12-Distance to rivers	0.3122	0.0679
C13-Lithology	0.1820	0.0396
C14-Distance to faults	0.4031	0.0877
C15-Land cover	0.2167	0.0471
C16-Distance to roads	0.1547	0.0336

[a] The normalized weights are calculated from the idealized weights. Idealized weights are the actual weights derived using fuzzy DEMATEL-ANP, but to prepare the landslide susceptibility map the normalized values were considered. The normalized weight was calculated by simply dividing the idealized value of a criterion with its summation value.

than the Naïve Bayes (NB) classifier. In case of the random forest (RF) classifier; the HS class has the largest coverage area (30.26%), followed by MS (24.12%), VHS (20.35%), LS (17.05%), and VLS (8.21%) classes. For the fuzzy DEMATEL- ANP model, the largest area also comes under HS class with 28.83%, followed by VHS, MS, LS, and VLS classes with the share of 23.42%, 22.74%, 16.67%, and 8.34% respectively. But for the Naïve Bayes (NB) classifier, HS and MS classes have the largest area with 27.89% and 26.53% respectively, followed by LS, VHS, and VLS classes with the share of 19.44%, 15.69%, and 10.46%, respectively. Thus, from a comparative analysis of the susceptibility model, the random forest (RF) classifier and the fuzzy DEMATEL-ANP are more similar in comparison to the Naïve Bayes (NB) classifier.

3.3. Validation of the models and their comparison

Performance of the applied models for predicting landslide susceptibility were evaluated and compared using the ROC curve and AUC. Additionally, different arithmetic measures were also estimated for determining optimal model in landslide susceptibility.

Fig. 13 shows the ROC curves of three landslide susceptibility models generated by the FDEMATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier models. The results confirmed that all three prediction models resulted in very high accuracy and well and thus accepted. The highest accuracy was recorded by the random forest (RF) classifier model (AUC = 0.954), followed by the FDEMATEL-ANP model (AUC = 0.924) and the Naïve Bayes (NB) classifier model (AUC = 0.909). As a result, the random forest (RF) classifier with AUC = 0.954 outperformed the FDEMATEL-ANP and Naïve Bayes (NB) classifier models.

Similar results were also obtained from the statistical measures. Different statistical measures showed that the random forest (RF) classifier had higher accuracy in comparison to the FDEMATEL-ANP and Naïve Bayes (NB) classifier model with higher values of OAC (92.2%) and Kappa index (0.8435) and lower values of MAE (0.1238) and RMSE (0.2555). On the other hand, the Naïve Bayes (NB) classifier had lower accuracy in comparison to the random forest (RF) and FDEMATEL-ANP models with lower values of OAC (76.35%) and Kappa index (0.5283) and higher values of MAE (0.2453) and RMSE (0.3457) (Table 5).

4. Discussion

Landslides are the complex natural disasters that need immediate preventive measures. Therefore, appropriate prediction and prevention

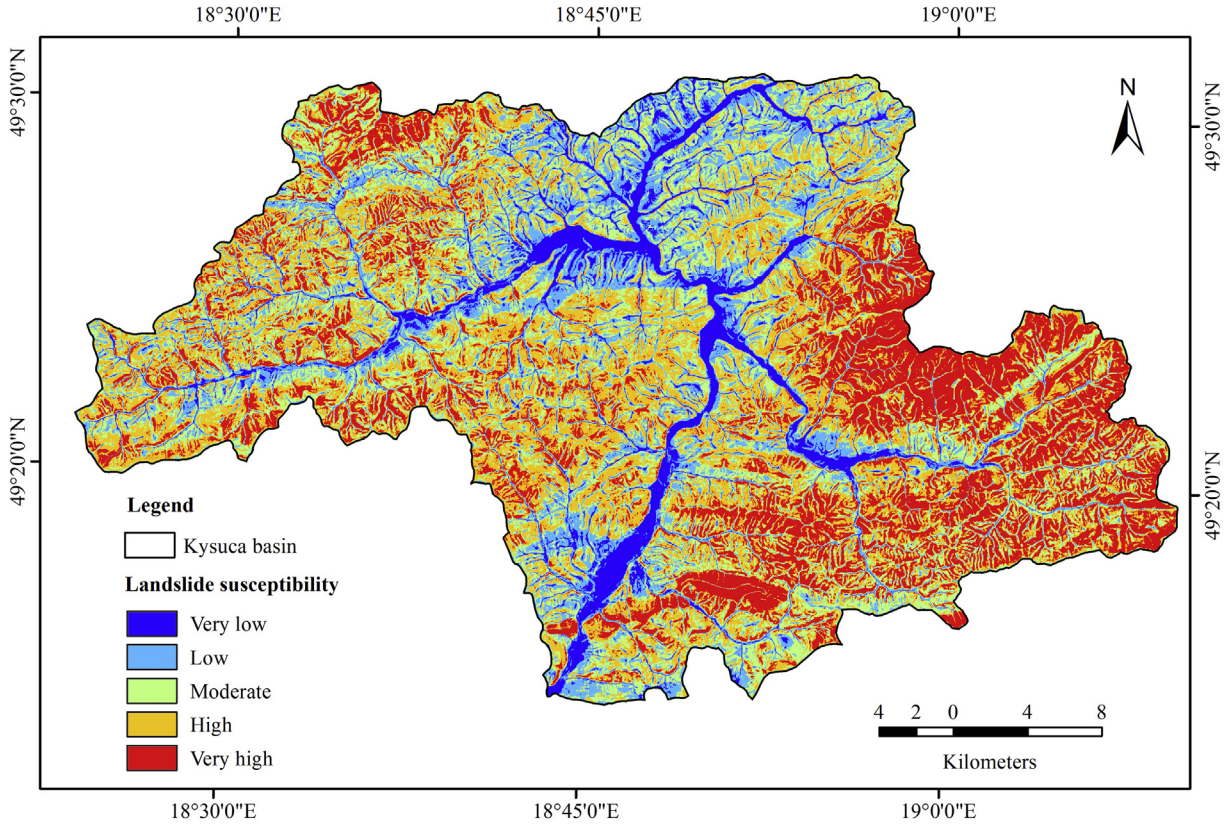

Fig. 8. Landslide susceptibility map via fuzzy DEMATEL-ANP model.

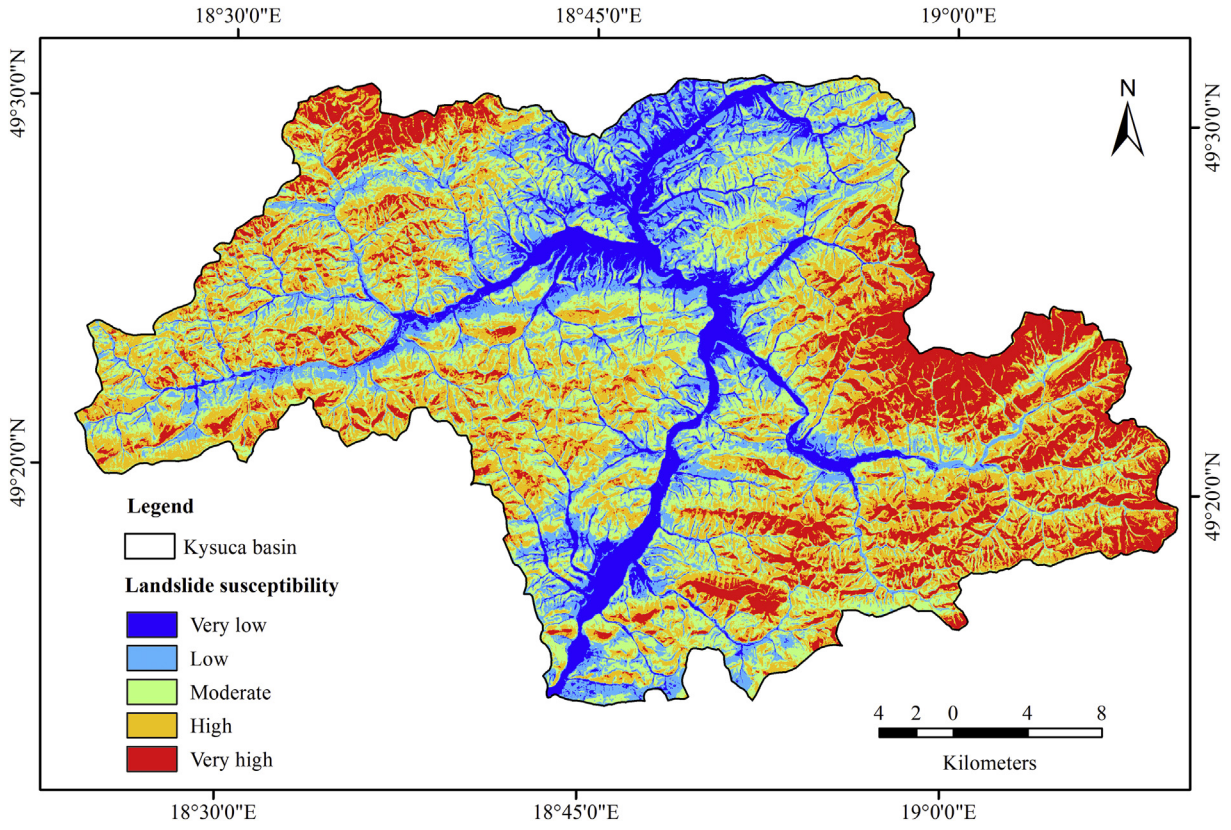

Fig. 9. Landslide susceptibility map via naïve Bayes (NB) classifier model.

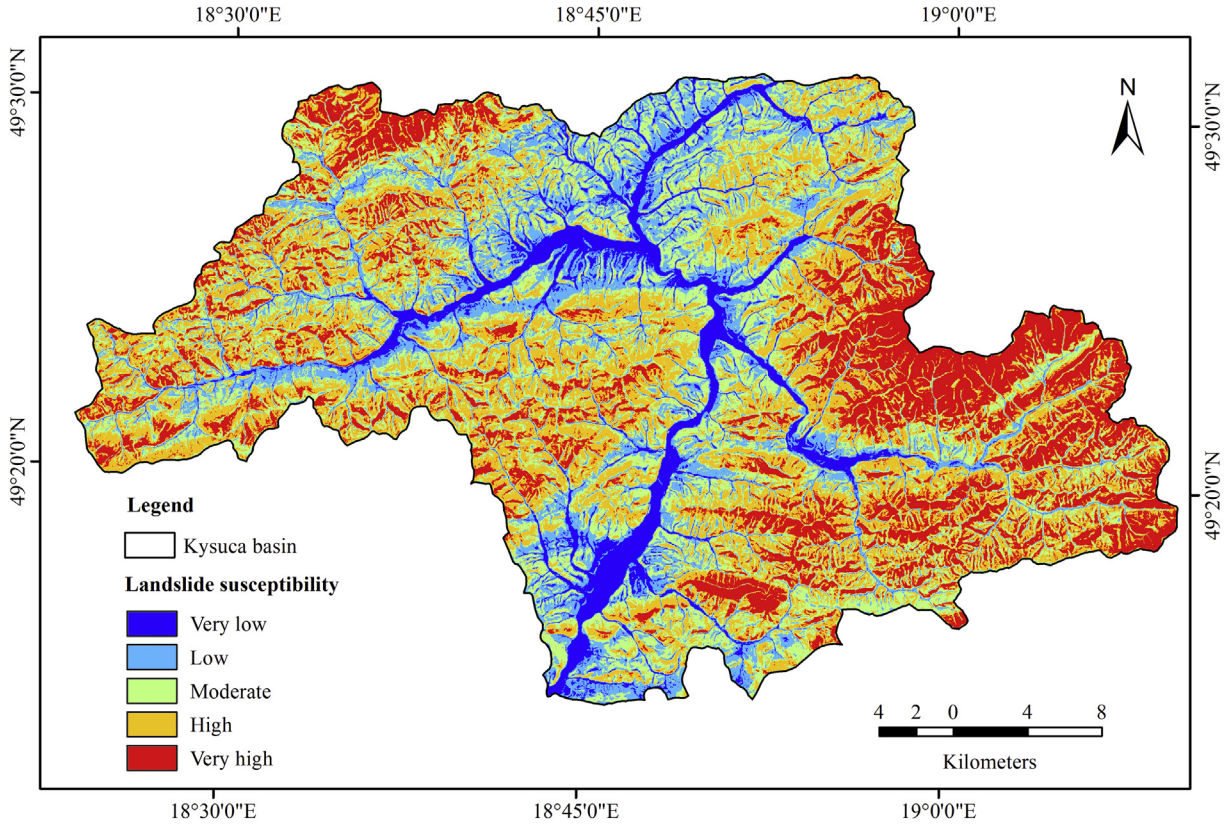

Fig. 10. Landslide susceptibility map via random forest (RF) classifier model.

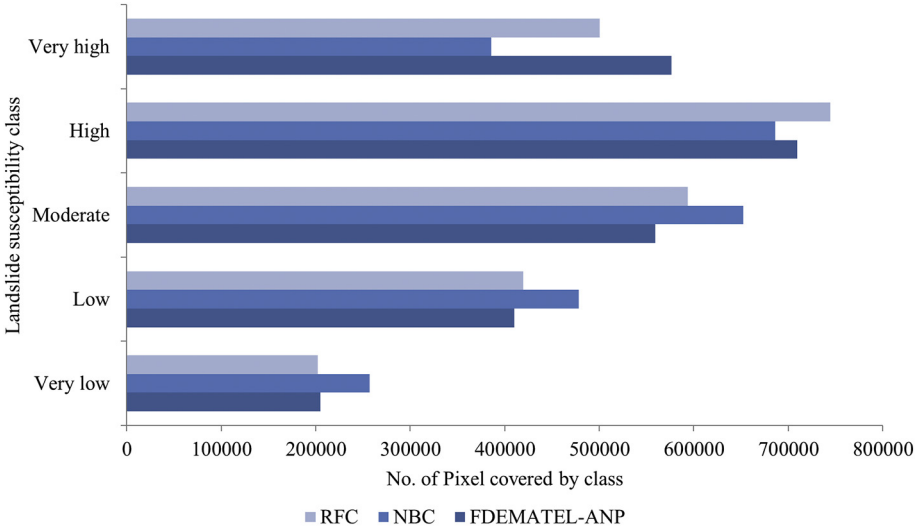

Fig. 11. Histogram of landslide susceptibility classes using different models.

techniques are becoming essential to minimize the damage caused by landslides. Applying suitable methods to produce LSMs are the key factor for risk assessment and management. Through spatial mapping the landslide and non-landslide prone areas can be identified and proper mechanisms can be applied to prevent the losses caused by the landslide. Currently, different models and approaches have been proposing to predict the spatial distribution of landslides at both micro and macro level. Certainly, optimal mapping of landslide susceptibility is now gaining great interest among the land geoscientists, but many of them focused on a single model for predicting landslides (Pourghasemi et al., 2012; Yilmaz et al., 2012; Dang et al., 2020). Recently, some of them emphasized the ensemble or hybrid learning due to their application advantages and sound accuracy (Trigila et al., 2015; Khosravi et al., 2016; Pham et al., 2017b; Vakhshoori et al., 2019). In this sense, the present study employed and compared the hybrid model incorporating different techniques, namely fuzzy multi-criteria (fuzzy DEMATEL-ANP) and machine learning methods (Naïve Bayes (NB) classifier and random forest (RF) classifier), for LSMs of the Kysuca river basin in Slovakia. Landslide susceptibility mapping was carried out by selecting 16 landslide conditioning factors.

An appropriate decision technique is also required for accurate prediction of models to gain insight into the landslide distribution characteristics (Brenning, 2012). Therefore, a statistically valid model with good predictive properties is needed to be incorporated to support the prediction. This technique provides easily interpretable results that highlight the landslide conditioning factors. Recently, ensemble techniques are widely applied to assess the performance of LSMs (Ada and San, 2018; Chen et al., 2018c; Bui Tien et al., 2012; Kavzoglu et al., 2019; Vakhshoori et al., 2019; Xiao et al., 2019). However, the comparison between different techniques, specifically multi-criteria evaluation and machine learning algorithms, in landslide susceptibility mapping is lack in literature. Thus, this study focused on the comparison between the fuzzy DEMATEL-ANP, and naïve Bayes and random forest classifier.

Machine learning methods are attracting more and more scientists in the field of model predictions (Brenning, 2005; Micheletti et al., 2014; Petschko et al., 2014; Kavzoglu et al., 2019). The random forest, as a classification technique, is considered to have the best predictive performance (Goetz et al., 2015a; Ließ et al., 2011). It is an ensemble mechanism that employs several classification trees to stimulate model predictions. Naïve Bayes mechanism is widely accepted for its fast supervised learning algorithm in data mining which offers reasonable results in large-scale prediction of difficult datasets. However, the main drawback is that it needs independence of criteria (Bui Tien et al., 2012).

For comparing three models applied in the landslide susceptibility assessment, sixteen landslide conditioning factors including slope, aspect, elevation, general curvature, plan curvature, profile curvature, STI, rainfall, annual solar radiation, SPI, TWI, distance to rivers, lithology, distance to faults, land cover, and distance to roads were selected in this study. The results obtained from the FDEMATEL-ANP model showed that TWI was the most effective factors of landslide occurrence with normalized weight of 0.2176. Other important factors were slope, STI, distance to faults, distance to rivers, elevation, plan curvature, and rainfall with the normalized weight of 0.1707, 0.0981, 0.0877, 0.0679, 0.0524, 0.0436, and 0.0463, respectively. For Naïve Bayes classifier model, the results revealed that slope (weight = 0.9636) had highest weight value, followed by TWI (weight = 0.9228), SPI (weight = 0.8471), elevation (weight = 0.8356), distance to faults (weight = 0.7714), annual solar radiation (weight = 0.6804), rainfall (weight = 0.4860), and land cover (weight = 0.4499) representing their significant role in landslide susceptibility modeling. In case of random forest classifier model the highest weight was found for TWI (0.6291), followed by slope (0.6097), SPI (0.5307), STI (0.529), rainfall (0.4863), elevation (0.3467), annual solar radiation (0.2925), and distance from faults (0.2221). As a whole of all three models, it was found that TWI, slope, and SPI were more significant for landslide susceptibility assessment while general curvature, plan curvature, and profile curvature had least important role with lower weights (Table 4 and Supplementary Table S8).

Based on performance level of these three models, the result showed that random forest classifier model gained highest accuracy (OAC = 92.2%, AUC = 95.4%) in comparison to FDEMATEL-ANP model (OAC = 80.8%, AUC = 92.4%) and Naïve Bayes classifier model (OAC = 76.35%, AUC = 90.9%). Thus, from the comparative analysis of these three models, it was found that random forest classifier model was highly accurate while Naïve Bayes classifier model had lower performance because it needs independence of criteria which makes it relatively robust. Therefore, the prediction capability of the Naïve Bayes classification is comparatively lower than other models applied in the present study (Soria et al., 2011; Bui Tien et al., 2012).

All in all, the present study applied the above-mentioned models for landslide susceptibility modeling in a multi-hazard region. The results evidenced that all of these models can be accepted for landslide susceptibility modeling in the study area as well as other similar vulnerable regions due to a high rate of prediction accuracy. But, from the viewpoint

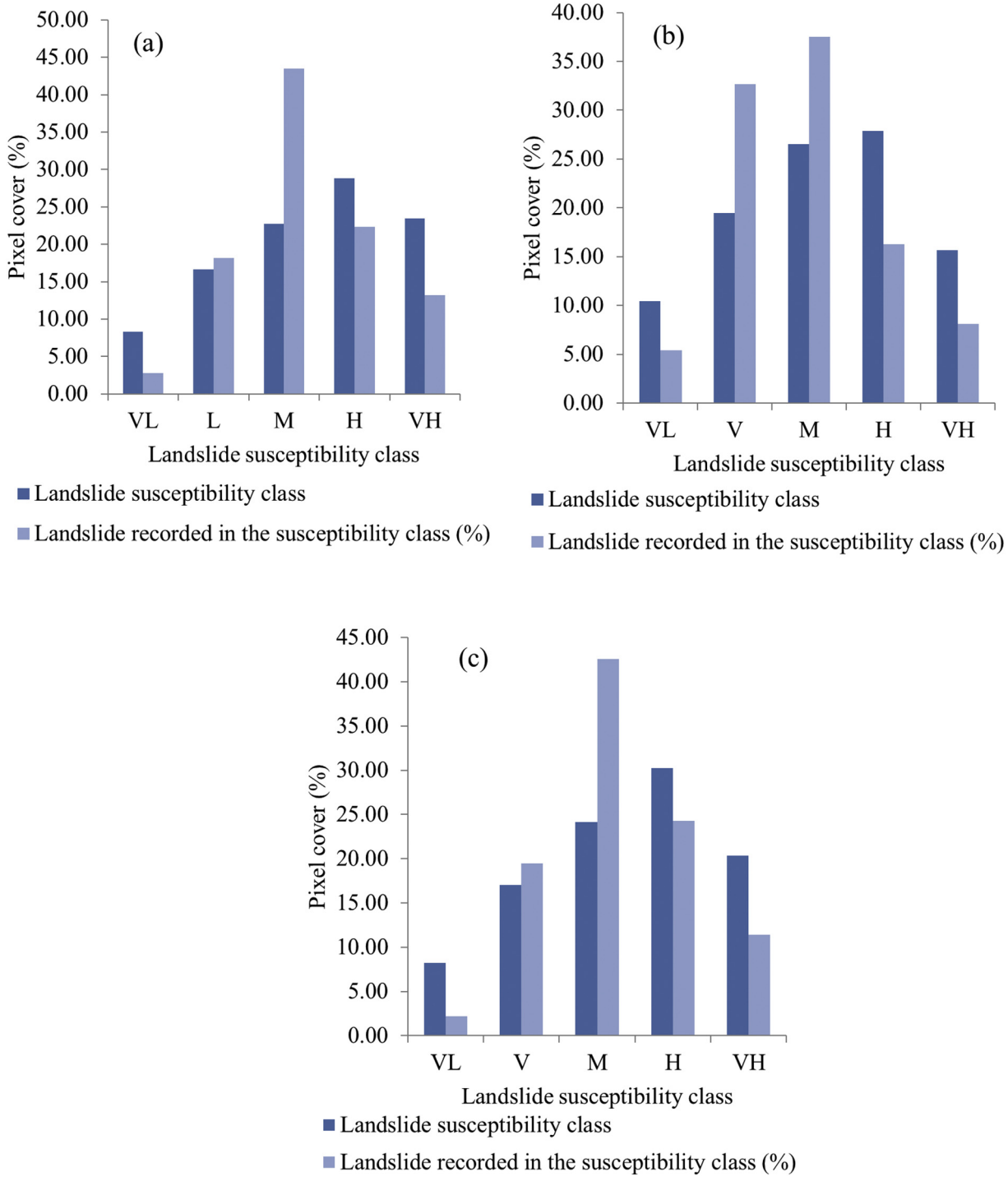

Fig. 12. Determination of optimal model based on landslide susceptibility and recorded landslides in respective classes (a) Fuzzy DEMATEL-ANP; (b) naïve Bayes (NB) classifier; (c) random forest (RF) classifier.

of comparative analysis, the random forest (RF) classifier model becomes an optimal model, as it outperformed the other two models in terms of ROC-AUC, ACC, Kappa index and other error estimation methods of MAE and RMSE, but these two models were also accurate and can be accepted for the landslide susceptibility modeling (as OAC and AUC is greater than 75%).

It is unanimously accepted that the methodological limitations are almost inevitable and could occur even in the research works whose central aim is the use of the state-of-the-art algorithms to estimate the natural hazards susceptibility. This fact was brought into discussion also in the present manuscript. Thus, a methodological limitation that could have a negative influence on the precision of the final results is represented by the lower accuracy of land cover predictor, extracted from CORINE Land Cover 2018 database, comparing to the other 15 predictors. Nevertheless, according to the reasons described in the following rows, it is considered that this fact has a minor negative influence on the accuracy of the final results. Thus, we consider that the presence of the other 15 landslide conditioning factors (from which 10 were

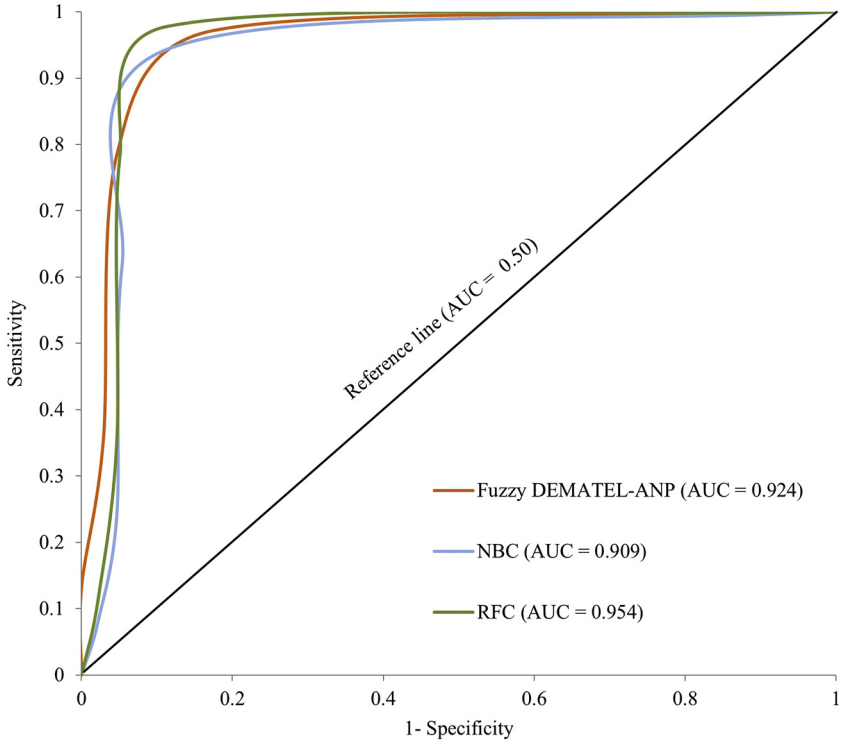

Fig. 13. ROC curves for assessing the model performance.

Table 5
Performance of the fuzzy DEMATEL-ANP, naïve Bayes (NB) classifier, and random forest (RF) classifier models for landslide susceptibility assessment.

Model accuracy	Models		
	FDEMATEL-ANP	NBC	RFC
MAE	0.1975	0.2453	0.1238
RMSE	0.3128	0.3457	0.2555
Kappa index	0.6185	0.5283	0.8435
TP Rate/Recall	0.616	0.586	0.973
TN Rate	1	0.941	0.871
FP Rate	0	0.059	0.129
FN Rate	0.384	0.414	0.027
OAC (%)	80.8	76.35	92.2

derived from Digital Elevation Model) with a higher precision in the analysis significantly contribute to obtaining reliable results associated to landslide susceptibility within the study area. Even though the land cover predictor has a lower precision than the other predictors, in order to avoid the loss of some valuable information regarding the spatial distribution of land cover classes, this conditioning factor was considered in the analysis. Another fact is also considered and that is a differentiated and high number (15) of land cover classes which are more than enough to determine the variability of the landslide susceptibility within the study area. Nevertheless, as mentioned at subsection 2.3, the CORINE Land Cover database is the most common data source from which the land cover predictor is extracted and then used in the assessment of landslide susceptibility across study areas with a surface area higher than 500 km^2.

5. Conclusions

The Kysuca river basin in Slovakia is a multi-hazard area where disaster like flood and landslide are very recurrent. Thus, in the present study, the FDEMATEL-ANP, Naïve Bayes (NB) classifier, and random forest (RF) classifier models were used for LSM in the study area. These three models have not been compared in the available literature so far in terms of landslide susceptibility modeling. An inventory map of landslide was produced based on the historical landslides recorded in the Atlas of Slope Stability Maps of the Slovak Republic and 16 conditioning factors were considered to construct the models. As a result, the landslide susceptibility models produced in the present study could help planners to prevent damage due to landsliding. The major findings of this study are concise as follows: (1) as per the ReliefF feature selection method, the slope was the most important factor of landslide susceptibility modeling in the study area, followed by lithological structure, slope aspect, land cover, distance from the faults, TWI, solar radiation, rainfall, and elevation. On the other hand, SPI was the least important factor in landslide susceptibility, followed by profile curvature, general curvature, plan curvature, distance from rivers, and distance from roads; (2) according to the results of susceptibility modeling and recorded landslide points, most of the landslides occurred at a slope angle between 17.33° and 63.14°, lithological structure of flyschoid rocks, land cover with coniferous forest and transitional woodland-shrub, distance from faults between 0 and 5333.69 m, TWI between 5.31 and 7.31, annual solar radiation between 177,867.40 and 956,475.24 WH/m^2, annual rainfall between 900 and 1250 mm, and at altitudes between 498.18 and 723.37 m; (3) all three models applied in this study resulted in very high accuracy and thus represent promising methods for LSM in the Kysuca river basin; (4) the ROC curves and statistical measures were used to assess the performance and comparison of the models. It was found that the random forest (RF) classifier model achieved the highest accuracy with the highest AUC value (0.954) and ACC value (92.2%) while the Naïve Bayes (NB) classifier model had lower accuracy in comparison to other applied models with lower AUC value (0.909) and ACC value (76.35%).

Applying to all models, it is noted that the eastern and south-eastern parts of the Kysuca river basin are more prone to landslides. Finally, it can be concluded that these models and results of the present study

may be beneficial to the land use planning in the study area as well as other regions with similar environmental settings for landslide risk mitigation and resilience.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gsf.2020.09.004.

References

Ada, M., San, B.T., 2018. Comparison of machine-learning techniques for landslide susceptibility mapping using two-level random sampling (2LRS) in Alakir Catchment area, Antalya, Turkey. Nat. Hazards 90, 237–263.

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

Akgun, A., Türk, N., 2010. Landslide susceptibility mapping for Ayvalik (western Turkey) and its vicinity by multi-criteria decision analysis. Environ. Earth Sci. 61 (3), 595–611. https://doi.org/10.1007/s12665-009-0373-1.

Akgun, A., Sezer, E.A., Nefeslioglu, H.A., Gokceoglu, C., Pradhan, B., 2012. An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. Comput. Geosci. 38 (1), 23–34. https://doi.org/10.1016/j.cageo.2011.04.012.

Althuwaynee, O.F., Pradhan, B., Lee, S.A., 2016. Novel integrated model for assessing landslide susceptibility mapping using CHAID and AHP pair-wise comparison. Int. J. Remote Sens. 37, 1190–1209.

Arabsheibani, R., Kanani Sadat, Y., Abedini, A., 2016. Land suitability assessment for locating industrial parks: a hybrid multi criteria decision-making approach using geographical information system. Geogr. Res. 54 (4), 446–460. https://doi.org/10.1111/1745-5871.12176.

Barančoková, M., Barančok, P., 2019. Landsliding as a limit to possible territorial development in the Kysuce region. Ekológia 38 (4), 301–317. https://doi.org/10.2478/eko-2019-0023.

Barančoková, M., Kenderessy, P., 2014. Assessment of landslide risk using GIS and statistical methods in Kysuce region. Ekologia 33 (1), 26–35. https://doi.org/10.2478/eko-2014-0004.

Błońska, E., Lasota, J., Piaszczyk, W., Wiecheć, M., Klamerus-Iwan, A., 2018. The effect of landslide on soil organic carbon stock and biochemical properties of soil. J. Soils Sediments 18, 2727–2737. https://doi.org/10.1007/s11368-017-1775-4.

Bochníček, O., Borsányi, P., Čepčeková, E., Faško, P., Chmelík, M., Jančovičová, Ľ., Kapolková, H., Labudová, L., Mikulová, K., Mišaga, O., Nejedlík, P., Pribullová, A., Snopková, Z., Šťastný, P., Švec, M., Turňa, M., 2015. Climate Atlas of Slovakia. Slovak Hydro Meteorological Institute, Bratislava.

Bonissone, P., Cadenas, J.M., Garrido, M.C., Díaz-Valladares, R.A., 2010. A fuzzy random forest. Int. J. Approx. Reason. 51, 729–747.

Borrelli, L., Cofone, G., Coscarelli, R., Gullà, G., 2015. Shallow landslides triggered by consecutive rainfall events at Catanzaro strait (Calabria–Southern Italy). J. Maps 11, 730–744.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat. Hazards Earth Syst. Sci. 5, 853–862. https://doi.org/10.5194/nhess-5-853-2005.

Brenning, A., 2012. Improved spatial analysis and prediction of landslide susceptibility: practical recommendations. In: Eberhardt, E., Froese, C., Turner, A.K., Leroueil, S. (Eds.), Landslides and Engineered Slopes: Protecting Society through improved Understanding. Proceedings of the 11th International and 2nd North American Symposium on Landslides and Engineered Slopes, vol. 1. Banff, Canada, 3–8 June 2012. CRC Press/Balkema Leiden, the Netherlands, pp. 789–794.

Brown, M.K., 2012. Landslide Detection and Susceptibility Mapping Using Lidar and Artificial Neural Network Modeling: A Case Study in Glacially Dominated Cuyahoga River Valley, Ohio. Bowling Green State University, Bowling Green, OH, USA.

Bui Tien, D., Pradhan, B., Lofman, O., Revhaug, I., 2012. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and Naïve Bayes models. Math. Probl. Eng., 1–26 https://doi.org/10.1155/2012/974638.

Buša, J., Tornyai, R., Bednarik, M., Greif, V., Rusnák, M., 2019. Hodnotenie zosuvného hazardu pomocou multivariačnej a bivariačnej štatistickej analýzy v Košickej kotline (Západné Karpaty). Geografický časopis 71, 383–405 (in Slovakm with English abstract). https://www.sav.sk/index.php?lang=en&doc=journal-list&part=article_response_page&journal_article_no=17766.

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D.T., Duan, Z., Ma, J.A., 2017a. Comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena 151, 147–160.

Chen, W., Panahi, M., Pourghasemi, H.R., 2017b. Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. Catena 157, 310–324.

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

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

Chen, W., Shahabi, H., Shirzadi, A., Li, T., Guo, C., Hong, H., Li, W., Pan, D., Hui, J., Ma, M., 2018c. A novel ensemble approach of bivariate statistical-based logistic model tree classifier for landslide susceptibility assessment. Geocarto Int. 1–23.

Chen, W., Xie, X., Peng, J., Shahabi, H., Hong, H., Bui, D.T., Duan, Z., Li, S., Zhu, A.X., 2018d. GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method. Catena 164, 135–149.

Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., Wang, X., Bian, H., Zhang, S., Pradhan, B., Ahmad, B.B., 2020. Modeling flood susceptibility using data-driven approaches of naïve Bayes tree, alternating decision tree, and random forest methods. Sci. Total Environ. 701, 134979. https://doi.org/10.1016/j.scitotenv.2019.134979.

Choi, K.Y., Cheung, R.W.M., 2013. Landslide disaster prevention and mitigation through works in Hong Kong. J. Rock Mech. Geotech. Eng. 5 (5), 354–365. https://doi.org/10.1016/j.jrmge.2013.07.007.

Cieslik, K., Shakya, P., Uprety, M., Dewulf, A., Russell, C., Clark, J., Dhital, M.R., Dhakal, A., 2019. Building resilience to chronic landslide hazard through citizen science. Front. Earth Sci. 7, 278. https://doi.org/10.3389/feart.2019.00278.

Colkesen, I., Sahin, E.K., Kavzoglu, T., 2016. Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. J. Afr. Earth Sci. 118, 53–64. https://doi.org/10.1016/j.jafrearsci.2016.02.019.

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

Dang, V.-H., Dieu, T.B., Tran, X.-L., Hoang, N.-D., 2018. Enhancing the accuracy of rainfall-induced landslide prediction along mountain roads with a GIS-based random forest classifier. Bull. Eng. Geol. Environ. 78, 2835–2849.

Dang, V.-H., Hoang, N.-D., Nguyen, L.-M.-D., Bui, D.T., Samui, P., 2020. A novel GIS-based random forest machine algorithm for the spatial prediction of shallow landslide susceptibility. Forests 11 (1), 118. https://doi.org/10.3390/f11010118.

Dehnavi, A., Aghdam, I.N., Pradhan, B., Varzandeh, M.H.M., 2015. A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive Neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. Catena 135, 122–148.

Devkota, K.C., Regmi, A.D., Pourghasemi, H.R., Yoshida, K., Pradhan, B., Ryu, I.C., Dhital, R.R., Althuwaynee, O.F., 2013. Landslide susceptibility mapping using certainty factor, index of entropy and logistic regression models in GIS and their comparison at Mugling–Narayanghat road section in Nepal Himalaya. Nat. Hazards 65, 135–165. https://doi.org/10.1007/s11069-012-0347-6.

Di Traglia, F., Bartolini, S., Artesi, E., Nolesini, T., Ciampalini, A., Lagomarsino, D., Marti, J., Casagli, N., 2018. Susceptibility of intrusion-related landslides at volcanic islands: the Stromboli case study. Landslides 15 (1), 21–29.

Ding, Q., Chen, W., Hong, H., 2017. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. Geocarto Int. 32 (6), 619–639.

Dou, Y., Sarkis, J., Bai, C., 2014. Government green procurement: a Fuzzy-DEMATEL analysis of barriers. In: Kahraman, C., Öztayşi, B. (Eds.), Supply Chain Management Under Fuzziness. Studies in Fuzziness and Soft Computing 313. Springer, Berlin, Heidelberg, pp. 567–589. https://doi.org/10.1007/978-3-642-53939-8_24.

Froude, M.J., Petley, D., 2018. Global fatal landslide occurrence from 2004 to 2016. Nat. Hazard Earth Sys. 18, 2161–2181.

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., Reichenbach, P., 2008. Comparing landslide inventory maps. Geomorphology 94, 268–289.

Ghorbanzadeh, O., Feizizadeh, B., Blaschke, T., 2018. Multi-criteria risk evaluation by integrating an analytical network process approach into GIS-based sensitivity and uncertainty analyses. Geomat. Nat. Hazards Risk 9 (1), 127–151.

Goetz, J., Brenning, A., Petschko, H., Leopold, P., 2015a. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput. Geosci. 81, 1–11.

Goetz, J.N., Brenning, A., Petschko, H., Leopold, P., 2015b. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Comput. Geosci. 81, 1–11. https://doi.org/10.1016/j.cageo.2015.04.007.

Gritzner, M.L., Marcus, W.A., Aspinall, R., Custer, S.G., 2001. Assessing landslide potential using GIS, soil wetness modeling and topographic attributes, Payette River, Idaho. Geomorphology 37, 149–165.

Haque, U., Da Silva, P.F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., Wilopo, W., Andersen, P., Lu, P., Lee, J., Yamamoto, T., Keellings, D., Wu, J.H., Glass, G.E., 2019. The human cost of global warming: Deadly landslides and their triggers (1995–2014). Sci. Total Environ. 682, 673–684.

Hong, H., Liu, J., Bui, D.T., Pradhan, B., Acharya, T.D., Pham, B.T., Zhu, A.-X., Chen, W., Ahmad, B.B., 2018. Landslide susceptibility mapping using J48 decision tree with ADAboost, bagging and rotation forest ensembles in the Guangchang area (China). Catena 163, 399–413.

Hussain, G., Singh, Y., Singh, K., Bhat, G.M., 2019. Landslide susceptibility mapping along national highway-1 in Jammu and Kashmir State (India). Innovat. Infrastruct. Solut. 4, 59. https://doi.org/10.1007/s41062-019-0245-9.

Iovine, G., Cohen, D., 2014. Advanced methods in landslide modelling. Nat. Hazards 73, 1–4. https://doi.org/10.1007/s11069-014-1320-3.

Iovine, G., Sheridan, M.F., 2009. Special issue 2007 in Natural Hazards on "Modelling and simulation of dangerous phenomena, and innovative techniques for hazard mapping and mitigation". Nat. Hazards 50 (3), 409–411.

Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F., Nasseri, M., 2019. A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. J. Hydrol. 572, 17–31. https://doi.org/10.1016/j.jhydrol.2019.02.034.

Kavoura, K., Sabatakakis, N., 2020. Investigating landslide susceptibility procedures in Greece. Landslides 17 (1), 127–145.

Kavzoglu, T., Colkesen, I., Sahin, E.K., 2019. Machine learning techniques in landslide susceptibility mapping: a survey and a case study. In: Pradhan, S., Vishal, V., Singh, T. (Eds.), Landslides: Theory, Practice and Modelling. Advances in Natural and Technological Hazards Research 50. Springer, Cham, pp. 283–301. https://doi.org/10.1007/978-3-319-77377-3_13.

Kayastha, P., Dhital, M.R., De Smedt, F., 2013. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: a case study from the Tinau watershed, West Nepal. Comput. Geosci. 52, 398–408. https://doi.org/10.1016/j.cageo.2012.11.003.

Khosravi, K., Nohani, E., Maroufinia, E., Pourghasemi, H.R., 2016. A GIS-based flood susceptibility assessment and its mapping in Iran: a comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. Nat. Hazards 83 (2), 947–987. https://doi.org/10.1007/s11069-016-2357-2.

Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. J. Hydrol. 573, 311–323.

Kijewska, K., Torbacki, W., Iwan, S., 2018. Application of AHP and DEMATEL methods in choosing and analysing the measures for the distribution of goods in Szczecin Region. Sustainability 10 (7), 2365. https://doi.org/10.3390/su10072365.

Kopecký, M., Martinčeková, T., Šimeková, J., Ondrášik, M., 2008. Landslide atlas – results of the geological project. Proceedings of the 6th Conference Geology and Environment, Bratislava, pp. 105–110.

Lee, S., Lee, M.-J., Jung, H.-S., 2017. Data mining approaches for landslide susceptibility mapping in Umyeonsan, Seoul, South Korea. Appl. Sci. 7 (7), 683. https://doi.org/10.3390/app7070683.

Leonardi, G., Palamara, R., Cirianni, F., 2016. Landslide susceptibility mapping using a fuzzy approach. Procedia Eng. 161, 380–387. https://doi.org/10.1016/j.proeng.2016.08.578.

Ließ, M., Glaser, B., Huwe, B., 2011. Functional soil-landscape modelling to estimate slope stability in a steep Andean mountain forest region. Geomorphology 132 (3), 287–299.

Lin, K.-P., Tseng, M.-L., Pai, P.-F., 2018. Sustainable supply chain management using approximate fuzzy DEMATEL method. Resour. Conserv. Recycl. 128, 134–142.

Liščák, P., Káčer, Š., 2013. Developments in landslides inventory and registry in Slovakia. In: Margottini, C., Canuti, P., Sassa, K. (Eds.), Landslide Science and Practice. Springer, Berlin, Heidelberg, pp. 65–71 https://doi.org/10.1007/978-3-642-31325-7_8.

Mazúr, E., Lukniš, M., 1986. Geomorfologické členenie SSR a ČSSR. Časť Slovensko. Slovenská Kartografia, Bratislava (in Slovak).

Meneses, B.M., Pereira, S., Reis, E., 2019. Effects of different land use and land cover data on the landslide susceptibility zonation of road networks. Nat. Hazards Earth Syst. Sci. 19 (3), 471–487.

Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., Kanevski, M., 2014. Machine learning feature selection methods for landslide susceptibility mapping. Math. Geosci. 46, 33–57. https://doi.org/10.1007/s11004-013-9511-0.

Ministry of Environment of the Slovak Republic, 2018. Predbežné hodnotenie povodňového rizika v čiastkovom povodí Váhu – aktualizácia 2018 (in Slovak). (Accessed 2020.01.18). https://www.minzp.sk/files/sekcia-vod/hodnotenie-rizika-2018/vah/phpr-vah.pdf

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

Moore, I.D., Wilson, J.P., 1992. Length-slope factors for the revised Universal Soil loss equation: simplified method of estimation. J. Soil Water Conserv. 47, 423–428.

Moore, I.D., Grayson, R.B., Ladson, A.R., 1991. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrol. Process. 5 (1), 3–30. https://doi.org/10.1002/hyp.3360050103.

Moosavi, V., Niazi, Y., 2016. Development of hybrid wavelet packet-statistical models (WP-SM) for landslide susceptibility mapping. Landslides 13 (1), 97–114. https://doi.org/10.1007/s10346-014-0547-0.

Myronidis, D., Papageorgiou, C., Theophanous, S., 2016. Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). Nat. Hazards 81 (1), 245–263.

Neaupane, K.M., Piantanakulchai, M., 2006. Analytic network process model for landslide hazard zonation. Eng. Geol. 85 (3), 281–294.

Nefeslioglu, H., Gokceoglu, C., Sonmez, H., 2008b. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. Eng. Geol. 97, 171–191.

Nefeslioglu, H.A., Duman, T.Y., Durmaz, S., 2008a. Landslide susceptibility mapping for a part of tectonic Kelkit Valley (Eastern Black Sea region of Turkey). Geomorphology 94 (3–4), 401–418. https://doi.org/10.1016/j.geomorph.2006.10.036.

Nsengiyumva, J.B., Luo, G., Nahayo, L., Huang, X., Cai, P., 2018. Landslide susceptibility assessment using spatial multi-criteria evaluation model in Rwanda. Int. J. Environ. Res. Public Health 15 (2), 243.

Nsengiyumva, J.B., Luo, G., Amanambu, A.C., Mind'je, R., Habiyaremye, G., Karamage, F., Ochege, F.U., Mupenzi, C., 2019. Comparing probabilistic and statistical methods in landslide susceptibility modeling in Rwanda/Centre-Eastern Africa. Sci. Total Environ. 659, 1457–1472. https://doi.org/10.1016/j.scitotenv.2018.12.248.

Othman, A.N., Naim, W.M., Noraini, S., 2012. GIS based multi-criteria decision making for landslide hazard zonation. Procedia Soc. Behav. Sci. 35, 595–602. https://doi.org/10.1016/j.sbspro.2012.02.126.

Papaioannou, G., Vasiliades, L., Loukas, A., 2015. Multi-criteria analysis framework for potential flood prone areas mapping. Water Resour. Manag. 29 (2), 399–418.

Petschko, H., Brenning, R., Bell, J., Goetz, T., 2014. Glade assessing the quality of landslide susceptibility maps – case study Lower Austria. Nat Hazards Earth Syst. Sci. 14, 95–118. https://doi.org/10.5194/nhess-14-95-2014.

Pham, B.T., Tien Bui, D., Indra, P., Dholakia, M.B., 2015. 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 (11), 338–344.

Pham, B.T., Bui, D.T., Prakash, I., Dholakia, M., 2016a. Rotation forest fuzzy rule-based classifier ensemble for spatial prediction of landslides using GIS. Nat. Hazards 83, 97–127.

Pham, B.T., Pradhan, B., Bui, D.T., Prakash, I., Dholakia, M.A., 2016b. Comparative study of different machine learning methods for landslide susceptibility assessment: a case study of Uttarakhand area (India). Environ. Model. Softw. 84, 240–250.

Pham, B.T., Bui, D.T., Pourghasemi, H.R., Indra, P., Dholakia, M.B., 2017a. 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 (1–2), 255–273.

Pham, B.T., Bui, D.T., Prakash, I., Dholakia, M., 2017b. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena 149, 52–63.

Pham, B.T., Prakash, I., Bui, D.T., 2018a. Spatial prediction of landslides using a hybrid machine learning approach based on random subspace and classification and regression trees. Geomorphology 303, 256–270.

Pham, B.T., Jaafari, A., Prakash, I., Bui, D.T., 2018b. A novel hybrid intelligent model of support vector machines and the multiboost ensemble for landslide susceptibility modeling. Bull. Eng. Geol. Environ. 1–22.

Pham, B.T., Bui, D.T., Prakash, I., 2018c. Bagging based support vector machines for spatial prediction of landslides. Environ. Earth Sci. 77, 146.

Pham, B.T., Shirzadi, A., Bui, D.T., Prakash, I., Dholakia, M.A., 2018d. Hybrid machine learning ensemble approach based on a radial basis function neural network and rotation forest for landslide susceptibility modeling: a case study in the Himalayan area, India. Int. J. Sediment Res. 33, 157–170.

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

Poudyal, C.P., Chang, C., Oh, H.-J., Lee, S., 2010. Landslide susceptibility maps comparing frequency ratio and artificial neural networks: a case study from the Nepal Himalaya. Environ. Earth Sci. 61, 1049–1064.

Pourghasemi, H.R., Pradhan, B., Gokceoglu, C., 2012. Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. Nat. Hazards 63, 965–996.

Pradhan, A.M.S., Kim, Y.-T., 2014. Relative effect method of landslide susceptibility zonation in weathered granite soil: a case study in Deokjeok-ri Creek, South Korea. Nat. Hazards 72, 1189–1217.

Quiroz, J.C., Mariun, N., Mehrjou, M.R., Izadi, M., Misron, N., Mohd Radzi, M.A., 2018. Fault detection of broken rotor bar in LS-PMSM using random forests. Measurement 116, 273–280.

Ramani, S.E., Pitchaimani, K., Gnanamanickam, V.R., 2011. GIS based landslide susceptibility mapping of Tevankarai Ar sub-watershed, Kodaikkanal, India using binary logistic regression analysis. J. Mt. Sci. 8 (4), 505–517.

Razavizadeh, S., Solaiman, K., Massiron, M., Kavian, A., 2017. Mapping landslide susceptibility with frequency ratio, statistical index, and weights of evidence models: a case study in northern Iran. Environ. Earth Sci. 76, 499. https://doi.org/10.1007/s12665-017-6839-7.

Robnik-Šikonja, M., Kononenko, I., 2003. Theoretical and empirical analysis of ReliefF and RReliefF. Mach. Learn. 53, 23–69.

Saaty, T.L., 1996. Decision Making with Dependence and Feedback: The Analytic Network Process. 4922. RWS Publications, Pittsburgh.

Saito, H., Uchiyama, S., Hayakawa, Y.S., Obanawa, H., 2018. Landslides triggered by an earthquake and heavy rainfalls at Aso volcano, Japan, detected by UAS and SfM-MVS photogrammetry. Progr. Earth Planet. Sci. 5, 15. https://doi.org/10.1186/s40645-018-0169-6.

Sangaiah, A.K., Gopal, J., Basu, A., Subramaniam, P.R., 2017. An integrated fuzzy DEMATEL, TOPSIS, and ELECTRE approach for evaluating knowledge transfer effectiveness with reference to GSD project outcome. Neural Comput. & Applic. 28 (1), 111–123.

Sarker, A.A., Rashid, A.K.M.M., 2013. Landslide and flashflood in Bangladesh. In: Shaw, R., Mallick, F., Islam, A. (Eds.), Disaster Risk Reduction Approaches in Bangladesh. Disaster Risk Reduction (Methods, Approaches and Practices). Springer, Tokyo, pp. 165–189 https://doi.org/10.1007/978-4-431-54252-0_8.

Schilirò, L., Montrasio, L., Mugnozza, G.S., 2016. Prediction of shallow landslide occurrence: validation of a physically-based approach through a real case study. Sci. Total Environ. 569, 134–144.

Segoni, S., Pappafico, G., Luti, T., Catani, F., 2020. Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization. Landslides 17, 2443–2453.

Shao, L., 2019. Geological disaster prevention and control and resource protection in mineral resource exploitation region. Int. J. Low-Carbon Technol. 14 (2), 142–146. https://doi.org/10.1093/ijlct/ctz003.

Shirzadi, A., Chapi, K., Shahabi, H., Solaimani, K., Kavian, A., Ahmad, B.B., 2017. Rock fall susceptibility assessment along a mountainous road: an evaluation of bivariate statistic, analytical hierarchy process and frequency ratio. Environ. Earth Sci. 76, 152.

Shirzadi, A., Soliamani, K., Habibnejhad, M., Kavian, A., Chapi, K., Shahabi, H., Chen, W., Khosravi, K., Pham, B.T., Pradhan, B., Ahmad, A., Ahmad, B.B., Tien Bui, D., 2018. Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. Sensors 18 (11), 3777. https://doi.org/10.3390/s18113777.

Šimeková, J., Martinčeková, T., Abrahám, P., Gejdoš, T., Grencíková, A., Grman, D., Hrašna, M., Jadroň, D., Záthurecký, A., Kotrčková, E., Lišťák, P., Malgot, J., Masný, M., Mokrá, M., Petro, Ľ., Polaščinová, E., Solčiansky, R., Kopecký, M., Žabková, E., Wanieková, D., 2006. Atlas of slope stability maps of the Slovak Republic 1:50,000. INGEO – ighp, spol. S r. o., Žilina, Slovakia (in Slovak).

Soni, J., Ansari, U., Sharma, D., Soni, S., 2011. Predictive data mining for medical diagnosis: an overview of heart disease prediction. Int. J. Comput. Appl. 17 (8), 43–48.

Soria, D., Garibaldi, J.M., Ambrogi, F., Biganzoli, E.M., Ellis, I.O., 2011. A 'non-parametric' version of the naive Bayes classifier. Knowl.-Based Syst. 24 (6), 775–784.

Sujatha, E.R., Kumaravel, P., Rajamanickam, V., 2012. Landslide susceptibility mapping using remotely sensed data through conditional probability analysis using seed cell and point sampling techniques. J. Ind. Soc. Remote Sens. 40 (4), 669–678.

Sumrit, D., Anuntavoranich, P., 2013. Using DEMATEL method to analyze the causal relations on technological innovation capability evaluation factors in Thai technology-based firms. Int. Trans. J. Eng. Manag. Appl. Sci. Technol. 4 (2), 81–103 Available at. http://TuEngr.com/V04/081-103.pdf.

Tehrany, M.S., Pradhan, B., Jebur, M.N., 2013. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. J. Hydrol. 504, 69–79.

Tevi, G., Stoica, A., 2019. Multi-criteria GIS based methodology used for landslide vulnerability evaluation, case study Prahova County, Romania. Geophysical Research Abstracts. 21. EGU Gen. Assemb.

Tien Bui, D., Shahabi, H., Shirzadi, A., Chapi, K., Hoang, N.D., Pham, B., Bui, Q.T., Tran, C.T., Panahi, M., Bin Ahamd, B., Saro, L., 2018. A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. Remote Sens. 10, 1538. https://doi.org/10.3390/rs10101538.

Tien Bui, D., Tsangaratos, P., Nguyen, V.T., Van Liem, N., Trinh, P.T., 2019. Comparing the prediction performance of a Deep Learning Neural Network model with conventional machine learning models in landslide susceptibility assessment. Catena 188, 104426.

Trigila, A., Iadanza, C., Esposito, C., Scarascia-Mugnozza, G., 2015. Comparison of Logistic Regression and Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri (NE Sicily, Italy). Geomorphology 249, 119–136.

Tseng, M.-L., Lin, Y.H., 2009. Application of fuzzy DEMATEL to develop a cause and effect model of municipal solid waste management in Metro Manila. Environ. Monit. Assess. 158 (1), 519–533.

Tzeng, G.H., Chiang, C.H., Li, C.W., 2007. Evaluating intertwined effects in e-learning programs: a novel hybrid MCDM model based on factor analysis and DEMATEL. Expert Syst. Appl. 32 (4), 1028–1044.

Vakhshoori, V., Pourghasemi, H.R., Zare, M., Blaschke, T., 2019. Landslide susceptibility mapping using GIS-based data mining algorithms. Water 11 (11), 2292. https://doi.org/10.3390/w11112292.

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

Vinodh, S., Balagi, T.S., Patil, A., 2016. A hybrid MCDM approach for agile concept selection using fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS. Int. J. Adv. Manuf. Technol. 83 (9–12), 1979–1987.

Vojteková, J., Vojtek, M., 2020. Assessment of landslide susceptibility at a local spatial scale applying the multi-criteria analysis and GIS: a case study from Slovakia. Geom. Nat. Hazards Risk 11 (1), 131–148.

Xiao, T., Yin, K., Yao, T., Liu, S., 2019. Spatial prediction of landslide susceptibility using GIS-based statistical and machine learning models in Wanzhou County, three Gorges Reservoir, China. Acta Geochim. 38, 654–669. https://doi.org/10.1007/s11631-019-00341-1.

Yang, Y.P., Shieh, H.M., Leu, J.D., Tzeng, G.H., 2008. A novel hybrid MCDM model combined with DEMATEL and ANP with applications. Int. J. Operat. Res. 5 (3), 160–168.

Yilmaz, C., Topal, T., Süzen, M.L., 2012. GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). Environ. Earth Sci. 65, 2161–2178.

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

Youssef, A.M., Pourghasemi, H.R., El-Haddad, B.A., Dhahry, B.K., 2016. Landslide susceptibility maps using different probabilistic and bivariate statistical models and comparison of their performance at Wadi Itwad Basin, Asir Region, Saudi Arabia. Bull. Eng. Geol. Environ. 75 (1), 63–87.

Zadeh, L.A., 1965. Fuzzy sets. Inf. Control. 8 (3), 338–353.

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

Zêzere, J., Pereira, S., Melo, R., Oliveira, S., Garcia, R., 2017. Mapping landslide susceptibility using data-driven methods. Sci. Total Environ. 589, 250–267.