

Spatial landslide hazard assessment along the Jelapang Corridor of the North-South Expressway in Malaysia using high resolution airborne LiDAR data

Norbazlan Mohd Yusof^{1,2} · Biswajeet Pradhan¹ · Helmi Zulhaidi Mohd Shafri¹ · Mustafa Neamah Jebur¹ · Zainuddin Yusoff¹

Received: 1 February 2015 / Accepted: 28 April 2015 / Published online: 8 May 2015
© Saudi Society for Geosciences 2015

Abstract Mapping landslide-prone regions are crucial in natural hazard management and urban development activities in hilly and tropical regions. This research aimed to delineate a spatial prediction of landslide hazard areas along the Jelapang Corridor of the North-South Expressway in Malaysia by using two statistical models, namely, logistic regression (LR) and evidential belief function (EBF). Landslides result in high economic and social losses in Malaysia, particularly to highway concessionaries such as PLUS Expressways Berhad. LR and EBF determine the correlation between conditioning factors and landslide occurrence. EBF can also be applied in bivariate statistical analysis. Thus, EBF can be used to assess the effect of each class of conditioning factors on landslide occurrence. A landslide inventory map with 26 landslide sites was recorded using field measurements. Subsequently, the landslide inventory was randomly divided into two data sets. Approximately 70 % of the data were used for training the models, and 30 % were used for validating the results. Eight landslide conditioning factors were prepared for landslide susceptibility analysis: altitude, slope, aspect, curvature, stream power index, topographic wetness index, terrain roughness index, and distance from river. The landslide probability index was derived from both methods and subsequently classified into five susceptible classes by using the quantile method. The

resultant landslide susceptibility maps were evaluated using the area under the curve technique. Results revealed the proficiency of the LR method in landslide susceptibility mapping. The achieved success and prediction rates for LR were 90 and 88 %, respectively. However, EBF was not successful in providing reasonable accurate results. The acquired success and prediction rates for EBF were 53 and 50 %, respectively. Hence, the LR technique can be utilized in landslide hazard studies for land use management and planning.

Keywords Landslide · Hazard · LiDAR · EBF · LR · GIS · Malaysia

Introduction

Landslides result in considerable commercial, social, and ecological damages worldwide. They are frequently triggered by other natural catastrophes, such as earthquakes and floods, and are hard to predict (Glenn et al. 2006). Moreover, strong or continuous rainfall can trigger land failure (Huang et al. 2012). Unplanned agricultural activities in landslide-prone areas can also cause significant soil erosion. Moreover, construction projects in hilly regions trigger land failures. Malaysia is a good example of this situation; this country is located in a tropical area that receives a huge amount of precipitation annually. Landslides have always posed serious threats to settlements and structures in Malaysia that support transportation, natural resources, and tourism. Landslide occurrences are prevalent in the hill complexes of Malaysia both in the highlands and lowlands, including along the highways. These occurrences have caused loss of lives and properties in recent years. Therefore, landslide susceptibility, hazard, and risk mapping should be implemented to reduce landslide occurrence (Pradhan and Lee 2010c). The first step in landslide

✉ Biswajeet Pradhan
biswajeet24@gmail.com; biswajeet@lycos.com

¹ Department of Civil Engineering, Geospatial Information Science Research Center (GISRC), Faculty of Engineering, University Putra Malaysia, UPM, 43400 Serdang, Malaysia

² Projek Lebuhraya Usahasama Berhad, Menara Korporat, Persada PLUS, Persimpangan Bertingkat Subang, KM 15, Lebuhraya Baru Lembah Klang, 47301 Petaling Jaya, Selangor, Malaysia

hazard and risk mapping is to prepare a landside susceptibility map. This map shows the areas with the same probability of landslide occurrence within a specific period (Pradhan and Youssef 2010). Therefore, comprehensive landslide prediction along highways reduces damages by providing preventive actions. Early warning and preparedness, assigning highly sensitive landslide zones as protected (non-developable) areas, as well as detecting and monitoring landslide occurrences are possible examples of these actions.

Landslide susceptibility mapping assesses the proneness of a terrain to land failure and landslide at a particular area or under specific conditioning factors (Pourghasemi et al. 2012). Remote sensing and geographic information system (GIS) are efficient techniques for landslide susceptibility mapping (Jebur et al. 2014b; Van Westen 2000). Several studies have focused on landslide susceptibility mapping. The most popular methods for such mapping are the analytic hierarchy process (AHP) (Pourghasemi et al. 2012), statistical approaches such as frequency ratio (FR) (Pradhan and Lee 2010b) and logistic regression (LR) (Lee and Pradhan 2007), support vector machine (SVM) (Wan and Lei 2009), neuro-fuzzy, fuzzy logic (Tien Bui et al. 2012), evidential belief function (EBF) (Lee et al. 2013; Pradhan et al. 2014), artificial neural network (ANN), and weight of evidence (WoE) (Lee and Choi 2004). However, only a few studies have compared the efficiency and performance of these methods (Pradhan 2013; Pradhan and Lee 2010a).

Qualitative methods such as AHP have been used in regional studies because of different opinions among experts. Several situations can affect the choice of experts and can consequently negatively influence their assessment (Ayalew and Yamagishi 2005). Moreover, two experts may have completely different views. Chen et al. (2011) believed that the use of qualitative methods such as AHP leads to subjective and uncertain outcomes. Some researchers attempted to overcome the disadvantages of AHP by choosing experts from different areas; nevertheless, the disadvantages of this approach cannot be completely overcome because human decision is involved. Pradhan (2013) recently utilized decision tree (DT), SVM, and an adaptive neuro-fuzzy inference system (ANFIS) to map prone areas in Penang Hill, Malaysia and compared their proficiency. He indicated that defining the rules for DT and choosing SVM parameters are difficult tasks. Moreover, these methods require considerable time to process data. ANFIS performs better than DT and SVM, but it involves several parameters. Moreover, all three methods require a high-speed processor that is capable of handling heavy performance and huge amount of spatial data. ANN is a common technique used in many domains and particularly in landslide studies (Wan et al. 2010). Some researchers call ANN as a black box with a complicated procedure and performance (Pradhan and Buchroithner 2010). Moreover, ANN cannot produce precise predictions when the validation data set

consists of values outside the range of those utilized for training (Wan et al. 2010). ANN is also considerably time consuming when a large number of factors are involved. Fuzzy logic is easier to understand than ANN. It has been employed in numerous landslide applications; however, the opinions of experts contribute some degrees of uncertainty in the outcomes of this method (Tilmant et al. 2002).

Statistical methods are simple, and their input, computation, and outcome processes are easily understood. FR and LR can be used to perform bivariate statistical analysis (BSA) and multivariate statistical analysis (MSA), respectively (Aleotti and Chowdhury 1999). The principles of statistical techniques are primarily based on the mathematical expressions of the association between conditioning factors and landslides. BSA allocates weights to the categories of each conditioning factor by assessing their role in landslide occurrence. The effect of each class can be recognized by examining the correlation between the landslide inventory map and landslide conditioning factors. The density of landslides in each category is the main factor considered during weight assignments. LR, which is frequently used in susceptibility mapping, examines the correlation between conditioning factors and landslide occurrence. Most MSA methods require strict assumptions that are defined prior to the study. LR can overcome this problem and provide an easy procedure for analysis without the need to define prior assumptions (Benediktsson et al. 1990). WoE is another BSA method that is less used in hazard studies than FR. Although each BSA method needs particular mechanisms for calculation, all these methods follow the same concept (Ozdemir and Altural 2013; Youssef et al. 2013, 2015; Zare et al. 2013; Regmi et al. 2014; Pourghasemi et al. 2013). BSA methods identify the effects of each class of conditioning factors on landslide occurrence but cannot evaluate the association between conditioning factors and landslide occurrence (Ayalew and Yamagishi 2005). The advantage of LR is that the relationship among conditioning factors is considered. Meanwhile, Althuwaynee et al. (2012) examined the efficiency of EBF in landside studies and found that the degree of belief (Bel), degree of disbelief (Dis), degree of uncertainty (Unc), and degree of plausibility (Pls) should be individually measured to calculate EBF (Tien Bui et al. 2012). Although both FR and EBF can perform BSA, EBF frequently provides better results than FR. EBF can also assess the correlation among conditioning factors. That is, EBF can perform both BSA and MSA. Tien Bui et al. (2012) compared the proficiency of EBF and fuzzy logic methods in landslide susceptibility map and found that the results derived from EBF exhibit the highest prediction capacity.

Statistical methods that are typically employed in landslide susceptibility mapping include LR and EBF (Althuwaynee et al. 2012; Lee and Pradhan 2007). In the present study, these two methods were chosen for landslide modeling along the

corridor of the North-South Expressway. Jelapang, a small stretch of PLUS Highway Berhad in Malaysia, is prone to landslides. PLUS Berhad holds the concession for a total of 987 km of toll expressways in Malaysia, the longest of which is the North-South Expressway or NSE. Acting as the “backbone” of the west coast of the peninsula, the NSE stretches from the Malaysian-Thai border in the north to the border with neighboring Singapore in the south, linking several major cities and towns along the way. North-South Expressway in Malaysia contributes to the country economic development through trade, social, and tourism sector. Presently, the highway is good in terms of its condition and connection to every state but some locations need urgent attention. Stability of slopes at these locations is of most concern as any instability can cause danger to the motorist. Therefore, landslide hazard assessment along the NSE is highly important. The achievement of this research will be considerably beneficial to landslide preparedness and damage mitigation for PLUS Highway Berhad.

Study area

The Jelapang Corridor of the North-South Expressway, also known as the PLUS Expressways in Malaysia, was selected for the implementation of the landslide susceptibility analysis because of the frequent occurrence of mass movements in this region (Fig. 1). The expressway links many major cities and towns in western Peninsular Malaysia, functioning as the backbone of the west coast of the peninsula. It is the longest expressway in Malaysia with a total length of approximately 772 km. This expressway passes through seven states on the peninsula,

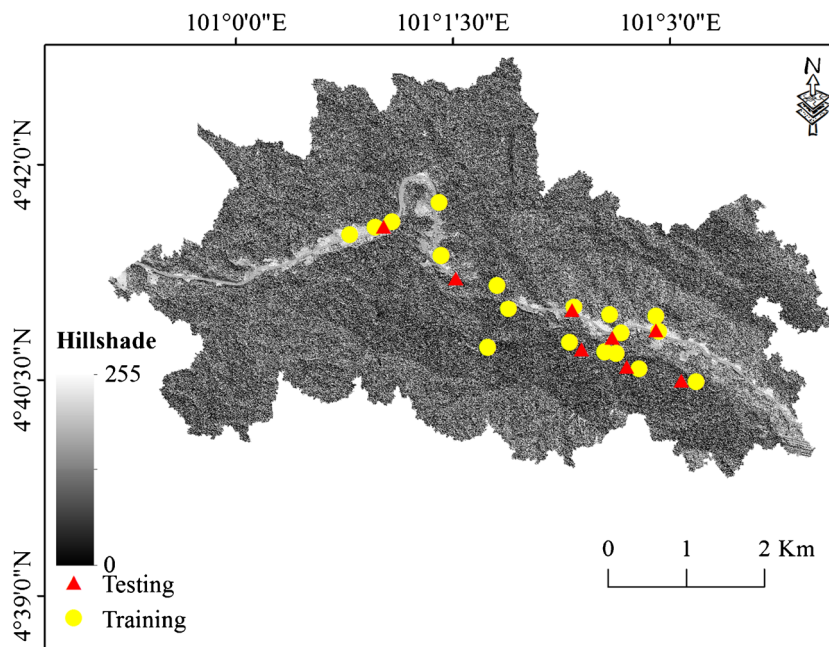
namely, Johor, Malacca, Negeri Sembilan, Selangor, Perak, Penang, and Kedah. This area is approximately located at the zone of 4°43' 7.605" N to 4°39' 18.038" N latitude and 101°4' 6.068" E to 100°59' 7.232" E longitude. The study area experiences frequent mass movements that cause erosion and landslides. Annual rainfall is very high, averaging between 2500 and 3000 mm per year. Two pronounced wet seasons occur from September to December and from February to May. Rainfall peaks between November to December and March to May. The geomorphology of the area consists of an undulating plateau and a hilly terrain. The geology of the area mostly consists of Quaternary and Devonian granite. In recent years, many landslides have been recorded along PLUS highways, roads, and streams, which scour the sides of streams.

Data used

Landslide inventory

Landslide inventory maps are the basis and first requirement of most landslide susceptibility mapping methods (Pradhan et al. 2014). Furthermore, inventory maps can be utilized to assess and decrease landslide occurrence and risk at a local scale (Umar et al. 2014). A landslide inventory map is crucial for assessing the correlation between landslide occurrence and conditioning factors. Multiple field surveys and observations were conducted in the study area to create a complete inventory map. A total of 26 landslides were found in the study area, which was subsequently divided into two data sets for training and testing. Based on literature, 70 % of the landslide

Fig. 1 Landslide location map with a hill-shaded map of Jelapang, Malaysia



inventory was used for training the models and 30 % was utilized for validation (Fig. 1). Training slope failure locations were used to create the dependent layer. The produced layer consists of two values, namely, 0 and 1, where 1 denotes the presence and 0 indicates the absence of slope failure in the study area. The remaining slope failure locations were utilized to test the outcomes. Both layers were made in ArcGIS and transformed into raster format.

Landslide conditioning factors

Landslide susceptibility is defined by utilizing reasonable qualitative and quantitative studies of the conditioning factors in affected areas (Domínguez-Cuesta et al. 2007). The conditions of conditioning factors differ per region. Providing the conditioning factor data set is a challenging task, and no exact rule is available to decide how many conditioning factors are sufficient for a particular susceptibility analysis (Nefeslioglu et al. 2010). These factors are typically selected based on expert knowledge or literature. In the current research, conditioning factors were chosen based on knowledge derived from literature. The conditioning factor data set contains altitude, slope, aspect, curvature, stream power index (SPI), topographic wetness index (TWI), terrain roughness index (TRI), and river factors. Table 1 lists the conditioning factors, and Fig. 2 shows the data layers. All conditioning factors were resized to a 1 m×1 m grid, and the grid of the Jelapang region was built by 9254 columns and 7067 rows (28,379,958 pixels; 28.37 km²).

Altitude, slope, aspect, curvature, SPI, TWI, and TRI maps were derived from a digital elevation model (DEM), as shown in Fig. 2a–g, respectively. To perform EBF, the quantile classification technique was utilized to categorize each conditioning factor. The logic behind such classification is to implement BSA; scale data should be categorized to evaluate the influence of each class on slope failure occurrence (Tehrany et al. 2013). At the altitude of 0–1339.24 m, 10 classes were established using the quantile classification method. This

method is commonly used in classification and various applications (Tehrany et al. 2014b). Slope is an influential conditioning factor in landslide occurrence. This factor directly affects slope failure occurrence and is typically used in landslide susceptibility analysis (Pradhan and Lee 2010a). The vertical component of gravity rises with the amount of slope. The slope in the study area ranges from 0 to 89.85° (Fig. 2b). Hence, the slope map of the study area was prepared and partitioned into 10 slope classes (Fig. 2b). Aspect is also a key landslide conditioning factor (Jebur et al. 2014a). The morphological situation of the study area and the extent of rainfall and sunlight are the meteorological conditions that influence the occurrence of slope failure. Aspect influences weathering, and therefore, indirectly affects the shear strength of rock mass. Although the relationship between aspect and slope failure occurrence has been discovered, no exact agreement is available on the effect of this factor on slope failure (Gokceoglu et al. 2005). The aspect map that was utilized to recognize the association between aspect and slope failure occurrence is displayed in Fig. 2c. Ten classes have been made for the aspect map (flat, north, northeast, east, southeast, south, southwest, west, northwest, and north). The effect of curvature on slope failure is the convergence or divergence of water during downhill movement (Oh and Pradhan 2011). Thus, this factor is another conditioning factor that is involved in landslide occurrence. Curvature was derived from DEM and subsequently categorized into three classes: concave, convex, and flat.

The hydrological factors SPI and TWI were calculated using Eqs. 1 and 2, respectively. Some researchers considered these two factors as secondary topographical characteristics in landslide susceptibility mapping (Gokceoglu et al. 2005).

$$\text{SPI} = A \tan \beta / b, \quad (1)$$

$$\text{TWI} = \log_e \left(\frac{A}{b \tan \beta} \right), \quad (2)$$

where A (m²) is the flow accumulation, b (m) is the cell width through which water flows, and β (radian) is the slope (Regmi et al. 2010). SPI demonstrates the power of water flow to create erosion based on the assumption that discharge is related to a particular catchment area. SPI predicts net erosion in the region of profile and tangential convexity (flow acceleration and convergence zones) and the net deposition in the areas of profile concavity (zones of decreasing flow velocity). TWI is the amount of water accumulation at a site. Ten classes were made for SPI (0–22.37) and TWI (0–16.41).

Table 1 Results of LR for each conditioning factor

Conditioning factors	Logistic coefficient	SIG
Altitude	0.003640	0.009817
Slope	−0.019112	0.054408
Aspect	−0.000193	0.850179
Curvature	0.000033	0.746136
SPI	0.384703	0
TWI	−0.433722	0.028179
TRI	−0.018751	0.016210
Distance from river	−0.000210	0.012215
Constant	0.214791	

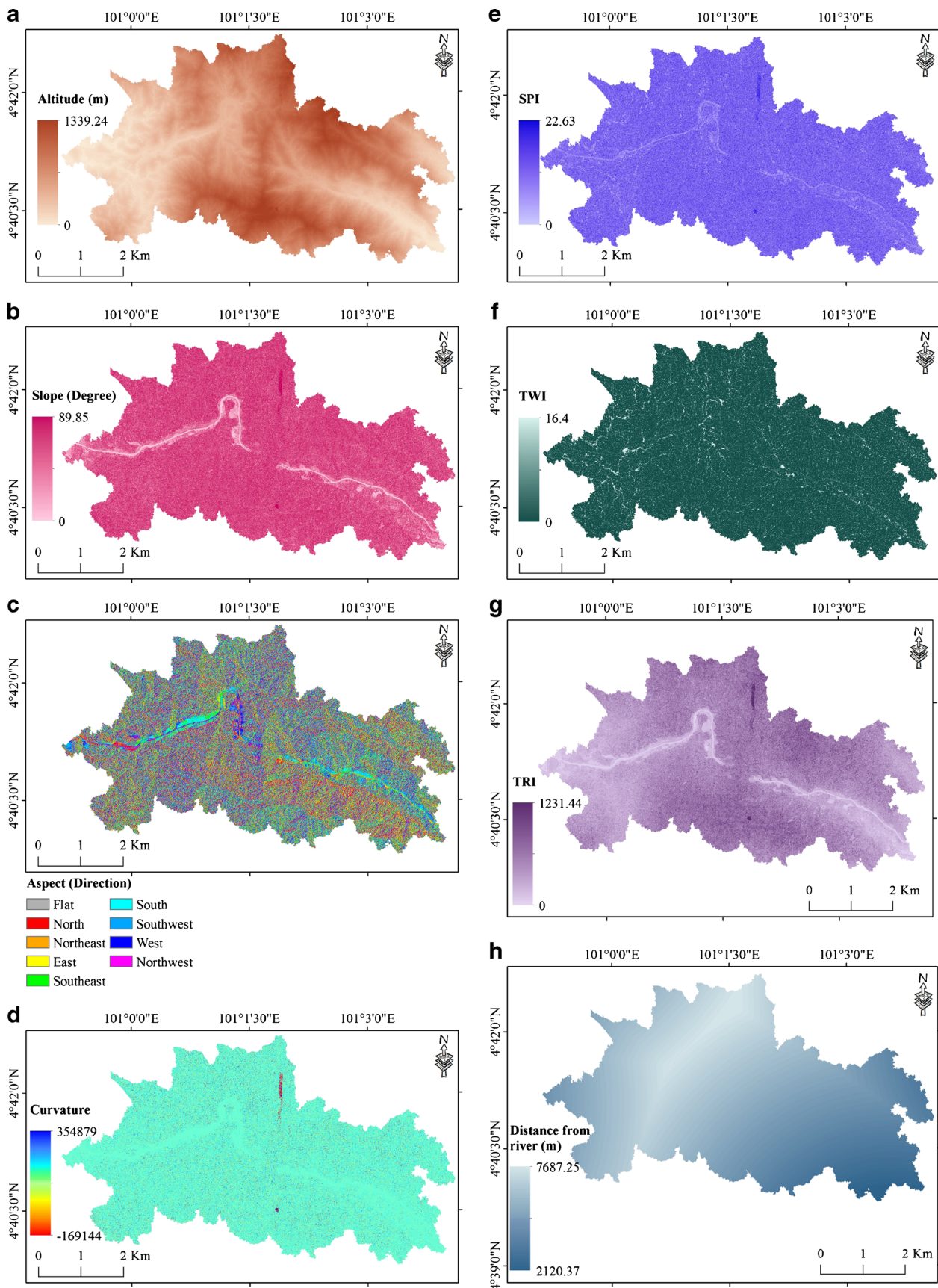


Fig. 2 Input thematic layers: **a** altitude, **b** slope, **c** aspect, **d** curvature, **e** SPI, **f** TWI, **g** TRI, and **h** distance from river

Another influential factor is TRI, which is broadly utilized in landslide studies. This factor was calculated using Eq. 3.

$$\text{TRI} = \sqrt{|x|(\max^2 - \min^2)}, \quad (3)$$

where max and min are the greatest and lowest values of the cells in the nine rectangular neighborhoods of altitude, respectively (Riley et al. 1999). TRI was also categorized into 10 classes using the quantile technique. In the case of distance from river, only the undercutting of the side slopes of rivers might cause slope failure initiation (Jebur et al. 2014a).

Methodology

Logistic regression (LR)

In some cases, landslide susceptibility maps are produced with a high level of uncertainty because of the non-efficiency of model performance and the limitation of data (Jebur et al. 2014a). LR is a commonly used multivariate statistical model. Recent studies have analyzed the performance on LR in landslide studies (Choi et al. 2012; Yesilnacar and Topal 2005). LR is a common MSA technique that assesses multivariate regression correlation between conditioning factors and landslide occurrence (Tehrany et al. 2014a). Akgün and Bulut (2007) evaluated the precision of the landslide susceptibility maps produced by weighted linear combination and LR. Brenning (2005) examined the efficiency of LR, SVM, and DT in probability mapping. Xu et al. (2012) compared the bivariate statistics model with ANN, LR, and SVM, where LR exhibited the best performance. Lee and Pradhan (2007) and Nandi and Shakoor (2010) also showed the capabilities of LR in the hazard domain.

LR was employed to recognize the possibility of landslide occurrence in the study area using Eqs. (4) and (5) produced by conditioning factors. The first requirement of this technique is having a dependent layer (landslide inventory) that consists of two values, namely, 0 and 1, which indicate the absence and presence of landslide, respectively. All the conditioning factors (Fig. 2) were converted from raster into ASCII format, as required in SPSS. SPSS software was utilized to implement MSA. The regression coefficients were calculated and listed in Table 1. When the LR coefficient is high, the probable influence on landslide occurrence is large. Using the measured LR coefficients, we calculated the landslide probability index as follows:

$$p = 1/(1 + e^{-z}), \quad (4)$$

where p is the landslide probability that is attained between 0 and 1 on an S-shaped curve. z is the linear combination. LR

involves fitting an equation with the following form into the data:

$$z = b^{\circ} + b_1x_1 + b_2x_2 + b_3x_3 + b_nx_n, \quad (5)$$

where b° is the intercept of the model, b_i ($i=0, 1, 2, \dots, n$) denotes the LR coefficients, and x_i ($i=0, 1, 2, \dots, n$) indicates the conditioning factors (Lee and Sambath 2006).

Evidential belief function

EBF follows Dempster–Shafer theory of evidence algorithms. This theory is a generalized Bayesian theory of subjective probability and has been used in the GIS environment in numerous applications (Carranza and Hale 2003). The main advantage of EBF is its flexibility, which is attributed to uncertainty acceptance; in addition, the belief of many sources can be integrated. Using EBF, the probability degree that shows the closeness of the probability to be true can be measured (Tien Bui et al. 2012). EBF has four basic mathematical functions, namely, Bel, Dis, Unc, and Pls (Althuwaynee et al. 2012). Pls and Bel present the upper and lower limits of the probability for the proposition, respectively. The variety between Pls and Bel is measured by Unc, which also shows the ignorance value. Finally, Dis is the belief in the false probability values for a particular case. In EBF, Bel+Unc+Dis=1. In the absence of landslide, Bel will be zero, and Dis should be reset to zero.

EBF can be calculated in two ways, namely, data-driven and subjective (Choi et al. 2012). Using data-driven analysis, the spatial correlation between landslide occurrence and conditioning factors is considered; moreover, the spatial correlation among the parameters themselves is taken. The relationship between landslide and a conditioning factor can be calculated by overlaying the landslide layer on the eight conditioning factors. The landslide conditioning factor $C=(C_i, i=1, 2, 3, \dots, n)$, which contains mutually exclusive and exhaustive factors of C_i , is adopted in this study. EBF analysis can be applied using Eq. 6 (Carranza and Hale 2003).

$$\text{Bel}(C_{ij}) = \frac{W_{C_{ij}(\text{landslide})}}{\sum_{j=1}^n W_{C_{ij}(\text{landslide})}}, \quad (6)$$

where the weight of C_{ij} is represented by $W_{C_{ij}(\text{landslide})}$ and supports the belief that the existence of landslides is more than their absence. $W_{C_{ij}(\text{non-landslide})}$ represents the weight of C_{ij} , which supports the belief that landslides are more absent than present. A detailed explanation of EBF measurement can be obtained from Tien Bui et al. (2012) and Jebur et al. (2014a). The EBF model was applied for the eight conditioning factors, and the resulting weight was used to reclassify each layer.

Validation

The prediction precision and efficiency of the methods used should be evaluated. In the current research, the results were assessed by comparing the generated landslide probability maps with existing landslide inventory data. The area under the curve (AUC) method was utilized to examine the outcomes quantitatively. AUC is commonly used to assess the reliability of outcomes, which defines the success and prediction rates (Jebur et al. 2014c). To evaluate the correctness and proficiency of the landslide probability maps, both success and prediction rate curves were measured. The success rate was attained using the training data set, which accounted for 70 % of the inventory landslide locations. The training data set

was used to produce the landslide model; hence, it cannot be used to validate the prediction capability of the method. The prediction rate reveals how well the model can predict slope failure in the study area. Hence, it was measured using the testing data set (30 % of the landslide events) that was not utilized in the training procedure. The acquired landslide probability index was arranged in descending order to calculate the relative ranks for each prediction pattern. Subsequently, the cell values were divided into 100 classes and set on the vertical axis (*y*), with accumulated 1 % intervals in the horizontal axis (*x*). The presence of the landslide events in each interval was measured, and the success and prediction rates were calculated (Tehrany et al. 2014b).

Fig. 3 Landslide probability maps derived from **a** LR method and **b** EBF method

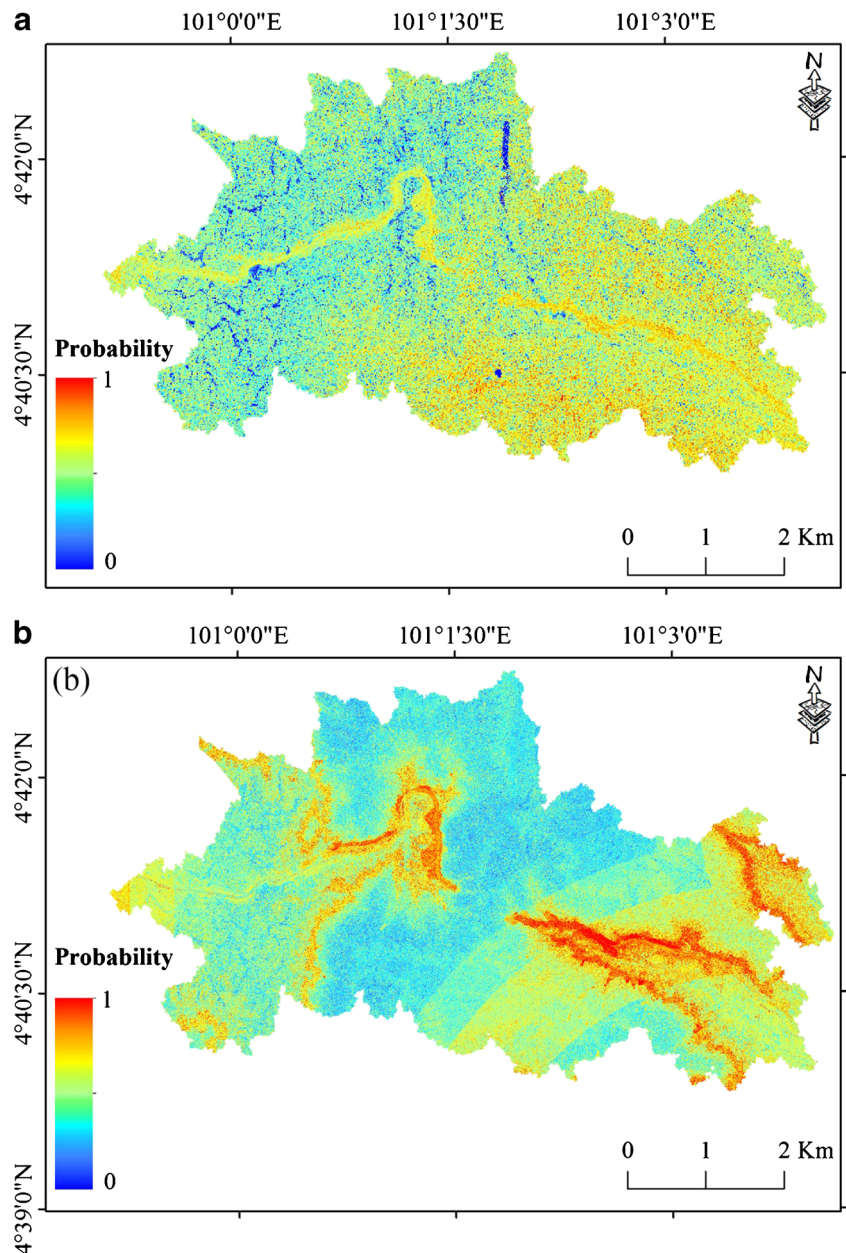


Table 2 Spatial correlation between conditioning factors and landslide occurrence extracted by EBF

Parameter	Classes	Belief	Disbelief	Uncertainty	Plausibility
Altitude	0–129	0	11	89	89
	129–185	5	10	85	90
	185–253	27	8	65	92
	253–314	24	8	68	92
	314–370	30	7	63	93
	370–419	6	10	84	90
	419–469	0	11	89	89
	469–524	0	11	89	89
	524–598	5	10	85	90
	598–1339	0	11	89	89
Slope	0–29.24	25	8	67	92
	29.24–42.99	17	9	74	91
	42.99–52.85	23	8	69	92
	52.85–59.91	12	9	79	91
	59.91–65.54	7	10	83	90
	65.54–70.12	4	10	86	90
	70.12–74.01	2	10	88	90
	74.01–77.52	0	11	89	89
	77.52–80.69	0	10	90	90
	80.69–89.85	4	10	86	90
Aspect	Flat	0	11	89	89
	North	4	12	84	88
	Northeast	21	9	70	91
	East	47	6	47	94
	Southeast	16	10	74	90
	South	4	12	84	88
	Southwest	0	12	88	88
	West	0	12	88	88
Curvature	Northwest	6	11	83	89
	Concave	12	27	61	73
	Flat	78	31	–9	69
Slope	convex	8	40	52	60
	SPI	0–3.81	5	10	85
Slope	3.81–4.61	11	9	80	91
	4.61–5.14	15	9	76	91
	5.14–5.59	14	9	77	91
	5.59–6.03	6	10	84	90
	6.03–6.47	15	9	76	91
	6.47–7.01	11	9	80	91
	7.01–7.72	8	10	82	90
	7.72–8.96	7	10	83	90
	8.96–22.37	3	10	87	90
	TWI	0–0.01	18	5	77
0.01–0.57		45	10	45	90
0.57–1.09		10	10	80	90
1.09–1.61		7	10	83	90
1.61–2.05		8	10	82	90
2.05–2.51		0	10	90	90
2.51–3.02		0	10	90	90
3.02–3.79		10	10	80	90

Table 2 (continued)

Parameter	Classes	Belief	Disbelief	Uncertainty	Plausibility	
TRI	3.79–5.33	0	10	90	90	
	5.33–16.41	0	10	90	90	
	0–28.97	27	8	65	92	
	28.97–43.46	26	7	67	93	
	43.46–57.95	14	9	77	91	
	57.95–67.61	13	9	78	91	
	67.61–77.26	5	10	85	90	
	77.26–86.92	2	10	88	90	
	86.92–101.41	2	11	87	89	
	101.41–115.91	1	10	89	90	
	115.91–135.21	1	10	89	90	
	135.21–1231.43	4	10	86	90	
	Distance	2120–3609	5	10	85	90
	from river	3609–4164	22	8	70	92
(meter)	4164–4716	19	8	73	92	
	4716–5281	17	9	74	91	
	5281–5771	6	10	84	90	
	5771–6141	0	11	89	89	
	6141–6461	5	10	85	90	
	6461–6778	0	11	89	89	
	6778–7104	11	9	80	91	
	7104–7687	11	9	80	91	
Constant						

Results and discussion

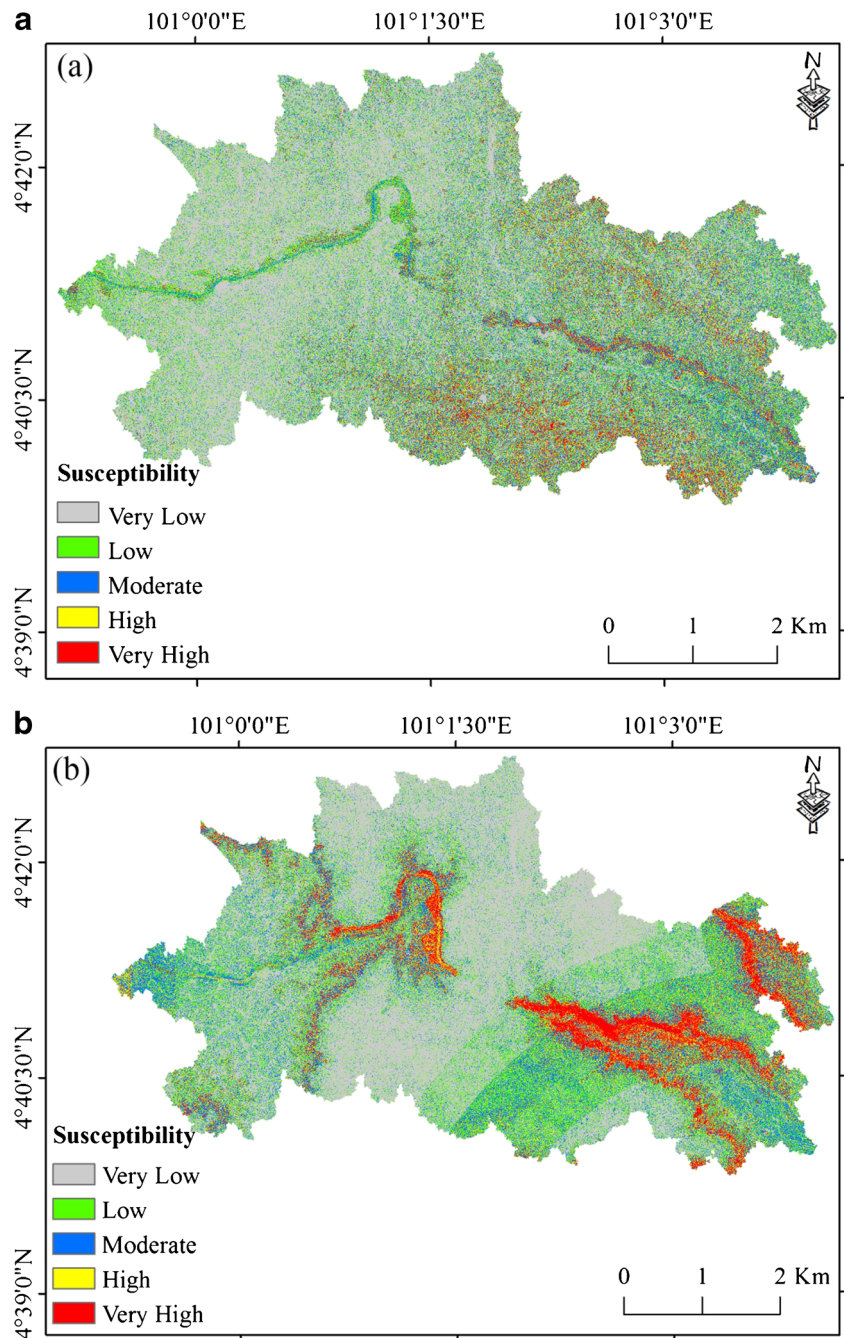
The MSA result was acquired through the LR method (Table 1). The correlation between landslide occurrence and landslide conditioning factors was evaluated using SPSS V.19. LR coefficients and significant probability (Sig) factors were calculated using LR and listed in Table 1. A Sig factor identifies the conditioning factors that significantly affect landslide occurrence (Papadopoulou-Vrynioti et al. 2013). Sig values less than 0.05 indicates that the conditioning factor exerts statistically significant effect on slope failure. As shown in Table 1, altitude, SPI, TWI, TRI, and distance from river were the most influential conditioning factors. Slope, aspect, and curvature attained Sig values over 0.05, which indicated that these factors did not exert significant effect on landslide occurrence. The reason is in some cases when two or more factors have similar impact, the regression analysis makes one of them significant and others are none. Based on the acquired LR coefficients, slope, aspect, TWI, TRI, and distance from river were negatively correlated with landslide occurrence. The LR coefficient for SPI (0.384703) revealed that this factor is strongly and positively correlated with landslide occurrence in the study area.

To obtain the landslide probability index, the LR coefficient for each factor was entered in Eq. 5, which yields the following:

$$z = (0.003640 \times \text{DEM}) - (0.019112 \times \text{slope}) - (0.000193 \times \text{aspect}) + (0.000033 \times \text{curvature}) + (0.384703 \times \text{SPI}) - (0.433722 \times \text{TWI}) - (0.018751 \times \text{TRI}) - (0.000210 \times \text{river}) + 0.214791.$$

In the next step, the landslide probability index was measured from Eq. 4. The association between conditioning factors and slope failure occurrence was recognized. The landslide probability index was calculated and categorized using the appropriate system. The thematic map of landslide probability is shown in Fig. 3a. This index denotes the predicted probabilities of slope failure for each pixel under the existence of a particular set of conditioning factors. As previously mentioned, EBF was used as a second method to produce a landslide

Fig. 4 Landslide susceptibility maps derived from **a** LR method and **b** EBF method



susceptibility map. The landslide probability map produced by EBF is shown in Fig. 3b.

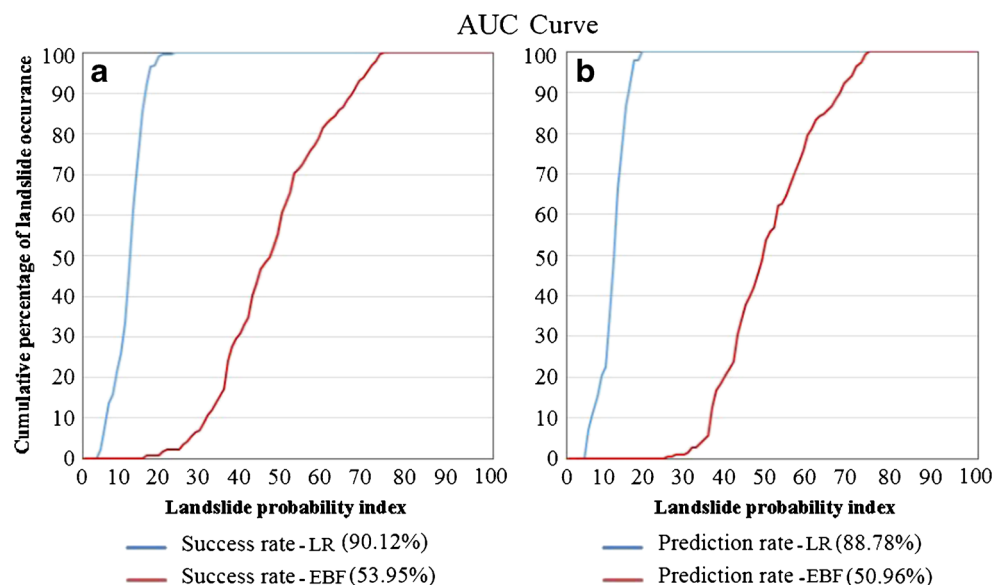
Table 2 shows the EBF values that were calculated for each class of each conditioning factor. As listed in Table 2, altitude significantly influenced the characteristics of the study area. Altitude ranges from 0 to 1339 m in the study area. It had three classes that revealed high landslide probability. These classes were 185–253 m, 253–314 m, and 314–370 m, with Bel values of 27, 24, and 30, respectively. This finding can be attributed to the instability of the terrain in high altitudes. Slope and aspect factors are related to the physical characteristics of the ground. In the case of slope, the classes (0–29.24°) and (42.99–52.85°) attained 25 and 23 successive Bel results, respectively. The “east” class in the aspect factor acquired the highest Bel value among the classes, which constituted 47. The highest Bel value of 78 was assigned to the class of “flat” in the curvature, which revealed the considerable effect of this class on landslide occurrence. Other concave and convex classes received 21 and 8 Bel values, respectively. In SPI, the classes (4.61–5.14) and (6.03–6.47) acquired the highest Bel value of 15. The two highest Bel values of 18 and 48 were associated with the classes of (0–0.01) and (0.01–0.57) of the TWI layer, respectively. Bel decreased as TRI increased. This result indicates that the first two classes of TRI acquired the highest Bel values of 27 and 26; however, the Bel values decreased as TRI ranges increased. In the case of distance from river, the areas closer to the river showed higher landslide probability than those regions far from the river. The distance from 3609 to 5281 m received the greatest Bel values, which presented an extremely high landslide probability.

To create a susceptibility map, the probability index must be divided into different classes (Ohlmacher and Davis 2003). Various classification techniques are available, each of which is suitable for a specific application. The most popular

techniques for classification include standard deviation, natural break, equal interval, and quantile (Ayalew and Yamagishi 2005). The quantile method was used in the current study because of its reputation in probability index classification. This method provides the same number of features for each class and has been employed by scientists such as Papadopoulou-Vrynioti et al. (2013), Umar et al. (2014), and Jebur et al. (2014c) in susceptibility studies. The technique produced appropriate outcomes on the comparison between the generated landslide susceptibility map and the spatial distribution of slope failure locations. Finally, landslide susceptibility maps were obtained using both methods, and the study area was divided into five classes of landslide susceptibility: very low, low, medium, high, and very high. The landslide susceptibility maps generated by LR and EBF are shown in Fig. 4.

As shown in Fig. 4, LR and EBF revealed significantly different results on highly susceptible areas. To understand which result is accurate and precise, accuracy assessment was performed. The success rate curves generated by AUC are shown in Fig. 5a. The acquired success rates for LR and EBF were 90.12 and 53.95 %, respectively. The prediction curve is shown in Fig. 5b. The AUC values of 88.78 and 50.96 % correspond to the prediction accuracies of LR and EBF, respectively. Based on the acquired validation outcomes, the performance of EBF was considerably less efficient than that of LR. However, LR produced a highly accurate landslide susceptibility map with high success and prediction rates. In the current research, the proficiency and strength of both LR and EBF methods were assessed and compared. EBF did not perform efficiently in this case study. The proficiency of this method might be improved by combining it with other methods, such as machine learning, which can be considered in future works.

Fig. 5 AUC: a success rate b prediction rate



Conclusion

Several techniques have been utilized to map landslide-prone regions. Each method has advantages and disadvantages. Among all the methods, statistical methods are simple, understandable, and accurate. The current research assessed the potential application of LR and EBF in landslide susceptibility mapping at an expressway corridor in Jelapang, Malaysia. LR and EBF were individually used to map the landslide-prone regions in the study area. Slope failure occurrence is related to several conditioning factors. Eight landslide conditioning factors were considered in the analysis, namely, altitude, slope, aspect, curvature, SPI, TWI, TRI, and river. Bel, Dis, Unc, and Pls were calculated for the EBF method, and the influence of the categories of each conditioning factor on slope failure occurrence was evaluated. Dempster's rule of combination was used to obtain the integrated EBF map for accuracy assessment. Subsequently, the influence of each conditioning factor on landslide occurrence was assessed using LR. Probability index maps were derived using both methods. The obtained probability index for each method was categorized using quantile classification. Two landslide susceptibility maps were created. The provided susceptibility maps produced spatial predictions of slope failures without producing any information on "when" and "how often" a future slope failure will occur. AUC graphs were used for validation, and the success and prediction rates were calculated. Comparison proved LR to be more efficient than EBF in the current research. The success rate and prediction rate of LR were 90.12 and 88.78 %, respectively. The acquired success and prediction rates for EBF were considerably low, which constituted 53.95 and 50.96 %, respectively. The validation outcomes revealed that the LR model had a reasonably higher prediction capability than the EBF model. The derived landslide susceptibility map using the LR method can help planners and the government control and avoid slope failures in the future.

Acknowledgments This research is supported by research project number CMD/2560/CM-CNT with UPM vote number 6300114.

References

- Akgün A, Bulut F (2007) GIS-based landslide susceptibility for Arsin-Yomra (Trabzon, North Turkey) region. *Environ Geol* 51:1377–1387
- Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Environ* 58:21–44
- Althuwaynee OF, Pradhan B, Lee S (2012) Application of an evidential belief function model in landslide susceptibility mapping. *Comput Geosci* 44:120–135
- Ayalew L, 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
- Benediktsson J, Swain PH, Ersoy OK (1990) Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Trans Geosci Remote Sens* 28:540–552
- Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat Hazard Earth Syst Sci* 5:853–862
- Carranza EJM, Hale M (2003) Evidential belief functions for data-driven geologically constrained mapping of gold potential, Baguio district, Philippines. *Ore Geol Rev* 22:117–132
- Chen YR, Yeh CH, Yu B (2011) Integrated application of the analytic hierarchy process and the geographic information system for flood risk assessment and flood plain management in Taiwan. *Nat Hazards* 59:1261–1276
- Choi J, Oh HJ, Lee HJ, Lee C, Lee S (2012) Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Eng Geol* 124:12–23
- Domínguez-Cuesta MJ, Jiménez-Sánchez M, Berrezueta E (2007) Landslides in the Central Coalfield (Cantabrian Mountains, NW Spain): geomorphological features, conditioning factors and methodological implications in susceptibility assessment. *Geomorphology* 89:358–369
- Glenn NF, Streutker DR, Chadwick DJ, Thackray GD, Dorsch SJ (2006) Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. *Geomorphology* 73:131–148
- Gokceoglu C, Sonmez H, Nefeslioglu HA, Duman TY, Can T (2005) The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Eng Geol* 81:65–83
- Huang R, Pei X, Fan X, Zhang W, Li S, Li B (2012) The characteristics and failure mechanism of the largest landslide triggered by the Wenchuan earthquake, May 12, 2008, China. *Landslides* 9:131–142
- Jebur M, Pradhan B, Tehrany M (2014a) Manifestation of LiDAR-derived parameters in the spatial prediction of landslides using novel ensemble evidential belief functions and support vector machine models in GIS. *IEEE J Sel Top Appl*. doi:10.1109/JSTARS.2014.2341276
- Jebur MN, Pradhan B, Tehrany MS (2014b) Detection of vertical slope movement in highly vegetated tropical area of Gunung pass landslide, Malaysia, using L-band InSAR technique. *Geosci J* 18:61–68
- Jebur MN, Pradhan B, Tehrany MS (2014c) Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. *Remote Sens Environ* 152:150–165
- Lee S, Choi J (2004) Landslide susceptibility mapping using GIS and the weight-of-evidence model. *Int J Geogr Inf Sci* 18:789–814
- Lee S, Pradhan B (2007) Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* 4:33–41
- Lee S, Sambath T (2006) Landslide susceptibility mapping in the Damrei Romel area, Cambodia using frequency ratio and logistic regression models. *Environ Geol* 50:847–855
- Lee S, Hwang J, Park I (2013) Application of data-driven evidential belief functions to landslide susceptibility mapping in Jinbu, Korea. *Catena* 100:15–30
- Nandi A, Shakoor A (2010) A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. *Eng Geol* 110:11–20
- Nefeslioglu H, Sezer E, Gokceoglu C, Bozkir A, Duman T (2010) Assessment of landslide susceptibility by decision trees in the metropolitan area of Istanbul, Turkey. *Math Probl Eng* 2010:1–15
- Oh HJ, Pradhan B (2011) Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput Geosci* 37:1264–1276

- Ohlmacher GC, Davis JC (2003) Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Eng Geol* 69:331–343
- Ozdemir A, Altural T (2013) A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *J Asian Earth Sci* 64:180–197
- Papadopoulou-Vrynioti K, Bathrellos GD, Skilodimou HD, Kaviris G, Makropoulos K (2013) Karst collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area. *Eng Geol* 158:77–88
- Pourghasemi HR, 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
- Pourghasemi HR, Pradhan B, Gokceoglu C, Mohammadi M, Moradi HR (2013) Application of weights-of-evidence and certainty factor models and their comparison in landslide susceptibility mapping at Haraz watershed, Iran. *Arab J Geosci* 6:2351–2365
- Pradhan B (2013) A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. *Comput Geosci* 51:350–365
- Pradhan B, Buchroithner MF (2010) Comparison and validation of landslide susceptibility maps using an artificial neural network model for three test areas in Malaysia. *Environ Eng Geosci* 16:107–126
- Pradhan B, Lee S (2010a) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environ Earth Sci* 60:1037–1054
- Pradhan B, 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. *Environ Model Softw* 25:747–759
- Pradhan B, Lee S (2010c) Regional landslide susceptibility analysis using back-propagation neural network model at Cameron Highland, Malaysia. *Landslides* 7:13–30
- Pradhan B, Youssef AM (2010) Manifestation of remote sensing data and GIS on landslide hazard analysis using spatial-based statistical models. *Arab J Geosci* 3:319–326
- Pradhan B, Abokharima MH, Jebur MN, Tehrany MS (2014) Land subsidence susceptibility mapping at Kinta Valley (Malaysia) using the evidential belief function model in GIS. *Nat Hazards* 73:1019–1042
- Regmi NR, Giardino JR, Vitek JD (2010) Modeling susceptibility to landslides using the weight of evidence approach: Western Colorado, USA. *Geomorphology* 115:172–187
- Regmi AD, Devkota KC, Yoshida K, Pradhan B, Pourghasemi HR, Kumamoto T, Akgun A (2014) Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab J Geosci* 7:725–742
- Riley SJ, DeGloria SD, Elliot R (1999) A terrain ruggedness index that quantifies topographic heterogeneity. *Int J Sci* 5:23–27
- Tehrany MS, Pradhan B, Jebur MN (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
- Tehrany MS, Lee M-J, Pradhan B, Jebur MN, Lee S (2014a) Flood susceptibility mapping using integrated bivariate and multivariate statistical models. *Environ Earth Sci* 72:1–15
- Tehrany MS, Pradhan B, Jebur MN (2014b) Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol* 512:332–343
- Tien Bui D, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. *Catena* 96:28–40
- Tilmant A, Vanclooster M, Duckstein L, Persoons E (2002) Comparison of fuzzy and nonfuzzy optimal reservoir operating policies. *J Water Resour Plan Manag* 128:390–398
- Umar Z, Pradhan B, Ahmad A, Jebur MN, Tehrany MS (2014) Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena* 118:124–135
- Van Westen CJ (2000) The modelling of landslide hazards using GIS. *Surv Geophys* 21:241–255
- Wan S, Lei TC (2009) A knowledge-based decision support system to analyze the debris-flow problems at Chen-Yu-Lan River, Taiwan. *Know-Based Syst* 22:580–588
- Wan S, Lei T, Chou TY (2010) An enhanced supervised spatial decision support system of image classification: consideration on the ancillary information of paddy rice area. *Int J Geogr Inf Sci* 24:623–642
- Xu C, Xu X, Dai F, Saraf AK (2012) Comparison of different models for susceptibility mapping of earthquake triggered landslides related with the 2008 Wenchuan earthquake in China. *Comput Geosci* 46:317–329
- Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng Geol* 79:251–266
- Youssef AM, Pradhan B, Maerz NH (2013) Debris flow impact assessment caused by 14 April 2012 rainfall along the Al-Hada Highway, Kingdom of Saudi Arabia using high-resolution satellite imagery. *Arab J Geosci* 7:2591–2601
- Youssef AM, Pradhan B, Pourghasemi HR, Abdullahi S (2015) Landslide susceptibility assessment at Wadi Jawrah Basin, Jizan region, Saudi Arabia using two bivariate models in GIS. *Arab J Geosci*. doi:10.1007/s12303-014-0065-z
- Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2013) Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab J Geosci* 6:2873–2888