



Applying different scenarios for landslide spatial modeling using computational intelligence methods

Alireza Arabameri¹ · Hamid Reza Pourghasemi² · Mojtaba Yamani³

Received: 14 September 2017 / Accepted: 11 December 2017 / Published online: 16 December 2017
© Springer-Verlag GmbH Germany, part of Springer Nature 2017

Abstract

Landslides every year impose extensive damages to human beings in various parts of the world; therefore, identifying prone areas to landslides for preventive measures is essential. The main purpose of this research is applying different scenarios for landslide susceptibility mapping by means of combination of bivariate statistical (frequency ratio) and computational intelligence methods (random forest and support vector machine) in landslide polygon and point formats. For this purpose, in the first step, a total of 294 landslide locations were determined from various sources such as aerial photographs, satellite images, and field surveys. Landslide inventory was randomly split into a testing dataset 70% (206 landslide locations) for training the different scenarios, and the remaining 30% (88 landslides locations) was used for validation purposes. To providing landslide susceptibility maps, 13 conditioning factors including altitude, slope angle, plan curvature, slope aspect, topographic wetness index, lithology, land use/land cover, distance from rivers, drainage density, distance from fault, distance from roads, convergence index, and annual rainfall are used. Tolerance and the variance inflation factor indices were used for considering multi-collinearity of conditioning factors. Results indicated that the smallest tolerance and highest variance inflation factor were 0.31 and 3.20, respectively. Subsequently, spatial relationship between classes of each landslide conditioning factor and landslides was obtained by frequency ratio (FR) model. Also, importance of the mentioned factors was obtained by random forest (RF) as a machine learning technique. The results showed that according to mean decrease accuracy, factors of altitude, aspect, drainage density, and distance from rivers had the greatest effect on the occurrence of landslide in the study area. Finally, the landslide susceptibility maps were produced by ten scenarios according to different ensembles. The receiver operating characteristics, including the area under the curve (AUC), were used to assess the accuracy of the models. Results of validation of scenarios showed that AUC was varying from 0.668 to 0.749. Also, FR and seed cell area index indicators show a high correlation between the susceptibility classes with the landslide pixels and field observations in all scenarios except scenarios 10_{RF} and 10_{SVM} . The results of this study can be used for landslides management and mitigation and development activities such as construction of settlements and infrastructure in the future.

Keywords Landslide spatial modeling · Scenario-based modeling · Statistical models · Computational intelligence methods

Introduction

Landslides are the movement of materials that form the slope including natural rocks, soil, artificial accumulations, or a mixture of them that move to lower parts by gravity force (Guzzetti 2015). The landslides are one of the most common catastrophic natural dangers that occur in any regions of world and cause to hundreds million dollar economic loss, soil erosion and hundreds thousand mortality and injuries yearly (Aleotti and Chowdhury 1999; De Sy et al. 2013; Lee et al. 2017). During recent years many governments and research institutions have invested for preparing the maps that indicate the landslides spatial distribution (Xie et al. 2005; Guo-liang

✉ Hamid Reza Pourghasemi
hr.pourghasemi@shirazu.ac.ir

¹ Department of Geomorphology, Tarbiat Modares University, Tehran, Iran

² Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz, Iran

³ Physical Geography (Geomorphology), Faculty of Geography, University of Tehran, Tehran, Iran

et al. 2017). Regardless the obtained progresses for identifying, measuring, forecast, and warning systems of landslide, but the losses result from landslide are increasingly in the worldwide, yet (Kincal et al. 2010). The landslides are the result of interconnected local temporal processes including hydrological processes (precipitation, evaporation, and underground waters), vegetation weight, root resistance, soil condition, mother stone, height from sea level, slope degree and direction, topography, and human activities (Youssef 2015; Myronidis et al. 2016). Over the last decades, different methods have been developed to evaluate landslide susceptibility in different parts of the world. These methods can be divided into two categories including qualitative and quantitative methods. Qualitative methods are based on field observations and knowledge of the experts. In these methods, weight of conditioning factors was determined by experts (Wen et al. 2017; Achour et al. 2017). Quantitative methods that in the last few decades have often been used for landslide susceptibility mapping are consisted from three main categories: deterministic approaches, statistical methods, and computational intelligence methods. Statistical methods can be divided into two categories including bivariate and multivariate methods. The bivariate methods are including information value method (Chen et al. 2016a, b; Ba et al. 2017; Achour et al. 2017), frequency ratio (Wu et al. 2016; Li et al. 2017), certainty factor (Wen et al. 2017; Hong et al. 2017a, b; Kornejady et al. 2017), and evidential belief function (Ding et al. 2016; Pourghasemi and Kerle 2016). In contrast, multivariate statistical analyses are known as logistic regression (Wang et al. 2015; Colkesen et al. 2016; Horafas and Gkeki 2017; Guo-liang et al. 2017). Computational intelligence models are artificial neural network (Dou et al. 2015; Moosavi and Niazi 2015; Wang et al. 2016; Chen et al. 2016a, b; Pourghasemi et al. 2017; Zeng et al. 2017), support vector machines (Ren et al. 2015; Hong et al. 2016a, b; Colkesen et al. 2016; Chen et al. 2017a, b), and random forest (Hong et al. 2016a, b; Zhang et al. 2017; Chen et al. 2017a, b; Kim et al. 2017; Lai et al. 2017; Pourghasemi and Rahmati 2017; Zhang et al. 2017). Several studies have also been done on various scenarios (Avolioa et al. 2000; Du et al. 2013; Mantovani et al. 2000; Prompera et al. 2014). Quantitative and qualitative methods have disadvantages and advantages in the literature reviews. The accuracy of qualitative methods is significantly dependent on the expertise of researcher (Feizizadeh et al. 2014), whereas deterministic models because they depend on the computation of the relevance between resisting and provocative forces, which requires precise data on the slope geometry, soils and rock's, and hydrological conditions are usually used in small region (Armas et al. 2014). In the statistical methods, bivariate models can obtain the impact of each conditioning factor class on landslide occurrence, but it does not consider variables importance, while multivariate statistical method is opposite to it (Guo-liang et al. 2017). Statistical models that

are proven to be suitable for studying large areas in terms of landslide susceptibility, but these models do not easily predict unforeseen relationships between large numbers of landslide conditioning factors and complex landslide systems (Pourghasemi et al. 2013). Computational intelligence models focus on appropriate learning approaches for recognition of the nonlinear relationship between conditioning factors and landslides (Gordan et al. 2015). These models have been successfully implemented for landslide susceptibility mapping.

So, the main objective of this study is to provide different scenarios for landslide susceptibility assessment in the Ghaemshahr Watershed, Mazandaran Province, Iran, using a combination of statistical and computational intelligence methods. In this regard, among the statistical methods, frequency ratio and among computational intelligence methods random forest and support vector machines were selected for applying ten scenarios on landslide modeling. By the way, the mentioned research aims to consider scenario-based landslide modeling in point and polygon formats.

Materials and methods

Study area

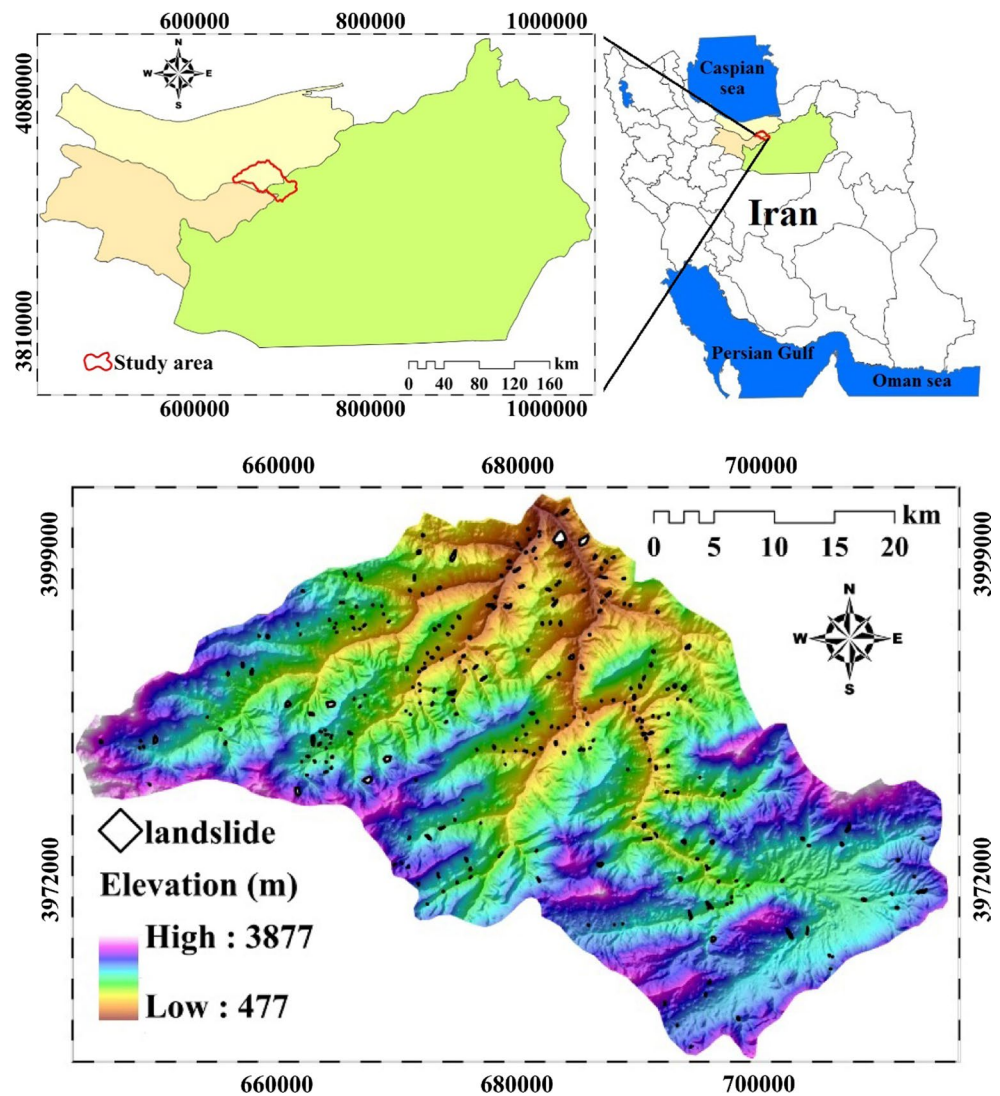
The Ghaemshahr Watershed with a total area of 1637 square kilometers is located approximately in 43 kilometers southwest of the city of Sari in Mazandaran Province. The study area lies between the latitudes of 35°44'–36°09'N, and longitudes of 52°36'–53°23'E (Fig. 1). The maximum height of the study area is located in the southwest with a height of 3877 m above sea level, and the minimum height is in the northeast of the area with a height of 476 m a.s.l. The monthly average rainfall of this basin is more than 500 mm, and maximum rainfall occurs during January to April according to nine rainfall stations (Alasht, Tale-Savadkoh, Alvand-Doab, Ori-Melk, Zardgol-Sorkhabad, Doabe-Savadkoh, Veresk, Docal, and Nesa) for the years 1985–2015 (Meteorological Organization, <http://www.irimo.ir/far/>). The study area is covered by various types of lithological formations including Triassic, Pliocene, Cretaceous, Eocene, Miocene, Cambrian, Devonian, Quaternary, and Paleozoic. Most of the study area is covered by forest (B), about 636 km². Other land use types are orchard (A), range (C), forest and dryfarming (D), irrigation agriculture and range (E), forest and range (F), and residential area (G).

Methodology

The flowchart of the methodology used in this study is shown in Fig. 2 and consists of five phases:

- (1) Preparation of data,

Fig. 1 Study area location



- (2) Multi-collinearity analysis among conditioning factors using tolerance and VIF indices,
- (3) Determination of the relationship between landslide occurrence and conditioning factors using FR model,
- (4) Running SVM and RF intelligent models by applying different scenarios on landslide point and polygon formats, and
- (5) Validation of the landslide susceptibility maps using the ROC curve.

Conditioning factors database

The tools used in this research are ArcGIS10.5, ENVI 4.8 (for extraction of LU/LC), SAGA-GIS 2.1.1, and Global Positioning System (GPS). The basic maps used were geological maps by scale of 1: 100,000, aerial photos (08/02/1964) on scale 1:40,000, topographic maps with scale of 1:50,000, satellite images of Landsat8, ASTER DEM, LISS-III, and rainfall data for a 30-year period

(from 1985 to 2015) and nine rainfall stations (Alasht, Tale-Savadkoh, Alvand-Doab, Ori-Melk, Zardgol-Sorkhabad, Doabe-Savadkoh, Veresk, Docal, and Nesa).

The first step for mapping landslide susceptibility and risk analysis is collecting data about landslides that have occurred in the past, so preparing landslide inventory map is prerequisite for such studies (Guzzetti et al. 2012). In the study area, a total of 294 landslides were mapped using aerial photograph with 1:40,000-scale, satellite imagers (IRS: LISS-III), and several field surveys. Most of the landslides are shallow rotational with a few translational. In this research, the landslide inventory was randomly split into a testing dataset 70% (206 landslide locations) for training the models and the remaining 30% (88 landslides locations) was used for validation purpose (Chen et al. 2017a, b). Field photographs of some identified landslides in the study area are shown in Fig. 3.

Identification and selection of conditioning factors on landslide occurrence is one of the most important steps for

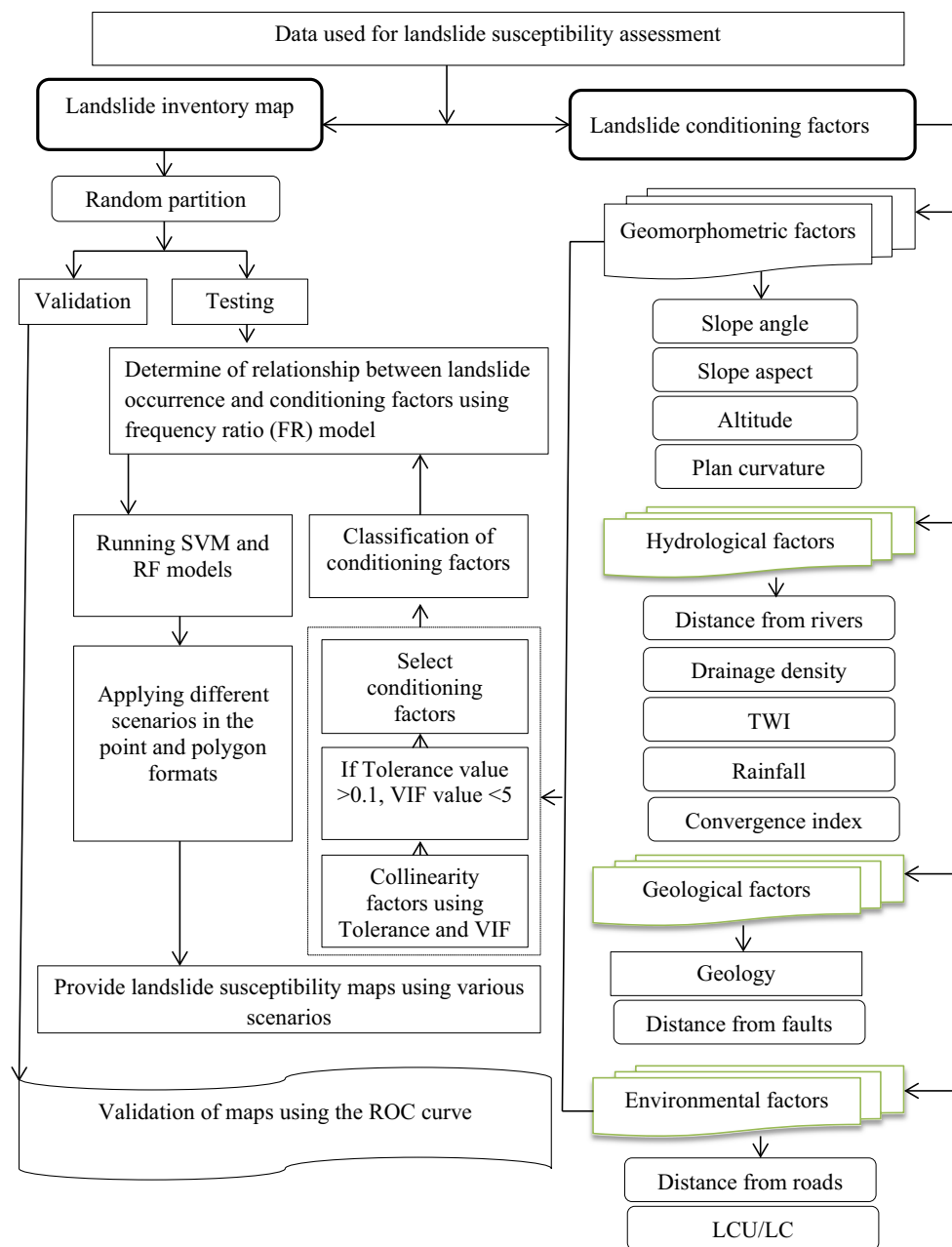


Fig. 2 Flowchart of research of in the study area

landslide susceptibility mapping (Ercanoglu and Gokceoglu 2002). In this research, based on the study of previous researches (Zhang et al. 2017; Zeng et al. 2017) and features of the study area, 13 conditioning factors affecting landslide such as slope aspect, slope degree, altitude, plan curvature, distance from river, drainage density, distance from fault, distance from road, LU/LC, TWI, annual rainfall, geology, and convergence index were selected (Fig. 4). Maps related to the effective factors were prepared in the ArcGIS 10.5 and prepared for processing. In order to prepare DEM (digital elevation model), slope aspect, slope degree, and

geomorphometric parameters such as TWI, plan curvature, and convergence index, ASTER DEM with 30 m spatial resolution are used. The elevation does not contribute directly to landslide occurrence, but in relation to the other parameters, like tectonics, erosion–weathering processes, and precipitation, the elevation contributes to landslide occurrence and influences the whole system (Rozos et al. 2011). The elevation map for study area with cell size 30 m × 30 m was produced from the ASTER DEM and classified into six classes of 477–1200, 1200–1600, 1600–2000, 2000–2400, 2400–2800, and > 2800 m (Fig. 4a). The slope map of the



Fig. 3 Field photographs of some identified landslides in the study area

study area is derived from the ASTER DEM using the slope function in ArcGIS 10.5. These slope values (in degree) are based on natural break scheme and divided into six different classes (Wu et al. 2016) including flat-gentle slope < 10, fair slope (10–15), low slope (15–20), moderate slope (20–30), steep slope (30–40), and very steep slope > 40 (Fig. 4b). Slope aspect strongly affects hydrological processes by means of evaporation-transpiration, direction of frontal precipitation, and thus affects weathering processes and vegetation and root development (Sidle and Ochiai 2006). Aspect layer has been categorized into nine classes including flat, north, northeast, east, southeast, south, southwest, west, and northwest (Fig. 4c). The convergence index (CI) gives a measure of how flow in a cell diverges (convergence index < 0) or converges (convergence index > 0) (Claps et al. 1994). CI map provided in SAGA-GIS 2.1.1

and divided into 5 classes: – 100 to – 22, – 22 to – 6, – 6 to 6, 6–21, and > 21 (Fig. 4d). The curvature of the surface (plan curvature) reflects the directional variations along a curve. The effect of plan curvature on the slope erosion process is the convergence and divergence of water along the flow direction (Ercanoglu and Gokceoglu 2002). The plan curvature map was produced using ArcGIS 10.5 and was classified into three categories (Pourghasemi and Kerle 2016): concave, flat, and convex (Fig. 4e). TWI is a combination of ups and downs that shows the ratio between slopes in the basin (Eq. 1). TWI index is an indicator of the spatial distribution of soil moisture along the landscape. Therefore, it is used for landslide susceptibility mapping (Pourghasemi et al. 2014; Naghibi et al. 2015).

$$TWI = \ln \left(\frac{S}{\tan \alpha} \right) \tag{1}$$

where *S* is the cumulative upslope area draining and *a* is the slope gradient in degrees (Moore et al. 1991). The TWI map divided into four classes (Hong et al. 2016a, b): – 5.4 to – 0.28, – 0.28 – 1.26, 1.26–4, and > 4 (Fig. 4f).

By applying the gradient formula of the rainfall region on the digital elevation model (Eq. 2), map of annual rainfall was prepared and classified into five classes (Chen et al. 2016a, b): 475–520, 520–580, 580–620, 620–670, and > 670 mm/yr (Fig. 4g).

$$Y = 0.5667X + 108.33 \tag{2}$$

where *Y* is annual rainfall and *X* is altitude. *R*² is 0.923.

The distance from linear factors such as river, road, and fault was calculated using the distance function available in the ArcGIS.10.5. Distance from roads (Fig. 4h) and distance from rivers (Fig. 4i) divided into six classes: 0–100, 100–200, 200–300, 300–400, 400–500, and > 500 m. Distance from fault also divided into five classes: 0–500, 500–1000, 1000–1500, 1500–2000, and > 2000 m (Fig. 4j). Drainage density is prepared by applying the density line function available in the ArcGIS.10.5 and divided into five classes (Chauhan et al. 2010): 0–0.5, 0.5–0.8, 0.9–1.1, 1.1–1.4, and > 1.4 mm² (Fig. 4k). Land use/land cover (LU/LC) map of the study area was prepared using Landsat 8 images. To create the land use map, a supervised classification using the maximum likelihood algorithm (Pourghasemi and Kerle 2016) was applied. Seven land use types were extracted such as orchard (A), forest (B), range (C), forest and dryfarming (D), irrigation agriculture and range (E), forest and range (F), and residential area (G) (Fig. 4l). The generated LC/LU was validated using 285 field verification points in the field. Kappa coefficient for the final map was estimated by Eq. 3. (Lo and Yeung 2002).

$$K = \frac{\{N \sum_{i=1}^r (X_{ii}) - N \sum_{i=1}^r (X_{i+} \cdot X_{+i})\}}{N^2 - \sum_{i=1}^r (X_{i+} \cdot X_{+i})} \tag{3}$$

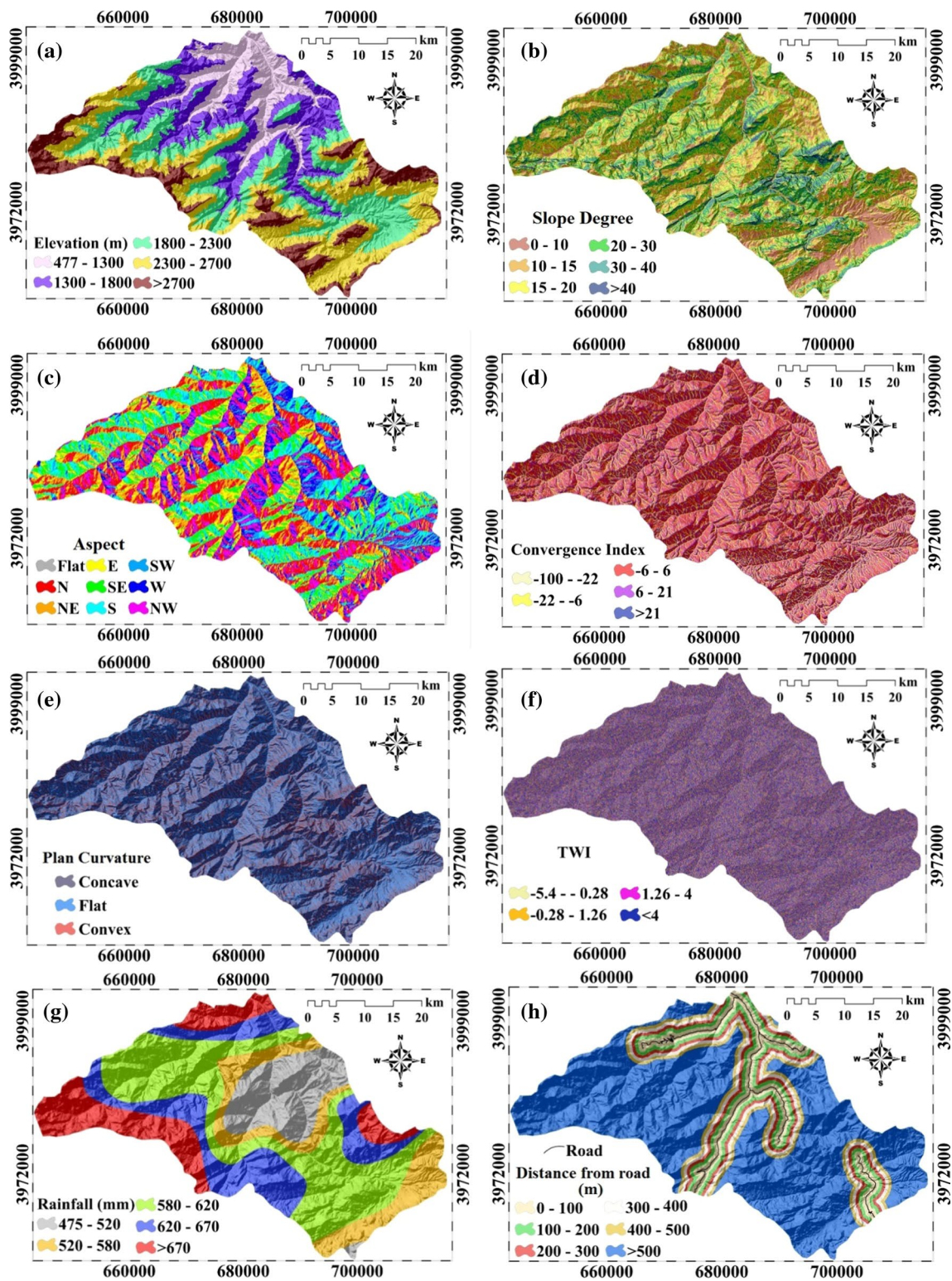


Fig. 4 Landslide conditioning factors

where r is number of rows in error matrix; X_{ii} is number of observations in row i and column i ; X_{i+} is total of observations in row i ; X_{+1} is total of observations in column i ; and

N is total number of observations included in the matrix. Kappa coefficient of generated LC/LU 97.65 was obtained.

Geological map of the region was prepared based on the digitization of the polygons of the lithological units in

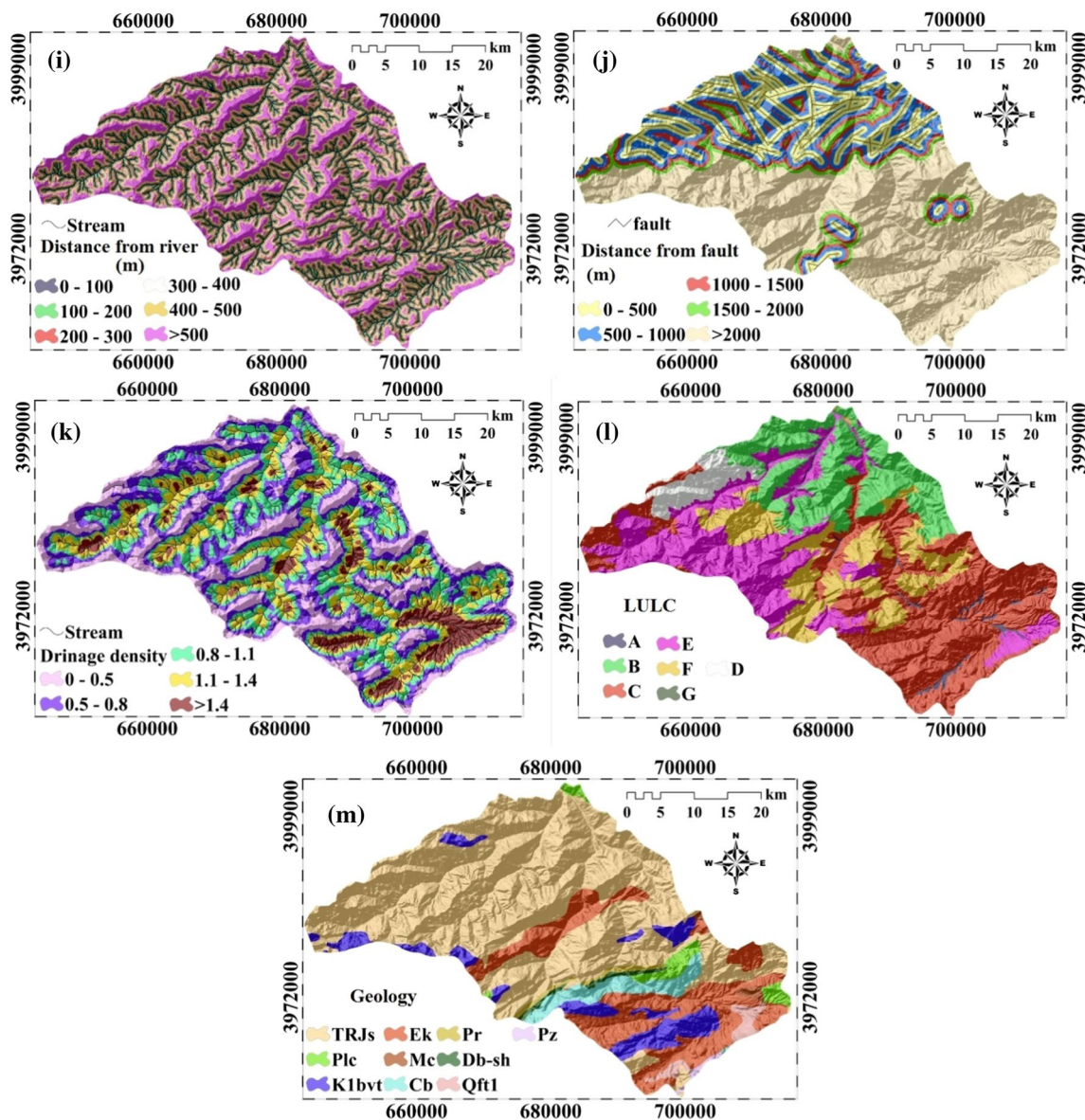


Fig. 4 (continued)

the geological map with a 1:100,000-scale in ArcGIS10.5. The lithological units were classified into ten categories according to formation and their susceptibility to landslide occurrence (Fig. 4m and Table 1).

For the classification of the conditioning factors, different methods such as manual, equal interval, and natural break were used.

Multi-collinearity analysis

An important issue in the use of conditioning factors in the preparation of a landslide susceptibility map is the effect of correlation among conditioning factors. When there is a high correlation between the two independent variables,

there is a problem called multi-collinearity. This high correlation reduces the accuracy of the results. Tolerance and the variance inflation factor (VIF) are two important indices for multi-collinearity recognition. A tolerance of less than 0.20 or 0.10 and/or a VIF of 5 or 10 and above indicates a multi-collinearity problem (Pourghasemi et al. 2013).

Determination of the relationship between landslide occurrence and conditioning factors using FR model

The FR model is considered as the most popular and the simplest approach for preparing landslide susceptibility maps (Wu et al. 2016). The FR is based on the observed relationship

Table 1 Lithology of the study area (GSI, 1997)

Code	Geological age	Lithology	Area (km ²)
TRJs	Triassic-Jurassic	Dark gray shale and sandstone (Shemshak formations)	1049.41
Plc	Pliocene	Polymictic conglomerate and sandstone	37.77
K1bvt	Early Cretaceous	Basaltic volcanic tuff	129.69
Ek	Eocene	Well-bedded green tuff and tuffaceous shale (Karaj Formation)	293.88
Mc	Miocene	Red conglomerate and sandstone	7.56
Cb	Cambrian	Alternation of dolomite, limestone, and variegated shale (Barut formation)	63.73
Pr	Permian	Dark gray medium-bedded to massive limestone (Ruteh limestone)	14.20
Db-sh	Devonian	Undifferentiated limestone, shale, and marl	12.17
Qft1	Quaternary	High-level piedmont fan and valley terrace deposits	23.52
Pz	Early Paleozoic	Undifferentiated lower Paleozoic rocks	4.89

between the distribution of landslides and each landslide-related factor (Tay et al. 2014). In determining the ratio of frequency ratios, the occurrence ratio of landslide in each class of conditioning factors toward the total of landslides is obtained and ratio of the surface of each class toward the total area of the region is also calculated. Finally, by dividing the occurrence ratio of landslides in each class by the rate of area of each class, relative to the entire study area, the frequency ratio of the classes of each factor is calculated. The calculation of the frequency ratio for each class of conditioning factors is expressed in Eq. 4 (Pradhan and Lee 2010):

$$FR = \frac{\left(\frac{A/B}{C/D}\right)}{\frac{E}{F}} \quad (4)$$

where A is the number of landslide pixels in each class, B is the total number of landslide pixels of the whole area, C is the number of pixels in each class of conditioning factors, D is the total number of pixels in the area, E is the percentage of landslide occurrences in each class of conditioning factors, and F is the relative percentage of the area of each class.

Support vector machine (SVM)

Support vector machine is a supervised learning method based on statistical learning theory (Vapnik 1995) and the principle of structural risk minimization (Chen et al. 2016a, b). It is based on the statistical approach in order to find an optimal hyperplane for separating two classes (Tien Bui et al. 2016). A more detailed of SVM algorithm for landslide assessment has recently been depicted by Marjanovic et al. (2011); Colkesen et al. (2016).

To perform the landslide susceptibility map using SVM, the “rminer” package (Cortez 2015) was used. Meanwhile, there are four types of kernels: linear, polynomial, radial

basis function (RBF), and sigmoid for SVM modeling, in this research used from RBF kernel.

Random forest (RF)

The random forest algorithm that developed by Breiman (2001) is based on a bunch of decision trees, and now it is one of the best learning patterns (Zhang et al. 2017). This model is based on the averaging of the results of all decision trees. The random forest method is widely used for data prediction and interpretation purposes and is suitable for non-linear high-dimensional landslide susceptibility modeling problems (Messenzehl et al. 2016). The RF algorithm tends to produce quite accurate models, because it decreases the variance of the model, without increasing the bias (Hastie et al. 2009). This algorithm needs two original parameters to be implemented by the user: the number of trees (T) and the number of variables (m).

The main advantage of this approach is that it can categorize a large number of input variables without variable deletion (Immitzer et al. 2012). Compared with other algorithms, this model has more efficiency in the classification of a large dataset; moreover, it can process high-dimensional datasets without feature selection and can rank the parameters in terms of importance after calculating (Zhang et al. 2017). This method uses unbiased estimation in model building, which means that the training is fast and simple to running. This method is suitable for regional-scale applications and is useful for the landslide susceptibility mapping (Zhang et al. 2017). For running random forest is used from R statistical and randomForest package (Briman and Cutler 2015).

Applying different scenarios for ensemble of intelligent techniques

After calculation of weight of classes of each factor using bivariate statistical method (FR), and running RF and SVM

computational intelligence methods, different scenarios were developed to provide a reasonable landslide susceptibility map in both landslide point and polygon formats According to Eqs. 5–16:

$$\text{Scenario1} = \frac{(\text{FR} + \text{SVM})}{2} \tag{5}$$

$$\text{Scenario2} = (2 \times \text{RF}) + \text{SVM} \tag{6}$$

$$\text{Scenario3} = (2 \times \text{SVM}) + \text{RF} \tag{7}$$

$$\text{Scenario4} = \text{SVM} + \text{RF} \tag{8}$$

$$\text{Scenario5} = \text{SVM} \times \text{RF} \tag{9}$$

$$\text{Scenario6} = \frac{(\text{SVM} + \text{RF})}{3} \tag{10}$$

$$\text{Scenario7} = \left(\frac{(\text{FR} \times \text{AUC}_{\text{RF}}) + (\text{SVM} \times \text{AUC}_{\text{SVM}})}{(\text{AUC}_{\text{RF}} + \text{AUC}_{\text{SVM}})} \right) \tag{11}$$

$$\text{Scenario8} = \text{RF weighted by FR} \tag{12}$$

$$\text{Scenario9}_{\text{RF}} = \text{Individual RF model} \tag{13}$$

$$\text{Scenario9}_{\text{SVM}} = \text{Individual SVM model} \tag{14}$$

$$\text{Scenario10}_{\text{RF}} = \text{Individual RF model} \tag{15}$$

$$\text{Scenario10}_{\text{SVM}} = \text{Individual SVM model} \tag{16}$$

where scenarios of 10_{SVM} and 10_{RF} are for landslide point format and in the rest of scenarios, landslides are in polygon format. AUC is area under the curve models.

Results and discussion

In the Ghaemshahr Watershed, landslides are the most serious natural problems that impress the economic development and cause loss of fertile soils and great damage to land and property. According to earlier studies by the Iranian Landslide Working Party (ILWP 2007), the highest frequency of landslide occurrence in Iran is in Mazandaran Province.

The soil is one of the most important components of the earth system as it controls erosional, hydrological, biological, and geochemical Earth cycles and provides a widespread range of services, goods, and resources to human kind (Keesstra et al. 2016, 2018; Comino et al. 2016; Vaezi et al. 2017). Soil erosion in agricultural areas because of loss of productivity and land degradation is a large problem worldwide (Kirchhoff et al. 2017) that must be solved by

means of nature-based strategies to be able to achieve sustainability (Cerdà et al. 2017). Also, soil erosion is a key factor of desertification and affects the goals for sustainability of the United Nations (Keesstra et al. 2016). Many of the 15 Sustainable Development Goals defined by UN have a strong relation to land and water management and demanding a sustainable use of resources, ecosystem restoration, biodiversity, and sustainable basin management (Keesstra et al. 2016). The development of human societies requires the prudently use of natural resources such as soil (Keesstra et al. 2016). Soil conservation not only depends on wise decisions by foresters, farmers, and land planners, but also on political decisions on rules and regulations (Keesstra et al. 2016).

Multi-collinearity among conditioning factors

In this study, the multi-collinearity test considered according to two indices such as VIF and tolerance (Table 2). According to Table 2, the smallest tolerance and highest VIF were 0.31 and 3.20, respectively. So, there is not any multi-collinearity between independent factors in the current research.

Spatial relationship between landslides and conditioning factor using FR model

Spatial relationship between landslides and conditioning factor by frequency ratio model is shown in Table 3. In the case of the relationship between landslide occurrence and altitude factors, most of landslide events are located in 477–2400 m including 477–1200 (2.53), 1200–1600 (3.39), 1600–2000 (1.21), and 2000–2400 (1.11), respectively. According to the results, at altitudes below 1200 m, due to the human activities such as agriculture and road construction, the most susceptibility to landslide has been shown. In contrast, at altitudes above 2800, due to rocky outcrops, the

Table 2 Multi-collinearity test

Factors	Collinearity statistics	
	Tolerance	VIF
Aspect	0.98	1.02
Convergence	0.68	1.46
Altitude	0.31	3.21
Drainage density	0.64	1.55
Distance from faults	0.48	2.07
Geology	0.70	1.43
LU/LC	0.73	1.36
Plan curvature	0.71	1.40
Rainfall	0.78	1.28
Distance from rivers	0.72	1.40
Distance from roads	0.42	2.38
TWI	0.99	1.00
Slope degree	0.97	1.03

probability of landslide is negligible. In the case of slope degree, most of landslide occurrences are in slope ranges of 20°–40°, including 20°–30° (1.085), and 30°–40° (1.289), respectively. The results showed that with the increasing in slope degree, the probability of landslide occurrence has also increased. While the slope degree increases, the shear stress on the slope material increases and the probability of landslide occurrence increases. Although steep slopes due to outcropping bedrock may not be susceptible to shallow landslides (Mohammady et al. 2012; Wu et al. 2016). In the case of slope aspect, aspect parameter on southwestern-facing slopes represents the highest probability (2.171) to landslide occurrence. In the case of convergence index, class of – 22 to – 6 has the highest FR value (1.043) and class of > 21 has the lowest FR value (0.000). Based on plan curvature parameters, concave class with FR (1.110) has shown the most susceptible to the occurrence of landslides, whereas convex and flat classes with (0.987, 0.894) are located in the next ranks, respectively. According to results of TWI factor, class of – 5.4 to – 0.28 with score of (1.02) is located in the first rank in terms of susceptibility to landslide and classes of (< 4, – 0.28 to 1.26 and 1.26–4) with FR values of (1.01, 0.99, and 0.98) are located in the next ranks. In the case of rainfall, class of > 670 mm with the highest rainfall compared to other classes, with FR (1.589) has been shown the most susceptibility to landslide. The results obtained from the distance from rivers and the distance from roads showed that with the increasing distance from these parameters, landslide susceptibility also decreased, and the class of 0–100 m has the highest FR (1.98 and 1.41) for distance from road and river, respectively. This is in line by results of Mohammady et al. (2012). In the case of distance from fault, class of 1500–2,000 m with FR (1.71) has shown a high susceptibility to landslides and classes of (1000–1500, 0–500, 500–1000, and > 2000 m) with FR (1.61, 1.21, 1.19, and 0.73) are in the next ranks, respectively.

Based on the results of the drainage density factor, there is a direct relationship between drainage density and landslide susceptibility; so, with increasing drainage density, landslide susceptibility has also increased. LU/LC analysis by FR indicated that forest (B) and irrigation agriculture and range (E) classes with the highest FR (1.85 and 1.12) in compared to other classes have more susceptibility to landslide. Result of geology factor explained that TRJs class with dark gray shale and sandstone (Shemshak formation) and FR of 1.20 has highly susceptible to landslide in the current study area.

Random forest model

The results of variables importance using random forest intelligence technique are shown in Fig. 5. The results show

according to mean decrease accuracy analysis, altitude, slope aspect, drainage density, and distance from rivers are the most important factors on landslide occurrence in the study area. Also, the other factors such as annual rainfall, slope degree, distance from roads, distance from faults, geology, convergence index, LU/LC, plan curvature, and TWI are in the next ranks, respectively. Out-of-bag (OOB) error rate in this study was 3.84% with 1000 trees and three variables tried at each split. Finally, the landslide susceptibility map (LSM) using the RF algorithm was provided and classified based on the natural break classification scheme in ArcGIS 10.5 (Pourghasemi and Kerle 2016) into five susceptibility classes: very low, low, moderate, high, and very high.

Support vector machine (SVM) model

In this research, the SVM model with radial basis function (RBF) was trained in R statistical software. The RBF kernel is one of the most powerful kernels and in many studies especially in nonlinear problems, RBF provides better prediction results for landslide susceptibility mapping than other kernels. The probability of landslide occurrence falls in the range between 0 and 1.

In this research, Hyper-parameter sigma and number of support vectors were 0.055 and 6557, respectively.

The results were then exported into the ArcGIS 10.5 software for visualization. Finally, landslide susceptibility map based on the natural break classification divided into five susceptibility classes: very low, low, moderate, high, and very high.

Applying difference scenarios for LSM provide

In general, the scenarios are in both landslide point and polygon formats. The landslide susceptibility maps produced by ten scenarios are represented in Fig. 6a–l. Implementing of different scenarios is done in the ArcGIS10.5 software environment using the Raster Calculator tool. The obtained pixel values from these scenarios were then classified based on the natural break classification scheme (Pourghasemi et al. 2013) into five susceptibility classes: very low, low, moderate, high, and very high. The results showed that most of the landslide area is located in very high susceptibility class. Furthermore, very high susceptibility class covers only low area of watershed (Youssef 2015). As in the scenarios 1, 2, 3, 4, 5, 6, 7, 8, 9 (FR), 9 (SVM), 10 (FR), and 10 (SVM) (8.98, 6.358, 11.28, 8.98, 1.76, 8.98, 8.66, 4.44, 3.11, 13.05, 12.05, and 13.92) percentage of the total area and (36.94, 31.68, 39.20, 36.94, 15.44, 36.94, 36.41, 22.99, 22.18, 36.81, 27.89, and 27.77) of landslide pixels are located in very high susceptibility classes, respectively.

Table 3 Relationship between landslide occurrence and conditioning factors using frequency ratio model

Factors	Classes	Total pixels	%	Landslide pixels	%	FR
Altitude (m)	477–1200	186519	10.26	2151	26.05	2.53
	1200–1600	311299	17.12	1532	18.55	1.08
	1600–2000	354520	19.50	1505	18.22	0.93
	2000–2400	417812	22.98	2175	26.34	1.14
	2400–2800	384275	21.14	894	10.82	0.51
	> 2800	163149	8.97	0	0.00	0
Slope (°)	0–10	79011	4.34	34	0.41	0.09
	10–15	210802	11.59	427	5.17	0.44
	15–20	267364	14.71	976	11.82	0.80
	20–30	365952	20.13	1803	21.83	1.08
	30–40	381482	20.98	2234	27.05	1.28
	> 40	512963	28.22	2783	33.70	1.19
Aspect	Flat	110	0.00	0	0.00	0.00
	N	318654	17.53	508	6.15	0.35
	NE	221334	12.17	794	9.61	0.79
	E	185126	10.18	1082	13.10	1.28
	SE	258023	14.19	1227	14.86	1.04
	S	246172	13.54	1392	16.85	1.24
	SW	149947	8.25	1479	17.91	2.17
	W	172065	9.46	1165	14.10	1.49
Convergence index (1/100 m)	NW	266143	14.64	610	7.38	0.50
	– 100 to – 22	5927	0.32	5	0.06	0.18
	– 22 to – 6	413886	22.77	1962	23.76	1.04
	– 6 to 6	1322371	72.75	6159	74.59	1.02
	6–21	69968	3.85	131	1.58	0.41
Plan curvature (1/100 m)	> 21	5422	0.29	0	0.00	0.00
	concave	780049	42.91	3933	47.63	1.11
	Flat	781140	42.97	3174	38.44	0.89
TWI	Convex	256385	14.10	1150	13.92	0.98
	– 5.4 to – 0.28	394167	21.66	1811	21.93	1.02
	– 0.28–1.26	342555	18.83	1542	18.67	0.99
	1.26–4	446786	24.56	2005	24.28	0.98
Rainfall (mm)	< 4	61338	3.37	283	3.42	1.01
	475–520	192159	10.56	1007	12.19	1.15
	520–580	180528	9.92	330	3.99	0.40
	580–620	562739	30.94	2371	28.71	0.92
	620–670	472727	25.99	1587	19.22	0.73
Distance from roads (m)	> 670	410567	22.57	2962	35.87	1.58
	0–100	152505	8.38	1371	16.60	1.98
	100–200	125014	6.87	1020	12.35	1.79
	200–300	105479	5.80	538	6.51	1.12
	300–400	115274	6.33	574	6.95	1.09
	400–500	99086	5.44	575	6.96	1.27
Distance from rivers (m)	> 500	1221362	67.15	4179	50.61	0.75
	0–100	275231	15.13	1756	21.26	1.40
	100–200	337909	18.57	1339	16.21	0.87
	200–300	278683	15.32	1398	16.93	1.10
	300–400	215971	11.87	1185	14.35	1.20
	400–500	214354	11.78	1122	13.58	1.15
	> 500	496572	27.30	1457	17.64	0.64

Table 3 (continued)

Factors	Classes	Total pixels	%	Landslide pixels	%	FR
Distance from fault (m)	0–500	296374	16.29	1622	19.64	1.20
	500–1000	247777	13.62	1343	16.26	1.19
	1000–1500	138402	7.61	1011	12.24	1.60
	1500–2000	112338	6.17	870	10.53	1.70
	> 2000	1023829	56.29	3411	41.31	0.73
Drainage density (km/km ²)	0–0.5	109686	6.03	66	0.79	0.13
	0.5–0.8	188861	10.38	199	2.41	0.23
	0.8–1.1	261005	14.35	1039	12.58	0.87
	1.1–1.4	353055	19.41	2241	27.14	1.39
	> 1.4	906113	49.82	4712	57.06	1.14
LU/LC	^a A	23154	1.27	5	0.06	0.04
	B	344741	18.95	2899	35.11	1.85
	C	650318	35.75	2018	24.44	0.68
	D	79680	4.38	313	3.79	0.86
	E	437190	24.03	2221	26.89	1.11
	F	282135	15.51	801	9.70	0.62
	G	1502	0.08	0	0.00	0.00
Geology	TRJs	1166028	64.11	6329	76.65	1.19
	Plc	41981	2.30	255	3.08	1.03
	K1bvt	144143	7.92	676	8.18	1.03
	Ek	326519	17.95	632	7.65	0.42
	Mc	8395	0.46	0	0.00	0.00
	Cb	70814	3.89	166	2.01	0.51
	Pr	15795	0.86	55	0.66	0.76
	Db-sh	13498	0.74	43	0.52	0.70
	Qft1	26126	1.43	101	1.22	0.85
Pz	5421	0.29	0	0.00	0.00	

^aOrchard (A), forest (B), range (C), forest and dryfarming (D), irrigation agriculture and range (E), forest and range (F), and residential area (G)

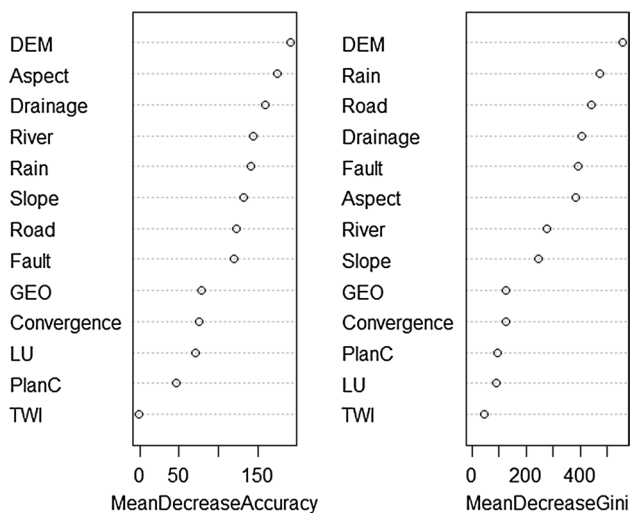


Fig. 5 Two measures of variable importance calculated by the random forest algorithm

Validation of landslide susceptibility maps

Validation of landslide susceptibility models (LSMs) is considered as one of the most important steps in assessment of landslide susceptibility. Furthermore, it is essential in order to assess the predictive capabilities of the landslide susceptibility maps. Thus, without validation, LSM will not have scientific significance (Wu et al. 2016). In this study, for considering the accuracy of the LSM maps provided using the different scenarios, the receiver operating characteristic (ROC) curve was used. In the ROC analysis, the area under the curve (AUC) value used to evaluate the model accuracy. The AUC value of 1.0 represents that the model performed perfectly; and the closer the AUC value to 1.0, the better the model is (Tien Bui et al. 2016). Also, using the frequency ratio (FR) and SCAI (seed cell area index), the accuracy of the separation between the susceptibility classes was verified and confirmed. In this context, the percentages of susceptibility are divided by the percentages of

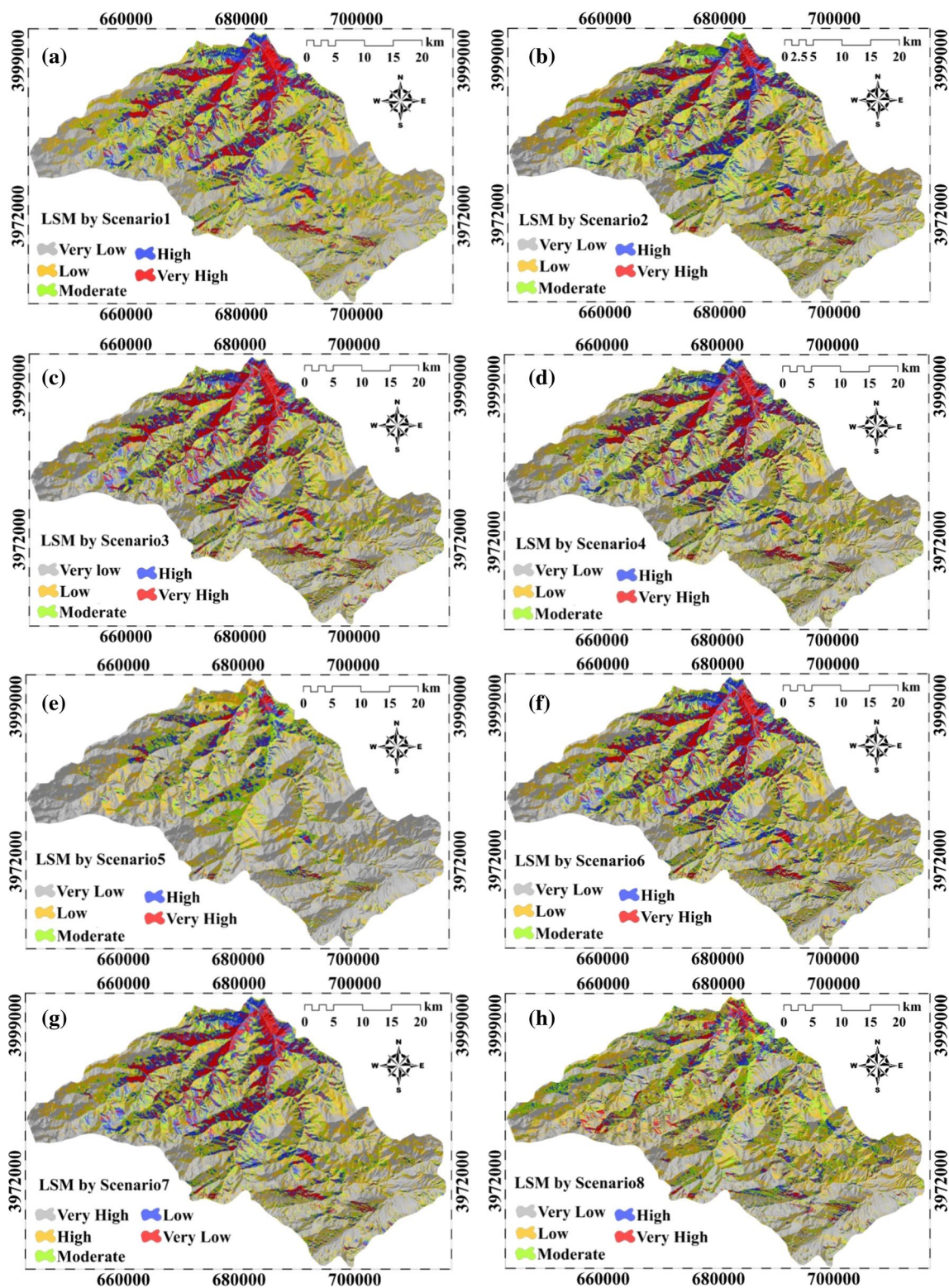


Fig. 6 Landslide susceptibility maps prepared using various scenarios

landslide cells in order to develop the SCAI density of landslides for the classes. Considering that the same landslides that are used in running of model cannot be used to

evaluate the built models (Komac 2006). As a result, the total landslides detected in the study area were randomly divided into two groups: 70% (206 landslide locations) for

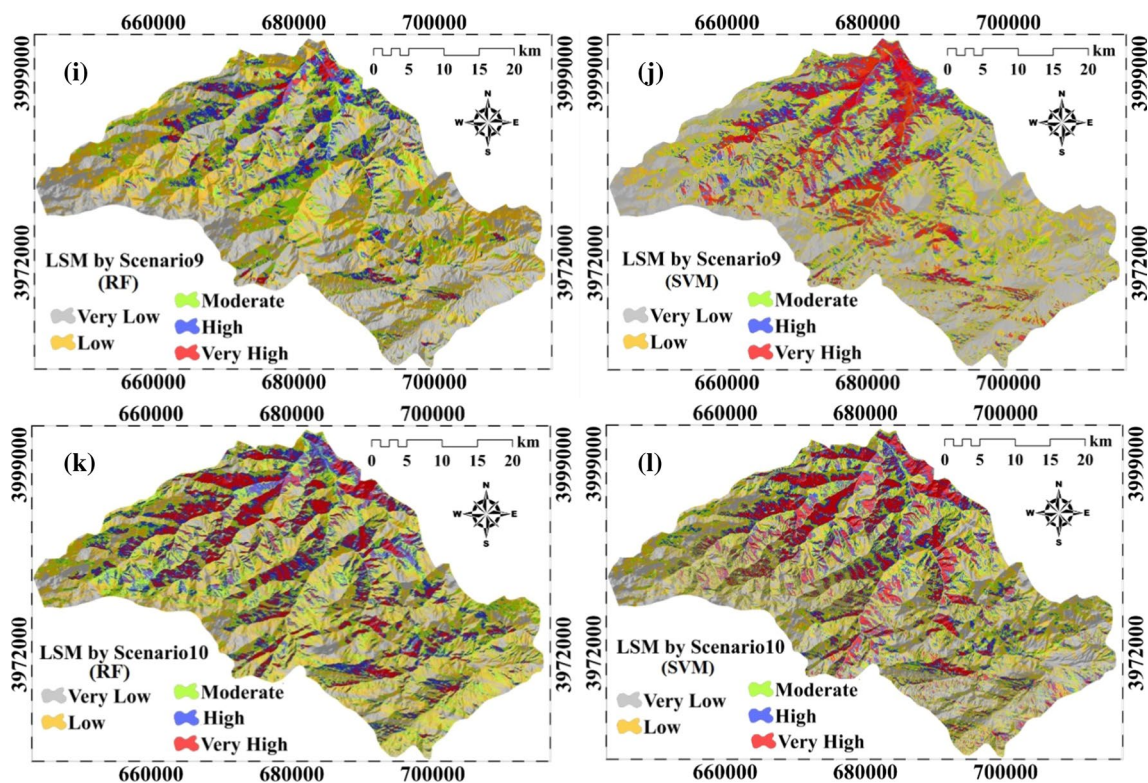


Fig. 6 (continued)

training the models and the remaining 30% (88 landslides locations) was used for validation purpose. According to the results of classification accuracy assessment using the SCAI and FR indicators (Table 4 and Figs. 7 and 8), in all models, with increasing the susceptibility of the risk from very low to very high, the FR is almost up trend, but the SCAI index shows a significant downward trend and indicates a high correlation between the susceptibility classes with the landslide pixels and field observations of the study area. Therefore, the separation order between the classes was evaluated in different scenarios, accurately. The results of the AUC evaluation showed in Table 5.

Results of scenarios showed that AUC was varying from 0.668 to 0.749. In general, maps produced by landslide polygon format (scenarios 1–9) represented the better prediction accuracy than another scenario (landslide point format) and can be used for the spatial prediction of landslide hazard analysis in the study area. Because when use from polygon format, certainly it consisted of several points in compared to a point as polygon centroid, toe of landslides, or landslide crown. So, the polygon format is better from sample points. Also, results of scenarios 1–9 indicated that accuracy of ensemble models is more from individual SVM model; meanwhile, random forest accuracy is similar with ensemble models. By the way,

comparison of scenarios 9 and 10 showed that SVM and RF models built by landslide polygon had the better accuracy than these models (SVM and RF) by landslide point format.

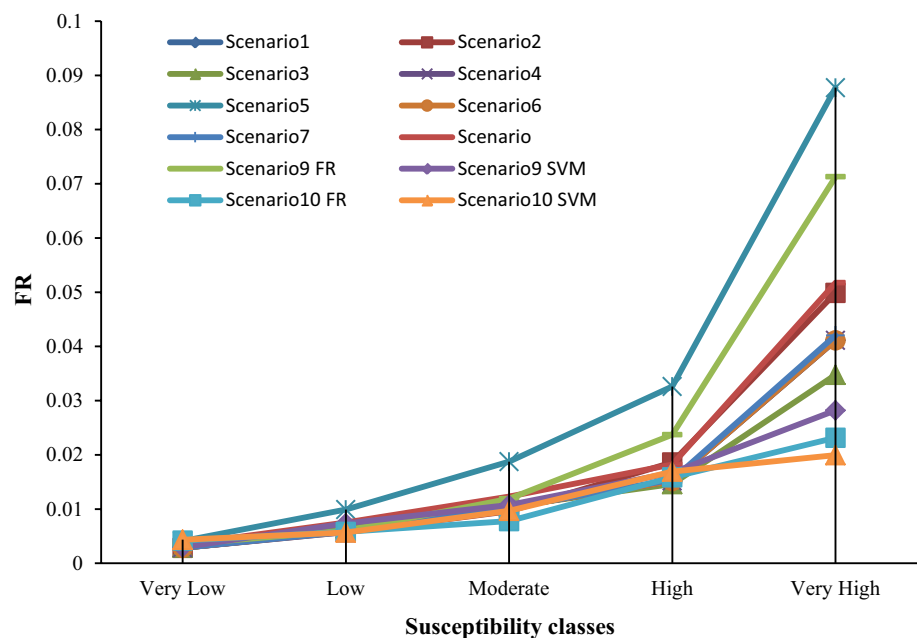
Pourghasemi and Kerle (2016) used random forests and evidential belief function-based models for landslide susceptibility assessment in Western Mazandaran, Iran, and stated that combination of these models with $AUC = 81.77$ has high ability to identify susceptible areas to landslides. This is in line with archived results from ensemble modeling in compared to individual SVM technique. Chen et al. (2016a, b) use support vector machine models for landslide susceptibility mapping in Qianyang County, China. Result of this research indicated that among four kernels, RBF with $AUC = 83.15$ has a high performance in providing a landslide susceptibility map. In our study, in both landslide formats (polygon and point), the SVM–RBF machine learning technique shows the lowest accuracy. Zhang et al. (2017) applied random forest and decision tree methods for landslide susceptibility mapping in the Three Gorges Reservoir area, China, and stated that RF with $AUC = 97.0$ is suitable for landslide susceptibility. Our results are in line with Zhang et al. (2017) as individual RF model in scenario 9 had a high accuracy.

Table 4 Values of FR and SCAI

Scenarios	Susceptibility classes	Total area of classes		Landslide in classes		No landslide area (km)	FR %	Seed	SCAI
		Area (km)	%	Area (km)	%				
Scenario 1	Very low	624.39	38.17	0.81	10.94	623.57	0.00	0.02	17.93
	Low	357.23	21.84	0.96	12.86	356.27	0.01	0.04	4.99
	Moderate	262.28	16.03	1.23	16.52	261.05	0.01	0.08	2.09
	High	245.05	14.98	1.69	22.74	243.36	0.02	0.11	1.33
	Very high	146.87	8.98	2.75	36.94	144.13	0.04	0.31	0.29
Scenario 2	Very low	610.89	37.34	0.79	10.63	610.10	0.00	0.02	17.65
	Low	392.78	24.01	1.03	13.79	391.75	0.01	0.04	5.62
	Moderate	293.84	17.96	1.28	17.22	292.56	0.01	0.07	2.52
	High	234.40	14.33	1.98	26.67	232.41	0.02	0.14	1.04
	Very high	103.91	6.35	2.35	31.68	101.55	0.05	0.37	0.17
Scenario 3	Very low	643.07	39.31	0.85	11.49	642.22	0.00	0.02	18.09
	Low	331.61	20.27	0.96	12.98	330.65	0.01	0.05	4.26
	Moderate	245.36	15.00	1.16	15.67	244.20	0.01	0.08	1.93
	High	231.31	14.14	1.53	20.65	229.78	0.01	0.11	1.30
	Very high	184.46	11.28	2.91	39.20	181.54	0.03	0.26	0.44
Scenario 4	Very low	624.39	38.17	0.81	10.94	623.57	0.00	0.02	17.93
	Low	357.23	21.84	0.96	12.86	356.27	0.01	0.04	4.99
	Moderate	262.28	16.03	1.23	16.52	261.05	0.01	0.08	2.09
	High	245.05	14.98	1.69	22.74	243.36	0.02	0.11	1.33
	Very high	146.87	8.98	2.75	36.94	144.13	0.04	0.31	0.29
Scenario 5	Very low	1022.31	62.50	1.91	25.74	1020.39	0.00	0.03	20.42
	Low	305.58	18.68	1.37	18.49	304.21	0.01	0.07	2.54
	Moderate	180.62	11.04	1.54	20.70	179.09	0.02	0.14	0.79
	High	98.51	6.02	1.46	19.63	97.05	0.03	0.24	0.25
	Very high	28.79	1.76	1.15	15.44	27.65	0.09	0.65	0.03
Scenario 6	Very low	624.39	38.17	0.81	10.94	623.57	0.00	0.02	17.93
	Low	357.23	21.84	0.96	12.86	356.27	0.01	0.04	4.99
	Moderate	261.42	15.98	1.22	16.48	260.20	0.01	0.08	2.09
	High	245.91	15.03	1.69	22.78	244.21	0.02	0.11	1.33
	Very high	146.87	8.98	2.75	36.94	144.13	0.04	0.31	0.29
Scenario 7	Very low	610.69	37.33	0.80	10.71	609.89	0.00	0.02	17.52
	Low	374.20	22.88	0.98	13.21	373.22	0.01	0.04	5.33
	Moderate	263.16	16.09	1.23	16.52	261.93	0.01	0.08	2.11
	High	246.08	15.04	1.72	23.16	244.36	0.02	0.11	1.32
	Very high	141.69	8.66	2.71	36.41	138.98	0.04	0.31	0.28
Scenario 8	Very low	654.23	39.99	1.03	13.88	653.20	0.00	0.03	15.51
	Low	409.34	25.02	1.40	18.84	407.93	0.01	0.06	4.47
	Moderate	313.61	19.17	1.75	23.57	311.86	0.01	0.09	2.10
	High	186.00	11.37	1.54	20.72	184.46	0.02	0.14	0.84
	Very high	72.65	4.44	1.71	22.99	70.94	0.05	0.38	0.12
Scenario 9 FR	Very low	648.10	39.62	1.02	13.75	647.08	0.00	0.03	15.37
	Low	462.58	28.28	1.27	17.03	461.31	0.01	0.04	6.32
	Moderate	300.43	18.37	1.62	21.85	298.81	0.01	0.09	2.08
	High	173.83	10.63	1.87	25.20	171.96	0.02	0.18	0.60
	Very high	50.87	3.11	1.65	22.18	49.23	0.07	0.53	0.06

Table 4 (continued)

Scenarios	Susceptibility classes	Total area of classes		Landslide in classes		No landslide area (km)	FR %	Seed	SCAI
		Area (km)	%	Area (km)	%				
Scenario 9 SVM	Very low	668.88	40.89	0.90	12.14	667.98	0.00	0.02	18.54
	Low	307.68	18.81	1.03	13.83	306.65	0.01	0.05	3.44
	Moderate	229.14	14.01	1.12	15.09	228.02	0.01	0.08	1.75
	High	216.65	13.24	1.65	22.14	215.00	0.02	0.12	1.07
	Very high	213.47	13.05	2.74	36.81	210.73	0.03	0.21	0.62
Scenario 10 FR	Very low	292.83	17.90	0.56	7.52	292.27	0.00	0.03	5.73
	Low	434.21	26.54	1.14	15.38	433.06	0.01	0.04	6.16
	Moderate	401.51	24.55	1.42	19.11	400.09	0.01	0.06	4.24
	High	310.17	18.96	2.24	30.10	307.93	0.02	0.12	1.61
	Very high	197.09	12.05	2.07	27.89	195.01	0.02	0.17	0.70
Scenario 10 SVM	Very low	431.06	26.35	0.85	11.47	430.21	0.00	0.03	8.15
	Low	378.05	23.11	0.97	13.12	377.07	0.01	0.04	5.48
	Moderate	323.83	19.80	1.42	19.17	322.40	0.01	0.07	2.75
	High	275.17	16.82	2.12	28.47	273.05	0.02	0.13	1.34
	Very high	227.69	13.92	2.06	27.77	225.63	0.02	0.15	0.94

Fig. 7 FR values in different scenarios

Chen et al. (2017a, b) introduced new ensembles of ANN, MaxEnt, and SVM machine learning techniques for landslide spatial modeling. They stated that ensemble models have a high performance for landslide susceptibility mapping. Our results showed that this scenario (ensemble modeling) can propose for other researchers, as ensemble modeling was accurately than some scenarios (scenario 9SVM, and scenario 10).

Conclusion

Landslides are one of the most important natural hazards in the world; so, providing of landslide susceptibility maps is very important that can help planners and decision makers in disaster management. The accuracy of landslide susceptibility maps mainly depends on the amount and the quality of data, the scale, and the methodology. In the present study,

Fig. 8 SCAI values in different scenarios

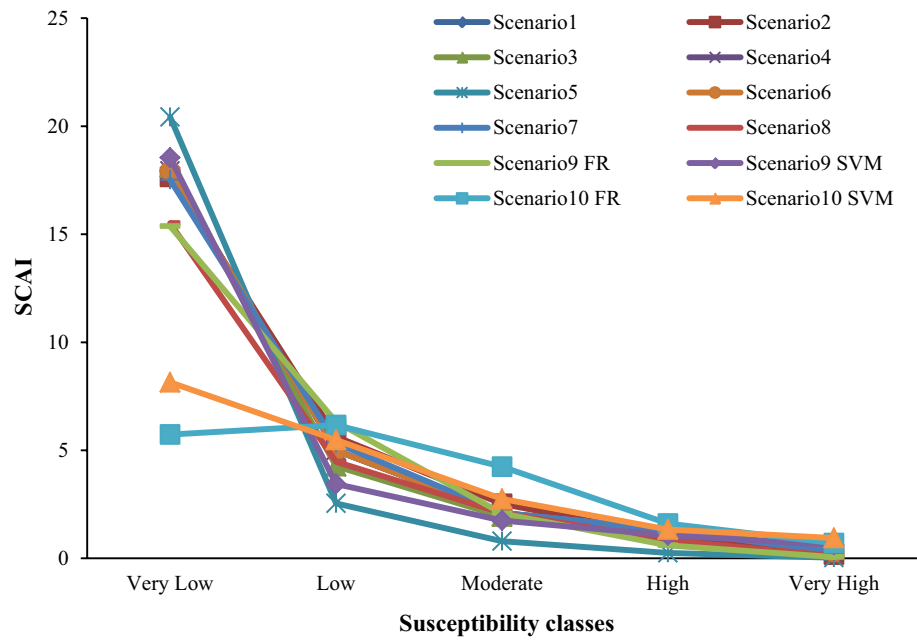


Table 5 Area under the curve

Models	Area	Standard error	Asymptotic significant	Asymptotic 95% CI	
				Lower bound	Upper bound
Scenario 1	0.743	0.007	0.000	0.729	0.756
Scenario 2	0.749	0.007	0.000	0.735	0.762
Scenario 3	0.735	0.007	0.000	0.721	0.749
Scenario 4	0.743	0.007	0.000	0.729	0.756
Scenario 5	0.749	0.007	0.000	0.736	0.762
Scenario 6	0.743	0.007	0.000	0.729	0.756
Scenario 7	0.743	0.007	0.000	0.730	0.757
Scenario 8	0.718	0.007	0.000	0.704	0.732
Scenario 9					
RF	0.743	0.007	0.000	0.729	0.756
SVM	0.715	0.007	0.000	0.701	0.729
Scenario 10					
RF	0.693	0.040	0.000	0.615	0.772
SVM	0.668	0.041	0.000	0.578	0.748

landslide susceptibility maps were prepared using combination of statistical method (FR) and computational intelligence methods (RF and SVM), by applying different scenarios using in landslide polygon and point formats. These maps will help planners and policy makers to mitigation dangers of landslides in construction of roads and settlement. In order to providing landslide susceptibility map, 13 conditioning factors including elevation, slope angle, plan curvature, slope aspect, topographic wetness index (TWI), lithology, LU/LC, distance from rivers, drainage density, distance from fault, distance from roads, convergence index and annual rainfall were used. The FR model was applied as a bivariate statistical method to evaluate the correlation

between the landslides and classes of each conditioning factors. Finally, the ROC curve is used for validation of LSMs. Results of validation indicated that AUC of individual and ensemble models was varying from 0.668 to 0.749. The result of landslide susceptibility maps showed that the high susceptibility areas are mainly distributed along the north to northeastern in the study area. Due to high residential density in this area, it is suggested that any construction operations in this area be made more cautiously. Also, due to the fact that landslides cause loss of fertile soil and land degradation, in order to soil conservation, it is recommended that farmers and foresters avoid unplanned actions on slopes that are sensitive to landslides.

Acknowledgements The authors would like to thank four anonymous reviewers for their helpful comments on the primary version of the manuscript.

References

- Achour Y, Boumezeur A, Hadji R, Chouabbi A, Cavaleiro V, Bendaoud EA (2017) Landslide susceptibility mapping using analytic hierarchy process and information value methods along a highway road section in Constantine, Algeria. *Arab J Geosci* 10:194
- Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Environ* 58:21–44
- Armas I, Vartolomei F, Stroia F, Bras ovanu L (2014) Landslide susceptibility deterministic approach using geographic information systems: application to Breaza town, Romania. *Nat Hazards* 70:995–1017
- Avolio MV, Gregorio SD, Mantovani F, Pasuto A, Rongo R, Silvano S, Spataro W (2000) Simulation of the 1992 Tessina landslide by a cellular automata model and future hazard scenarios. *Sci Direct* 2:41–50
- Ba Q, Chen Y, Deng S, Wu Q, Yang J, Zhang J (2017) An improved information value model based on gray clustering for landslide susceptibility mapping. *Geo-Inf* 6:18
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32
- Briman L, Cutler A (2015) Package ‘randomForest’, pp 29. Date/Publication 2015-10-07
- Cerdà A, Rodrigo-Comino J, Giménez-Morera A, Keesstra SD (2017) An economic, perception and biophysical approach to the use of oat straw as mulch in Mediterranean rain fed agriculture land. *Ecol Eng* 108:162–171
- Chauhan S, Sharma M, Arora MK, Gupta NK (2010) Landslide Susceptibility Zonation through ratings derived from Artificial Neural Network. *Int J Appl Earth Obs Geoinf* 12:340–350
- Chen W, Li W, Chai H, Hou E, Li X, Ding X (2016a) GIS-based landslide susceptibility mapping using analytical hierarchy process (AHP) and certainty factor (CF) models for the Baozhong region of Baoji City, China. *Environ Earth Sci* 75:310
- Chen W, Pourghasemi HR, Zhao Z (2016b) A GIS-based comparative study of Dempster-Shafer, logistic regression, and artificial neural network models for landslide susceptibility mapping. *Geocarto Int* 32:367–385
- Chen W, Pourghasemi HR, Kornejady A, Zhang N (2017a) Landslide spatial modeling: introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma* 305:314–327
- Chen W, Xie X, Wang J, Pradhan B, Hong H, Bui DT, Duan Z, Ma J (2017b) 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
- Claps P, Fiorentino M, Oliveto G (1994) Informational entropy of fractal river networks. *J Hydrol* 187(1–2):145–156
- Colkesen I, Sahin EK, Kavzoglu T (2016) Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines and logistic regression. *J Afr Earth Sci* 118:53–64
- Comino JR, Quiquerez A, Follain S, Damien R, Le Bissonnais Y, Casali J, Gimenez R, Cerda A, Keesstra SD, Brevik EC, Pereira P, Senciales JM, Seeger M, Sinoga JDR, Ries JB (2016) Soil erosion in sloping vineyards assessed by using botanical indicators and sediment collectors in the Ruwer-Mosel valley. *Agr Ecosyst Environ* 233:158–170
- Cortez P (2015) Package ‘rminer’, pp 59. Date/Publication 2015-07-18
- De Sy V, Schoorl JM, Keesstra SD, Jones KE, Claessens L (2013) Landslide model performance in a high resolution small-scale landscape. *Geomorphology* 190:73–81
- Ding Q, Chen W, Hong H (2016) Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. *Geocarto Int* 32:1–21
- Dou J, Yamagishi H, Pourghasemi HR, Song X, Ali YP, Xu Y, Zhu Z (2015) An integrated model for the landslide susceptibility assessment on Osado Island, Japan. *Nat Hazards* 78:1749–1776
- Du J, Yin K, Nadim F, Lacasse S (2013) Quantitative vulnerability estimation for individual landslides. In: Proceedings of the 18th international conference on soil mechanics and geotechnical engineering, Paris, pp 2181–2184
- Ercanoglu M, Gokceoglu C (2002) Assessment of landslide susceptibility for a landslide prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environ Geol* 41(6):720–730
- Feizizadeh B, Blaschke T, Nazmfar H (2014) GIS-based ordered weighted averaging and Dempster–Shafer methods for landslide susceptibility mapping in the Urmia Lake Basin, Iran. *Int J Digit Earth* 7:688–708
- Geology Survey of Iran (GSI) (1997) http://www.gsi.ir/Main/Lang_en/index.html
- Gordan B, Armaghani DJ, Hajihassani M, Monjezi M (2015) Prediction of seismic slope stability through combination of particle swarm optimization and neural network. *Eng Comput (Germany)* 32:85–97
- Guo-liang D, Yong-shuang Z, Javed I, Xin Y (2017) Landslide susceptibility mapping using an integrated model of information value method and logistic regression in the Bailongjiang watershed, Gansu Province, China. *J Mt Sci* 14(2):249–268
- Guzzetti F (2015) Forecasting natural hazards, performance of scientists, ethics, and the need for transparency. *Toxicol Environ Chem* 112:42–66
- Guzzetti F, Mondini AC, Cardinali M, Fiorucci F, Santangelo M, Chang KT (2012) Landslide inventory maps: new tools for an old problem. *Earth Sci Rev* 112:42–66
- Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning, 2nd edn. Springer, New York
- Hong H, Pourghasemi HR, Pourtaghi ZS (2016a) 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
- Hong H, Pradhan B, Jebur MN, Tien Bui D, Xu C, Akgun A (2016b) Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines. *Environ Earth Sci* 75:1–14
- Hong H, Chen W, Xu C, Youssef AM, Pradhan B, Tien Bui D (2017a) Rainfall-induced landslide susceptibility assessment at the Chongren area (China) using frequency ratio, certainty factor, and index of entropy. *Geocarto Int* 32:139–154
- Hong H, Xu C, Chen W (2017b) Providing a landslide susceptibility map in Nancheng County, China, by implementing support vector machines. *Am J Geogr Inf Syst* 6(1A):1–13
- Horafas D, Gkeki T (2017) Applying logistic regression for landslide susceptibility mapping. The case study of Krathis Watershed, North Peloponnese, Greece. *Am J Geogr Inf Syst* 6(1A):23–28
- Immitzer M, Atzberger C, Koukal T (2012) Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sens (Basel)* 4:2661–2693
- Keesstra SD, Bouma J, Wallinga J, Tittonell P, Smith P, Cerda A, Montanarella L, Quinton J, Pachepsky Y, van der Putten WH, Bardgett RD, Moolenaar S, Mol G, Fresco LO (2016) FORUM paper: the significance of soils and soil science towards realization of the UN sustainable development goals (SDGs). *SOIL Discuss* 2:111–128
- Keesstra S, Nunes J, Novara A, Finger D, Avelar D, Kalantari Z, Cerda A (2018) The superior effect of nature based solutions in

- land management for enhancing ecosystem services. *Sci Total Environ* 610:997–1009
- Kim J, Lee S, Jung H, Lee S (2017) Landslide susceptibility mapping using random forest and boosted tree models in Pyeong-Chang, Korea. *Geocarto Int*. doi:<https://doi.org/10.1080/10106049.2017.1323964>
- Kincal C, Singleton A, Li Z, Drummond J, Hoey T, Muller J, Qu W, Zeng Q, Zhang J, Du P (2010) Mass movement susceptibility mapping using satellite optical imagery compared with INSAR monitoring: Zigui county, three gorges region, China. *Dragon-2 Symposium*: 1–5
- Kirchhoff M, Rodrigo Comino J, Seeger M, Ries JB (2017) Soil erosion in sloping vineyards under conventional and organic land use managements (Saar-Mosel valley, Germany). *Cuadernos de Investigación Geográfica* 43:119–140
- Komac M (2006) A landslide susceptibility model using the analytical hierarchy process method and multivariate statistics in perialpine Slovenia. *Geomorphology* 74(1–4):17–28
- Kornejady A, Ownegh M, Rahmati O, Bahremand A (2017) Landslide susceptibility assessment using three bivariate models considering the new topo-hydrological factor: HAND. *Geocarto Int*. doi:<https://doi.org/10.1080/10106049.2017.1334832>
- Lai C, Chen X, Wang Z, Xu C, Yang B (2017) Rainfall-induced landslide susceptibility assessment using random forest weight at basin scale. *Hydrol Res* 48(4):1–16
- Lee S, Hong SM, Jung HS (2017) A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea. *Sustainability* 9:48
- Li L, Lan H, Guo C, Zhang Y, Li Q, Wu Y (2017) A modified frequency ratio method for landslide susceptibility assessment. *Landslides* 14:727–741
- Lo CP, Yeung AKW (2002) Concepts and techniques of geographic information system. Pearson Education Inc, New Jersey
- Mantovani A, Pasuto A, Silvano S, Zannoni A (2000) Collecting data to define future hazard scenarios of the Tessina landslide. *Int J Appl Earth Obs Geoinf* 2:33–40
- Marjanovic M, Kovacevic M, Bajat B, Vozenflek V (2011) Landslide susceptibility assessment using SVM machine learning algorithm. *Eng Geol* 123(3):225–234
- Messenzehl K, Meyer H, Otto J, Hoffmann T, Dikau R (2016) Regional-scale controls on the spatial activity of rockfalls (Turtmann valley, Swiss Alps)—a multivariate modeling approach. *Geomorphology* 287:29–45
- (Meteorological Organization. <http://www.irimo.ir/far/>)
- Mohammady M, Pourghasemi HR, Pradhan B (2012) Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. *J Asian Earth Sci* 61:221–236
- Moore ID, Grayson RB, Ladson AR (1991) Digital terrain modeling: a review of hydrological, geomorphological, and biological applications. *Hydrol Process* 5(1):3–30
- Moosavi V, Niazi Y (2015) Development of hybrid wavelet packet-statistical models (WP-SM) for landslide susceptibility mapping. *Landslides* 13:97
- Myronidis D, Papageorgiou C, Theophanous S (2016) Landslide susceptibility mapping based on landslide history and analytic hierarchy process (AHP). *Nat Hazards* 81:245–263
- Naghbi SA, Pourghasemi HR, Pourtaghi ZS, Rezaei A (2015) Groundwater qanat potential mapping using frequency ratio and Shannon's entropy models in the Moghan Watershed, Iran. *Earth Sci Inform* 8(1):171–186
- Pourghasemi HR, Kerle N (2016) Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environ Earth Sci* 75:185
- Pourghasemi H, Pradhan B, Gokceoglu C, Mohammady M, Moradi H (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(7):2351–2365
- Pourghasemi HR, Moradi HR, Fatemi Aghda SM (2014) GIS-based landslide susceptibility mapping with probabilistic likelihood ratio and spatial multi-criteria evaluation models (North of Tehran, Iran). *Arab J Geosci* 7(5):1857–1878
- Pourghasemi HR, Rahmati O (2017) Prediction of the landslide susceptibility: Which algorithm, which precision?. *CATENA*
- Pourghasemi HR, Yousefi S, Kornejady A, Cerda A (2017) Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci Total Environ* 609:764–775
- Pradhan B, Lee S (2010) 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
- Prompera C, Puissant A, Malet JP, Glade T (2014) Analysis of land cover changes in the past and the future as contribution to landslide risk scenarios. *Appl Geogr* 53:11–19
- Ren F, Wu X, Zhang K, Niu R (2015) Application of wavelet analysis and a particle swarm-optimized support vector machine to predict the displacement of the Shuping landslide in the Three Gorges, China. *Environ Earth Sci* 73:4791–4804
- Rozos D, Bathrellos GD, Skilodimou HD (2011) Comparison of the implementation of Rock Engineering System (RES) and Analytic Hierarchy Process (AHP) methods, based on landslide susceptibility maps, compiled in GIS environment. A case study from the Eastern Achaia County of Peloponnesus, Greece. *Environ Earth Sci* 63(1):49–63
- Sidle RC, Ochiai H (2006) Landslides: Processes, Prediction, and Land Use. American Geophysical Union, Water Resources Monograph, No 18, Washington
- Tay LT, Lateh H, Hossain MK, Kamil, AA (2014) Landslide hazard mapping using a poisson distribution: a case study in Penang Island, Malaysia. *Landslide Sci Safer Geoenviron* 521–525
- Tien Bui D, Pham BT, Nguyen QP, Hoang ND (2016) Spatial prediction of rainfall-induced shallow landslides using hybrid integration approach of least-squares support vector machines and differential evolution optimization: a case study in Central Vietnam. *Int J Digital Earth* 9:1077–1097
- Vaezi AR, Abbasi M, Keesstra S, Cerdà A (2017) Assessment of soil particle erodibility and sediment trapping using check dams in small semi-arid catchments. *CATENA* 157:227–240
- Vapnik VN (1995) The nature of statistical learning theory. Springer, New York
- Wang Y, Seijmonsbergen AC, Bouten W, Chen QT (2015) Using statistical learning algorithms in regional landslide susceptibility zonation with limited landslide field data. *J Mt Sci* 12(2):268–288
- Wang LJ, Guo M, Sawada K, Lin J, Zhang JC (2016) A comparative study of landslide susceptibility maps using logistic regression, frequency ratio, decision tree, weights of evidence and artificial neural network. *Geosci J* 20:117–136
- Wen F, Xin-sheng W, Yan-bo C, Bin Z (2017) Landslide susceptibility assessment using the certainty factor and analytic hierarchy process. *J Mt Sci* 14(5):906–925
- Wu Y, Li W, Wang Q, Yan S (2016) Landslide susceptibility assessment using frequency ratio, statistical index and certainty factor models for the Gangu County, China. *Arab J Geosci* 9(2):84
- Xie QM, Bian X, Xia YY (2005) Systematic analysis of risk evaluation of landslide hazard (in Chinese). *Rock Soil Mech* 26(1):71–74
- Youssef AM (2015) Landslide susceptibility delineation in the Ar-Rayth area, Jizan, Kingdom of Saudi Arabia, using analytical hierarchy process, frequency ratio, and logistic regression models. *Environ Earth Sci* 73(12):8499–8518

- Zeng B, Xiang W, Rohn J, Ehret D, Chen X (2017) Assessment of shallow landslide susceptibility using an artificial neural network in Enshi region. *Nat Hazards Earth Syst Sci Discuss*, China. <https://doi.org/10.5194/nhess-2017-176>
- Zhang K, Wu X, Niu R, Yang K, Zhao L (2017) The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. *Environ Earth Sci* 76:405