



## Comparative Study among Bivariate Statistical Models in Landslide Susceptibility Map

YUKNI ARIFANTI<sup>1</sup>, PAMELA<sup>1</sup>, FITRIANI AGUSTIN<sup>2</sup>, and DICKY MUSLIM<sup>3</sup>

<sup>1</sup>Center for Volcanology and Geological Hazard Mitigation, Geological Agency

Jln. Diponegoro No. 57 Bandung, Indonesia

<sup>2</sup>Geological Survey Institute, Geological Agency

Jln. Diponegoro No. 57 Bandung, Indonesia

<sup>3</sup>Faculty of Geological Engineering, Padjadjaran University

Jln. Raya Bandung-Sumedang Km.21 Jatinangor, Indonesia 45363

Corresponding author: [yukni.arifanti@esdm.go.id](mailto:yukni.arifanti@esdm.go.id), [pamela@esdm.go.id](mailto:pamela@esdm.go.id), [fitriani.agustin@esdm.go.id](mailto:fitriani.agustin@esdm.go.id)

Manuscript received: December, 3, 2018; revised: May, 29, 2019;

approved: November, 6, 2019; available online: March, 6, 2020

**Abstract** - The main purpose of this paper is to compare the performance of bivariate statistical models *i.e.* Frequency Ratio, Weight of Evidence, and Information Value for landslide susceptibility assessment. These models were applied in Cianjur Regency, West Java Province (Indonesia), in order to map the landslide susceptibility and to rate the importance of landslide causal factors. In the first stage, a landslide inventory map and the input layers of the landslide conditioning factors were prepared in the Geographic Information System (GIS) supported by field investigations and remote sensing data. The 298 landslides were randomly divided into two groups of modeling/training data (70%) and validation/test data sets (30%). The landslide conditioning factors considered for the studied area were slope angle, elevation, slope aspect, lithological unit, and land use. Subsequently, the thematic data layers of conditioning factors were integrated by frequency ratio (FR), weight of evidence (WofE), and information value (IV). Model performance was tested with receiver operator characteristic analysis. The validation findings revealed that the three models showed promising results since the models gave good accuracy values. The success rates of FR, WofE, and IV models were 0.920, 0.926, and 0.930, while the prediction rates of the three models were 0.913, 0.912, and 0.895, respectively. However, the FR model was proved to be relatively superior in estimating landslide susceptibility throughout the studied area.

**Keywords:** frequency ratio, weight of evidence, information value, Cianjur

© IJOG - 2020. All right reserved

### How to cite this article:

Arifianti, Y., Pamela, Agustin, F., and Muslim, D., 2020. Comparative Study among Bivariate Statistical Models in Landslide Susceptibility Map. *Indonesian Journal on Geoscience*, 7 (1), p.51-63. DOI: [10.17014/ijog.7.1.51-63](https://doi.org/10.17014/ijog.7.1.51-63)

### INTRODUCTION

According to Landslide Inventory Database of Indonesia, from 2011 to 2015, almost 40% of landslides in Indonesia occur in West Java Province. Cianjur Regency with its prominent factors of landslide, highly weathered material (lithology), and the steep morphology is one of hotspots for landslide in West Java (Arifianti and Agustin,

2017). The accelerated population growth towards the landslide-prone areas caused the increasing of casualties by human-induced landslide hazard each year. A significant effort to reduce the number of losses was then carried out through landslide disaster mitigation. One of its activities is to conduct Landslide Susceptibility Assessment (LSA) as the basis of Landslide Susceptibility Map. LSA plays a significant part of landslide disaster mitigation,

and has received more attention with the highest number of publications in international journals (Gokceoglu and Sezer, 2009).

Many studies have been carried out to assess landslide susceptibility, with increasing application of GIS using different models. Numerous methods have been used for landslide susceptibility assessment and mapping, which can be classified into two categories (i) qualitative and (ii) quantitative methods. Qualitative method is based on field observations and prior knowledge of experts in identifying judgment rules or assign weighted values for conditioning factor maps and which overlay them to produce a landslide susceptibility map, such as analytical hierarchy process (Ghosh, 2011; Kayastha *et al.* 2012; Mondal and Maiti, 2012). The quantitative method primarily refers to several statistical analyses, which can be categorized into bivariate statistical and multivariate analysis. This study was only using bivariate statistical analysis such as frequency ratio (Lee and Pradhan, 2006; Vijith and Madhu, 2007; Constantin *et al.* 2011; Mezughi *et al.* 2011; Regmi *et al.* 2014), information value model (Yin and Yan, 1988; Lin and Tung 2004; Sarkar *et al.* 2008; Conforti *et al.* 2011; Wang *et*

*al.* 2014; Zhu *et al.* 2014), and weight of evidence model (Poli and Sterlacchini, 2007; Dahal *et al.* 2008; Sharma and Kumar, 2008; Kayastha *et al.* 2012; Chen and Li, 2014; Teerarungsikul *et al.* 2015). Geomatics by taking advantage of modern tools, such as Geographic Information System (GIS) and Remote Sensing (RS) provide a perfect opportunity for using, validating, and comparing different methods to produce a landslide susceptibility map (Vakhshoori and Zare, 2016). Some studies have applied and compared two or more methods to the same region (Pradhan and Lee, 2010a; Ercanoğlu and Temiz, 2011; Yalcin *et al.*, 2011; Regmi *et al.*, 2014; Vakhshoori and Zare, 2016; Akıncı *et al.* 2017; Chen *et al.* 2019).

### Studied Area

The area under investigation is located in Cianjur Regency between  $106^{\circ}46'56''$  and  $107^{\circ}29'23''$  E latitudes and  $6^{\circ}33'14''$  and  $7^{\circ}26'49''$  S longitudes (WGS84 UTM 48 S), covering about 3,730 km<sup>2</sup> area (Figure 1). The altitude of the region ranges from 0 to 2,961 m above the mean sea level.

Based on its physiography, West Java Province is divided into four zones, viz; Jakarta Coastal Plain, Bogor, Bandung or Central Depression

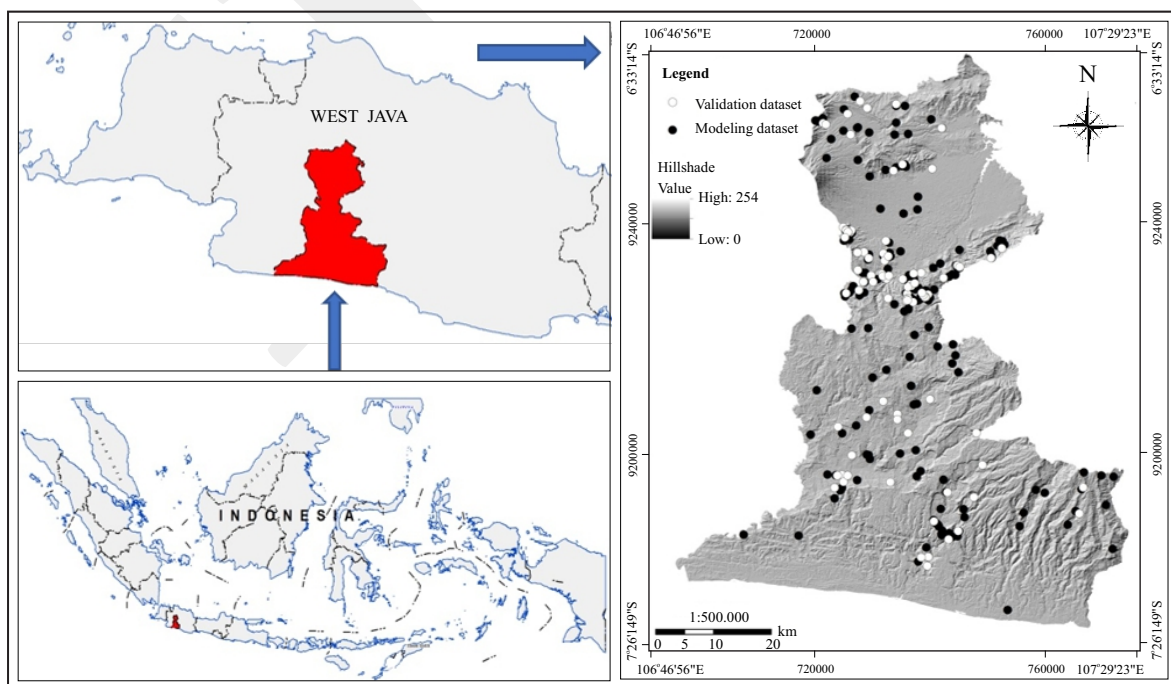


Figure 1. Location of the studied area in West Java Province and the location of landslides.

Zone, and West Java Southern Mountain Zone (Van Bemmelen, 1949). Cianjur area is situated in Bandung Zone with mainly morphological features of steep hills, and the predominant lithology is Quaternary volcanic products.

## METHODS AND MATERIALS

The landslide sampled as a homogen geo-referenced point (Poli and Sterlacchini, 2007; Neuhäuser *et al.*, 2011; Ozdemir, 2011; Tien Bui *et al.*, 2012; Xu *et al.*, 2014). An approach called seed cell was used to indicate the occurrence or non-occurrence of landslides. Seed cell is a neighborhood analysis (spatial analysis tool) to select landslide pixels within a buffer zone along the crown and flanks. It is a method to describe a pre-failure conditions, the undestroyed morphological conditions before the landslides occurred (Süzen and Doyuran, 2004; Nefeslioglu *et al.*, 2008; Bai *et al.*, 2010; Dou *et al.*, 2015; Hussin *et al.*, 2015).

The total of 298 landslide points in Cianjur area were compiled and mapped into the landslide inventory map. The landslide points as the seed cells were used to build the models. The points were randomly divided into 196 points (70%) as a training dataset for building process model. The other 89 points (30%) as a test/validation dataset were not used in building process model, but were used for validation purposes.

In this study, conditions considered as the primary factors were selected in the occurrence of a landslide in the studied area. There were the set of five landslide-related factors which were used and defined as conditioning factors. These conditioning factors are slope angle, slope aspect, elevation, lithological unit, and land use (Table 1). The factors were converted to raster maps of grid size of 15 x 15 m with a spatial resolution of 15 x 15 m. The relevant data and its analysis for this study were collected and processed in a GIS-environment using ArcGis 10.6 programmes.

In this study, Frequency Ratio, Weight of Evidence, and Information Value models were applied

Table 1. Details of Data Used in the Study

Category	Factors	Data Type	Scale/ resolution
Topographic map	Slope angle Slope aspect Elevation	Grid	15 x 15 m
Geological map	Lithology	Polygon	1 : 50.000
Land Use Map	Land Use	Polygon	1 : 50.000

on landslide susceptibility assessment to generate Landslide Susceptibility Maps (LSMs) of the studied area using the five landslide conditioning factors. All LSMs were classified into four landslide susceptibility zones based on the landslide distribution percentage of the total populated as very low (0% - 5%), low (5% - 10%), moderate (10% - 75%), and high (> 75%) (SNI, 2016).

### Frequency Ratio (FR)

The FR is one of probability models which is based on observed spatial relationships between landslide distribution and each conditioning factor related to landslides (Pradhan and Lee, 2010a, 2010b; Choi *et al.*, 2012; Mohammady *et al.*, 2012; Park *et al.*, 2013; Pardeshi *et al.*, 2013). FR is the ratio of landslides (the probability of an occurrence and a nonoccurrence) in a desired class (given attributes) as a percentage of all landslides (%Ld) to the area of the class as a percentage of the entire map (%Cd):

$$FR_d = (\%L_d) / (\%C_d) \dots\dots\dots (1)$$

Where:

$FR_d$  is the FR weight of the desired class.

The landslide susceptibility index (LSI) for each pixel or each factor ratio (Lee and Min 2001) is the summation of total overlapped pixels. It is formulated as:

$$LSI = \sum_{d=1}^n FR_d \dots\dots\dots (2)$$

### Weight of Evidence (WofE)

The theory of evidence (Weight of Evidence) is a log-linear version on the theorem of Bayes

used to calculate probability based on the concept of prior (P) and posterior probability (Agterberg *et al.*, 1993; Elmoulat *et al.*, 2015). This approach is based on the information obtained from the interrelation between landslide conditioning factors and the landslide distribution (Barbieri and Cambuli, 2009; Pardeshi *et al.*, 2013). The landslide conditioning factors are the input parameters for the WofE approach and to provide the information which may control the occurrence of areas prone to landslides (Arifianti and Agustin, 2017). The WofE calculates the spatial relationship between the conditioning factors with the distribution of landslides (VM), in the form of positive (W+) and negative weights (W-). These positive and negative weights are calculated from the ratios of the natural logarithms (Bonham-Carter, 1994; Elmoulat *et al.*, 2015), as below:

$$W^+ = \ln \frac{P \{VP / VM\}}{P \{VP / \overline{VM}\}} \dots\dots\dots (3)$$

$$W^- = \ln \frac{P \{\overline{VP} / VM\}}{P \{\overline{VP} / \overline{VM}\}} \dots\dots\dots (4)$$

The contrast of the weight (C) is added to define how significant the overall spatial association between the landslide conditioning factors and the landslide distribution (Dahal *et al.*, 2008; Neuhauser *et al.*, 2011). The contrast value is calculated as the difference of positive and negative weights (Ozdemir, 2011):

$$C = W^+ - W^- \dots\dots\dots (5)$$

#### Information Value Model (IVM)

The IVM is a statistical approach that has the advantage of assessing landslide susceptibility in an objective way. The IVM is used to calculate the weight for each class of factor layer by rationing landslide density of each class to the landslide density of the total area. In general, the landslides will occur in the future that has the same condition as the past landslides (Lee and Pradhan, 2006).

The IVM model is used to evaluate the spatial relationship between the conditioning factor classes and the probability of landslide occurrence. The higher value of IVM corresponds to the stronger relationship between the probability of landslide occurrence and the conditioning factor class. The IV model can be calculated as follows (Yin and Yan, 1988; Zhu *et al.*, 2004; Wang *et al.*, 2014):

$$I_i = \log_2 \frac{S_i / A_i}{S / A} \dots\dots\dots (6)$$

where:

$S_i$  is the number of landslides containing factor class (i),  
 $A_i$  is the area of factor class (i),  
 $S$  is total number of landslides, and  
 $A$  is the total area of the entire study.

#### Validation of Landslide Susceptibility Models

The validation of LSMs based on statistical methods reveals the reliability of the modelling processes. It is to compare the accuracy of different models and the choice of their parameter variables. The 'Area Under Curve' (AUC) of the 'Receiver Operating Characteristics' (ROC) method was performed for the validation. The success rate curve used the training dataset (70% of the whole set) to determine how well the resultant maps had classified the areas of existing landslides (Chung and Fabbri, 1999; Chen *et al.*, 2017). The prediction rate curve using the validation dataset (30% of the whole set) can explain how well the models and conditioning factors predict the future landslides (Chung and Fabbri, 2003; Pradhan and Lee, 2010a). The model accuracy ratings are usually given as 0.9 - 1.0 = excellent, 0.8 - 0.9 = good, 0.7 - 0.8 = acceptable, 0.6 - 0.7 = poor, and 0.5 - 0.6 = failed (Yilmaz, 2009).

## RESULTS AND DISCUSSION

### Spatial Relationship between Conditioning Factors and Landslides

The conditioning factors classified into several classes and weights were assigned to



them for FR, WofE, and IV methods as shown in Figure 2. The spatial relationship between the

conditioning factors and landslides is presented in Table 2.

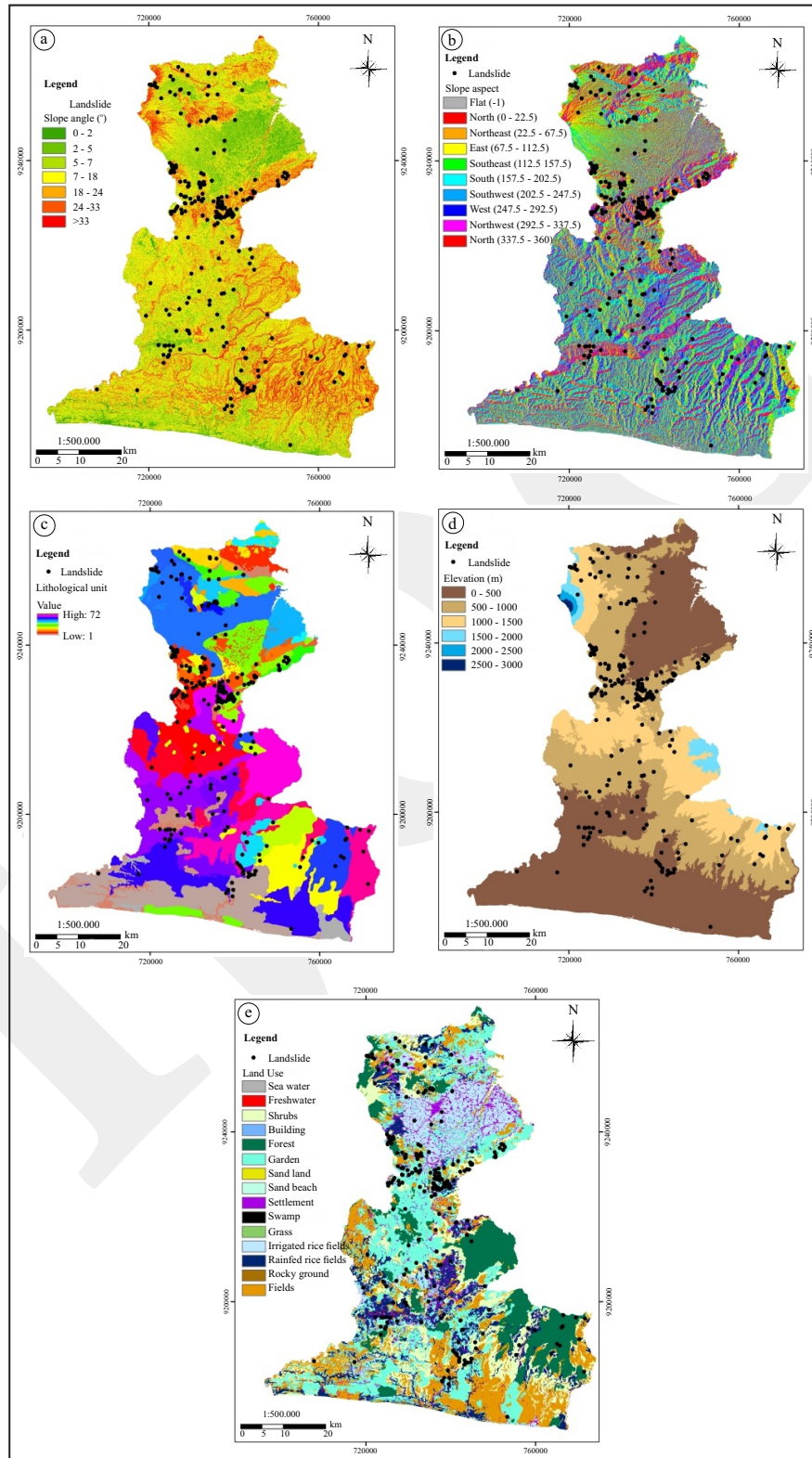


Figure 2. Landslide contributing-factor layers produced for the studied area: (a) slope angle, (b) slope aspect, (c) elevation, (d) lithological unit, and (e) land use.

Table 2. Spatial Relationship between Each Conditioning Factor and the Landslides for the FR, WofE, and IV Models

Conditioning Factor	Area Total (pixel)	Landslide points (n)	FR	WofE	IV
<b><i>Slope angle (°)</i></b>					
0 – 2	11225	4	309	-1207	-1174
2 – 5	19816	6	262	-1412	-1337
5 – 8	21774	20	796	-251	-227
8 – 18	61431	60	847	-243	-165
18 – 24	29537	45	1321	351	278
24 – 33	19021	39	1778	681	575
> 33	7184	22	2655	1055	976
<b><i>Slope aspect</i></b>					
Flat	4941	2	351	-1064	-1046
North	10996	19	1498	431	404
Northeast	18552	37	1729	632	547
East	19719	28	1231	231	208
Southeast	25710	31	1045	45	44
South	23396	10	370	-1086	-992
Southwest	19135	17	770	-295	-260
West	19963	15	651	-478	-428
Northwest	18320	20	946	-67	-54
North	9256	17	1592	491	465
<b><i>Elevation (m)</i></b>					
0-500	83104	47	490	-1224	-712
500-1000	54708	121	1918	1098	651
1000-1500	28107	26	802	-378	-220
1500-2000	3396	2	510	-798	-671
2000-2500	451	0	0	-5407	-17975
2500-3000	222	0	0	-5406	-18684
<b><i>Lithology</i></b>					
Limestone member	2458	0	0	-5302	-16279
Black sandstone, tuffaceous sandstone, siltstone	16049	7	378	-1030	-972
Claystone, siltstone, sandstone, tuffaceous sandstone, tuff	3093	1	280	-1276	-1271
Loose materials, blocks, boulder, gravel from sand, clay	3519	1	246	-1405	-1400
Andesite basalt intrusive	257	2	6749	1916	1909
Limestone	308	0	0	-5289	-18356
Andesite lava	239	0	0	-5289	-18610
Hornblende andesite Intrusive	93	2	18651	2947	2925
Breccia, tuffaceous sandstone, calcareous sandstone	227	2	7641	2041	2033
Old volcanic deposits: Breccia, lava	4098	22	4655	1628	1538
Mount Gede Volcanic Slide; basalt	1194	1	726	-323	-319
Andesite lava	1200	3	2168	777	773
Old volcanic lava deposit	594	6	8760	2198	2170
tuffaceous breccia, lava, sandstone, conglomerate	934	1	928	-77	-74
Lava, Beser sand	130	0	0	-5288	-19219
Breccia, lava, Beser sand	74	0	0	-5288	-19782
Old volcanic deposits: breccia, lava	436	3	5967	1796	1786
Alluvial fan: breccia, lava, lahar	1043	0	0	-5294	-17137
Old volcanic lava deposit	402	1	2157	766	768
Intrusive: vitrophyre, porphyry, dolerite	603	0	0	-5291	-17685
Sandstone	728	1	1191	171	175
Claystone	872	2	1989	687	687
Pyroclastic, lahar	339	0	0	-5290	-18260
Breccia andesite, Tuffaceous breccia	208	0	0	-5289	-18749
Brownish sandstone, Tuffaceous sandstone, andesite	7118	1	121	-2111	-2105
Pyroxene andesite	556	1	1559	441	444
Lahar deposits; breccia, tuf andesite	557	2	3114	1138	1135
Alluvial fan: breccia, lava, lahar	184	1	4713	1550	1550
Old Volcanic Sediment of Pasir Menteng: Clay, marl, quartz sandstone	1208	12	8615	2209	2153
Alluvial fan: breccia, lava, lahar	5003	1	173	-1757	-1752
pyroclastic, lahar	790	1	1097	89	93
Breccia	523	0	0	-5291	-17827
Quartz sandstone	983	5	4411	1500	1484
Breccia, Mount Manengge lava	719	1	1206	183	187
Quater limestone	1580	0	0	-5297	-16721

Table 2. continued...

Conditioning Factor	Area Total (pixel)	Landslide points (n)	FR	WofE	IV
Sandstone, siltstone	1399	8	4959	1631	1601
Mount Manengge lava	18	0	0	-5288	-21196
Breccia, lava, lahar	3881	2	446	-816	-805
Breccia, lava, Beser sand	468	0	0	-5290	-17938
loose material, clay, sand, boulder, gravel and the mix	1151	3	2260	819	815
Breccia, lava, lahar	319	0	0	-5290	-18321
Marl, quartz sandstone	1178	0	0	-5295	-17015
Loose material, clay, sand, gravel, blocks	2659	3	978	-25	-21
Basalt	9	0	0	-5288	-21889
Pyroclastic, lahar	1500	6	3469	1261	1243
Cirata Lake	8	0	0	-5288	-22007
Breccia, lava, lahar	3423	0	0	-5308	-15948
Mount Pangrango Volcanic deposits	135	0	0	-5288	-19181
Breccia, lava	357	0	0	-5290	-18209
Mount Gede lava	290	0	0	-5289	-18417
Mount Gede pyroclastic deposits	14567	10	595	-556	-518
Breccia, Mount Limo lava	1356	5	3197	1175	1162
Claystone and siltstone	603	3	4314	1470	1462
Mount Patuha breccia	5715	3	455	-803	-786
Breccia, Mount Balukbuk lava	723	1	1199	178	181
Breccia, lava, pyroclastic, lahar	494	2	3511	1258	1255
Andesite breccia, tuffaceous sandstone, tuff, lapilli, conglomerate	13756	3	189	-1717	-1665
Old terrace deposits	1541	0	0	-5297	-16746
Brownish sandstone, Tuffaceous sandstone, andesite	5918	7	1025	22	25
Claystone	8243	12	1262	242	233
Andesite breccia	5325	4	651	-441	-428
Claystone Beser Formation	500	0	0	-5291	-17872
Old volcanic sediment of Pasir Menteng: breccia, lava, tuff, conglomerate	1853	21	9828	2389	2285
Sandstone	972	0	0	-5293	-17207
Mount Kendeng lahar and lava	7254	0	0	-5331	-15197
Breccia, lava, lahar	4924	5	880	-133	-127
Breccia, lava, lahar	4932	5	879	-135	-128
Mount Patuha lava	1218	2	1424	351	353
Black sandstone	789	0	0	-5292	-17416
Breccia, lava, lahar	2305	0	0	-5301	-16344
Tuffaceous breccia, crystal tuff	6717	2	258	-1374	-1353
Pyroxene andesite	5169	9	1510	423	412
<b>Land Use</b>					
Sea water	16	0	-5298	0	-21314
Freshwater	879	0	-5303	0	-17308
Shrubs	24350	53	780	1887	635
Building	46	0	-5298	0	-20258
Forest	23426	22	-247	814	-205
Garden/agriculture land	44747	51	-29	988	-11
Sand land	3	0	-5298	0	-22988
Sand beach	2	0	-5298	0	-23393
Settlement	9986	14	194	1215	195
Swamp	13	0	-5298	0	-21522
Grass	653	1	270	1328	283
Irrigated rice fields	14456	8	-790	479	-734
Rain-fed rice fields	18052	25	192	1201	183
Rocky ground	10	0	-5298	0	-21784
Fields	33349	22	-667	572	-558

The spatial relationship between landslide occurrence and its conditioning factors using the three models indicates a relative similar susceptibility of each class. The most susceptible classes of the slope angle are 7° - 18°,

18° - 24°, and 24° - 33°. The models show that the landslide probability increases with the slope angle. This defined as a strong correlation between the slope angle and the landslide occurrence.

In the case of slope aspect and elevation factors, the models depicted the highest susceptible classes is the northeast facing slope with the elevation of 500 - 1,000 m a.s.l. The frequency of landslides is relatively lower on the south direction, with the exception in the flat areas. This means the two factors have less correlation with the landslide occurrence and elevation than the slope angle.

The result from lithology factor indicates that the most susceptible classes were (1) breccia, lava, tuff, and conglomerate from old volcanic sediments of Pasir Menteng, (2) clay, marl, and quartz sandstone from Rajamandala Formation, (3) breccia and lava from old volcanic deposits,

and (4) old volcanic lava deposits. These four lithological units are most prone to landslides in the studied area. The land use factor has an approximately similar susceptibility on the three models. It shows the highest susceptible is in the vicinity of settlement, shrubs, rain-fed rice fields, agricultural areas, and forestry region.

### Comparison and Validation

Landslide susceptibility maps were constructed from bivariate statistical analysis using the FR, WofE, and IV models. The LSMs obtained from three models were divided into four zones using the quantile method in ArcGis: very low, low, moderate, and high (Figure 3).

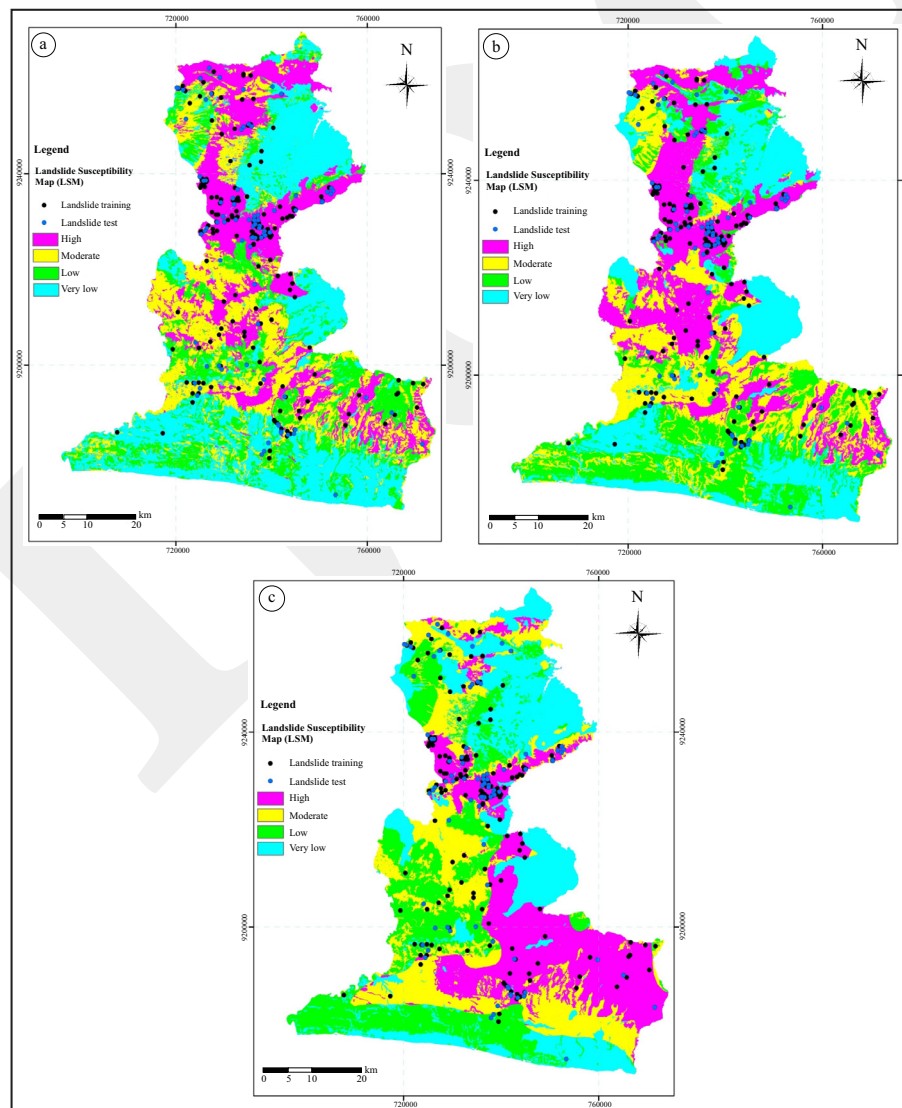


Figure 3. Landslide susceptibility map derived from the models of: (a) FR, (b) WofE, (c) IV.



The areas and the seed cells in the very low, low, moderate, and high in LSM of each models are shown in Table 3. Most of the landslides in

Table 3. Densities of Landslides among the Classified Susceptibility Zones of the Three Models

	Zones	Area %	Seed %
<b>FR</b>	Very Low	31.53	8.61
	Low	21.65	6.22
	Moderate	23.06	12.44
	High	23.77	72.73
<b>WofE</b>	Very Low	32.99	8.61
	Low	20.88	6.7
	Moderate	22.46	16.27
	High	23.66	68.42
<b>IV</b>	Very Low	24.91	6.22
	Low	25.07	10.05
	Moderate	24.97	22.49
	High	25.06	61.24

the whole studied area with the FR, WofE, and IV models have 85.17 %, 84.69%, and 83.73 % seed cells respectively occurring in areas with susceptibility zones of moderate to high. The three models produced acceptable results as it

shows the majority of the seed cells which are in moderate to high susceptibility zones.

Finally, the AUC of the ROC method was applied in order to reveal which model is more accurate in this study. The AUC was obtained for both the training dataset and the validation dataset (Figure 4). The AUCs value of success rates based on training dataset are 0.92 for the FR model, 0.926 for the WofE model, and 0.93 for the IV. The AUCs value of the prediction rate based on the validation dataset for the FR, WofE, and the IV models are 0.913, 0.912, and 0.895, respectively.

The result for the success rate and prediction rate curve shows that all the three models exhibit a similar performance. The models are found to have an excellent fit to the data with a slight difference where the IV model is the best one with the model accuracy of 93%, followed by WofE with 92.6%, and FR with 92%. It means the IV model produced the most accurate landslide susceptibility map in the studied area. In contrast, the model with the highest prediction ability is FR model with the prediction accuracy of 91.3%, followed by WofE with 91.2%, and IV model with 89.5%. It means the FR model showed the best accuracy in predicting the landslide susceptibility of the studied area.

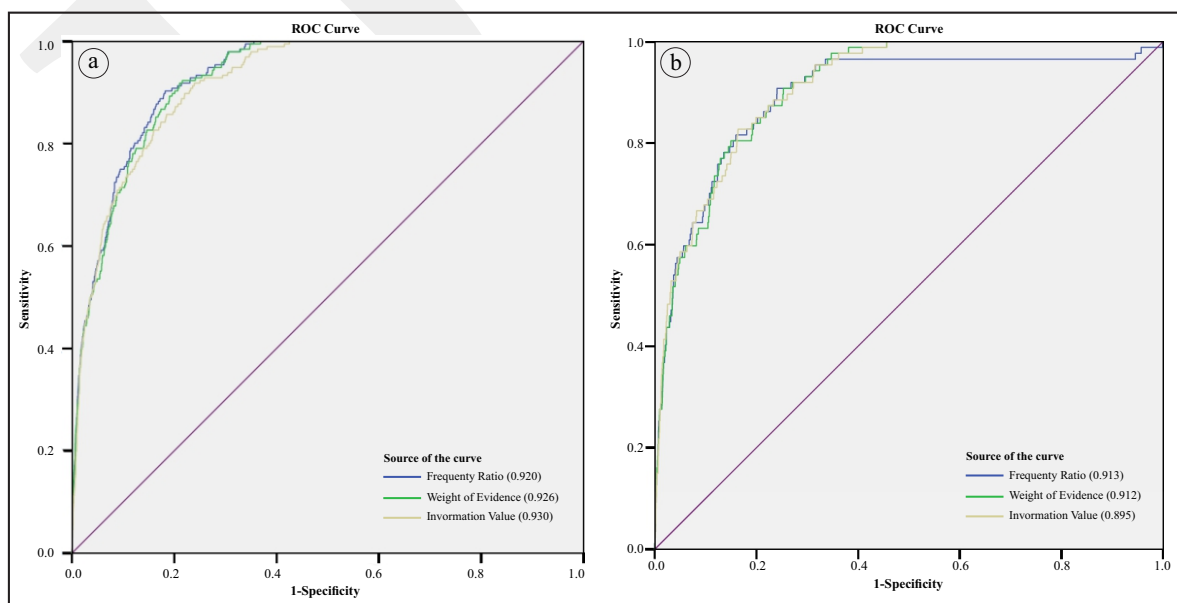


Figure 4. Success rate curve (a) and prediction rate curve (b) for different models by ROC curve.

## CONCLUSIONS

It is observed in Table 3, that the moderate to high susceptible zones of the LSMs produced by the FR, WofE, and IV model cover 46.83%, 46.12%, and 50.03% of the studied area, respectively. These covered areas are the most landslide-prone regions that should be considered in a susceptibility management. Preferably in the vicinity of settlement, shrubs, rain-fed rice fields, agricultural, and forestry area, with a slope angle between 18° - 33° and the elevation of 500 - 1,000 m a.s.l.

Although this bivariate statistical models, using the term “favourability values” by Chung *et al.* (1995) were applied in the conditioning factors (*e.g.*, slope units, lithological units, etc.) for better values to the expert’s opinion, the selection of the models and the landslide related factors was based on a consideration within expert’s scientific knowledge. This knowledge-base component was applied for finding the relevance, availability, and scale of data for the studied area.

According to the result given, the success rates and prediction rates of the three models are above 89% (Figure 4). The result reveals that the landslide susceptibility map of each model in this study has successfully achieved a high degree of reliability. The LSMs of the models will provide spatial-based decision making for the of government Cianjur Regency and other associated authorities and agencies.

## ACKNOWLEDGEMENTS

We thank our colleagues from Center for Volcanology and Geological Hazard Mitigation (CGVHM) and Center for Geological Survey - Geological Agency of Indonesia, who provided data and related documents. All authors are the main contributors. All read and approved with all of the results and conclusions of this paper.

## REFERENCES

- Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., and Wright, D.F., 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. *In: Davis JC, Herzfeld UC (eds). Computers in geology, 25 years of progress.* Oxford University Press, Oxford, p.13-32.
- Akıncı, H., Dogan, S., and Kiliçoğlu, C., 2017. Landslide susceptibility mapping of canik (samsun) district using bayesian probability and frequency ratio models. *The Selcuk International Science Conference on Applied Science (ISCAS)*, 2016, p.283-299. DOI: 10.15317/Scitech.2017.89
- Arifianti, Y. and Agustin, F., 2017. An assessment of the effective geofactors of landslide susceptibility: case study Cibeber, Cianjur, Indonesia. *In: Yamagishi, H. and Bhandary, N.P. (eds.), GIS Landslide.* Springer, Tokyo, p.183-195.
- Bai, S.B., Wang, J., and Zhou, P.G., 2010. GIS-based and logistic regression for landslide susceptibility mapping of Zhongxian segment in the Three Gorge area, China. *Geomorphology*, 115 (1-2), p.23-31. DOI: 10.1016/j.geomorph.2009.09.025 .
- Barbieri, G. and Cambuli, P., 2009. The weight of evidence statistical method in landslide susceptibility mapping of the Rio Pardu Valley (Sardinia, Italy). *18<sup>th</sup> World IMACS/ MODSIM Congress*, Cairns, Australia, p.2658-2664.
- Bonham-Carter, G.F., 1994. Geographic information systems for geoscientists: modeling with GIS. *In: Bonham-Carter F (ed), Computer methods in the geosciences.* Pergamon, Oxford, p.398.
- Chen, W. and Li, W., 2014. Application of weights-of-evidence model in landslide susceptibility mapping at Baozhong Region in Baoji, China. *Electronic Journal of Geotechnical Engineering*, 19, p.791-810.
- Chen, Z., Liang, S., Ke, Y., Yang, Z., and Zhao, H., 2019. Landslide susceptibility assessment using evidential belief function, certainty factor and frequency ratio model at Baxie River basin, NW China. *Geocarto International*, 34 (4), p.348-367. DOI: 10.1080/10106049.2017.1404143

- Choi, J., Oh, H.J., Lee, H.J., Lee, C., and Lee, S., 2012. Combining landslide susceptibility maps obtained from frequency ratio, logistic regression, and artificial neural network models using ASTER images and GIS. *Engineering Geology*, 124, p.12 - 23.
- Chung, C.J.F. and Fabbri, A.G., 1999. Probabilistic prediction models for landslide hazard mapping. *Photogrammetric Engineering and Remote Sensing*, 65 (12), p.1389-1399. DOI: 10.1002/9780470012659.ch4.
- Chung, C.J.F. and Fabbri, A.G., 2003. Validation of Spatial Prediction Models for Landslide Hazard Mapping. *Natural Hazards*, 30 (3), p.451-472. DOI: 10.1023/B:NHAZ.0000007172.62651.2b
- Chung, C.J.F., Fabbri, A.G., and van Westen, C.J., 1995. Multivariate regression analysis for landslide hazard zonation. In: Carrara, A. and Guzzetti, F. (eds.). *Geographical Information Systems in Assessing Natural Hazards*. Kluwer Academic Publisher, Dordrecht, The Netherlands, p.107-133.
- Conforti, M., Aucelli, P.P., Robustelli, G., and Scarciglia, F., 2011. Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (northern Calabria, Italy). *Natural Hazards*, 56 (3), p.881-898.
- Constantin, M., Bednarik, M., Jurchescu, M.C., and Vlaicu, M., 2011. Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). *Environmental Earth Science*, 63, p.397-340.
- Dahal, R.K., Hasegawa, S., Nonomura, A., Yamanaka, M., Masuda, T., and Nishino, K., 2008. GIS-based weights-of-evidence modelling of rainfall-induced landslides in small catchments for landslide susceptibility mapping. Springer Verlag. *Environmental Geology*, 54, p.311-324. DOI: 10.1007/s00254-007-0818-3
- Dou, J., Tien Bui, D.P., Yunus, A., Jia, K., Song, X., and Revhaug, I., 2015. Optimization of causative factors for landslide susceptibility evaluation using remote sensing and GIS data in parts of Niigata, Japan. *PLOS One*, 10 (7), e0133262.
- Elmoulat, M., Brahim, L.A., Mastere, M., and Jemmah, A.I., 2015. Mapping of mass movements susceptibility in the Zoumi Region using satellite image and GIS technology (Moroccan Rif). *International Journal of Scientific and Engineering Research*, 6 (2), p.210-217.
- Ercanoğlu, M., and Temiz, F.A., 2011. Application of logistic regression and fuzzy operators to landslide susceptibility assessment in Azdavay (Kastamonu, Turkey). *Environmental Earth Sciences*, 64, p.949-964. DOI: 10.1007/s12665-011-0912-4
- Ghosh, S., 2011. *Knowledge guided empirical prediction of landslide hazard*. Enschede: University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), the Netherlands.
- Gokceoglu, C. and Sezer, E., 2009. A statistical assessment on international landslide literature (1945 - 2008). *Landslides*, 6, p.345-351. DOI: 10.1007/s10346-009-0166-3.
- Hussin, H.Y., Zumpano, V., Reichenbach, P., Sterlacchini, S., Micu, M., van Westen, C., and Bălteanu, D., 2015. Different landslide sampling strategies in a grid-based bi-variate statistical susceptibility model. *Geomorphology*, 253, p.508-523. DOI: 10.1016/j.geomorph.2015.10.030.
- Kayastha, P., Dhital, M., and Smedt, F.D., 2012. Landslide susceptibility mapping using the weight of evidence method in the Tinau Watershed, Nepal. *Natural Hazards*, 63 (2), p.479-498. DOI: 10.1007/s11069-012-0163-z
- Lee, S. and Min, K., 2001. Statistical analysis of landslide susceptibility at Yongin, Korea. *Environmental Geology*, 40 (9), p.1095-1113. DOI: 10.1007/s002540100310.
- Lee, S. and Pradhan, B., 2006. Probabilistic landslide hazards and risk mapping on Penang Island, Malaysia. *Journal of Earth System Science*, 115 (6), p.661-667. DOI: 10.1007/s12040-006-0004-0

- Lin, M.L. and Tung, C.C., 2004. A GIS-based potential analysis of the landslides induced by the Chi-Chi earthquake. *Engineering Geology*, 71 (1-2), p.63-77. DOI: 10.1016/S0013-7952(03)00126-1.
- Mezoughi, T., Akhir, J.M., Rafek, A.G., and Abdullah, I., 2011. A multi-class weight of evidence approach for landslide susceptibility mapping applied to an area along the E-W Highway (Gerik - Jeli), Malaysia. *Electronic Journal of Geotechnical Engineering*, 16, p.1259-1273. DOI: 10.1007/s11053-007-9043-8.
- Mohammady, M., Pourghasemi, H.R., and Pradhan, B., 2012. Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. *Asian Journal of Earth Sciences*, 61, p.221-236. DOI: 10.1016/j.jseaes.2012.10.005.
- Mondal, S. and Maiti, R., 2012. *Landslide susceptibility analysis of Shiv-Khola Watershed, Darjiling: a remote sensing and GIS based analytic hierarchy process*. *Journal of Asian Earth Sciences*, 3, p.483-496.
- Nefeslioglu, H.A., Gökceoglu, C., and Sonmez, H., 2008. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Engineering Geology*, 97 (3), p.171-191. DOI: 10.1016/j.enggeo.2008.01.004.
- Neuhäuser, B., Damm, B., and Terhorst, B., 2011. GIS-based assessment of landslide susceptibility on the base of the Weights-of-Evidence model. *Landslides*, 9 (4), p.511-528. DOI: 10.1007/s10346-011-0305-5.
- Ozdemir, A., 2011. Landslide susceptibility mapping using bayesian approach in the Sultan Mountains (Aksehir, Turkey). *Natural Hazards*, 59 (3), p.1573-1607. DOI: 10.1007/s11069-011-9853-1.
- Pardeshi, S.D., Autade, S.E., and Pardeshi, S.S., 2013. Landslide hazard assessment: recent trends and techniques. *SpringerPlus*, 2 (1), p.523. DOI: 10.1186/21931801-2-523
- Park, S., Choi, C., Kim, B., and Kim, J., 2013. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environmental Earth Sciences*, 68, p.1443-1464.
- Poli, S. and Sterlacchini, S., 2007. Landslide representation strategies in susceptibility studies using weights-of evidence modeling technique. *Natural Resources Research*, 16 (2), p.121-134.
- Pradhan, B. and Lee, S., 2010a. Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60 (5), p.1037-1054.
- Pradhan, B. and Lee, S., 2010b. Landslide susceptibility assessment and factor effect analysis: back-propagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modeling. *Environmental Modelling and Software*, 25 (6), p.747-75. DOI: 10.1016/j.envsoft.2009.10.016
- 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 weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arabian Journal of Geosciences*, 7 (2), p.725-742. DOI: 10.1007/s12517-012-0807-z.
- Sarkar, S. and Anbalagan, R., 2008. Landslide hazard zonation mapping and comparative analysis of hazard zonation maps. *Journal of Mountain Science*, 5, p.232-240. DOI: 10.1007/s11629-008-0172-2
- Sharma, M. and Kumar, R., 2008. GIS-based landslide hazard zonation: a case study from the Parwanoo area, Lesser and Outer Himalaya, H.P., India. *Bulletin of Engineering Geology and the Environment*, 67, p.129-137. DOI: 10.1007/s10064-007-0113-2
- Süzen, M.L. and Doyuran, V., 2004. Data driven bivariate landslide susceptibility



- assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Engineering Geology*, 71 (3), p.303-321. DOI: 10.1016/s00137952(03)00143-1.
- Teerarungsikul, S., Torizin, J., Fuchs, M., Kühn, F., and Chonglakmani, C., 2015. An integrative approach for regional landslide susceptibility assessment using weight of evidence method: a case study of Yom River Basin, Phrae Province, Northern Thailand. *Landslides*, 13, p.1151-1165. DOI: 10.1007/s10346-015-0659-1.
- Tien Bui, D., Pradhan, B., Lofman, O., and Revhaug, I., 2012. Landslide Susceptibility Assessment in Vietnam Using Support Vector Machines, Decision Tree, and Native Bayes Models. *Mathematical Problems in Engineering*, 2012, p.1-26. DOI: 10.1155/2012/974638.
- Vakhshoori, V. and Zare, M., 2016. Landslide susceptibility mapping by comparing weight of evidence, fuzzy logic, and frequency ratio methods. *Geomatics, Natural Hazards and Risk*, 7 (5), p.1731-1752, DOI: 10.1080/19475705.2016.1144655 .
- Van Bemmelen, R.W., 1949. *The geology of Indonesia*, Government Printing Office, The Hague, 732pp.
- Vijith, H. and Madhu, G., 2007. Application of GIS and frequency ratio model in mapping the potential surface failure sites in the Poonjar subwatershed of Meenachil River in Western Ghats of Kerala. *Journal of the Indian Society of Remote Sensing*, 35 (3), p.275-285. DOI: 10.1007/BF03013495.
- Wang, J., Yin, K., and Xiao, L., 2014. Landslide susceptibility assessment based on GIS and weighted information value: a case study of Wanzhou district, Three gorges reservoir. *Chinese Journal of Rock Mechanics and Engineering*, 33, p.797-808.
- Xu, C., Xu, X., Yao, X., and Dai, F., 2014. Three (nearly) complete inventories of landslides triggered by the May 12, 2008 Wenchuan Mw 7.9 earthquake of China and their spatial distribution statistical analysis. *Landslides*, 11 (3), p.441-461. DOI: 10.1007/s10346-013-0404-6.
- Yalcin, A., Reis, S., Aydinoglu, A.C., and Yomralioglu, T., 2011. A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. *Catena*, 85 (3), p.274-287. DOI: 10.1016/j.catena.2011.01.014
- Yilmaz, I., 2009. Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey). *Computers & Geosciences*, 35, p.1125-1138. DOI: 10.1016/j.cageo.2008.08.007
- Yin, K.L. and Yan, T.Z., 1988. Statistical prediction models for slope instability of metamorphosed rocks. *Proceedings of the 5<sup>th</sup> International Symposium on Landslides, Lausanne, Switzerland*, 2, p.1269-1272.
- Zhu, A.X., Wang, R., Qiao, J., Qin, C.Z., Chen, Y., Liu, J., Du, F., Yang, L., and Zhu, T., 2014. An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic. *Geomorphology*, 214, p.128-138. DOI: 10.1016/j.geomorph.2014.02.003