Spatial prediction of soil erosion susceptibility: an evaluation of the maximum entropy model

Maryam Pournader[1] · Hasan Ahmadi[1] · Sadat Feiznia[2] · Haji Karimi[3] · Hamid Reza Peirovan[4]

Abstract
Soil erosion is considered as the most widespread form of soil degradation which causes serious environmental problems. This study investigates the performance of the maximum entropy (ME) in mapping rill erosion susceptibility in the Golgol watershed, Ilam province, Iran. To this end, ten rill erosion conditioning factors were selected to be employed in the modelling process based on an investigation of the literature. These layers are: elevation, slope percent, aspect, stream power index, topographic wetness index, distance from streams, plan curvature, lithology, land use, and soil. Then, a training dataset of rill erosion locations was used for modelling this phenomenon. The area under receiver operating characteristics curve was used for evaluating the performance of the ME model. In addition, Modified Pacific South-West Inter Agency Committee (MPSIAC) framework was applied and sediment yield was determined for different hydrological units in the study area. At last, Jackknife test was implemented to show the contribution of the factors in the modelling process. The results depicted that area under ROC curve for training and validation datasets were 0.867, and 0.794, respectively. Therefore, this conclusion can be achieved that ME worked well and could be a good tool for generating rill erosion susceptibility maps and its output could be employed for soil conservation in similar areas.

Introduction

The soil is regarded as an important natural resource in each country that different types of erosion cause its loss. Risk assessment and quantification of the soil erosion are two main activities to provide better political plans for natural resources, agriculture and environment (Feng et al. 2010; Mandal and Sharda 2013; Zhao et al. 2013). Soil erosion and sediment yield are complex issues which are affected by different factors

✉ Sadat Feiznia
sfeiz@ut.ac.ir

[1] Department of Watershed Management, Science and Research Branch, Islamic Azad University, Tehran, Iran

[2] Department of Reclamation of Arid and Mountainous Regions, University of Tehran, Karaj, Iran

[3] Department of Watershed and Range Management, Ilam University, Ilam, Iran

[4] Soil Conservation and Watershed Management Research Institute (SCWMRI), Tehran, Iran

(Choubin et al. 2018). Several studies have described on-site (e.g., loss of productive capacity and land degradation) and off-site (e.g., sedimentation of reservoirs, damage to the infrastructure, etc.) effects of soil erosion which often cause environmental and economic problems (Riksen and De Graaff 2001; Ledermann et al. 2010; Shi et al. 2012; Mullan 2013; Mekonnen et al. 2015). Quantification of the relationship between soil erosion and its conditioning factors is a big challenge and have attracted many researchers to itself (Bakker et al. 2005; Koulouri and Giourga 2007). In fact, quantification of the soil erosion is a process with complex and unstructured decision nowadays; therefore, it is necessary to conduct a comprehensive and systematic method to gain this goal (Renschler and Harbor 2002; Bahadur 2009; Conoscenti et al. 2013).

In the recent past, different methods have been introduced for soil erosion and sediment yield estimation. One of the most famous methods is Universal Soil Loss Equation (USLE) that is applied in many countries and different characteristics (Harmon and Doe 2001). Some of its applications have been done in developing countries such as Meusburger et al. (2013), and Csáfordi et al. (2012). The revised USLE (RUSLE) was introduced as a modification of the USLE model to calculate the mean yearly erosion (Duarte et al. 2016;

Renard et al. 1997; Abdullah et al. 2017). Another model which has been used in other researches is Erosion Potential Method (EPM), and empirical model that was produced to predict soil erosion and sediment yield (Amiri 2010). Modified Pacific South-west Inter-Agency Committee (MPSIAC) framework was produced to estimate erosion in the USA (Ilanloo 2012). Application of the MPSIAC framework has been reported acceptable in different papers in Iran under arid and semi-arid features (Shahzeidi et al. 2012; Bagherzadeh and Daneshvar 2013; Taheri et al. 2013). In a study, Abdullah et al. (2017) used MPSIAC, EPM, and RUSLE models to predict soil erosion in Umm Nigga area, Kuwait. The findings of this study elaborated that MPSIAC had better performance than the other models, followed by EPM and RUSLE models.

On the other hand, other researchers have presented a new approach to spatial prediction of susceptibility to different kinds of erosion (e.g. gully, landslide, rill, and interrill) by using statistical and machine learning models. For instance, Conoscenti et al. (2008) analyzed relationships between geo-environmental factors and the spatial distributions of the erosion landforms using a geostatistical multivariate approach. Rahmati et al. (2016a) evaluated the capability of frequency ratio (FR) and weights-of-evidence (WofE), two common statistical models, for spatial prediction of gully erosion susceptibility. They demonstrated that FR and WofE models are efficient tools in gully susceptibility assessment. In another work, Rahmati et al. (2017a) also applied a conditional probability (CP) model to model the susceptibility of gully erosion in a semi-arid region, Iran. Angileri et al. (2016) generated a susceptibility map of water erosion by implementing stochastic gradient Treeboost in a study area in Italy. In this work, rill-interrill and gully types of erosion were studied. For their investigation, they regarded altitude, landform, aspect, and land-use as the most necessary factors affecting the rill and interrill erosion. Conforti et al. (2011) used an information value model to map gully erosion susceptibility that land use, stream power index (SPI), slope, aspect, topographic wetness index (TWI), slope length factors were considered for modelling. Their findings implied that 88% of the gullies fell in high and very high susceptibility categories. Märker et al. (2011) used stochastic gradient boosting and bootstrap aggregation models to predict the potential spatial distribution of erosion processes. They stated these models provides insights into factors controlling erosion processes and also are valuable tools in soil erosion conservation and geomorphology. Conoscenti et al. (2014) applied a GIS-based logistic regression model in central-northern Sicily, Italy, for susceptibility mapping of gully erosion. Maximum entropy is a data mining model which has been conducted in several fields of study including ecological modeling (Phillips et al. 2006; Phillips and Dudík 2008; Kleidon et al. 2010; Elith et al. 2011; Harte and Newman 2014), groundwater potential

mapping (Rahmati et al. 2016b), and landslide susceptibility mapping (Vorpahl et al. 2012; Felicísimo et al. 2013; Park 2015; Kornejady et al. 2017a).

In the current study, two major objectives were considered which are: (i) assessing the capability of maximum entropy model for susceptibility mapping of rill erosion which is being applied for the first time, and (ii) application of the MPSIAC framework for estimating soil erosion rate in the study area. For conducting this research, Golgol watershed, Ilam province, Iran was chosen since it is one of the sub-watersheds of the Ilam Dam and the results of this study could be very helpful for management of the dam. Determining the most susceptible areas to rill erosion could be an initial step for soil conservation plans resulting in lower costs and required time.

Material and methods

Study area

The Golgol watershed, Ilam province, Iran is located between 46° 27′ to 38° 46′ eastern longitudes and 33° 25′ to 33° 38′ northern latitudes (Fig. 1). Elevation in the Golgol watershed ranges from 1013 to 2156 m. The average rainfall and temperature in the study area are measured as 580 mm, and 16.9 °C, respectively. The main species in forest lands of the Golgol watershed is *Quercus brantii*. There is a reservoir dam at low lands of the study area for supplying drinking water of Ilam city. Produced sediments from the sub-watersheds can fill the dam and are regarded as crucial threats to sustainable development in the study area. In a geological view point, this area is located in the folded Zagros zone with different lithological units such as limestone, shale, valley terrace deposits or a combination of them.

Methodology

Application of MPSIAC framework

The MPSIAC framework needs 9 factors to be applied which are surface geology, land cover, soil, climate, land use, channel erosion, present erosion, topography, and runoff (Daneshvar and Bagherzadeh 2012). These layers were provided by employing data obtained from Ministry of Agriculture-Jahad, Iran (MAJ 2014).

Surface geology (Y_1) The surface geology (Y_1) was calculated using X_1 which is surface geology in PSIAC model. The X_1 can be defined based on stone types, hardness and fracture (Daneshvar and Bagherzadeh 2012). It needs to be mentioned that scores of the units were defined between a range of 1 to 10 based on the local condition of Iran (Feyznia 1995).
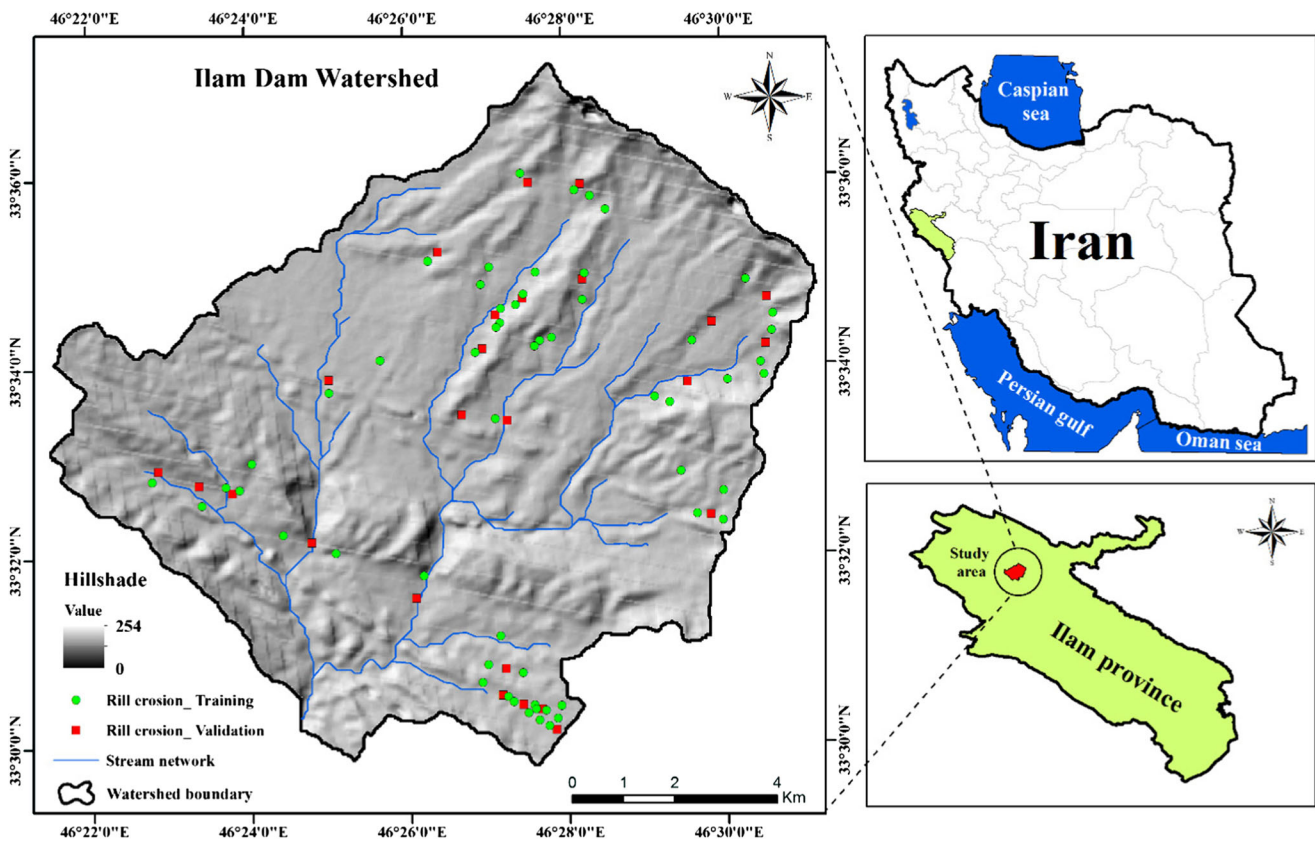
Fig. 1 Location of the study area, training and validation rill erosion locations in Ilam province, Iran

Soil (Y$_2$) The soil map was obtained from Ministry of Agriculture-Jahad, Iran (MAJ 2014). This factor can be calculated by erodibility factor (K) as follows:

$$Y_2 = 16.67K \tag{1}$$

where, Y$_2$ shows soil variable in MPSIAC, K denotes erodibility factor which can be calculated by RUSLE model (Benzer 2010; Dumas and Printemps 2010).

Climate factor (Y$_3$) Climate influences the soil and vegetable cover and affects the runoff at watershed-scale. This factor in MPSIAC can be computed as below:

$$Y_3 = 0.2P_2 \tag{2}$$

where, Y$_3$ denotes climate variable, and P shows two-year returning period six-hour rainfall (mm).

Runoff (Y$_4$) Runoff has a close relationship with the climate of the watershed. An intense flood which rarely incidents has a high role in yearly sediment yield of the watershed. This variable can be obtained as below:

$$Y_4 = 0.006R + 10Q_p \tag{3}$$

where, Y$_4$ denotes runoff variable, R shows runoff (mm) and Q_p specific peak discharge calculated as (m^3.km^{-2}.s^{-1}).

Topography (Y$_5$) This factor was obtained from DEM of the study area. This variable can be calculated as below:

$$Y_5 = 0.33S \tag{4}$$

where, Y$_5$ denotes topography variable, and S represents mean slope (%) which was extracted from a 1: 50,000 scale DEM of the study area.

Ground cover (Y$_6$) This factor includes any cover on the ground which influences the effect of rainfall on the ground such as vegetable cover litter. This factor can be calculated as below:

$$Y_6 = 0.2Pb \tag{5}$$

where, Y$_6$ denotes land cover variable, and Pb represents the bare lands at each land unit (%).

Land use (Y$_7$) The land use map was obtained from MAJ (2014). This factor can be obtained as follows:

$$Y_7 = 20 - 0.2Pc \tag{6}$$

where, Y$_7$ denotes land use variable, and Pc shows the canopy covering at each land unit (%).

Upland erosion (Y$_8$) Upland erosion was determined by employing the US Bureau of Land Management BLM

method (Abdullah et al. 2017). This factor can be obtained as follows:

$$Y_8 = 0.25 SSF \tag{7}$$

where, Y_8 denotes current erosion amount, and SSF shows soil surface factor. The SSF is affected by 7 conditioning factors based on BLM.

Channel erosion (Y_9) Regarding the channel erosion in any watershed, type, shape, geomorphology, and bank erosion of the rivers are essential factors to be considered. This variable could be calculated as below:

$$Y_9 = 1.67 SSF.g \tag{8}$$

where, Y_9 represents channel erosion, SSF.g shows the gully erosion amount based on BLM model.

Rill erosion susceptibility mapping

In order to create the rill erosion susceptibility map of the study area, first the location of this erosion type was determined in the study area by extensive field surveys and 79 locations were detected. Based on the literature review, these locations were categorized into two classes of training and validation with a 70:30 ratios. In the other word, out of the 79 rill erosion occurrences, 55 cases were randomly selected for training the model, and the remaining 24 cases were used for validation purpose. The training dataset was used for training the maximum entropy (ME) model. Then, rill erosion susceptibility conditioning factors were provided and plotted in ArcGIS 10.2 software. These factors include elevation, slope percent, slope aspect, SPI, topographic wetness index (TWI), distance from the stream, plan curvature, lithology, land use, and soil. In fact, these conditioning factors were used as independent variables, while rill erosion inventory was considered as a dependent variable. In the next step, the ME model was applied by employing the conditioning factors and training locations. Lastly, receiver operating characteristics (ROC) curve was calculated to determine its performance (Chang-Jo and Fabbri 2003; Pourghasemi and Rahmati 2018).

Rill erosion conditioning factors In this work, ten rill erosion conditioning factors were selected to be employed in the modelling process based on an investigation of the literature (Lu et al. 2001; Cerdan et al. 2002; Govers et al. 2007; Hancock et al. 2008; Auerswald et al. 2009; Wirtz et al. 2012; Angileri et al. 2016). However, a standard methodology and specific framework have not yet been established for the choice of conditioning factors for modeling the rill erosion susceptibility (Conoscenti et al. 2014; Angileri et al. 2016). First, a topographic map of the study watershed with 100,000-scale was prepared.

From this DEM, some layers such as elevation, slope percent, slope aspect, SPI, TWI, and plan curvature were extracted (Razandi et al. 2015; Falah et al. 2017; Siahkamari et al. 2017). Elevation in this watershed changes between 1013 to 2156. Slope percent influences the water flow speed and subsequently, the higher slopes would be more susceptible to rill erosion. Slope percent was calculated in this study ranging from 0 to 352.8%. Slope aspect was prepared for the study area and categorized into 9 classes of main, sub-main, and flat. Stream power index signifies the erosion power of the flow in the studied region. This factor was calculated in SAGA-GIS software and changes from 0 to 81.75. TWI is another topo-hydrological factor which was developed by Moore et al. (1991) and can be obtained as below:

$$TWI = \ln(\alpha/\tan\beta) \tag{9}$$

where, α represents the accumulative area which flows to a point and β shows its slope angle. Plan curvature is the last topographic factor which was considered in this study. This factor could be regarded as a contour line that is built by intersection of the horizontal plane and the surface (Yilmaz et al. 2012; Ghorbani Nejad et al. 2017).

Lithology was obtained from a geology map of the study area with 1: 100,000- scale (GSI 1997). As it can be seen, there are four categories of lithology in the study area KEpd-gu, Kbgp, Qft2, and OMas (Table 1).

Landuse of the study area was prepared by the Enhanced Thematic Mapper Plus (ETM+) images and supervised classification using maximum likelihood algorithm method in ENVI 4.2 software. Five classes of land use exist in the study watershed which are agriculture, rangeland, residential, and fragmented forest (Fig 2i). The soil map of the study area is comprised of Entisols, Inceptisols, and Vertisols (Fig 2j). In addition, distance from rivers map of the study area was calculated by implementing Euclidean distance function in ArcGIS 10.2 (Fig. 2f).

Rill erosion susceptibility mapping by maximum entropy model The ME model was developed by Phillips et al. (2006) which was first employed for ecological studies in modelling species distribution modelling (Rahmati et al. 2016b). This approach could be implemented by only having incident data (i.e. erosion locations data in this study). The basis of the ME is machine learning technique which makes it possible to predict the incident from deficient data (Medley 2010). The probability distribution in this model considers a set of limitations which are gained from the incident data by investigating the conditioning factors (Felicísimo et al. 2012, 2013). The output of this model is a map that represents the probability of rill erosion incident at each pixel of the study area. This model was applied in MaxEnt software.

Table 1 Lithology of the Ilam Dam watershed, Iran

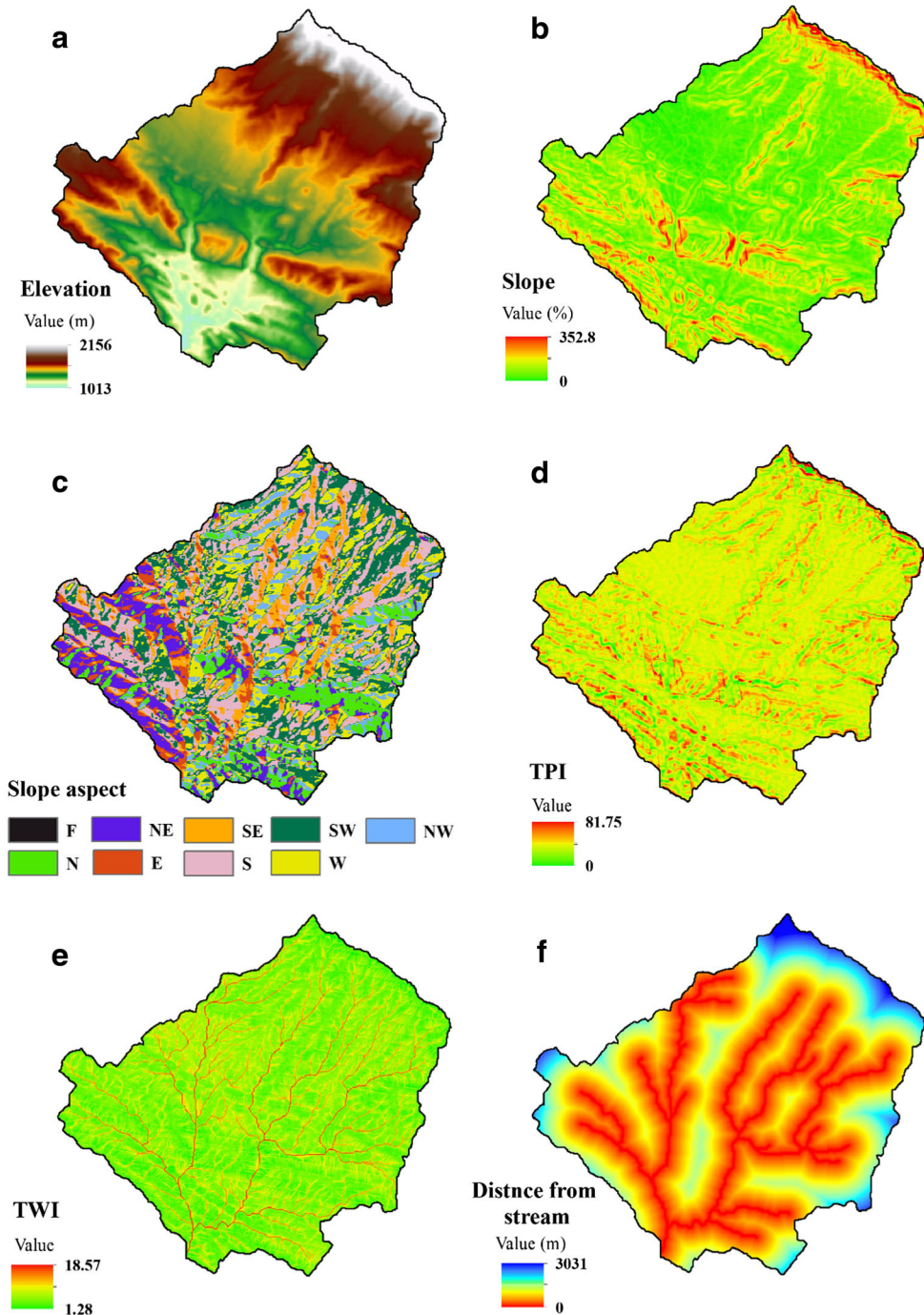| Code	Lithology	Formation	Geological age
KEpd-gu	Grey and brown, massive fossiliferous limestone	Kazhdumi	Late Cretaceous
Kbgp	Undivided Bangestan Group, mainly limestone and shale	Sarvak, Surgah and Ilam	Cretaceous
Qft2	Low level piedmont fan and valley terrace deposits	–	Quaternary
OMas	Cream to brown, weathering, limestone with intercalations of shale	Asmari	Miocene

Fig. 2 Rill erosion conditioning factors

g

Plan curvature

Value 9.75
- 11.16

h

Lithology

KEpd-gu   Kbgp   OMas   Qft2

i

Landuse

Agriculture   Rangeland   Other
Fragmented forest   Residential
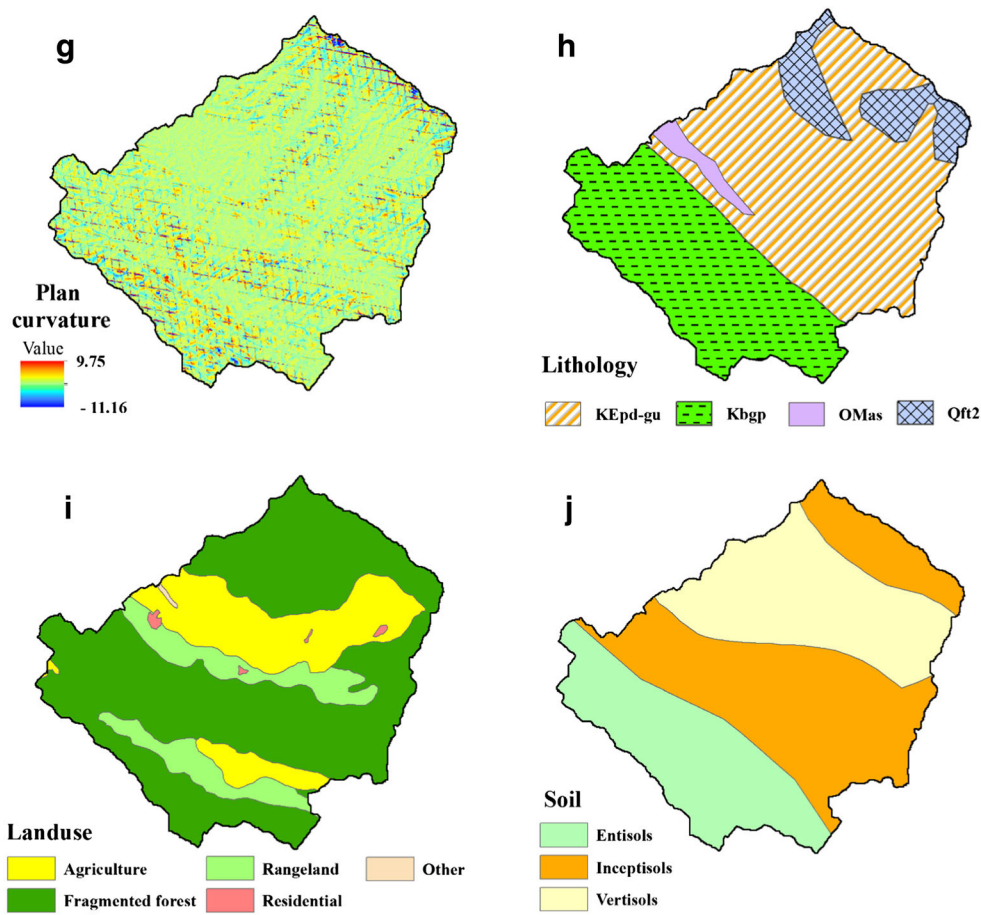
j

Soil

Entisols
Inceptisols
Vertisols

Fig. 2 continued.

Results

MPSIAC framework

The results of MPSIAC are represented in Table 2. The factors of this model are explained below. In the case of surface geology factor, X_1 values, as well as Y1 values, range from 3.65 to 9.89 with an average of 7.67. In the respect of soil factor, Y2 changes from 6.83 to 8.67 with an average of 4.51. The results of runoff showed that Y4 ranges from 3.52 to 5.28 having an average of 4.51. In the respect of topography, y5 has maximum, minimum, and average values of 4.32, 5.5, and 4.91, respectively. Ground cover ranges from 3.21 to 4.82 having an average value of 4.02. In the case of land use, minimum, maximum and average values are 8.96, 10.76, and 9.841, respectively. In the respect of upland erosion, Y8 ranges between 5.95 and 8.2 having an average of 7.18. In the respect of channel erosion, Y8 ranges from 6.91 to 10.18 with an average value of 9.282.

Annual sediment yield of hydrological units

Yearly sediment yield (Q_s) ($m^3/km^2/y$) was calculated for each hydrological unit as well as the whole area (Table 3).

According to the results of MPSIAC, predicted soil loss changes from 124.36 $m^3/km^2/y$ to 200.97 $m^3/km^2/y$ with for hydrological unit 10 and 13, respectively. The soil loss predicted for the whole watershed is calculated as 170.67 $m^3/km^2/y$.

Susceptibility map of rill erosion

The results of the response curve in ME model are represented in Fig. 3. In the case of elevation, most of the erosion locations are concentrated at 1200-1600. The results depict that slope percent of higher than 25 is the location of most rill erosions occurred. In the case of slope aspect, south, south-west, and west-facing aspects have more rill erosion incidents. The results of SPI depicts that SPIs more than 6 have a higher frequency of rill erosion occurrence in the studied area. In the respect of TWI, areas with TWI values between 8 and 13 had the highest occurrence of rill erosion. Distances from the river of 0 to 1000 had the highest concentration of the rill erosion incident. In addition, it was observed that a reverse relationship exists between this factor and rill erosion occurrence. In the respect of plan curvature, values lower than 0 had a high concentration of this erosion. The results of response curve for

Table 2 The scores of different factors in MPSIAC model

Hydrological Unit /Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X_1	5.89	8.95	8.85	6	3.94	4	8.93	8.99	8.86	3.65	8.98	7.56	9.75	9	9.89
Y_1	5.89	8.95	8.85	6	3.94	4	8.93	8.99	8.86	3.65	8.98	7.56	9.75	9	9.89
X_2	0.45	0.43	0.41	0.52	0.46	0.42	0.47	0.51	0.42	0.49	0.47	0.44	0.51	0.48	0.43
Y_2	7.50	7.17	6.83	8.67	7.67	7.00	7.83	8.50	7.00	8.17	7.83	7.33	8.5	8.00	7.17
X_3	23.15	26.4	24.1	24.65	20.1	21.55	24.35	25.15	26.15	17.6	21.1	19.8	21.7	22.18	20.45
Y_3	4.63	5.28	4.82	4.93	4.02	4.31	4.87	5.03	5.23	3.52	4.22	3.96	4.34	4.44	4.09
X_4	31.2	35.2	33.1	29.5	33.00	28.2	26.6	31.15	33.55	34.1	32.3	30.71	34.71	32.31	34.65
Y_4	6.24	7.06	6.62	5.90	6.60	5.64	5.32	6.23	6.71	6.82	6.46	6.14	6.94	6.46	6.93
X_5	16.4	13.7	14.7	15.3	14.1	13.1	16.7	14.4	15.5	16.4	14.1	14.5	15.1	13.9	15.8
Y_5	5.41	4.52	4.84	5.04	4.66	4.32	5.50	4.74	5.12	5.41	4.65	4.78	4.98	4.58	5.21
X_6	16.4	13.7	14.7	15.3	14.1	13.1	16.7	14.4	15.5	16.4	14.1	14.5	15.1	13.9	15.8
Y_6	5.41	4.52	4.84	5.04	4.66	4.32	5.50	4.74	5.12	5.41	4.65	4.78	4.98	4.58	5.21
X_7	50.02	48.3	46.2	50.2	55.2	48.3	47.3	46.2	54.3	55.2	54.2	49.5	51.8	52.9	52.1
Y_7	9.96	10.34	10.76	9.96	8.96	10.34	10.54	10.76	9.14	8.96	9.16	10.1	9.64	9.42	9.58
X_8	25	31.7	30.8	24.2	31.2	32.1	27.2	29.4	25.1	23.8	30.3	29.8	32.8	27.9	30.1
Y_8	6.25	7.92	7.7	6.05	7.8	8.02	6.80	7.35	6.27	5.95	7.57	7.45	8.2	6.97	7.52
X_9	4.52	5.90	5.78	4.14	5.87	5.94	5.84	6	5.78	4.20	5.83	5.84	6.1	5.65	6.02
Y_9	7.54	9.85	9.65	6.91	9.80	9.91	9.75	10.02	9.65	7.02	9.73	9.75	10.18	9.42	10.05

lithology showed that Kbgp had the highest erosion incident, while the lowest amount was seen in OMas class. In the case of land use, fragmented forest and agriculture were seen to have the highest erosion occurrence, respectively. The results of response curve for soil factor represented that Vertisols and Entisols had the highest amount of rill erosion.

In addition, the importance of rill erosion conditioning factors was assessed by using a Jackknife test as shown in Table 4. Accordingly, it can be seen that land use, slope percent, aspect and SPI have been selected as the most important erosion conditioning factors (ECFs) with contribution percent values of 14.5, 13.58, 13.3, and 12.2, respectively. On the other hand, TWI, plan curvature, and soil were identified as the least important factors with contribution percent values of 3.2, 7.1, and 7.5, respectively (Table 4).

Table 3 The results of sediment yield calculation by MPSIAC model for each hydrological unit and the whole area

Hydrological Unit	Qs (m³/km²/y)	Hydrological Unit	Qs (m³/km²/y)
1	144.05	10	124.36
2	187.38	11	176.47
3	189.27	12	169.55
4	147.4	13	200.97
5	150.38	14	178.24
6	148.88	15	189.27
7	178.25	Basin	170.67
8	200.97	Minimum	124.36
9	174.71	Maximum	200.97

Erosion susceptibility map (ESM) produced by ME model is represented in Fig. 4. The ESM was classified into low, moderate, high and very high classes. As it can be seen, low, moderate, high, very high classes consist 19.61, 34.22, 28.76, and 17.39% of the studied region, respectively (Table 5, Fig. 5).

Evaluation of the model performance

In several work, receiver operating characteristic (ROC) curve has been used for classification and validation purposes in different issues such as landslide, flood, groundwater, and forest fire (Rahmati and Melesse 2016; Hong et al. 2017; Chen et al. 2017a, b; Naghibi et al. 2017c; Rahmati et al. 2017b; Rahmati and Pourghasemi 2017). Thus, in this investigation, ROC curve was utilized to evaluate the ESM produced by ME. The area under this curve depicts the how efficient is the model in classifying incident and non-incident of the erosion (Yesilnacar and Topal 2005; Tahmassebipoor et al. 2016; Haghizadeh et al. 2017). The area under the curve (AUC-ROC) value of close to 1 depicts high performance of the model, while a lower value represents the weakness of the model in classifying the event (Naghibi et al. 2017a, b). For conducting this curve, the same number of non-erosion locations similar to the erosion locations were regarded and the values of the ME were extracted for them. Then, these values were entered in SPSS 20 and ROC plot was generated. Fig. 6 shows AUC-ROC for the ME model regarding training and validation datasets. Accordingly, AUC-ROC
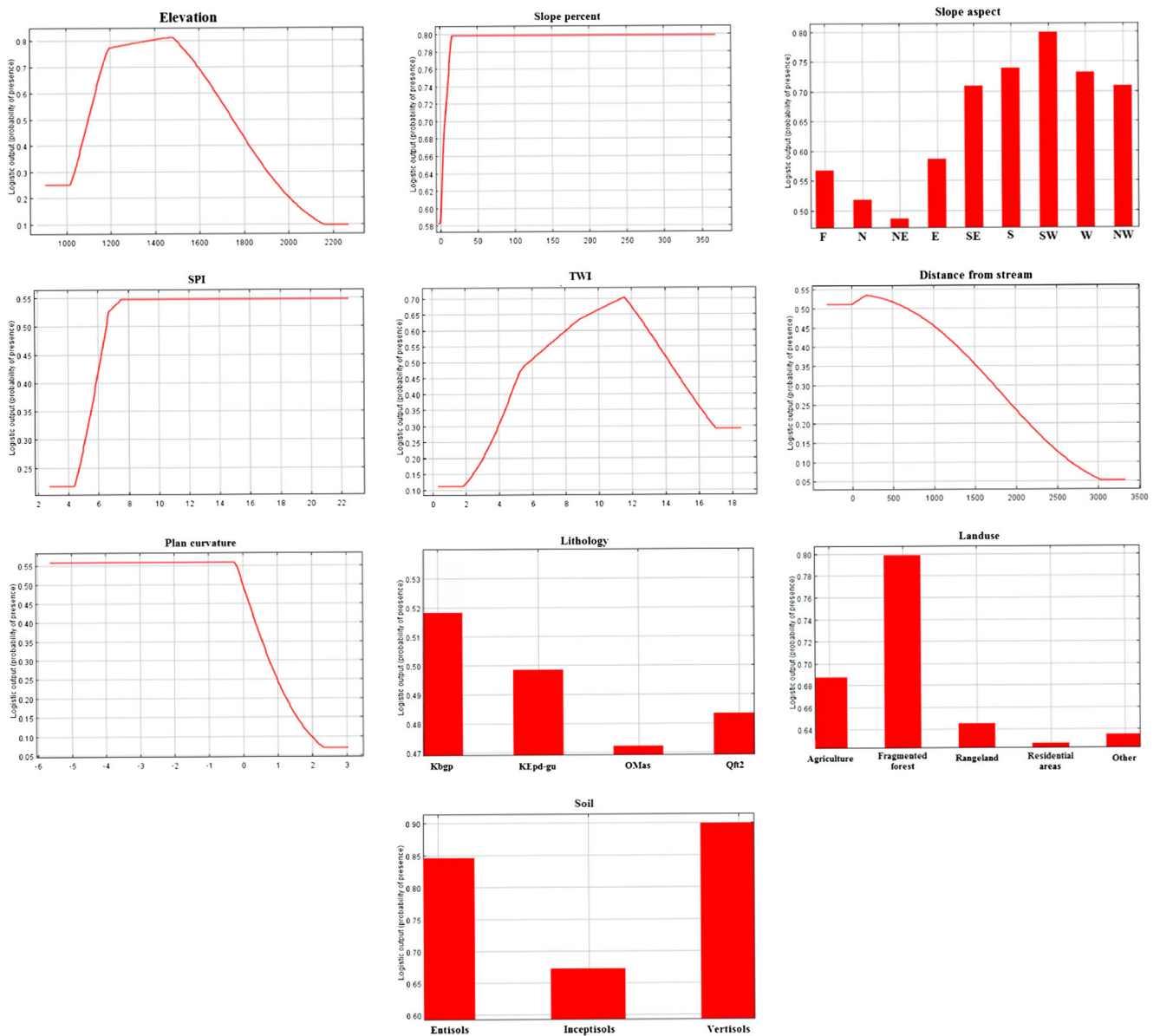
Fig. 3 Response curves for each erosion conditioning factor

values for training and validation datasets are 0.867 (86.7%), and 0.794 (79.4%), respectively.

Discussion and conclusion

There are various kinds of erosion such as gully erosion and rill erosion, each one affecting by different conditioning factors; hence, it is necessary to investigate them separately and consider their specific characteristics (Vandekerckhove et al. 2000; Kornejady et al. 2017b). Considering these differences, the current study was conducted to investigate the parameters affecting rill erosion and their importance by Jackknife test, generate rill erosion susceptibility map by the ME model, and

Table 4 The contribution of each rill erosion conditioning factor in the modelling process

Factor	Contribution (%)
Land use	14.5
Slope percent	13.8
Aspect	13.3
SPI	12.2
Altitude	10.1
Lithology	9.7
Distance from rivers	8.6
Soil	7.5
Plan curvature	7.1
TWI	3.2

Fig. 4 Soil erosion susceptibility
map produced by ME model
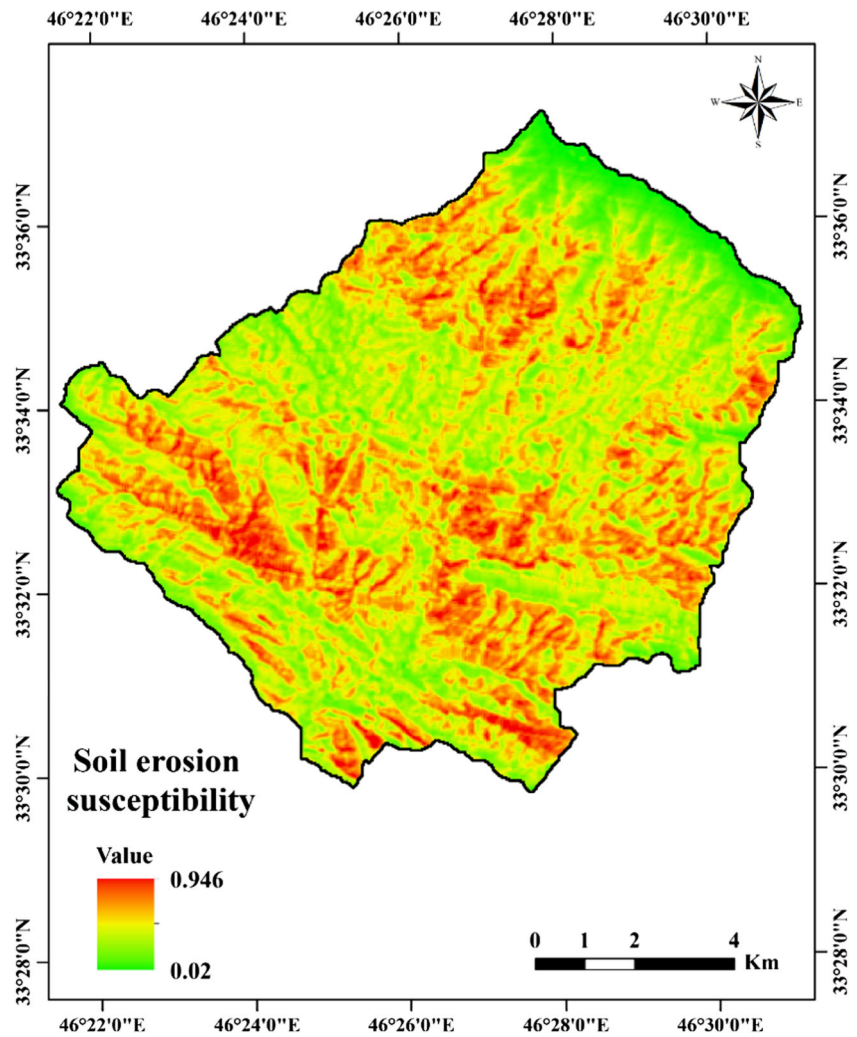
define the sediment yield at each hydrological unit of the watershed.

According to the results of MPSIAC, the highest and lowest predicted values of soil loss were observed in hydrological units 13 and 10, respectively. The soil loss predicted for the whole watershed was calculated as 170.67 m^3/km^2/y. The higher soil loss in hydrological unit 13 could be related to its lithological characteristics which are comprised of marl, and anhydrite with a calcite layer in between. The higher susceptibility of marl and anhydrite causes higher susceptibility of this

hydrological unit to rill erosion. In the case of hydrological unit 10, it is formed by calcite which is a resistant geological unit to erosion representing the lowest value of erosion.

Furthermore, the findings of this study showed that the ME model was successful in predicting the probability of rill erosion occurrence in both training and validation datasets. The higher value of AUC-ROC than 0.70 (70%) shows that the ESM produced in this study is trustable and could be implemented in other studies with similar conditions. The ME model is suitable for modelling natural phenomena such as rill erosion (Rahmati et al. 2016a). It needs to be mentioned that ME machine learning algorithm does not need earlier outlier removal. Another capability of this model is that it is able to predict incident of an event which is complicated and has a nonlinear structure (Phillips et al. 2006; Phillips and Dudík 2008; Kornejady et al. 2017a).

In addition, the importance of ECFs was assessed using Jackknife test. The high contribution of land use, slope percent, aspect and SPI and low contribution of TWI, plan curvature, and soil were identified in the study area. Two highly

Table 5 Area percentage related to different classes of the ESM produced by ME model

Susceptibility class	No. pixel	Area (ha)	Area (%)
Low	28,486	81.38	19.61
Medium	49,715	142.1	34.22
High	41,789	119.4	28.76
Very high	25,268	72.19	17.39

Fig. 5 Classified soil erosion susceptibility map of the study area produced by ME model
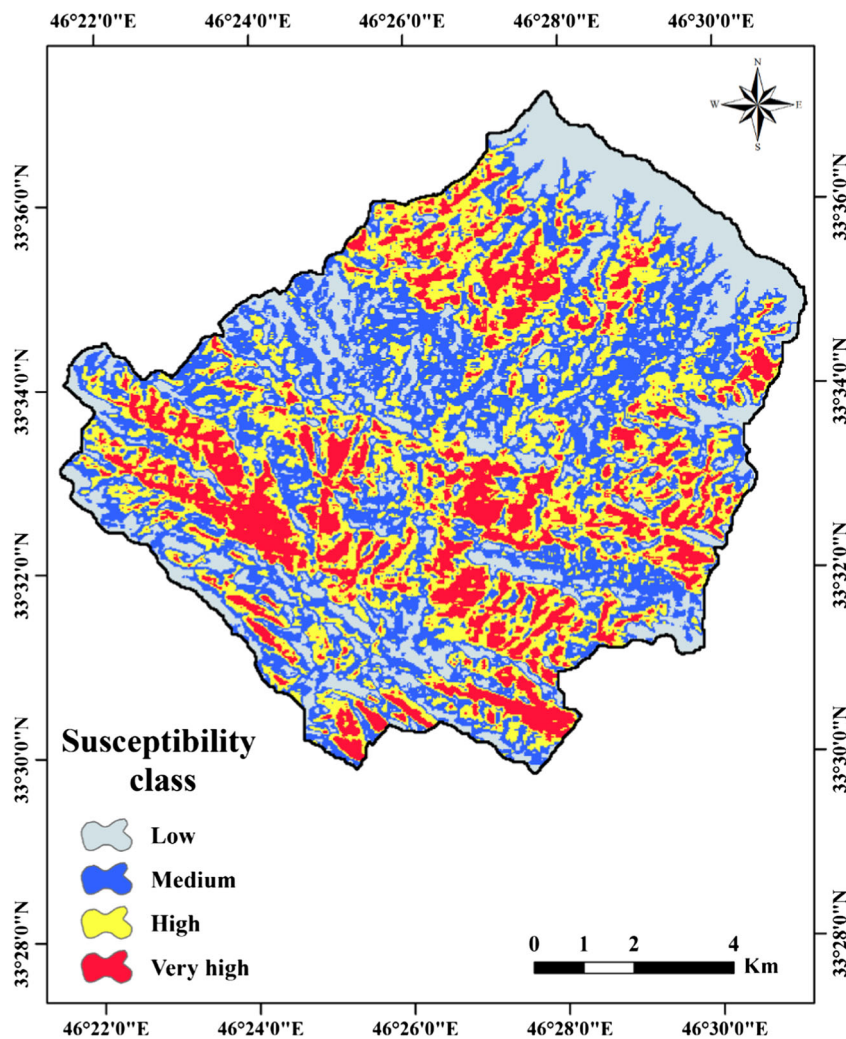
important factors of slope percent and SPI strongly influence the erosion power of water flow in a watershed, and their high
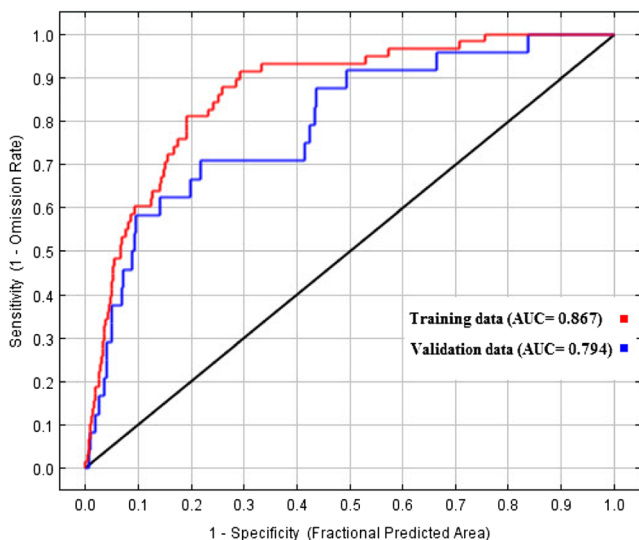

Fig. 6 ROC plot for training and validation data by ME model

importance could be justified with this point of view (Shrimali et al. 2001; Yesilnacar and Topal 2005; Sharma and Tiwari 2009; Conforti et al. 2011).

Finally, the findings of this study confirmed the acceptable performance of the ME model in producing rill erosion susceptibility map validated by training, validating data sets accompanied by ROC curve. The result is in agreement with findings reported by Pourghasemi et al. (2017), who applied the ME model to predict the susceptibility of gully erosion in Iran. The ME is a general-purpose machine learning model which presence-only property of the model can be considered as a strong advantage in remote and inaccessible areas. This feature is mostly important to soil erosion and landslide studies since one cannot reject the possibility of erosion/landslide occurrence even in the absence of the phenomena (Kornejady et al. 2017a). However, as a main disadvantage of the ME model, inadequacy of geo-environmental factors and lack of attention to the erosion process can threaten the prediction accuracy. As another conclusion, the high contribution of land use, slope percent, aspect and SPI and low contribution of

TWI, plan curvature, and soil were identified in the study area. Investigating the results of MPSIAC and ME models show that, in total, the southern part of the study area have a higher susceptibility to rill erosion compared to the northern part; however, in some places located at the northern part of the study area, high susceptible areas exist. This finding can be a useful tool for soil erosion control and conservation plans in the studied area. A deeper investigation of the results enlightens that high susceptible areas to rill erosion are often located in fragmented forest and agriculture land use classes. Inappropriate wood logging and transportation strategies as well as unsuitable agricultural plans and management may have caused this relationship between land use and rill erosion susceptibility. Considering the acceptable application of the ME model in this research, its utilization in other areas with different characteristics can be suggested in order to validate this methodology for a more general and wider application. At last, it could be suggested to use newer data mining models and coupling statistical and data mining models to get better results in mapping rill erosion susceptibility.

References

Abdullah M, Feagin R, Musawi L (2017) The use of spatial empirical models to estimate soil erosion in arid ecosystems. Environ Monit Assess. https://doi.org/10.1007/s10661-017-5784-y

Amiri F (2010) Estimate of erosion and sedimentation in semi- arid basin using empirical models of erosion potential within a geographic information system. Air, Soil and Water Research 3:37

Angileri SE, Conoscenti C, Hochschild V, Märker M, Rotigliano E, Agnesi V (2016) Water erosion susceptibility mapping by applying Stochastic Gradient Treeboost to the Imera Meridionale River Basin (Sicily, Italy). Geomorphology 262:61–76

Auerswald K, Fiener P, Dikau R (2009) Rates of sheet and rill erosion in Germany—A meta-analysis. Geomorphology 111(3):182–193

Bagherzadeh A, Daneshvar MRM (2013) Evaluation of sediment yield and soil loss by the MPSIAC model using GIS at Golestan watershed, northeast of Iran. Arab J Geosci 6(9):3349–3362

Bahadur KK (2009) Mapping soil erosion susceptibility using remote sensing and GIS: a case of the Upper Nam Wa Watershed, Nan Province, Thailand. Environ Geol 57(3):695–705

Bakker MM, Govers G, Kosmas C, Vanacker V, Van Oost K, Rounsevell M (2005) Soil erosion as a driver of land-use change. Agric Ecosyst Environ 105(3):467–481

Benzer N (2010) Using the geographical information systemand remote sensing techniques for soil erosion assessment. Pol J Environ Stud 19(5):881–886

Cerdan O, Le Bissonnais Y, Souchère V, Martin P, Lecomete V (2002) Sediment con- centration in interrill flow: interactions between soil surface conditions, vegetation and rainfall. Earth Surf Process Landf 27(2):193–205. https://doi.org/10.1002/esp.314

Chang-Jo FC, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. Nat Hazards 30(3):451–472

Chen W, Pourghasemi HR, Naghibi SA (2017a) A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China. Bull Eng Geol Environ:1–18. https://doi.org/10.1007/s10064-017-1010-y

Chen W, Pourghasemi HR, Naghibi SA (2017b) Prioritization of landslide conditioning factors and its spatial modeling in Shangnan County, China using GIS-based data mining algorithms. Bull Eng Geol Environ:1–19. https://doi.org/10.1007/s10064-017-1004-9

Choubin B, Darabi H, Rahmati O, Sajedi-Hosseini F, Kløve B (2018) River suspended sediment modelling using the CART model: A comparative study of machine learning techniques. Sci Total Environ 615:272–281

Conforti M, Aucelli PPC, Robustelli G, Scarciglia F (2011) Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). Nat Hazards 56(3):881–898

Conoscenti C, Di Maggio C, Rotigliano E (2008) Soil erosion susceptibility assessment and validation using a geostatistical multivariate approach: a test in Southern Sicily. Nat Hazards 46(3):287–305

Conoscenti C, Agnesi V, Angileri S, Cappadonia C, Rotigliano E, Märker M (2013) A GIS-based approach for gully erosion susceptibility modelling: a test in Sicily. Italy Environ Earth Sci 70(3):1179–1195

Conoscenti C, Angileri S, Cappadonia C, Rotigliano E, Agnesi V, Märker M (2014) Gully erosion susceptibility assessment by means of GIS-based logistic regression: a case of Sicily (Italy). Geomorphology 204:399–411

Csáfordi P, Pődör A, Bug J, Gribovsyki Z (2012) Soil erosion analysis in a small forested catchment supported by ArcGIS Model Builder. Acta Silvatica et Lignaria Hungarica 8(1):39–56

Daneshvar MRM, Bagherzadeh A (2012) Evaluation of sediment yield in PSIAC and MPSIAC models by using GIS at Toroq Watershed, Northeast of Iran. Front Earth Sci 6(1):83–94

Duarte L, Teodoro A, Gonçalves J, Soares D, Cunha M (2016) Assessing soil erosion risk using RUSLE through a GIS open source desktop and web application. Environ Monit Assess 188(6):1–16

Dumas P, Printemps J (2010) Assessment of soil erosion using USLE model and GIS for integrated watershed and coastal zone management in the South Pacific Islands. In: Proceedings Interpraevent, International Symposium in Pacific Rim, Taipei, pp. 856–866

Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explana- tion of Maxent for ecologists. Divers Distrib 17:43–57

Falah F, Ghorbani Nejad S, Rahmati O, Daneshfar M, Zeinivand H (2017) Applicability of generalized additive model in groundwater potential modelling and comparison its performance by bivariate statistical methods. Geocarto Int 32(10):1069–1089

Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2012) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10(2):175–189

Felicísimo ÁM, Cuartero A, Remondo J, Quirós E (2013) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. Landslides 10(2):175–189

Feng X, Wang Y, Chen L, Fu B, Bai G (2010) Modeling soil erosion and its response to land-use change in hilly catchments of the Chinese Loess Plateau. Geomorphology 118(3):239–248

Feyznia S (1995) Rocks strength against erosion factors in different climates of Iran. Journal of the Natural Resources of Iran 47:95–116

Geology Survey of Iran (GSI) (1997) Geology map of the Birjand Township. http://www.gsi.ir/Main/Lang_en/index.html. Accessed September 2000

Ghorbani Nejad S, Falah F, Daneshfar M, Haghizadeh A, Rahmati O (2017) Delineation of groundwater potential zones using remote

sensing and GIS-based data-driven models. Geocarto Int 32(2):167–187

Govers G, Giménez R, Van Oost K (2007) Rill erosion: Exploring the relationship between experiments, modelling and field observations. Earth-Sci Rev 84(3–4):87–102

Haghizadeh A, Siahkamari S, Haghiabi AH, Rahmati O (2017) Forecasting flood-prone areas using Shannon's entropy model. J Earth Syst Sci. https://doi.org/10.1007/s12040-017-0819-x

Hancock GR, Crawter D, Fityus SG, Chandler J, Wells T (2008) The measurement and modelling of rill erosion at angle of repose slopes in mine spoil. Earth Surf Process Landf 33(7):1006–1020

Harmon RS, Doe WW (2001) Landscape erosion and evolution modeling: Springer Science & Business Media

Harte J, Newman EA (2014) Maximum information entropy: a foundation for ecological theory. Trends Ecol Evol 29(7):384–389

Hong H, Naghibi SA, Moradi Dashtpagerdi M, Pourghasemi HR, Chen W (2017) A comparative assessment between linear and quadratic discriminant analyses (LDA-QDA) with frequency ratio and weights-of-evidence models for forest fire susceptibility mapping in China. Arab J Geosci 10(7):167. https://doi.org/10.1007/s12517-017-2905-4

Ilanloo M (2012) Estimation of soil erosion rates usingMPSIAC models (Case Study Gamasiab basin). International Journal of Agriculture and Crop Sciences 4(16):1154–1158

Kleidon A, Malhi Y, Cox PM (2010) Maximum entropy production in environmental and ecological systems. Philos T Roy Soc B 365(1545):1297–1302

Kornejady A, Ownegh M, Bahremand A (2017a) Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. Catena 152:144–162

Kornejady A, Ownegh M, Rahmati O, Bahremand A (2017b) Landslide susceptibility assessment using three bivariate models considering the new topo-hydrological factor: HAND. Geocarto Int. https://doi.org/10.1080/10106049.2017.1334832

Koulouri M, Giourga C (2007) Land abandonment and slope gradient as key factors of soil erosion in Mediterranean terraced lands. Catena 69(3):274–281

Ledermann T, Herweg K, Liniger HP, Schneider F, Hurni H, Prasuhn V (2010) Applying erosion damage mapping to assess and quantify off-site effects of soil erosion in Switzerland. Land Degrad Dev 21(4):353–366

Lu H, Gallant J, Prosser IP, Moran C, Priestley G (2001) Prediction of sheet and rill erosion over the Australian continent, incorporating monthly soil loss distribution. Land and Water Technical Report. CSIRO, Canberra

Mandal D, Sharda VN (2013) Appraisal of soil erosion risk in the Eastern Himalayan region of India for soil conservation planning. Land Degrad Dev 24(5):430–437

Märker M, Pelacani S, Schröder B (2011) A functional entity approach to predict soil erosion processes in a small Plio-Pleistocene Mediterranean catchment in Northern Chianti, Italy. Geomorphology 125(4):530–540

Medley KA (2010) Niche shifts during the global invasion of the Asian tiger mosquito, Aedes albopictus Skuse (Culicidae), revealed by reciprocal distribution models. Glob Ecol Biogeogr 19:122–123

Mekonnen M, Keesstra SD, Stroosnijder L, Baartman JE, Maroulis J (2015) Soil conservation through sediment trapping: a review. Land Degrad Dev 26(6):544–556

Meusburger K, Mabit L, Park J, Sandor T, Alewell C (2013) Combined use of stable isotopes and fallout radionu- clides as soil erosion indicators in a forested mountain site, South Korea. Biogeosciences 10(8):5627–5638

Ministry of Agriculture-Jahad (MAJ) (2014) The evaluation of land and resource planning, Report II, 283 p. http://www.maj.ir/Portal/Home/Default.aspx?CategoryID=92dd15a0-2395-43be-9859-c79116ab05da

Moore ID, Grayson RB, Ladson AR (1991) Digital terrain modeling: a review of hydrological, geomorphological and biological applications. Hydrocarb Process 5:3–30

Mullan D (2013) Soil erosion under the impacts of future climate change: Assessing the statistical significance of future changes and the potential on-site and off-site problems. Catena 109:234–246

Naghibi SA, Ahmadi K, Daneshi A (2017a) Application of Support Vector Machine, Random Forest, and Genetic Algorithm Optimized Random Forest Models in Groundwater Potential Mapping. Water Resour Manag. https://doi.org/10.1007/s11269-017-1660-3

Naghibi SA, Moghaddam DD, Kalantar B, Pradhan B, Kisi O (2017b) A Comparative Assessment of GIS-Based Data Mining Models and a Novel Ensemble Model in Groundwater Well Potential Mapping. J Hydrol. https://doi.org/10.1016/j.jhydrol.2017.03.020

Naghibi SA, Pourghasemi HR, Abbaspour K (2017c) A comparison between ten advanced and soft computing models for groundwater qanat potential assessment in Iran using R and GIS. Theor Appl Climatol. https://doi.org/10.1007/s00704-016-2022-4

Park NW (2015) Using maximum entropy modeling for landslide susceptibility mapping with multiple geoenvironmental data sets. Environ Earth Sci 73(3):937–949

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31(2):161–175

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modelling of species geographic distributions. Ecol Model 190:231–259

Pourghasemi HR, Rahmati O (2018) Prediction of the landslide susceptibility: which algorithm, which precision? CATENA 162:177–192

Pourghasemi HR, Yousefi S, Kornejady A, Cerdà A (2017) Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. Sci Total Environ 609:764–775

Rahmati O, Melesse AM (2016) Application of Dempster–Shafer theory, spatial analysis and remote sensing for groundwater potentiality and nitrate pollution analysis in the semi-arid region of Khuzestan. Iran Sci Total Environ 568:1110–1123

Rahmati O, Pourghasemi HR (2017) Identification of Critical Flood Prone Areas in Data-Scarce and Ungauged Regions: A Comparison of Three Data Mining Models. Water Resour Manag 31(5):1473–1487

Rahmati O, Haghizadeh A, Pourghasemi HR, Noormohamadi F (2016a) Gully erosion susceptibility mapping: the role of GIS-based bivariate statistical models and their comparison. Nat Hazards 82(2):1231–1258

Rahmati O, Pourghasemi HR, Melesse AM (2016b) Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at Mehran Region, Iran. Catena 137:360–372

Rahmati O, Tahmasebipour N, Haghizadeh A, Pourghasemi HR, Feizizadeh B (2017a) Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework. Sci Total Environ 579:913–927

Rahmati O, Tahmasebipour N, Haghizadeh A, Pourghasemi HR, Feizizadeh B (2017b) Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. Geomorphology 298:118–137

Razandi Y, Pourghasemi HR, Neisani NS, Rahmati O (2015) Application of analytical hierarchy process, frequency ratio, and certainty factor models for groundwater potential mapping using GIS. Earth Sci Inf 8(4):867–883

Renard KG, Foster G, Weesies G, McCool D, Yoder D (1997) Predicting soil erosion by water: a guide to conser- vation planning with the Revised Universal Soil Loss Equation (RUSLE), vol 703. United States Department of Agriculture, Washington, DC

Renschler CS, Harbor J (2002) Soil erosion assessment tools from point to regional scales—the role of geomorphologists in land management research and implementation. Geomorphology 47(2):189–209

Riksen MJPM, De Graaff J (2001) On-site and off-site effects of wind erosion on European light soils. Land Degrad Dev 12(1):1–11

Shahzeidi S, Entezari M, Gholami M, Dadashzadah Z (2012) Assessment rate of soil erosion by GIS (case study Varmishgan, Iran). Journal of Basic and Applied Scientific Research 2(12):13115–13121

Sharma SK, Tiwari KN (2009) Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley Catchment. J Hydrol 374(3):209–222

Shi ZH, Fang NF, Wu FZ, Wang L, Yue BJ, Wu GL (2012) Soil erosion processes and sediment sorting associated with transport mechanisms on steep slopes. J Hydrol 454:123–130

Shrimali SS, Aggarwal SP, Samra JS (2001) Prioritizing erosion-prone areas in hills using remote sensing and GIS—a case study of the Sukhna Lake catchment, Northern India. Int J Appl Earth Obs 3(1): 54–60

Siahkamari S, Haghizadeh A, Zeinivand H, Tahmasebipour N, Rahmati O (2017) Spatial prediction of flood-susceptible areas using frequency ratio and maximum entropy models. Geocarto Int:1–15. https://doi.org/10.1080/10106049.2017.1316780

Taheri M, Landi A, Archangi B (2013) Using RS, GIS systems and MPSIAC model to produce erosion mapand to estimate

sedimentation. International Journal of Agriculture: Research and Review 3(4):881–886

Tahmassebipoor N, Rahmati O, Noormohamadi F, Lee S (2016) Spatial analysis of groundwater potential using weights-of-evidence and evidential belief function models and remote sensing. Arab J Geosci 9(1):79. https://doi.org/10.1007/s12517-015-2166-z

Vandekerckhove L, Poesen J, Oostwoud-Wijdenes D, Nachtergaele J, Kosmas D, Roxo MJ, Figueiredo TD (2000) Thresholds for gully initiation and sedimentation in Mediterranean Europe. Earth Surf Process Landf:1201–1220

Vorpahl P, Elsenbeer H, Märker M, Schröder B (2012) How can statistical models help to determine driving factors of landslides? Ecol Model 239:27–39

Wirtz S, Seeger M, Ries JB (2012) Field experiments for understanding and quantifica- tion of rill erosion processes. Catena 91:21–34

Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a mediumscale study, Hendek region (Turkey). Eng Geol 79:251–266

Yilmaz C, Topal T, Suzen ML (2012) GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak Turkey). Environ Earth Sci 65(7):2161–2178

Zhao G, Mu X, Wen Z, Wang F, Gao P (2013) Soil erosion, conservation, and eco-environment changes in the loess plateau of China. Land Degrad Dev 24(5):499–510