Article
Vol. 22, No. 6, p. 1053–1067, December 2018
https://doi.org/10.1007/s12303-018-0052-x
pISSN 1226-4806 eISSN 1598-7477
Geosciences Journal GJ

Analysis of the relationships between topographic factors and landslide occurrence and their application to landslide susceptibility mapping: a case study of Mingchukur, Uzbekistan

Azam Kadirhodjaev[1], Prima Riza Kadavi[2], Chang-Wook Lee[2], and Saro Lee[3,4]*

[1]The State Committee of the Republic of Uzbekistan on Geology and Mineral Resources (Goscomgeology), 11, T. Shevchenko, Tashkent 100060, Uzbekistan
[2]Division of Science Education, Kangwon National University, 1 Kangwondaehak-gil, Chuncheon-si, Gangwon-do 24341, Republic of Korea
[3]Geological Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124 Gwahang-no, Yuseong-gu, Daejeon 34312, Republic of Korea
[4]Department of Geophysical Exploration, Korea University of Science and Technology, 217 Gajeong-roYuseong-gu, Daejeon 34113, Republic of Korea

ABSTRACT: This paper uses a probability-based approach to study the spatial relationships between landslides and their causative factors in the Mingchukur area, Bostanlik districts of Tashkent, Uzbekistan. The approach is based on digital databases and incorporates methods including probability analysis, spatial pattern analysis, and interactive mapping. First, an object-oriented conceptual model for describing landslide events is proposed, and a combined database of landslides and environmental factors is constructed by integrating various databases within a unifying conceptual framework. The frequency ratio probability model and landslide occurrence data are linked for interactive, spatial evaluation of the relationships between landslides and their causative factors. In total, 15 factors were analyzed, divided into topography, hydrology, and geology categories. All analyzed factors were also divided into numerical and categorical types. Numerical factors are continuous and were evaluated according to their R^2 values. A landslide susceptibility map was constructed based on conditioning factors and landslide occurrence data using the frequency ratio model. Finally, the map was validated and the accuracy showed the satisfactory value of 83.3%.

Key words: landslide susceptibility, topography, GIS, frequency ratio, Uzbekistan

Manuscript received August 5, 2018; Manuscript accepted September 11, 2018

1. INTRODUCTION

Currently, the eight most widespread and dangerous exogenous geological processes in Uzbekistan are landslides, avalanches, debris break up in alpine lakes, mudflows, ravine erosion, subsidence, dips, and flooding. Overall, 21.3% (90,000 km^2) of Uzbekistan is mountainous, and 40% of the mountainous territory is at risk from landslides, avalanches, and mudflows and associated processes There is a danger of damage from landslides in a 15,000 km^2area of mountainous territory in Uzbekistan, wherein there are 2,000 landslide areas; in the past 50 years, 8,300 landslide displacements have been recorded. The characteristics of these displacements divided into some case: 70% were small surface landslides; 10–12% were large landslides (> 100,000 m^3); 18–20% were average-sized landslides; 77% of the landslides developed in loess, and 23% in clayey rocks; 65% of the landslides were caused by snowmelt, atmospheric precipitation, and groundwater; 15–20% of the landslides were triggered by earthquakes; 20–25% of the landslides were caused by technogenic factors.

Landslides, characterized by rapid development and geographical and temporal unpredictability, represent the most dangerous exogenous geological processes in Uzbekistan. The increasing

***Corresponding author:**
Saro Lee
Geological Research Division, Korea Institute of Geoscience and Mineral Resources (KIGAM), 124 Gwahang-no, Yuseong-gu, Daejeon 34312, Republic of Korea
Tel.:+82-42-868-3057, E-mail: leesaro@kigam.re.kr

scale and intensity, and the annual recurrence pattern, of landslides now represents a constant threat to populated areas, sites of importance to the national economy, and agricultural lands; this has increased the frequency of emergency situations and significant material damage. The State Service of the Republic of Uzbekistan for Monitoring Dangerous Geological Processes, which monitors these processes in the mountainous and foothill areas of the country, is part of the Goscomgeology System and carries out its activities in accordance with the Decree of the Cabinet of Ministers, of November 13, 2017 (No. 909),via seven tracking stations. The monitoring efforts extend to 870 sites, including 382 settlements, 47 health facilities, 343 motor and railroad facilities, 51 irrigation canals, and 47 mining and hydraulic facilities. In the past 17 years (2001–2017), there have been 840 slope processes with a volume exceeding 1,000 m^3 in foothill areas. Surface displacements of 1,000 to 3,000 m^3 accounted for 37.2% of these processes, while cracks with lengths of 80–120 m accounted for a further 21%, and large landslides of 100,000–500,000 m^3 for 9.4%; landslides exceeding 1.0 million m^3 in extent represent 3.3% of all documented processes.

Remote sensing and GIS tools are very useful for analyzing various natural disasters in the world (Park and Kyriakidis, 2008; Oh et al., 2009; Lee et al., 2015; Oh and Lee, 2017; Lee et al., 2018). Analyzing the relationships between landslides and their various causal factors provides insight into landslide mechanisms, and can also serve as a basis for future landslide predictions and landslide hazard assessments. Over the past few decades, many researchers have introduced efficient and reliable methods for creating landslide susceptibility maps. These methods include the frequency ratio (FR) (Lee et al., 2015, 2016; Hong et al., 2016; Oh et al., 2017), logistic regression (Chen et al., 2016; Lee et al., 2016; Lee and Lee, 2017), decision trees (Lee and Park, 2013; Chen at al., 2018b; Kim et al., 2018; Truong et al., 2018), support vector machines (Lee et al., 2017; Tien Bui et al., 2017; Chen et al., 2018a; Mezaal and Pradhan, 2018), and artificial neural networks (Pradhan and Lee, 2010a; Lee et al., 2017; Pham et al., 2017; Polykretis and Chalkias, 2018).

In this study, the relationships between landslides and their causal factors were analyzed using the FR model. However, currently available approaches have limitations in describing and analyzing temporal-spatial patterns of the relationships between landslides and their antecedents. Therefore, there is a need to improve existing methods and develop new approaches. This study aimed to develop a probability analysis approach for interactive, exploratory analysis of the relationships between landslides and their causal factors, including in terms of their strength. In this paper, environmental and landslide conditions in the study area are described first. Then, the principles and methods used for the analysis are described in detail. Finally, the results of a case study (Mingchukur) are presented and analyzed, and the advantages and disadvantages of the analysis approach, as well directions for further research, are discussed.

2. STUDY AREA

The relief of Bostanlik District, located in the Tashkent region of Uzbekistan, is relatively uniform and comprised mainly of hills, mountains, and highlands (Fig. 1). The most hazardous area in the district is the Mingchukur landslide area, located in the center of the Brichmulla Depression on the northern shore of the Charvak Reservoir, 3.5 km from the Charvak hydropower station dam.

The Mingchukur landslide area covers the lower part of the slope of an ancient landslide, also on the northern shore of the Charvak Reservoir. Data pertaining to the formation of the landslide area indicate that it is ancient (QIII) but has continued to develop up to the present. Its height above the Chirchik River channel is 300 m. The upper part of the landslide area is at an elevation of 1,080 m, and the lower part is at 760 m. The shape of the landslide area is concave-convex, and its average slope angle is 10–70°. According to kinematic analysis, the landscape in this area is blocked and multistage. The parameters of the landslide are as follows: overall length = 700 m, length of mass surface = 500 m, length of disruption surface = 200 m, depth of drifted mass = 30 m, depth of surface = 50 m, maximum width of drifted mass = 1,100 m, total area = 770,000 m^2, volume = 38,500,000 m^3.

The slope formed at the landslide surface consists of conglomerate and gravel covered by loamy soil. The thickness of the conglomerate and gravel layer is 5–20 m, and that of the loam layer is 5–10 m. After onset of the formation of continental deposits, thick continental strata of Neogene molasse consisting of siltstone, sandstone, clay, and gravelites were deposited. The geologic and tectonic structure of the landslide is attributable to the flexural rupture zone of the Pskem Fault. In the height interval of 1,000–1,200 m, more than 10 springs emerge, testament to a groundwater horizon distributed in the middle Quaternary conglomerate-gravels. Slope steepness ranges from 17–200 m. Precipitation of the Mingchukur landslide began in the first year of operation of the Charvak Reservoir (Rakhmatullaev et al., 2013). The aforementioned Tashkent region is the most landslide-prone area (Fig. 2), and attention should be paid to the Mingchukur landslide because further development thereof could result in negative outcomes.

3. DATA

3.1. Landslide Inventory Map

Landslide inventory maps are necessary to assess the

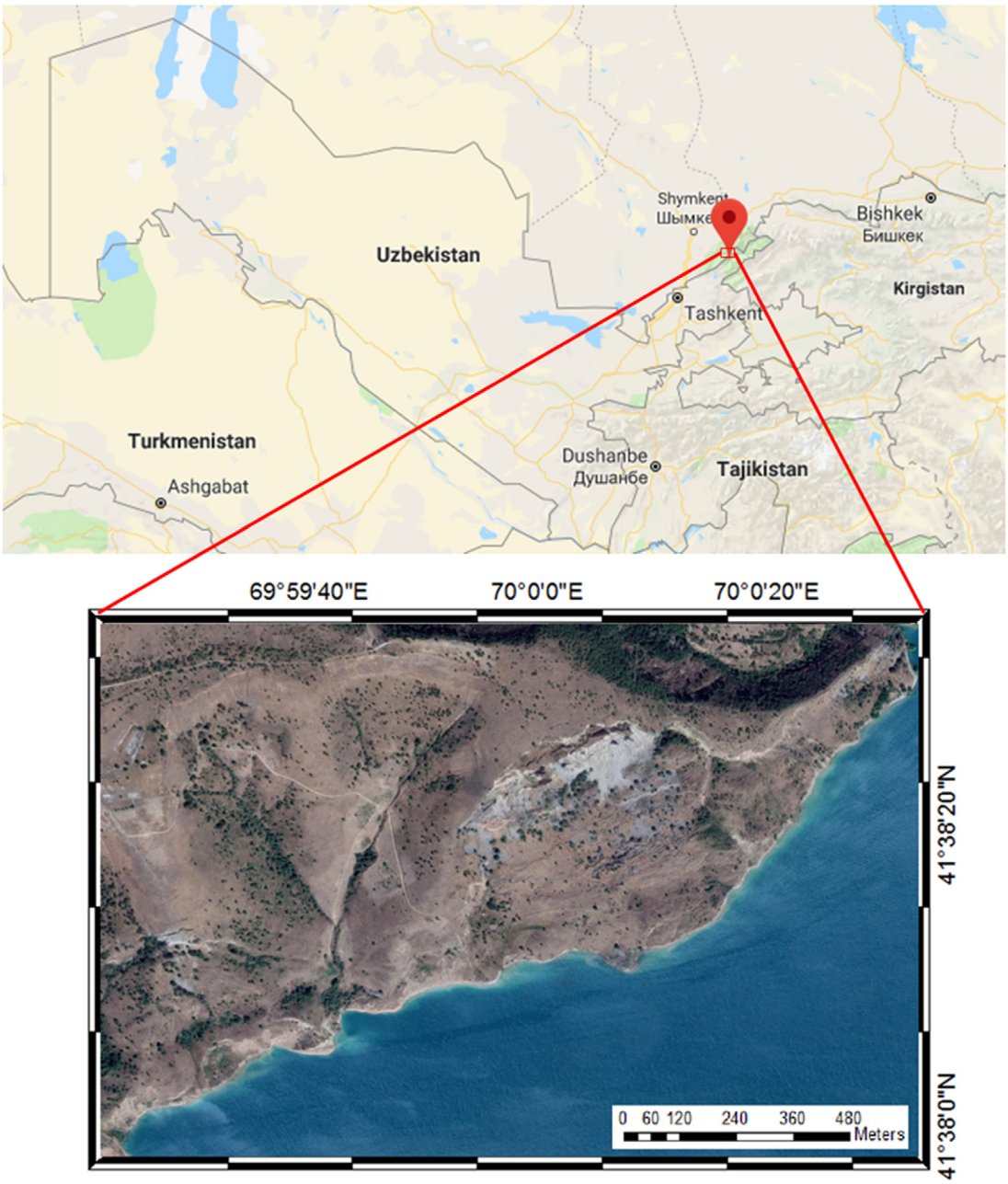

Fig. 1. Location of the study area from Google map.

relationships between landslide occurrences and their predisposing factors. A landslide inventory map was constructed using 5,520 landslide points, which were identified by interpretation of aerial photographs and survey data (Oh, 2010). Given this high number of landslide points, the map is very detailed, because the Mingchukur area is small. The landslide point data formed a boundary indicating the landslide area. The landslide points were plotted using a digital elevation model (DEM) for spatial processing (Fig. 3).

3.2. Landslide Conditioning Factors

Environmental factors are important for the determination of landslide-susceptible areas. In this study, 15 conditioning factors related to landslide occurrence were considered. However, although important, the spatial distributions of these factors were difficult to determine. The locations of landslides and areas in which environmental factors exert an effect are denoted by 1×1 m pixels. Topographic and hydrologic factors were extrapolated from a DEM by exploiting the terrain analysis

https://doi.org/10.1007/s12303-018-0052-x

Fig. 2. Landslide occurrences in the study area from field photograph.

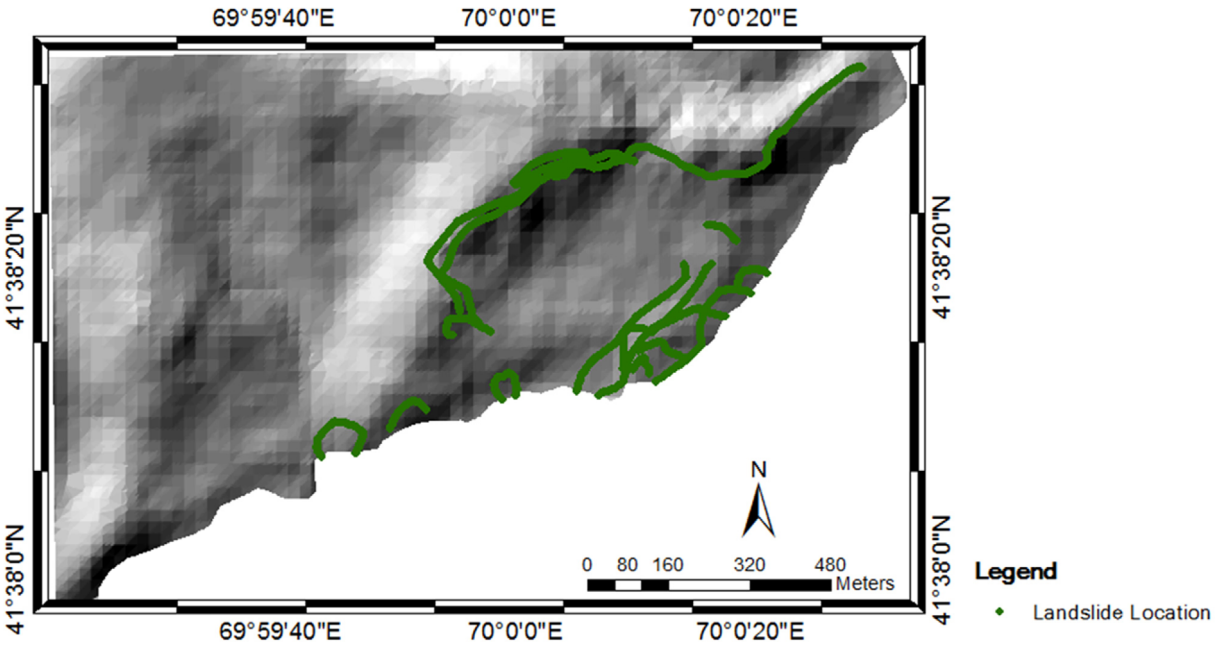

Fig. 3. Digital elevation model (DEM) and landslide occurrences in the study area.

capabilities of the SAGA GIS module (Table 1), and geologic factors were extracted from geological maps. The factors are listed in Table 1.

Slope aspect, which indicates the direction of a slope, has been considered in several relevant studies (Park and Lee, 2014; Lee et al., 2017; Oh and Lee, 2017). Slope aspect refers to the maximum

Table 1. Data layer related to landslide of study area

Category	Factors	Data Type	Scale
Topographic factors	Slope aspect Digital elevation model (DEM) Diurnal anisotropic heating (DAH) Gradient Local downslope curvature (LDC) Local upslope curvature (LUC) Mid slope position (MSP) Slope angle Surface area Valley depth	Grid	1: 1,000
Hydrologic factors	Catchment area Flow path length LS factor Topographic wetness index (TWI)	Grid	1: 1,000
Geology	Lithology	Polygon	1:25,000

rate of change in the downslope direction from a given cell to its neighbors; it can be thought of as the slope direction (Fig. 4a). The values of each cell in the output raster indicate the compass direction of the surface faces at a given location; this is measured clockwise from 0° (due north) to 360° (also due north), thus coming full circle. Flat areas having no downslope direction are assigned a value of −1.

The Diurnal Anisotropic heating (DAH) index was developed for application to the Northern Hemisphere (Fig. 4b). The DAH provides an index of the effects of both temperature and topographic solar radiation, which together determine snowmelt on particular slopes. In relation to landslides, DAH indicates the slope direction that is exposed to heat from the sun, where such exposure affects the material on the slope surface (Cristea et al., 2017).

A DEM is a specialized database that can represent the relief of a surface between points of known elevation (Fig. 4c). By interpolating known elevation data based on ground surveys and photogrammetric data capture, a rectangular DEM grid can be created. DEMs are typically used to represent bare-earth terrain, i.e., that devoid of vegetation and manmade features.

Gradient is related to slope. Gradient can refer to the change in a two-point slope within a certain range (Fig. 4d). The gradient change within a specific area on a map corresponds to the lay of the land, allowing geologists and environmentalists to determine any effect of the gradient of a specific area on those that surround it. Erosion constitutes a good example of why it is important to ascertain the gradient of specific areas (Burrough and McDonell, 1998).

The curvature function gives the shape, or curvature, of the slope. A given part of a surface can be concave or convex, and this can be determined by examining the curvature value. The curvature is calculated by computing the second derivative of the surface. The output of the curvature function can be used to describe the physical characteristics of a drainage basin, to understand erosion and runoff processes. The curvature value can be used to determine soil erosion patterns and the distribution of water on land (Moore et al., 1991). The local curvature of a cell is given by the sum of the gradients of its neighboring cells. Downslope curvature is the distance-weighted average local curvature in a cell's downslope area, based on multiple flow direction algorithms (Fig. 4e) (Freeman, 1994). Similar to local downslope curvature, upslope curvature is the distance-weighted average local curvature in a cell's upslope area, based on multiple flow direction algorithms (Fig. 4f) (Freeman, 1994).

Mid-slope position (MSP) is the intermediate slope position between high and low positions (Fig. 4g). MSP is often related to the topographic position index (TPI), which is being increasingly used to measure topographic slope positions and automate landform classifications. TPI measures the relative topographic position of the central point (MSP), as the difference between the elevation at this point and the mean elevation within a predetermined neighborhood (Wilson, 2000).

Slope angle indicates the steepness of a hill (Fig. 4h). The slope angle at a particular location is the maximum rate of change in elevation between that location and its surrounding areas. Slope angle can be expressed in degrees or as a percentage. The Slope angle GIS tool calculates the maximum rate of change in value from a given cell to its neighbors. Essentially, the maximum change in elevation between the cell and its eight neighbors corresponds to the steepest downhill descent from that cell (Burrough and McDonell, 1998).

Surface area is a fundamental parameter derived from terrain analysis by geographical information systems (GIS), used for modeling real world terrain (Fig. 4i). Surface area is also the basis for measuring the topographic roughness of a landscape (Jeff, 2004).

Valley depth is the difference between the elevation and an interpolated ridge level (Fig. 4j). The ridge level interpolation in

https://doi.org/10.1007/s12303-018-0052-x

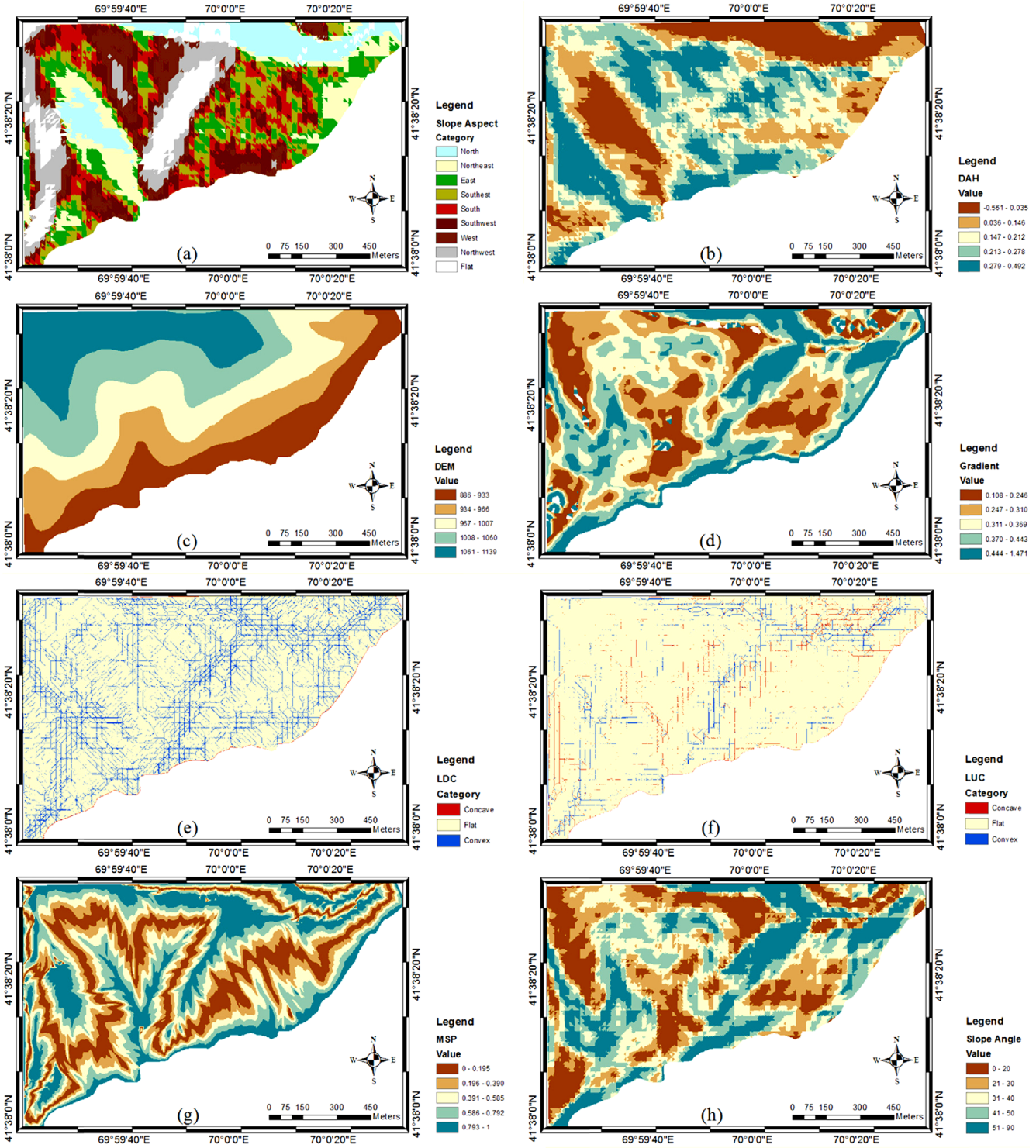

Fig. 4. Spatial database of the landslide causative factors. (a) slope aspect, (b) DAH, (c) DEM, (d) gradient, (e) local downslope curvature, (f) local upslope curvature, (g) MSP, (h) slope angle, (i) surface area, (j) valley depth, (k) catchment area, (l) flow path length, (m) LS factor, (n) TWI, (o) geology.

this study was based on the algorithm of the Vertical Distance to Channel Network tool (Conrad et al., 2015) and involves the following steps: Definition of the ridge cells (using the Strahler order of the inverted DEM); Interpolation of the ridge level; Subtraction of the original elevations from the calculated ridge level.

The catchment area is a hydrological unit. All drops of precipitation falling in a catchment area eventually join the same

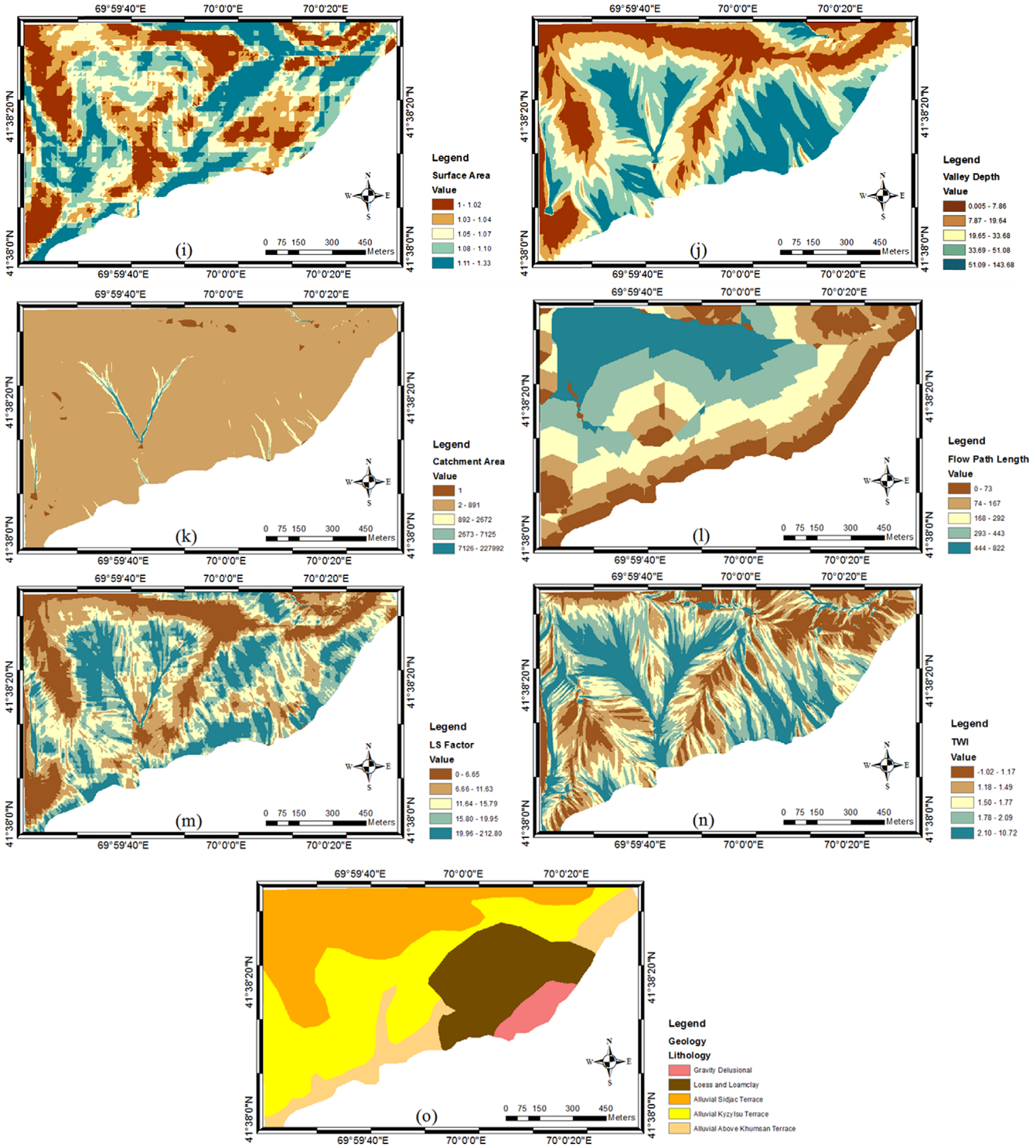

Fig. 4. (continued).

river flowing toward the sea, if the drop does not evaporate; however, this can take a very long time. Catchment areas are separated from each other by watersheds (Fig. 4k). A watershed is a natural dividing line running along the highest points in an area. Catchments can be divided into sub-catchments, also according to lines of elevation. Recursive processing of cells can be conducted for calculation of flow accumulation and related parameters. The set of algorithms used herein recursively processes all upwardly connected cells, until all cells in the DEM have been processed (Carillo et al., 2011).

A primary function of the Flow Length GIS tool is to calculate the length of the longest flow path within a given basin (Fig. 4l).

https://doi.org/10.1007/s12303-018-0052-x

This measure is often used to calculate the time-concentration profile of a basin. This is done using the Upstream GIS tool. This tool can also be used to create distance-area diagrams of hypothetical rainfall and runoff events, using the weight raster as an impedance analyzer of downslope flow (Freeman, 1991; Quinn et al., 1991).

The LS factor measures the effect of slope steepness, and the L factor describes the impact of slope length. The combined LS factor describes the effect of topography on soil erosion (Fig. 4m). Here in, the LS calculation was performed using the equation of Desmet and Govers (1996).

The topographic wetness index (TWI) is a steady-state wetness index commonly used to quantify the influence of topography on hydrological processes (Fig. 4n). The TWI has been applied to compound terrain derivatives in hydrology and ecology studies (Buchanan et al., 2014).

The geology of a rock or lithological unit refers to its physical characteristics, such as color, texture, grain size, and composition, which are visible in outcrop, hand or core samples, or under low-magnification microscopy, (Fig. 4o) (Bates and Jackson, 1984). The physical characteristics of rocks has a correlation to porosity and permeability, the physical properties of rocks also affect weathering resistance. Because various lithological units have different susceptibilities to active geomorphologic processes, such as landslides, lithology is important in landslide occurrence (Lee et al., 2016; Chen et al., 2018a). Therefore, lithology was considered as the main geological parameter in this study.

4. METHOD

The FR model was applied to determine the correlations between topographic factors and landslide occurrence in the study area. It is commonly assumed that landslide occurrence is determined by landslide related factors, and that future landslides will occur under the same conditions. Under this assumption, the relationship between landslides occurring in an area and landslide-related factors can be distinguished from that between landslide non-occurrence and landslide related factors. The FR model is based on the assumption that future landslides will occur under similar conditions as previous landslides (Lee and Pradhan, 2007). The FR approach is based on the observed relationships between the distribution of landslides and landslide-causing factors; thus, it reveals correlations between landslide locations and these factors in a given area (Lee and Pradhan, 2007). To calculate the FR in this study, the ratio between landslide occurrence and non-occurrence (Regmi et al., 2014) was calculated for each factor class. The FR was used to quantitatively assess landslide probability in the study area. Because this model is useful and easy to apply, it has been used widely by numerous researchers (Youssef et al., 2015; Ding et al., 2017; Oh et al., 2017; Kim et al., 2018). The FR model correlates landslide locations with conditioning factors in a given area (Pradhan and Lee, 2010b).

The coefficient of determination, denoted by R^2 or r^2 and pronounced "R squared", is a statistical measure that represents the proportion of the variance of a dependent variable that is explained by an independent variable (Kvalseth, 1985). The R^2 statistic provides information about the goodness-of-fit of a model or factor (Cameron and Windmeijer, 1997). Landslide susceptibility analysis uses categorical and numerical factors. Categorical factors in this study include slope aspect, curvature, and geology. Numerical factors are continuous, for example slope angle, gradient, surface area, catchment area, etc. In this study, the dependent variable was the FR and the independent variables were the continuous numerical factors. The R^2analysis yielded a linear regression trend line of the FR for each factor. The R^2 was calculated by dividing the residual sum of squares by the total sum of squares, and subtracting the product from 1 (Kvalseth, 1985). R^2 values range from 0 to 1 and are commonly stated as percentages from 0 to 100%.

5. RESULTS

5.1. Spatial Relationships between Landslides and Conditioning Factors

In total, 15 landslide conditioning factors were included in our analysis of the relationships between causative factors and landslide occurrence in Mingchukur, Uzbekistan. Those 15 factors were classified as topographic, hydrological, and geological factors. The FR was used to denote the relationships between the causative factors and landslide occurrence.

5.1.1. Spatial relationships between landslide and topographic factors

Topographic factors, such as slope aspect, curvature (local downslope and upslope curvature), elevation, DAH, gradient, MSP, slope angle, surface area, and valley depth, were derived from the DEM using SAGA GIS (Table 1). Calculated spatial relationship between landslide and topographic factor were used frequency ratio method and the result showed in Table 2. In the case of slope aspect, landslides were most frequent on east- and southeast-facing hill slopes. The frequency of landslides was lowest on northwest- and west-facing hill slopes.

Curvature values denote the morphology of the topography. In this case, curvature was divided into local downslope and upslope curvature. A positive curvature indicates that the surface is upwardly convex at a given cell and a negative curvature

Table 2. Calculation frequency ratio value of topographic factor

Factor	Type	Domain	Total (%)	Landslide pixel	Landslide (%)	Ratio
Slope Aspect	F	45394	3.91	59	1.07	0.27
	N	69891	6.02	109	1.98	0.33
	NE	150957	13.00	572	10.39	0.80
	E	254202	21.89	2186	39.72	1.81
	SE	304824	26.25	1875	34.07	1.30
	S	152388	13.12	549	9.97	0.76
	SW	110060	9.48	147	2.67	0.28
	W	57149	4.92	6	0.11	0.02
	NW	16387	1.41	1	0.02	0.01
	Total	1161252	100.00	5504	100.00	1.00
Digital Elevation Model	886–933	228426	19.53	2598	47.07	2.41
	934–966	237605	20.32	634	11.49	0.57
	967–1007	237460	20.30	581	10.53	0.52
	1008–1060	231401	19.79	1578	28.59	1.44
	1061–1139	234683	20.07	129	2.34	0.12
	Total	1169575	100.00	5520	100.00	1.00
Diurnal Anisotropic Heating	–0.561–0.035	225896	19.31	362	6.56	0.34
	0.036–0.146	221101	18.90	1318	23.88	1.26
	0.147–0.212	243471	20.82	1076	19.49	0.94
	0.213–0.278	244117	20.87	1395	25.27	1.21
	0.279–0.492	234990	20.09	1369	24.80	1.23
	Total	1169575	100.00	5520	100.00	1.00
Gradient	0.108–0.246	218752	18.93	520	9.54	0.50
	0.247–0.310	241538	20.91	1334	24.46	1.17
	0.311–0.369	232943	20.16	608	11.15	0.55
	0.370–0.443	238044	20.60	740	13.57	0.66
	0.444–1.471	224087	19.40	2251	41.28	2.13
	Total	1155364	100.00	5453	100.00	1.00
Local Downslope Curvature	Concave	6046	0.52	76	1.38	2.66
	Flat	997546	85.29	4530	82.07	0.96
	Convex	165983	14.19	914	16.56	1.17
	Total	1169575	100.00	5520	100.00	1.00
Local Upslope Curvature	Concave	33513	2.87	183	3.32	1.16
	Flat	1093628	93.51	5040	91.30	0.98
	Convex	42434	3.63	297	5.38	1.48
	Total	1169575	100.00	5520	100.00	1.00
Mid-Slope Position	0–0.195	230527	19.71	287	5.20	0.26
	0.196–0.390	233643	19.98	534	9.67	0.48
	0.391–0.585	232927	19.92	1023	18.53	0.93
	0.586–0.792	235383	20.13	2030	36.78	1.83
	0.793–1	237095	20.27	1646	29.82	1.47
	Total	1169575	100.00	5520	100.00	1.00
Slope Angle	0–20	231427	19.79	963	17.45	0.88
	21–30	237378	20.30	1062	19.24	0.95
	31–40	232324	19.86	765	13.86	0.70
	41–50	233675	19.98	656	11.88	0.59
	51–90	234771	20.07	2074	37.57	1.87
	Total	1169575	100.00	5520	100.00	1.00
Surface Area	1–1.02	231228	19.77	963	17.45	0.88
	1.03–1.04	229110	19.59	1020	18.48	0.94
	1.05–1.07	245777	21.01	827	14.98	0.71
	1.08–1.10	231534	19.80	651	11.79	0.60
	1.11–1.13	231926	19.83	2059	37.30	1.88
	Total	1169575	100.00	5520	100.00	1.00
Valley depth	0.005–7.86	219494	18.77	845	15.31	0.82
	7.87–19.64	246820	21.10	1143	20.71	0.98
	19.65–33.68	238791	20.42	1233	22.34	1.09
	33.69–51.08	233447	19.96	642	11.63	0.58
	51.09–143.68	231023	19.75	1657	30.02	1.52
	Total	1169575	100.00	5520	100.00	1.00

https://doi.org/10.1007/s12303-018-0052-x

indicates that the surface is upwardly concave at that cell. A value of zero indicates that the surface is flat. There is little difference between local downslope and upslope curvature. Lower negative values indicate a higher probability of landslides. Based on our results, curvature is related to landslide occurrence in Mingchukur. This is because an upwardly concave slope holds more water, and retains it for longer, following heavy rainfall, such that the FR exceeds 1. It is not possible to determine any such correlation in the case of concave topography, for which FR is high in hilly and mountainous areas and low in flat areas. These results are associated with the slope angle, because the type of topography is related to the slope angle.

With steeper slopes landslide probability is greater. The slope angle 0–50 have a FR below 1, indicating a low probability of landslide, but with a trend toward higher FR with a steeper angle. For the slope angle 51–90, the FR is almost 2, indicating a very high probability of landslide. The slope angle is an essential component of slope stability analysis. As the slope angle increases, the shear stress in soil or other unconsolidated material generally increases as well. Gentle slopes are expected to have a low frequency of landslides because of the generally lower shear stresses associated with such gradients. Gradient is consistent with slope and very steep gradients have the highest landslide probabilities. Moreover, the MSP accords with landslide probability.

Landslides occurred mostly at the elevations 886–933 m, because there were numerous geological activities at these elevations such as weathering and erosion. DAH is considered to be an indicator of the effects of both temperature and topographic solar radiation, which influence the degree of snowmelt on particular slopes. A slope with high temperature has a high probability of landslide occurrence.

Surface area provides the basis for measuring the topographic roughness of a landscape. The results showed that the roughest (1.11–1.33) surface area had the highest landslide probability. Valley depth is the difference between the elevation and an interpolated ridge level. Landslides mostly occurred at very large valley depths (51.09–143.68). Based on the results, a deeper valley has a higher landslide probability.

5.1.2. Spatial relationships between landslides and hydrologic factors

Hydrological factors play an important role in landslide susceptibility analyses. Groundwater includes free groundwater (surface groundwater at a depth of 8–11 m), aquifers in Neogene sediments at intervals of 8–32 m (due to coastal infiltration of the reservoir), and surface seepage water (four springs with a discharge rate of 0.1–5.0 L/sec). Motion measurements taken above the surface indicated that, as long as the level of the Charvak Reservoir drops by 70–100 mm/day, with a rate of increase of 3–10 mm/day, deep motion does not occur. Based on these data, the hydrological factors of catchment area, flow path length, the LS factor, and TWI were analyzed by reference to DEM data using SAGA GIS (Table 3).

Within a given catchment area, each drop of precipitation that falls onto the land eventually joins the same river flowing toward the sea, if it does not evaporate. If the catchment area cannot

Table 3. Calculation frequency ratio value of hydrologic factor

Factor	Type	Domain	Total (%)	Landslide pixel	Landslide (%)	ratio
Catchment Area	1	9103	0.78	19	0.34	0.44
	2–891	1124340	96.13	5280	95.65	1.00
	892–2672	23614	2.02	197	3.57	1.77
	2673–7125	7642	0.65	21	0.38	0.58
	7126–227992	4876	0.42	3	0.05	0.13
	Total	1169575	100.00	5520	100.00	1.00
Flow Path Length	0–73	352562	30.14	2468	44.71	1.48
	74–167	278578	23.82	1048	18.99	0.80
	168–292	227156	19.42	525	9.51	0.49
	293–443	183710	15.71	1479	26.79	1.71
	444–822	127569	10.91	0	0.00	0.00
	Total	1169575	100.00	5520	100.00	1.00
LS Factor	0–6.65	216266	18.49	777	14.08	0.76
	6.66–11.63	251335	21.49	965	17.48	0.81
	11.64–15.79	258423	22.10	1057	19.15	0.87
	15.80–19.95	223803	19.14	1195	21.65	1.13
	19.96–212.80	219748	18.79	1526	27.64	1.47
	Total	1169575	100.00	5520	100.00	1.00
Topographic Wetness Index	–1.02–1.17	201007	17.19	1268	22.97	1.34
	1.18–1.49	247543	21.17	920	16.67	0.79
	1.50–1.77	264570	22.62	892	16.16	0.71
	1.78–2.09	232574	19.89	1066	19.31	0.97
	2.10–10.72	223881	19.14	1374	24.89	1.30
	Total	1169575	100.00	5520	100.00	1.00

accommodate the water, it will flow down slopes and cause landslide, especially during heavy rain events. The results showed that very high catchment areas (7126–227992) have the lowest landslide probability; very low (0–1) catchment areas also have low landslide probability, while moderate catchment areas (892–2672) have the highest landslide probability. Furthermore, the probability of landslide occurrence is very high in areas with high flow path length and is lowest in those with moderate flow path length. For very low flow path length areas (0–73), the FR was more than 1, which indicates high landslide probability.

The LS factor explains the effect of topography on soil erosion. Soil erosion is generally caused by water and wind, among other factors; the greater the influence of topography on soil erosion (very high LS factor class), the higher the probability that landslides will occur. The TWI is a steady-state wetness index commonly used to quantify topographic influences on hydrological processes. In the present study area, very low TWI values (−1.02–1.17) correspond to the highest landslide probability, with moderate TWI values (1.50–1.77) being associated with the lowest landslide probability.

5.1.3. Spatial relationships between landslides and geological factors

In the study area, the lithology is divided into five units: alluvial Sidjac terrace, alluvial Kyzylsu terrace, alluvial above Khumsan terrace, loess and loamclay, and gravity delusional (Table 4). Alluvial Sidjac terrace is alluvial-proluvial sediments of average quaternary terraces of the Tashkent complex, it is contained clay loams, pebbles including boulders and sandy gravel close grained aggregate. Alluvial Kyzylsu terrace is almost similar with alluvial Sidjac terrace, but in alluvial Kyzylsu terrace, there is not include pebbles in the sediment. Alluvial above Khumsan terrace is alluvial-proluvial sediments of upper quaternary terraces of the Tashkent complex, it is contained clay loams, pebbles including boulders and sandy gravel close grained aggregate, overlapped by clay loamy. Gravity delusional is sediments of nowadays landslides, it is contained crumbly clay and sand loams including clastic material (State service of the Republic of Uzbekistan on monitoring for hazard geologic processes). Gravity delusional has a higher likelihood of landslides than all other units. Based on the FR calculation, alluvial Sidjac terrace has no possibility of landslides and alluvial Kyzylsu terrace has a very low possibility. The landslide probability likely increases from alluvial above Khumsan terrace, loess and loamclay, and gravity delusional, respectively.

Table 5. R^2 values for each factor

Factor		R^2
Topography	Digital Elevation Model	0.405
	Diurnal Anisotropic Heating	0.496
	Gradient	0.402
	Mid-Slope Position	0.821
	Slope Angle	0.256
	Surface Area	0.262
	Valley Depth	0.209
Hydrology	Catchment Area	0.067
	Flow Path Length	0.214
	LS factor	0.866
	Topographic Wetness Index	0.004

5.2. Evaluating Topographic Properties as Landslide Conditioning Factors

The goodness-of-fit of the model was evaluated by calculation of R^2, with 0.4 being the cut-off for good fit. Many researchers have used R^2 to evaluate regression models (Zhu and Huang, 2006; Domínguez-Cuesta et al., 2010; Du et al., 2017). We calculated R^2 to evaluate the goodness-of-fit for the landslide conditioning factors, especially the continuous numerical factors. In this study, R^2 was derived from the FR as a linear regression trend line.

Table 5 shows the R^2 value for each continuous numerical factor. Six such factors were important in landslide susceptibility, i.e., showed R^2 values greater than 0.4. These were DEM, DAH, gradient, MSP, and the LS factor. Other factors, such as slope angle, surface area, valley depth, catchment area, flow path length, and TWI, had R^2 values lower than 0.4, i.e., were only weakly correlated with landslide occurrence.

5.3. Landslide Susceptibility Mapping Using the Frequency Ratio Model

5.3.1. Landslide susceptibility map

A landslide susceptibility map was constructed by calculating and classifying landslide susceptibility indexes (LSIs) for the

Table 4. Calculation frequency ratio value of Geological factor

Factor	Type	Domain	Total (%)	Landslide pixel	Landslide (%)	ratio
Lithology	Alluvial Sidjac Terrace	358008	30.60	0	0.00	0.00
	Alluvial Kyzylsu Terrace	416879	35.64	440	7.97	0.22
	Alluvial Above Khumsan Terrace	127257	10.88	746	13.51	1.24
	Loess and Loamclay	225693	19.29	2650	48.01	2.49
	Gravity Delusional	42016	3.59	1684	30.51	8.49
	Total	1169853	100.00	5520	100.00	1.00

https://doi.org/10.1007/s12303-018-0052-x

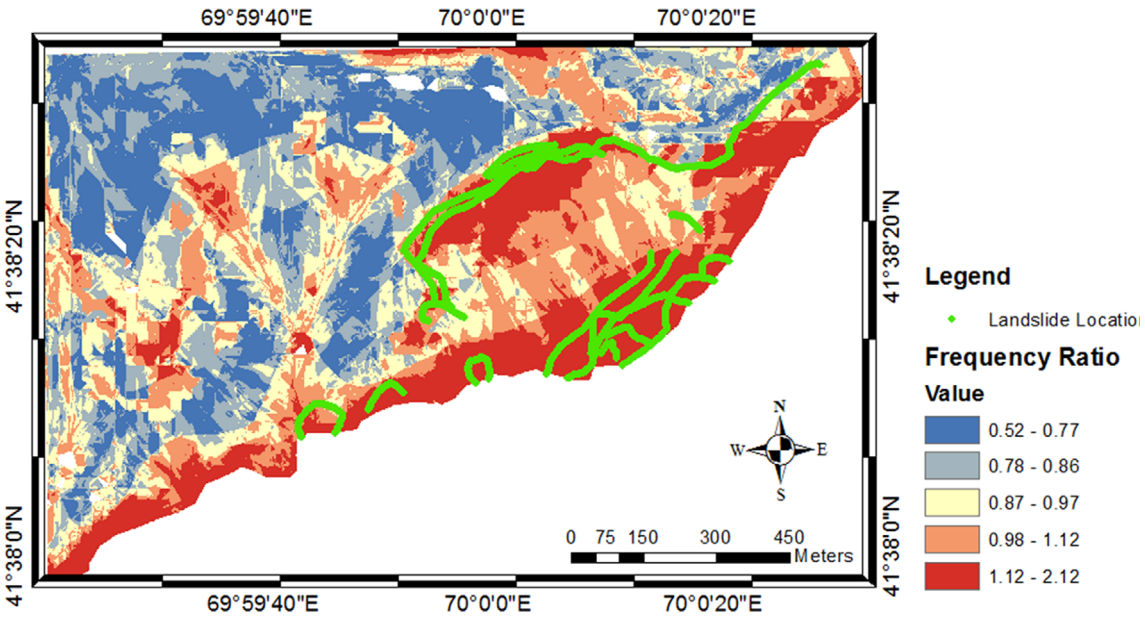

Fig. 5. Landslide susceptibility map of the Mingchukur area using frequency ratio model.

whole study area. The LSI denotes the susceptibility of an area to landslides. Areas with smaller LSIs are less susceptible to landslides. LSIs were calculated based on the FR values determined during model training (Tables 2–4). The LSI is calculated using Equation (1). In the first stage, the FR value for each factor is calculated. Then, the FR values are summed using the following equation (Lee and Talib, 2005):

$$LSI = \sum FR. \quad (1)$$

Many methods can be employed to classify LSIs, such as the equal interval, natural break, and standard deviation methods (Ayalew and Yamagishi, 2005). Among these, the natural break method is the most widely used (Irigaray et al., 2007), and thus was employed for classifying the LSIs in the present study. Using this method, the LSIs were classified into five susceptibility classes: very low (0.52–0.77), low (0.78–0.86), moderate (0.87–0.97), high (0.98-1.12), and very high (1.12–2.12). The landslide susceptibility map of the study area developed using the FR model is shown in Figure 5.

5.3.2. Validation of landslide susceptibility map

A landslide susceptibility map should be able to predict possible future landslide areas effectively, and it should also be validated by combining existing landslide location data with data on new landslide locations as they occur. In this study, the results from the landslide susceptibility analysis were confirmed using the validation data. The landslide susceptibility map produced by the FR models was validated using the actual landslide data

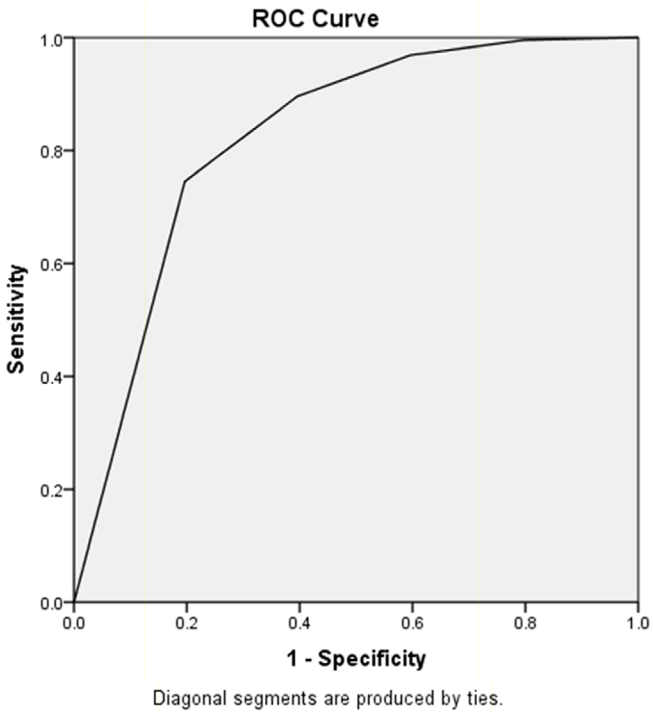

Fig. 6. The ROC curves of the landslide susceptibility maps based on frequency ratio in the Mingchukur area.

for this, 100% (5,520) of the landslide data were used. The landslide data were not randomly separated into two datasets; we used all of the data in both the training and validation sets.

To measure how well the model performed, receiver operating characteristic (ROC) curve and area under the curve (AUC) analyses were used (Fig. 6): the FR model had an AUC value of

0.823, corresponding to a landslide prediction accuracy of 82%. Therefore, FR models can be used for spatial prediction of landslide hazards in Mingchukur, and to determine the influence of landslide-related factors.

6. DISCUSSION AND CONCLUSION

Landslide studies are multi-disciplinary, encompassing the fields of geoscience and engineering, among others, and include many aspects, such as the causative mechanism, spatial-temporal distribution, conditioning factors, socioeconomic losses, environmental impact, and hazard reduction methods and management. In recent years, various methods have been proposed for landslide susceptibility assessment and mapping, such as machine learning, decision trees, logistic regression, and the FR. For the landslide susceptibility assessment performed in this study, conditioning factors were needed. The correlation analysis between landslides and topographic properties proved very useful for evaluating landslide conditioning factors. In total, 15 conditioning factors, pertaining to topography, hydrology, and geology, were analyzed. All factors were classified as categorical or continuous. The categorical factors were slope aspect, local downslope curvature, local upslope curvature, and lithology and all other factors were continuous numerical factors. Based on the trend lines of FR values, the relationships between landslide occurrence and topographic and hydrological factors were quantified.

The importance of each factor in the landslide susceptibility analysis was given by the R^2 statistic, which provides information about the goodness-of-fit of a model. Six continuous numerical factors showed R^2 values greater than 0.4, the cut-off for a good fit. These factors were DEM, DAH, gradient, MSP, and the LS factor. The other factors, i.e., slope angle, surface area, valley depth, catchment area, flow path length, and TWI, had R^2 values lower than 0.4, denoting weak correlations with landslide occurrence in Mingchukur.

Integrating a landslide incidence database with conditioning factors, such as those within a database of topographic factors, using GIS, avoids the need to collect primary data and results in a resource that can inform further digital and quantitative analyses of the relationships between landslides and topographic properties. In the case study of Mingchukur, Uzbekistan reported in this paper, rainfall intensity influenced the occurrence of landslides. Both slope gradient and absolute elevation influence the landslide distribution, and regional variation in the relationship between landslides and topography is present. Hydrology and geology related factors also influence the spatial distribution of the landslides in this region.

The FR method has many advantages for analysis landslide susceptibility mapping. The FR-based probability analysis method can be used both to study the general characteristics of landslides and their association with environmental factors, as well as to reveal the spatial variation in such relationships. The FR model is easy model to understand and interpret, this approach is a good choice for landslide susceptibility modeling and mapping and should be of great use to planners and engineers. Landslides susceptibility maps are of great help to planners and engineers for choosing suitable locations to implement developments. These results can be used as basic data to assist slope management and land use planning. The disadvantages of using FR models are the methods used in the study are valid for generalized planning and assessment purposes, although they may be less useful at the site-specific scale where local geological and geographical heterogeneities may prevail.

ACKNOWLEDGMENTS

This research was part of a Basic Research Project of the Korea Institute of Geoscience and Mineral Resources (KIGAM) funded by the Ministry of Science and ICT and The National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2016K1A3A1A09915721 and No. 2017R1A2B4003258) and part of a Basic Research Project of the State Committee of the Republic of Uzbekistan on Geology and Mineral Resources with the cooperation and assistance of Dr. Lee, Saro, Korea Institute of Geoscience and Mineral Resources (KIGAM). The authors wish to express their gratitude to Mr. Gani Bimurzaev, Director of the State Service for Monitoring of Dangerous Geological Processes, Goscomgeology for providing the necessary information and numerous data in this text.

REFERENCES

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

Bates, R.J. and Jackson, J.A., 1984, Dictionary of Geological Terms (3rd edition). American Geological Institute, New York, 299 p.

Buchanan B.P., Fleming, M., Schneider, R.L., Richards, B.K., Archibald, J., Qiu, Z., and Walter, M.T., 2014, Evaluating topographic wetness indices across central New York agricultural landscapes. Hydrology and Earth System Sciences, 18, 3279–3299.

Burrough, P.A., McDonell, R.A., and Lloyd, C.D., 1998, Principles of Geographical Information Systems (3rd edition). Oxford University Press, New York, 190 p.

Cameron, A.C. and Windmeijer F.A.G., 1997, An R-squared measure of goodness of fit for some common nonlinear regression models. Journal of Econometrics, 77, 329–342.

Carillo, G., Torch, P.A., Sivapalan, M., Wagener, T., Harman, C., and Sawicz, K., 2011, Catchment classification: hydrological analysis of catchment behaviour through process-based modelling along a cli-

 https://doi.org/10.1007/s12303-018-0052-x

mate gradient. Hydrology and Earth System Science, 15, 3411–3430.

Chen, W., Pourghasemi, H.R., and Naghibi, S.A., 2018a, A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China. Bulletin of Engineering Geology and the Environment, 77, 647–664.

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Tien Bui, D., Duan, Z., and Ma, J., 2017, A 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., Zhang, S., Li, R., and Shahabi, H., 2018b, Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naive bayes tree for landslide susceptibility modeling. Science of the Total Environment, 644, 1006–1018.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner J., 2015, System for automated geoscientific analyses (SAGA) v. 2.1.4. Geoscientific Model Development, 8, 1991–2007.

Cristea, N.C., Breckheimer, I., Raleigh, M.S., HilleRisLambers, J., and Lundquist, J.D., 2017, An evaluation of terrain-based downscaling of fractional snow covered area data sets based on LiDAR-derived snow data and orthoimagery. Water Resources Research, 53, 6802–6820.

Desmet, P.J.J. and Govers, G., 1996, A GIS procedure for automatically calculating the USLE LS factor on topographically complex landscape units. Journal of Soil and Water Conservation, 51, 427–433.

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

Dominguez-Cuesra, M.J., Jimenez-Sanchez, M., and Gonzalez-Rogriguez, G., 2010, Modelling shallow landslide susceptibility: a new approach in logistic regression by using favourability assessment. International Journal of Earth Sciences, 99, 661–674.

Du, G.L., Zhang, Y.S., Iqbal, J., Yang, Z.H., and Yao, X., 2017, Landslide susceptibility mapping using integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu province, China. Journal of Mountain Science, 14, 249–268.

Freeman, G.T., 1991, Calculating catchment area with divergent flow based on a regular grid. Computers and Geosciences, 17, 413–422

Hong, H., Pourghasemi, H.R., and Pourtaghi, Z.S., 2016, Landslide susceptibility assessment in Lianhua County (China): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology, 259, 105–118.

Irigaray, C., Fernandez, T., El Hamdouni, R., and Chacon, J., 2007, Evaluation and validation of landslide susceptibility maps obtained by a GIS matrix method: examples from the Betic Cordillera (southern Spain). Natural Hazards, 41, 61–79.

Jeff, S.J., 2004, Calculating landscape surface area from digital elevation models. Wildlife Society Bulletin, 32, 829–839.

Kim, J.C., Jung, H.S., and Lee, S., 2018, Groundwater productivity potential mapping using frequency ratio and evidential belief function and artificial neural networks: focus on topographic factors. Journal of Hydroinformatics. https://doi.org/10.2166/hydro.2018.120

Kim, J.C., Sunmin, L., Jung, H.S., and Lee, S., 2018, Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. Geocarto International, 33, 1000–1015.

Kvalseth, T.O., 1985, Cautionary Note about R2. The American Statistician, 39, 279–285.

Lee, M.J., Park, I., and Lee, S., 2015, Forecasting and validation of landslide susceptibility using an integration of frequency ratio and neuro-fuzzy models: a case study of Seorak mountain area in Korea. Environmental Earth Sciences, 74, 413–429.

Lee, M.J., Park, I., Won, J.S., and Lee, S., 2016, Landslide hazard mapping considering rainfall probability in Inje, Korea. Geomatics, Natural Hazards and Risk, 7, 424–446.

Lee, S. and Lee, M.J., 2017, Susceptibility mapping of Umyeonsan using logistic regression (LR) model and post-validation through field investigation. Korean Journal of Remote Sensing, 33, 1047–1060.

Lee, S. and Park, I., 2013, Application of decision tree model for the ground subsidence hazard mapping near abandoned underground coal mines. Journal of Environmental Management, 127, 166–176.

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

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

Lee, S., Hong, S.M., and Jung, H.S., 2017, A support vector machine for landslide susceptibility mapping in Gangwon province, Korea. Sustainability, 9, 48.

Lee, S., Jeon, S.W., Oh, K.Y., and Lee, M.J., 2016, The spatial prediction of landslide susceptibility applying artificial neural network and logistic regression models: a case study of Inje, Korea. Open Geosciences, 8, 117–132.

Lee, S., Lee, M.J., and Jung, H.S., 2017, Data mining approaches for landslide susceptibility mapping in Umyeonsan, Seoul, South Korea. Applied Sciences, 7, 683.

Lee, S., Lee, S., Lee, M.J., and Jung, H.S., 2018, Spatial assessment of urban flood susceptibility using data mining and geographic information System (GIS) tools. Sustainability, 10, 648.

Lee, S., Won, J.S., Jeon, S.W., Park, I., and Lee, M.J., 2015, Spatial landslide hazard prediction using rainfall probability and a logistic regression model. Mathematical Geosciences, 47, 565–589.

Mezaal, M.R. and Pradhan, B., 2018, Data mining-aided automatic landslide detection airborne laser scanning data in densely forested tropical areas. Korean Journal of Remote Sensing, 34, 45–74.

Moore, I.D., Grayson, R.B., and Landson, A.R., 1991, Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. Hydrological Processes, 5, 3–30.

Oh, H.J., Lee, S., and Hong, S.M., 2017, Landslide susceptibility assessment using frequency ratio technique with iterative random sampling. Journal of Sensors, 2017, 21.

Oh, C.Y., Kim, K.T., and Chou, C.U., 2009, Analysis of landslide characteristics of Inje area using SPOT5 image and GIS analysis. Korean Journal of Remote Sensing, 25, 445–454.

Oh, H.J. and Lee, S., 2017, Shallow landslide susceptibility modeling using the data mining models artificial neural network and boosted tree. Applied Sciences, 7, 1000.

Oh, H.J., 2010, Landslide detection and landslide susceptibility mapping using aerial photos and artificial neural network. Korean Journal of remote sensing, 26, 47–57.

Park, I. and Lee, S., 2014, Spatial prediction of landslide susceptibility using a decision tree approach: a case study of the Pyeongchang area, Korea. International Journal of Remote Sensing, 35, 6089–6112.

Park, N.W. and Kyriakidis, P.C., 2008, Gestatistical integration of different sources of elevation and its effect on landslide hazard mapping. Korean Journal of Remote Sensing, 24, 453–462.

Pham, B.T., Tien Bui, D., Pourghasemi, H.R., Indra, P., and Dholakia, M.B., 2017, Landslide susceptibility assessment in the Uttarakhand area (India) using GIS: a compariosn study of prediction capability of naive bayes, multilayer perceptron neural networks, and functional trees methods. Theoretical and Applied Climatology, 128, 255–273.

Polykretis, C. and Chalkias, C., Comparison and evaluation of landslide susceptibility maps obtained from weight of evidence, logistic regression, and artificial neural network. Natural Hazards, 93, 249–274.

Pradhan, B. and Lee, S., 2010a, Delineation of landslide hazard areas using frequency ratio, logistic regression, and artificial neural network model at Penang Island, Malaysia. Environmental Earth Sciences, 60, 1037–1054.

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

Quinn, P., Beven, K., Chevallier, P., and Planchon, O., 1991, The prediction of hill-slope flow paths for distributed hydrological modelling using digital terrain models. Hydrological Processes, 5, 59–79.

Rakhmatullaev, S., Huneau, F., Celle-Jeanton, H., Le Coustumer, P., Motelica-Heino, M., and Bakiev, M., 2013, Water reservoirs, irrigation and sedimentation in Central Asia: a first-cut assessment for Uzbekistan. Environmental Earth Sciences, 68, 985–998.

Regmi, A.D., Devkota, K.C., Yoshida, K., Pradhan, B., Pourghasemi, H.R., Kumamoto, T., and Akgun, A., 2014, Application of frequency ratio, statistical index, and weight of evidence models and their comparison in landslide susceptibility mapping in central Nepal Himalaya. Arabian Journal of Geosciences, 7, 725–742.

Tien Bui, D., Tuan, T.A., Hoang, N.D., Thanh, N.Q., Nguyen, D.B., Liem, N.V., and Pradhan, B., 2017, Spatial prediction of rainfall-induced landslides for the Lao Cai area (Vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization. Landslides, 14, 447–458.

Truong, X.L., Mitamura, M., Kono, Y., Raghavan, V., Yonezawa, G., Truong, X.Q., Hang Do, T., Tien Bui, D., and Lee, S., 2018, Enhancing prediction performance of landslide susceptibility model using hybrid machine learning approach of bagging ensemble and logistic model tree. Applied Sciences, 8, 1046.

Wilson, J.P. and Gallant, J.C., 2000, Terrain Analysis: Principles and Applications. John Wiley and Sons, Inc., New York, 479 p.

Youssef, A.M., Al-Kathery, M., and Pradhan, B., 2015, Landslide susceptibility mapping at Al-Hasher area, Jizan (Saudi Arabia) using GIS-based frequency ratio and index of entropy models. Geosciences Journal, 19, 113–134.

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

 https://doi.org/10.1007/s12303-018-0052-x