

Article

Gully Erosion Susceptibility Mapping Using Multivariate Adaptive Regression Splines—Replications and Sample Size Scenarios

Narges Javidan ¹, Ataollah Kavian ^{1,*}, Hamid Reza Pourghasemi ² , Christian Conoscenti ³ and Zeinab Jafarian ⁴

¹ Department of Watershed Management, Faculty of Natural Resources, Sari Agricultural Sciences and Natural Resources University (SANRU), Sari 48441-74111, Iran; narges.javidan20@gmail.com

² Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz 71441-65186, Iran; hr.pourghasemi@shirazu.ac.ir

³ Department of Earth and Marine Sciences (DISTEM), University of Palermo, Via Archirafi 22, 90123 Palermo, Italy; christian.conoscenti@unipa.it

⁴ Department of Range Management, Sari Agricultural Sciences and Natural Resources University (SANRU), Sari 48441-74111, Iran; z.jafarian@sanru.ac.ir

* Correspondence: a.kavian@sanru.ac.ir

Received: 14 September 2019; Accepted: 29 October 2019; Published: 6 November 2019



Abstract: Soil erosion is a serious problem affecting numerous countries, especially, gully erosion. In the current research, GIS techniques and MARS (Multivariate Adaptive Regression Splines) algorithm were considered to evaluate gully erosion susceptibility mapping among others. The study was conducted in a specific section of the Gorganroud Watershed in Golestan Province (Northern Iran), covering 2142.64 km² which is intensely influenced by gully erosion. First, Google Earth images, field surveys, and national reports were used to provide a gully-hedcut evaluation map consisting of 307 gully-hedcut points. Eighteen gully erosion conditioning factors including significant geoenvironmental and morphometric variables were selected as predictors. To model sensitivity of gully erosion, Multivariate Adaptive Regression Splines (MARS) was used while the Area Under the Receiver Operating Characteristic (ROC) Curve (AUC), drawing ROC curves, efficiency percent, Yuden index, and kappa were used to evaluate model efficiency. We used two different scenarios of the combination of the number of replications, and sample size, including 90%/10% and 80%/20% with 10 replications, and 70%/30% with 5, 10, and 15 replications for preparing gully erosion susceptibility mapping (GESM). Each one involves a various subset of both positive (presence), and negative (absence) cases. Absences were extracted as randomly distributed individual cells. Therefore, the predictive competency of the gully erosion susceptibility model and the robustness of the procedure were evaluated through these datasets. Results did not show considerable variation in the accuracy of the model, with altering the percentage of calibration to validation samples and number of model replications. Given the accuracy, the MARS algorithm performed excellently in predictive performance. The combination of 80%/20% using all statistical measures including SST (0.88), SPF (0.83), E (0.79), Kappa (0.58), Robustness (0.01), and AUC (0.84) had the highest performance compared to the other combinations. Consequently, it was found that the performance of MARS for modelling gully erosion susceptibility is quite consistent while changes in the testing and validation specimens are executed. The intense acceptable prediction capability of the MARS model verifies the reliability of the method employed for use of this model elsewhere and gully erosion studies since they are qualified to quickly generating precise and exact GESMs (gully erosion sensitivity maps) to make decisions and management edaphic and hydrologic features.

Keywords: gully erosion susceptibility; GIS; robustness; MARS algorithm

1. Introduction

Soil erosion through water is found to be a drastic soil destruction process, which accounts for about one billion hectares worldwide [1], resulting in low plant development, filling reservoirs and valleys, geo-environmental degradation, degradation of a large part of the soil, and siltation of watercourses [2–4]. One of the soil degradation processes is gully erosion and serves as the most intricate erosion phenomena [5], usually stimulated or exacerbated integrating extreme rainstorms and unwise land exploitation [6]. Such erosion contains wide varieties of small processes, including head-cut, fluting, piping, continuous cracking progress, and mass flow [7,8]. Generally, the increasing attention towards analysis of gully erosion indicates the necessity to enhance our awareness regarding its consequences and condition agents that change due to various factors [6]. Gullies involve complex pathways adjusted through the variability of correlated variables including soil texture, lithology, land use and plant canopy, climate, and topography [9]. As gully erosion is a threshold phenomenon [8], various studies have emphasized characterizing the topographic as well as hydraulic conditions to forecast and assess the starting gullies susceptibility mapping [10], to a threshold approach where characterization and positions of erosion processes might as well as be anticipated by using bivariate to multivariate methods. Such techniques provide scholars the ability to describe soil erosion processes, through evaluating space distribution of the gully erosion forms (the consequences) compared to some predictors (geological and environmental factors). As for geomorphology, actuarial methods are enormously used to evaluate landslide susceptibility mapping [11–17].

The large number of studies have at the same time used the probabilistic method to map erosion sensitivity and other hazards. As well as bivariate methods [18–20], various multivariate actuarial approaches were considered to satisfy this end including logistic regression [21,22], classification and regression trees [23,24], and multivariate adaptive regression splines [25–28]. Gutiérrez et al. [27] used the Multivariate Adaptive Regression Splines (MARS) model to predict gully creation locations. The results showed that this model has good performance in geomorphic research.

Gully erosions have the highest sediment production potential. In recent decades, gully erosion has developed in most watersheds of Iran. Gullies are an important sediment source and often cause environmental problems [29]. Due to their damages, such as loss of productive capacity and significant land degradation, high sediment discharge and sediment yields, which can transport both pollutants and nutrients, reducing the water capacity of the reservoirs and damage to the infrastructure and transport routes, prediction of susceptible areas is, therefore, considered essential in management of watersheds [30]. The result of gully erosion study demonstrates the susceptibility to erosion over a country, providing beneficial information for remediation strategies and establishing land use plans [29].

In ongoing research, we adopted the multivariate adaptive regression splines [31] as a multivariate actuarial method for analyzing, assessing, and forecasting the local incidence of gully erosion pathways. The MARS model serves as the most common actuarial method that previously proved to offer credible patterns of gully sensitivity. The reason behinds choosing MARS models for the forecasting gully erosion are as follows: (1) possibility for modeling curvilinear association among the conditioning factors and gully incidence; (2) it allows working with various outcome variables and may manipulate data from different measures; (3) according to previous studies in this area, that any studies have applied this model for evaluating its ability and robustness for gully susceptibility.

This area in Gorganrood has witnessed gully erosion that caused many issues in this area and led organizations to reassess the Weiridness of adopted constant genesis strategies (CONRWMGP 2009). Astonishingly, the major initial place for the erosion phenomenon is cutting down trees positioned at the upper part of Gorganrood watershed and land-use changes (CONRWMGP 2009). From 1990 to 2005 due to the presence of loess soils in the northern of Golestan Province, 430,000 ha of these areas were affected by erosion. Soil erosion in this province is 5–6 tons/ha/year in forest areas (Department of Golestan Natural Resources and Watershed management). Gully erosion in Maravetape and Kalale counties leads to the loss of soil, the imposition of large costs, reduced agricultural potential, and has caused the migration of people in the villages of this region and exacerbated soil erosion pathways

influence farming system efficiency. Consequently, to describe a robust model to evaluate the sensitivity of the territory to the development of gully pathways is necessary and Gorganrood watershed as the susceptible area will be the focus of the present study.

The difference between this study and previous studies who used the MARS model is applying two scenarios of the combination of number of replications and sample size, including 90%/10% and 80%/20% with 10 replications, and 70%/30% with 5, 10, and 15 replications for preparing gully erosion susceptibility mapping (GESM) and assessing their performance by MARS algorithm. Each one involves a various subset of both positive and negative cases. Absences are extracted as randomly distributed individual cells. Therefore the predictive competency of the gully erosion susceptibility model and the robustness of the procedure were evaluated through these datasets.

Therefore, the main scope of the current study is gully erosion modeling based on the MARS technique and explores the ability and robustness of the MARS model to forecast the incidence of gully erosion by various data sets and assessment measures. This study enriches the systematic assessment of the MARS model to map gully erosion susceptibility among others. This study tried to investigate gully erosion susceptibility through the MARS model and analyzed the performance and accuracy of this technique for zoning gully erosion.

The result of this study demonstrates ways of offering beneficial insight for remediation strategies and founding land exploitation projects by erosion susceptibility.

2. Materials and Methods

2.1. Study Area

The area under research belongs to Gorganrood basin, related to Golestan Province located in north-eastern Iran. This area accounts for 2142.64 km² and is intensely influenced by gully erosion. Its coordinates span over 37°18'–37°52' N and 55°18'–56°10' E (Figure 1). Topographically, this region is considered as plain. The mean height ranges between 56 and 2165 m. Its prevalent soil textures are Silty-loamy (approximately 53.7%) and Silty-clay-loamy (nearly 45.3%) soils. The major land uses in this region include pasturelands (40.1%), agriculture (farming) (29.4%), and forest (18.4%). The mean annual rainfall is approximately between 460 and 603 mm. The average minimum and maximum temperatures are 11 and 18.5 °C, respectively [32]. In the last decade, this area has been challenged with natural hazards and has faced intensive gully erosion. Therefore, this study area was selected as a potential gully erosion-prone area. Figure 2 presents some Google Earth images of gully erosion in the research field.

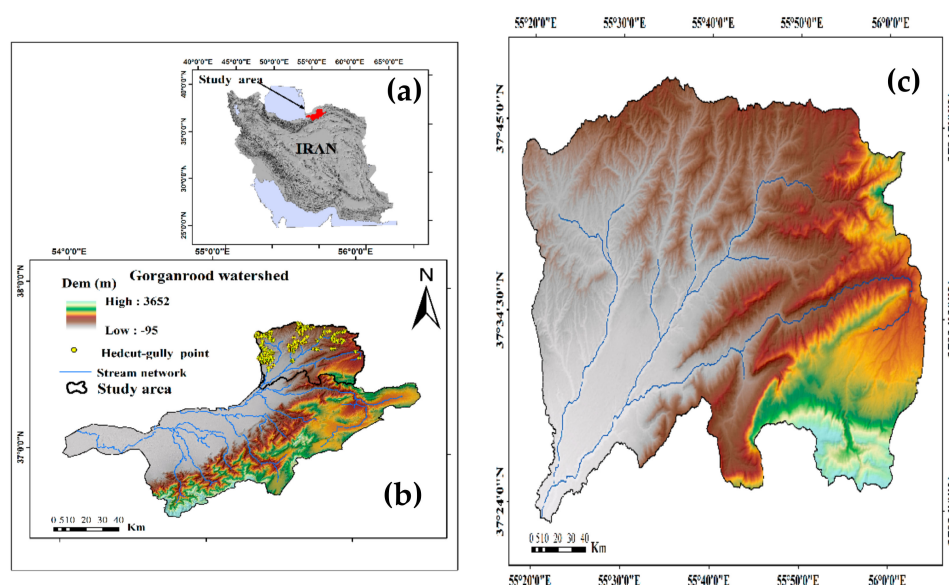


Figure 1. Study area and topographical characteristics. (a) Iran, (b) Gorganrood Watershed, (c) Study area.



Figure 2. Some photographs of the gully erosion zoned on study area.

2.2. Methodology

Figure 3 indicates a flow-diagram for the method applied for the application of multivariate adaptive regression splines model for gully erosion sensitivity zoning developed for this specific research in the north of Iran. As shown, the flowchart consists of five steps: (1) preparing thematic layers or 18 effective conditioning factors; (2) selection of factors using a multi collinearity test; (3) applying two scenarios of the combination of the number of replications, and sample size, including 90%/10% and 80%/20% with 10 replications, and 70%/30% with 5, 10, and 15 replications; (4) gully erosion susceptibility modeling using the MARS technique; (5) validation of the susceptibility maps using the ROC-AUC (Area Under the Receiver Operating Characteristic) curve, efficiency percent, Yuden index, and kappa.

2.2.1. Gully Erosion Inventory Mapping

An essential stage to zoning is to create a hazard evaluation for hazard zones [22]. The gully erosion inventory for the present section of Gorganrood Watershed was provided by Google Earth images, field surveys, and national and regional reports from different organizations. The present map constitutes a set of incidences (307 hedcut points). While designing statistical plans, the training set must differ from sets applied in the validation part [33]. To distinguish training points from the validation points a random dividing algorithm [29,34] was used. In this research, two scenarios were used: these scenarios were selected after altering different sample sizes and the number of replications, including 90%/10% and 80%/20% with 10 replications. To assess the robustness of the model's data sensitivity [8,22,35], 5, 10, and 15 sample data sets, (replicates) for 70%/30% sample size, were prepared through randomly multi-extracting of various data sets in the calibration and validation subsets [36]. Every set was adjusted through addition to positives (i.e., pixels having hedcut points) an equal number of randomly selected negative points, corresponding to pixels without hedcut [37].

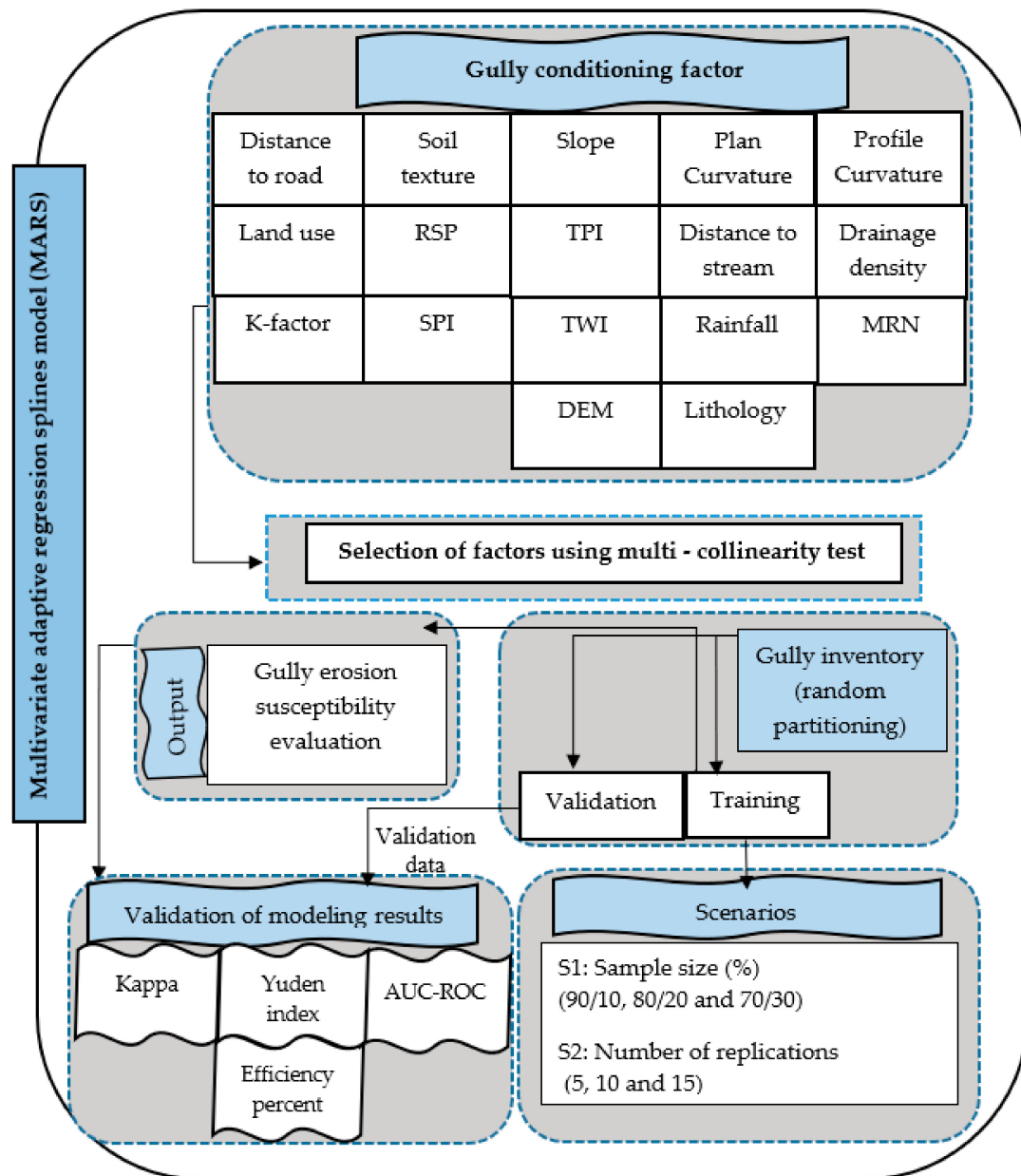


Figure 3. Flow diagram for the research method.

2.2.2. Gully Erosion Predictor Variables (GEPV)

It is essential to determine the effective factors on natural hazards and man-made fatalities in order to performing gully erosion susceptibility maps have great importance [38]. Good knowledge of the main gully erosion-related factors is required to recognize the susceptible areas. Such contributors always apply in studies analyzing gully erosion. Therefore, such factors were chosen from past studies [34,39]. In this study, to generate and exhibit such data grid, ArcGIS 10.5 and a system for automated geoscientific analyses (SAGA) software were used. For the application of the MARS model, all agents were transformed into a raster network with 30 × 30 m grid pixel. All conditioning factors were primarily continuous, and some of them (litology, soil, and land use) were classified within different categories based on expert knowledge and literature review [40–42].

The predicting factors used in this work are (a) digital elevation model (m), (b) aspect map, (c) slope percent, (d) curvature of profile, (e) curvature of plan, (f) land use (LU), (g) soil texture, (h) Topographic wetness index (TWI), (i) distance to streams (m), (j) distance to roads (m), (k) drainage

density, (l) annual rainfall (mm), (m) stream power index, (n) relative slope position, (o) lithological formation, (p) K factor, (q) Melton ruggedness number, (r) topographic position index.

Digital contour data obtained from the Department of Natural Resources Management of Iran was applied. A DEM (Digital Elevation Model) (Figure 4a) of the research field characterized with a 30 m pixel was generated. Drawing upon DEM, physiographical and geomorphological grids such as aspect (Figure 4b), slope percent (Figure 4c) as well as curvature layers were extracted using ArcGIS 10.5. The slope percent includes a large section of the intrinsic views and is an important factor as it affects drainage density, surface runoff, influences, vegetation structure the soil erosion, soil moisture, and geomorphological processes [9,43–47]. Slope aspect is another important factor related to precipitation, snow meltwater, land cover, soil moisture patterns, and physiographic trends [48–52]. The suitable geomorphological data may be inferred via curvature assessment [53–55]. Three categories, convex, concave, and flat, were used to develop the slope curvature map. Positive curvature exhibits convex ($>+0.1$), negative curvature depicts concave (<-0.1), and zero curvature represents flat ($(-0.1)-(+0.1)$). In addition, profile and plan curvatures (Figure 4d,e) include various negative as well as positive quantities and indicate various description in every measure. Negative as well as positive ones in profile curvature denote convexity (increasing flow velocity) and concavity (reducing flow velocity), respectively. In contrast, negative and positive values in the plan curvature imply concavity (flow convergence) as well as the convexity (flow divergence) [54,56]. Those near to zero denote neutral curvature in both conditions.

LU has a substantial contribution in geomorphological and hydrological pathways through direct or indirect influences on evapotranspiration, infiltration, run-off generation, and sediment dynamics [52,57]. The other hand, agricultural activities have an important impact on gully erosion development as well as genesis [58]. The land-use map in the region in 1:100,000 scale was prepared by the Natural Resources department of Golestan Province and manipulated by Google Earth images. The land-use of the region consists of residential areas (urban), range and farming, forestlands, rangelands, farming, and lake (Figure 4f).

Soil texture (Figure 4g) usually is recognized as a substantial limiting factor in the process of infiltration and overflow and it is effective on gully occurrence [59–61]. Through digitizing the soil texture map of Golestan Province (1:100,000 scale) obtained from the Agriculture Department, Iran the aforementioned layer was prepared. The soil texture in the study area consists of sandy-loam, clay-loam, sandy-clay-loam, silty-clay, silty-clay-loam, as well as silty-loam.

Moore, Grayson, and Ladson [62] and Grabs et al. [63] mentioned that TWI (Topographic wetness index) represents the spatial distribution of wetness conditions, and inclination of gravitation to transport water to downstream. This factor was prepared using Equation (1):

$$TWI = \ln\left(\frac{\alpha}{\tan \beta}\right), \quad (1)$$

where α denotes the aggregated upstream area leaving a point (contour points length) and $\tan \beta$ is point angle. Here, GIS 10.5 software was used for TWI mapping and ranges from 1.20 to 22.92 (Figure 4h).

Distance to streams (Figure 4i) serves as a key determinant because of its important effect on flow magnitude as well as gully erosion [29]. Based on field studies, gully erosions are diffused typically close the linear aspects in particular roads. Undoubtedly, road making imposes an adverse effect on hill sustainability at which flow may be appropriate for gullies [22,64].

Layers of the proximity were produced using the Euclidean metric function in ArcGIS 10.5 and ranged from 0 to 11,720 m for roads (Figure 4j), and 0–15,080 m for streams. The national topographic map at the scale of 1:50,000 was used to extract routes and streams.

The drainage density (Figure 4k) is found to be a great factor which plays an important role in the incidence of many hazards [65]. Several factors such as the structure and nature of the soil characteristics, geological beds, infiltration rate, plant cover condition, and slope degree [66] affect

drainage networks. To turn the drainage network model to a reasonable value, the drainage density was specified via an extension of “line density” in ArcGIS 10.5 software.

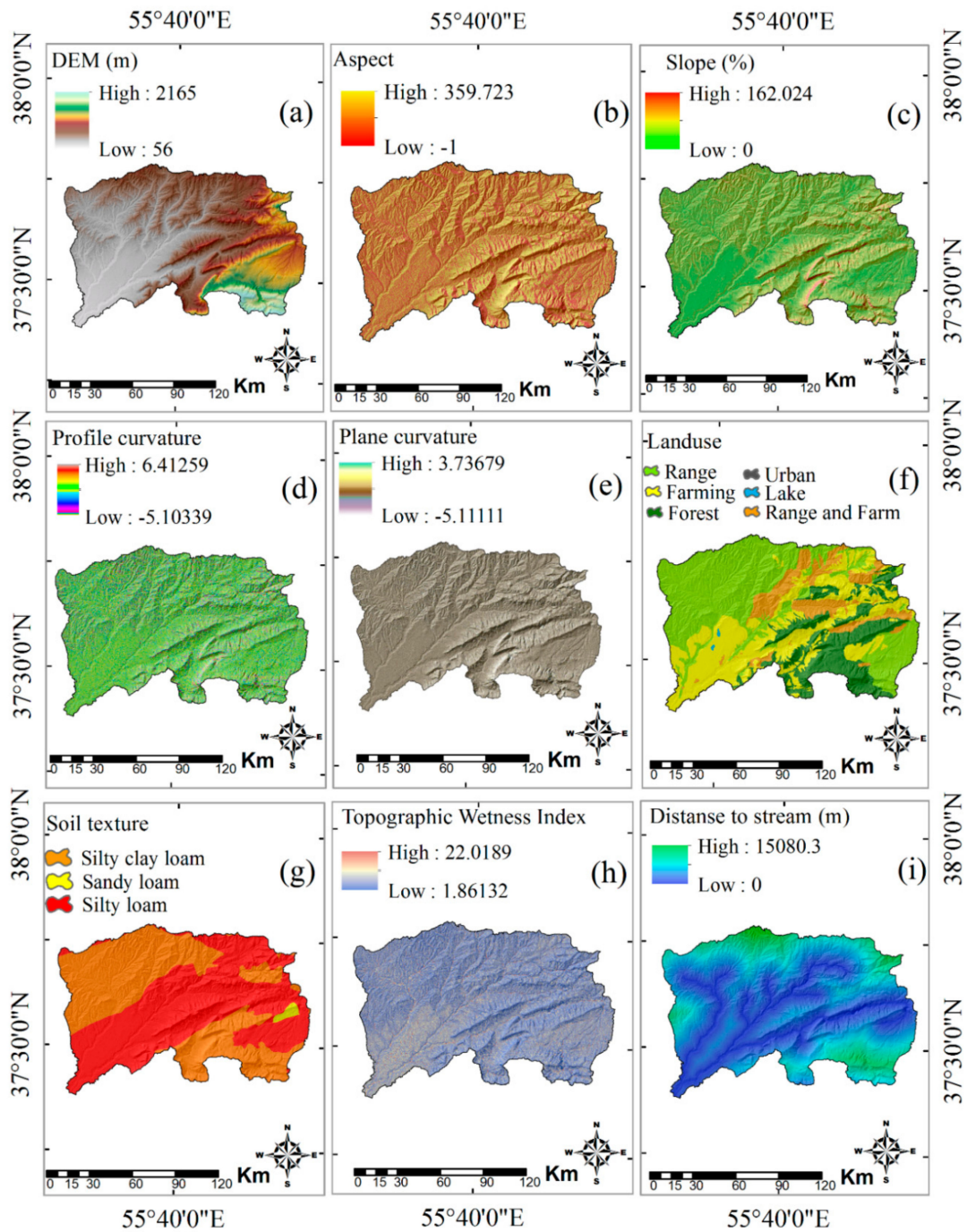


Figure 4. Cont.

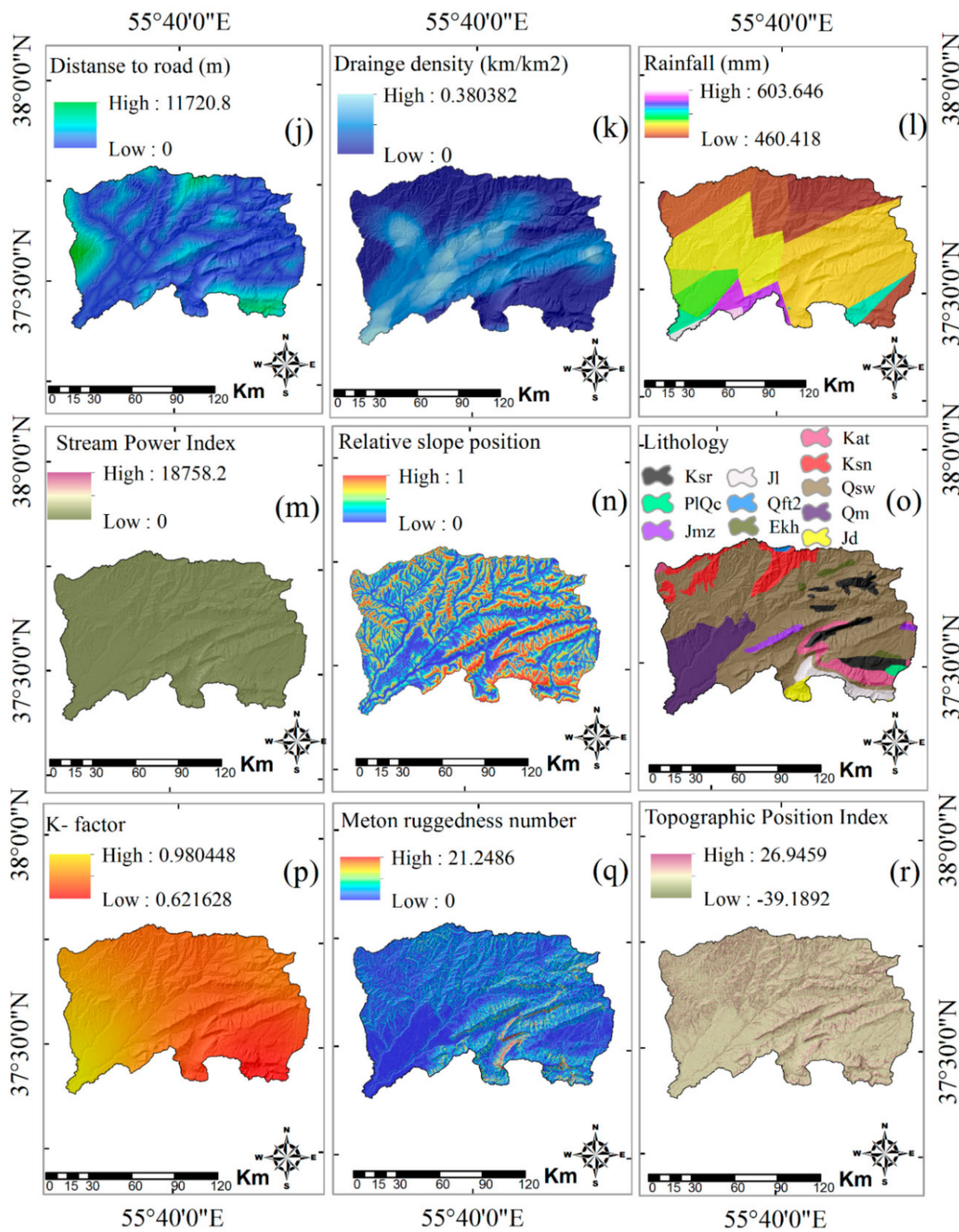


Figure 4. Gully conditioning factor maps. (a) Digital Elevation Model (DEM) (m), (b) Aspect, (c) Slope (%), (d) Profile curvature, (e) Plan curvature, (f) Land use, (g) Soil texture, (h) Topographic Wetness Index, (i) Distance to stream (m), (j) Distance to roads (m), (k) Drainage density (l) Rainfall (mm), (m) Stream Power Index, (n) Relative slope position (o) lithology, (p) K factor, (q) Melton ruggedness number, (r) Topographic Position Index.

The yearly average precipitation map (Figure 4l) of Gorganrood Basin was developed according to precipitation data obtained from the Golestan province Regional Water Organization. The above-mentioned map was developed using the 53 gauges and statistical period of 2009–2016 based on Inverse Distance Weight (IDW) interpolation method [67] (Equation (2)). This map ranges from 460 to 603 mm/year. The rainfall map was developed in 30 × 30 m in ArcGIS 10.5 as an input layer for susceptibility assessment of gully erosion.

$$\lambda_i = \frac{D_i^{-\alpha}}{\sum_{i=1}^n D_i^{-\alpha}}, \quad (2)$$

where, λ_i represents point i weight, D_i denotes the space between the point i and the point of unknown, and α implies weighing power [67].

The stream power index (SPI) (Figure 4m) is an index of the water flow erosive power according to the hypothesis that flow is relative to the particular watershed [68].

$$SPI = A_s \times \tan \sigma, \quad (3)$$

where, A_s is the particular watershed area per meter and r is the gradient of slope in degrees. The SPI index is the most important factor adjusting slope erosion processes, as erosive power of flow directly affect s river cutting and slope to erosion [68]. The regions with great river power measures have excessive erodibility because it indicates potential energy for limiting sediment [69].

Relative slope position (RSP), (Figure 4n), as a tool, could calculate several terrain indices from the digital elevation model. General information on the computational concept can be found in [70].

Lithology units (Figure 4o), have a dominant contribution in specifying gully erosion sensitivity [9,50,52,71,72] as gully erosion relies on the lithology properties and different lithological units display significant differences in erosion instability. In this research, the lithological map for the region was generated as the present geological maps with a 1:100,000 scale obtained from the Geological Survey Department, Iran. This area of Gorganrood Watershed is full of various types of outcrop formations and divided into 10 classes (Table 1).

Table 1. Lithology of the Gorganrood watershed.

Group	Code	Explanation	Formation
1	Ksr	Shale containing Ammonite with interaction of orbitolin limestone	Sarcheshmeh
2	PIQc	Fluvial conglomerate, Piedmont conglomerate, and sandstone.	-
3	Jmz	Grey thick-fluvial limestone and dolomite	Mozduran
4	Ksn	Brown to block shale and thin layers of siltstone and sandstone	Sanganeh
4	Murm	Light-red to brown marl and gyps marl with sandstone intercalations	-
4	Murmg	Gypsiferous marl	-
4	E1m	Marl, gypsiferous marl and limestone	-
5	Mur	Red marl, gypsiferous marl, sandstone and conglomerate	Dalichai
5	Kad-ab	Usual unit comprising argillaceous limestone, marl and shale	-
5	Jd	Well-bedded to thin-bedded, greenish-grey argillaceous limestone with intercalations of calcareous shale	-
6	Qft1	Concentrated piedmont fan and valley terrace deposits	-
6	Qft2	Low level piedment fan and valley terrace sedimentation	-
6	Qal	River channel, braided drainage and flood plain sedimentation	-
6	Qs,d	Loose loess sand sedimentation such as dunes	-
7	Jl	Light brown, thin-bedded to massive limestone	Lar
8	Ekh	Olive-green shale and sandstone	Khangiran
9	Kat	Green glauconitic sandstone and shale	Aitamir
10	Qsw	Swamp	-
10	Qm	Swamp and marsh	-

Soil erosion potential (k-factor) (Figure 4p) influences soil persistence to run-off force or rainfall impact [73]. The soil erodibility (K) in Universal Soil Loss Equation (USLE), is measured by the texture, organic matter, infiltration, and soil structure.

A simplified flow accumulation measure, computed as discrepancy among max and min elevation in the watershed divided by the square root of the watershed area is Melton ruggedness number (MRN) (Figure 4q). The measurement is done for every grid pixel, hence minimum elevation equals elevation in the cell's location. Flow measurement is easily performed with Deterministic 8 given the inconsistent nature of a single maximum elevation [74–76].

Topographic position index (TPI) (Figure 4r) serves as an approach widely applied to evaluate topographic slope location, and to zone ordination automation. This function produces single-band raster characterized quantities measured upon elevation. TPI is an abbreviation for Topographic Position Index, in turn, described as the difference between the main pixel and the average of its adjacent pixels [77].

2.3. Multi-Collinearity Test

The above-mentioned factors were used to consider the effect of correlation among them as the independent variable. If both predictor variables are intensely related, it is a problem in the modelling process. The issue is named collinearity. The VIF (variance inflation factor) and Tolerance include both significant measures of multi-collinearity recognition. Indeed, VIF is simply the reciprocal of tolerance, on the other hand, tolerance is $1-R^2$ for the variable regression versus predictors, deprived of the dependent variable [78]. A VIF of five or 10 and above and/or a tolerance of less than 0.20 or 0.10 indicates a multi-collinearity problem [79,80].

2.4. Multivariate Adaptive Regression Splines (MARS Model)

MARS can adapt complicated, non-linear associations among the independent and dependent measures although providing an explainable model [81]. The MARS algorithm has been applied in geomorphology to forecast and mapping the incidence of gullies [25–27] and landslides [17,38,82].

This technique functions via dividing magnitudes of the explanatory predictors into areas and through the fitting, for every area, a linear regression equation. “Knots” divides quantities among regions, whereas the phrase “basis function” (BF) denotes implying every various condition factor interval (dependents). BFs are functions of the following:

$$\begin{aligned} &\max(0, x - k) \text{ or} \\ &\max(0, k - x), \end{aligned}$$

where k denotes constant equal to a knot and x is an independent variable. The MARS may be stated generally as below:

$$y = f(x) = \alpha + \sum_{n=1}^N \beta_n h_n(x), \quad (4)$$

where y is predictor variable anticipated via function $f(x)$, N denotes terms number, who is shaped by a coefficient β_n , α is a constant, and $h_n(x)$ represents a separate basis function or multiplication of BFs. MARS analysis was carried out by the Earth version of the R software [83–85]

Evaluation of the Model

The evaluation pathway comprises the fitting degree assessment (goodness of fit), robustness, and prediction skills of the model [22,86]. The goodness of fit demonstrations capability of the approach in forecasting the training subset, whereas the predictive efficiency (prediction skills) is a fundamental step for model precision to forecast a validation set (the percent of gully points that do not use the training process) [39,87]. While changing training and validation points, the precision of the forecasting model is determined as the consistency of outputs of the model in respect to model precision. Here, predictive performance and goodness of fit of the model are subjected to assessment using both threshold-driven and non-threshold-dependent methods. ROC curve, as a threshold-driven method,

was drawn for all datasets as well as afterwards goodness of fit (i.e., degree of fitting) and forecasting efficacy of the algorithms were studied, respectively [88].

The Kappa coefficient, Youden index, and efficiency (as threshold dependent performance) were calculated based on the components of the confusion matrix [89]. Furthermore, sensitivity (SST) and specificity (SPF) are common statistical indexes applied for validation of each model performance [90]. Comparison of the model observed data and results are demonstrated through a contingency matrix (Table 2). According to Table 2, the true negative (TN) and true positive (TP) are the numbers of pixels that are correctly and appropriately classified, whereas false negative (FN) and false positive (FP) are the numbers of pixels fallaciously classified.

$$SST = \frac{TP}{TP + FN} \tag{5}$$

$$PF = \frac{TN}{TN + FP} \tag{6}$$

Efficiency (E) is the ratio of gully points and non-gully cells which output model accurately divided:

$$E = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

The Kappa factor (K) describes the potential of the employed model to categorize the gully point cells [1], and can be expressed as the ratio of given consistency beyond that expected by chance:

$$K = \frac{P_{obs} - P_{exp}}{1 - P_{exp}} \tag{8}$$

where P_{obs} represents the ratio of cells that are properly divided as gully occurrence or non-gully and P_{exp} denotes the ratio of pixels whose consistency is random [91]. P_{obs} and P_{exp} can be measured as below:

$$P_{obs} = TP + TN, \tag{9}$$

$$P_{exp} = [(TP + FN)(TP + FP) + (FP + TN)(FN + TN)]. \tag{10}$$

The performance of model given the Kappa factor may be categorized as ≤ 0 (weak), 0–0.2 (slight), 0.2–0.4 (fair), 0.4–0.6 (moderate), 0.6–0.8 (high), and 0.8–1 (almost perfect) [1,2].

Table 2. Contingency matrix applied for the Multivariate Adaptive Regression Splines (MARS) model assessment.

Observed	Predicted	
	%Gully (+)	%Non-Gully (–)
Gully (+)	(+ +) True positive (TP)	(– +) False negative (FN)
Non-gully (–)	(+ –) False positive (FP)	(– –) True negative (TN)

The ROC (receiver operating characteristics) curve, as a non-threshold-driven method, is widely used to quantify the measure classification [92–94]. The receiver operating characteristics is the area under the curve (AUC-ROC) that describes the performance of a model to precisely forecast non-incidence or an incidence [95,96]. The forecasting precision of the model according to the AUC value may be categorized as three classes of accuracy following the classification proposed by [97]: 0.7, 0.8, and 0.9. AUC thresholds were used for acceptable, superior, and substantial performance, respectively [22,37].

Since the admission of a forecasting model entails, for assessment of its precision, trifle variations of the input data (i.e., data susceptibility), a gully erosion sensitivity model was developed on 10 various

samples of mapping measures. Therefore, robustness in forecasting models and their consistency were furthermore evaluated when the training and validation samples are altered (i.e., a replicate method) [8,22,41,96,98]. The model was employed to given data, and subsequently was tested by the validation datasets. As for the ROC curve, the calculated event map by the validation dataset were compared.

Through subtracting the maximum and minimum accuracy values according to each assessment measures precision of the model was calculated [22,96]:

$$R_{AUC-ROC} = AUC - ROC_{max} - AUC - ROC_{min}, \quad (11)$$

where $R_{AUC-ROC}$ is model precision as per AUC-ROC measures, and $AUC-ROC_{max}$ includes maximum precisions within all sets. As well, least accuracy values are represented as $AUC-ROC_{min}$ within whole datasets. In addition, in this research, ROC curves were applied to designate the optimum benchmark point of the model rate, via considering Youden's index (J) [99], with J relative to the maximum perpendicular space between the ROC and the former bisector as follows:

$$J = \text{Maximum} (\text{sensitivity} + \text{specificity} - 1). \quad (12)$$

3. Results

3.1. Gully Erosion Susceptibility Model

First, Google Earth images, field surveys, and national reports were used to provide a gully-hedcut evaluation map consisting of 307 gully-hedcut points. Here, the MARS model was employed on balanced given sets (positives/negatives), and each one comprises all the positive (gully point) pixels and same randomly inferred negative (non-gully spots) cells, that was for two processes of integrating some repeats and samples, involving 90%/10% and 80%/20% with 10 replications. To assess the robustness of the model's data sensitivity, 5, 10 and 15 sample datasets (replicates) for 70%/30% sample size, were prepared through random selection of different data sets in the calibration and validation subsets. Figure 5 indicates the relative diffusion of the mean of gully erosion susceptibility categorizes for 10 groups in the sample data sets (70%/30%).

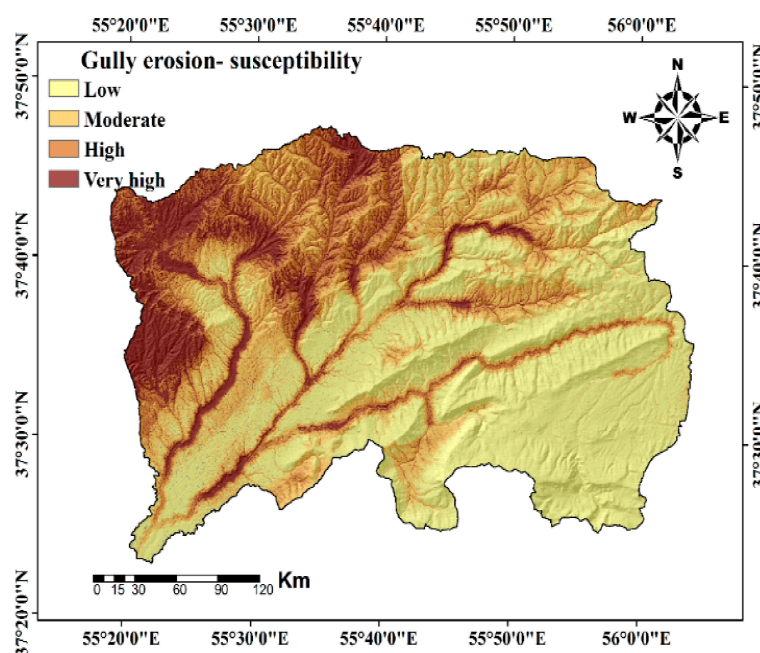


Figure 5. Sensitivity maps of gully erosion of the study area by the MARS model for 70%/30%, 10 replications.

The results of the relative distribution of the average of gully erosion susceptibility classes for other models are presented in Table 3. However, in the other combinations, the results were almost similar together. As well, the actuarial features of the probabilistic forecasting of the gully erosion of 10 sample data sets and replicates are shown in Table 4.

Table 3. Relative distributions of the gully susceptibility classes.

Relative Distributions of the Gully Susceptibility Classes					
MARS Model	70%/30%			80%/20%	90%/10%
	5 rep *	10 rep	15 rep	10 rep	10 rep
Low	47.14	44.72	45.86	47.55	47.65
Medium	22.83	23.63	22.85	22.16	21.94
High	15.70	17.10	16.27	15.20	16.17
Very high	14.34	14.55	15.02	15.09	14.25

* rep: Replicate.

Table 4. Actuarial features of the probability values inferred from the MARS model.

Probabilistic Prediction Values					
MARS Model	70%/30%			80%/20%	90%/10%
	5 rep	10 rep	15 rep	10 rep	10 rep
Mean	0.277	0.279	0.283	0.277	0.275
SD	0.281	0.270	0.280	0.285	0.273
Minimum	0.000	0.000	0.000	0.000	0.000
Maximum	0.999	0.997	0.996	0.999	0.998

SD: Standard deviation.

3.2. Evaluation of the Susceptibility in Gully Erosion

The results of the goodness-of-fit are shown in Table 5. The results did not show considerable variation in the accuracy of the model, with altering the percentage of calibration to validation samples and number of model replications.

Table 5. Forecasting efficiency of the model given 10 data sets.

MARS Model	70%/30%			80%/20%	90%/10%
	5 rep	10 rep	15 rep	10 rep	10 rep
Sensitivity	0.86	0.79	0.85	0.88	0.86
Specificity	0.72	0.81	0.66	0.83	0.72
(Negative predictive value)	0.70	0.78	0.72	0.74	0.75
(Positive predictive value)	0.83	0.73	0.82	0.85	0.84
Efficiency (%)	79.0	76.0	76.0	79.0	77.9
Kappa	0.58	0.51	0.52	0.58	0.58
AUC Mean	0.80	0.82	0.83	0.84	0.83
Robustness	0.03	0.08	0.11	0.01	0.15

The MARS algorithm performed excellently both in predictive performance. It can be observed that the MARS model for combination of 80%/20% had the highest performance in terms of SST (0.88), SPF (0.83), E% (0.79), Kappa (0.58), Robustness (0.01) and AUC (0.84) compared to the other combinations. The combination of 70%/30% with five replicates only had the highest performance in terms of E% (0.79). Figure 6 illustrates the mean ROC curves for the MARS model through 10 replicates. The results of AUC as a threshold-driven method for other scenarios are as follows: (70%/30% with 5, 10 and 15 replicates: 0.80, 0.82, 0.83 respectively; 90%/10%: 0.83). The MARS model revealed from acceptable to excellent performances, with AUC values above the 0.77 and 0.8 thresholds. By adopting

the J index, 0.386 probability cut-off values were generated for the susceptibility gully erosion model (Figure 6). Based on these cut-off values, the probability of gully erosion occurrence for each cell was adapted into a binary (positive/negative) prediction to achieve the spatial distribution of cases correctly categorized (true positives and negatives (TP, TN)) and incorrectly classified (false positives and negatives (FP, FN)) for the MARS susceptibility model (Table 6). Figure 7 shows the distribution of true positive (TP), false positive (FP), true negative (TN), and false-negative (FN) cases within the study area. A larger true positive prediction is produced by the model, indicating that the conditions for gully erosion are widespread.

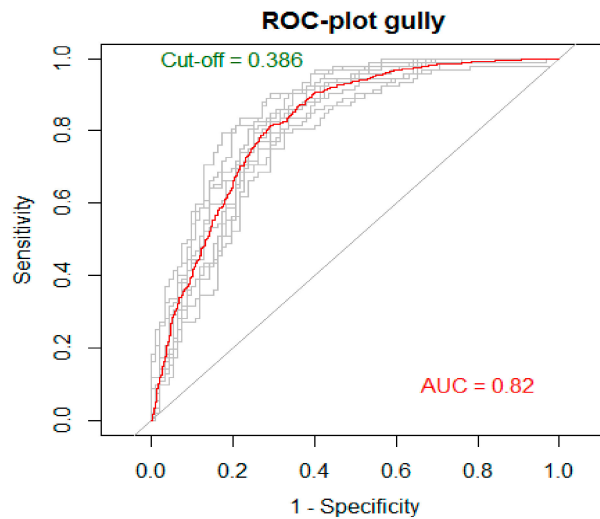


Figure 6. Average Receiver Operating Characteristic (ROC) curves for the MARS model through 10 replicates.

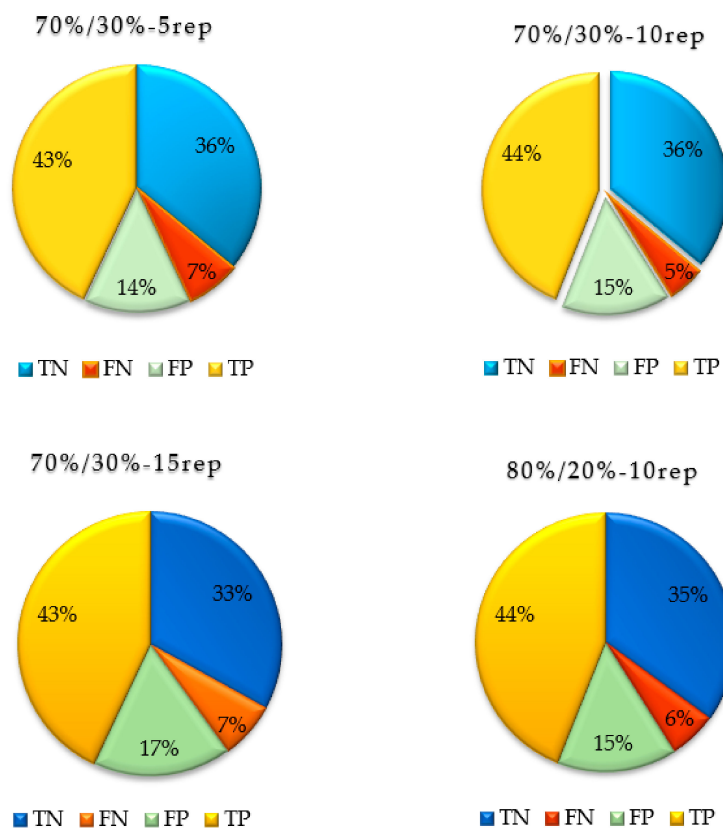


Figure 7. Contingency matrix applied for assessing the MARS model for all scenarios.

Table 6. Contingency matrix applied for assessment of the MARS model.

Observed	Predicted	
	%Gully (+)	%Non-gully (−)
Gully (+)	(+ +) 40% (TP)	(− +) 10% (FN)
Non-gully (−)	(+ −) 14% (FP)	(− −) 36% (TN)

4. Discussion

The results of the model (based on all sample data sets) show different ranges of susceptibility values of gullying. The output map was divided into four classes of poor, high, and very high for each model by the Natural breaks classification approach [96,100,101]. As the regions of the area were categorized as high and very high, gully erosion susceptibility maps were complementary with the sections of the watershed with low slope and near to roads. For an average of gully susceptibility, more than 31% of the study area has a high (HGES) and very high sensitivity (VHGES). Approximately 23.63% of the research field was classified as moderate classes (MGES). A total of 44.72% of the cells in the study region classified into low susceptibility groups (LGES). On the other hand, the low density of gullies was observed in forest areas with high slopes. In the forest areas, roughness due to vegetation covering may lead to medium runoff factors on this area as well, hence, low degradation force of centralized flow [96]. Areas nearer to roads and streams, with sparse vegetation and higher drainage density than other areas, have more potential for gully occurrence. These findings are in line with [4]. The parts of the basin with low slope and near to roads classified as high and very high gully erosion susceptibility maps (HGES and VHGES), and high density of gully erosions occurred in these parts of the catchment. The distribution of gullies over lithological units by other studies observed that poorly sorted materials are favorable conditions for gully erosion development [102–104].

The MARS model precision for gully erosion susceptibility was assessed by both threshold-driven and non-threshold-driven methods as pointed out in the methodology section. To assess the robustness of the model's data sensitivity similar with references [8,22,35], 5, 10 and 15 sample data sets, (replicates) for 70%/30% sample size, were prepared through random selection of different data sets in the calibration and validation subsets. This method is in agreement with studies of Rotigliano et al. [36]. The MARS model reproduces non-linear relationships using several linear regressions. Hence, this allows MARS to generate models with a better fit to the training data set while maintaining high predictive power [5].

Given the accuracy, the MARS algorithm performed excellently in predictive performance. It can be observed that the MARS model for combination of 80%/20% had the highest performance in terms of SST (0.88), SPF (0.83), E% (0.79), Kappa (0.58), Robustness (0.01), and AUC (0.84) compared to the other combinations. Since the accuracy values are very similar for the all sample data sets, the MARS model was robust when the validation group changed [2]. As well, in the research of Rahmati et al. [29], three various classes of training samples were applied to accompany different machine learning models for predicting the susceptibility of gully erosion. Three different sample data sets (S1, S2, and S3), were randomly prepared to evaluate the robustness of the models. Their results illustrated accurate predictions. Additionally, it was found that performance of RF and RBF-SVM for modelling gully erosion occurrence is quite stable when the learning and validation samples are changed. Regarding forecasting efficiency, the results of AUC as a threshold-driven method for scenarios are as follows: (70%/30% with 5, 10 and 15 replicates: 0.80, 0.82, 0.83 respectively; 90%/10%: 0.83, 80%/20%: 0.84). The MARS model revealed from acceptable to excellent performances, with AUC values above the 0.77 and 0.8 thresholds [97]. This result demonstrated a strong agreement between the distribution of the existing gully erosion points and the final predictive susceptibility map. Additionally, the AUC values for all data sets are approximately similar and the modelling method can be considered as robust to changes in learning points. Our results similar and agree with studies of Gómez-Gutiérrez et al. [26], who also applied the MARS model to predict gully erosion occurrence, and obtained AUC in the range of 0.75–0.98. The other MARS application to gully erosion susceptibility evaluation was made

by Gómez-Gutiérrez et al. [27], who achieved a mean AUC of 0.826 and 0.859 in Spain and Sicily respectively. Accordingly, Conoscenti et al. [17] pointed out that even for the worst validation AUC value inferred from five datasets, MARS is much more than the best calibration AUC value calculated compared with the logistic regression (LR) model. In addition, Conoscenti et al. [102] evaluated gully erosion sensitivity in both surrounding farmed basins of Sicily (Italy) by using multivariable adaptive regression splines. Model assessment on the whole basins indicates the outstanding predictive performance of models. This finding supports our results. Based on cut-off values, the probability of gully erosion occurrence for each cell was adapted into a binary (positive/negative) prediction to achieve the spatial distribution of cases correctly categorized (true positives and negatives (TP, TN)) and incorrectly classified (false positives and negatives (FP, FN)) for the MARS susceptibility model (Table 6). A larger true positive prediction is produced by the model, indicating that the conditions for gully erosion are widespread. As Rahmati et al. [96] pointed out, random ordination of datasets is the main origin of uncertainty in spatial modelling. From the validation result, it is clear that the MARS model provided acceptable to excellent performance in predicting the probability of gully erosion occurrence based on independent and dependent assessment measures. In light of abovementioned results, it is obvious that the MARS model can be used as an efficient statistical model for the prediction gully erosion susceptibility map. This is in line with the other studies that applied this model to landslides and gully erosion susceptibility mapping [9,25–27,102]. This is a relevant issue to achieve sustainable land management where gully erosions must be restored when they have developed as a result of human mismanagement, and for this it is necessary to use nature-based solutions [6]. To achieve success in gully erosion control, the strategies must find a way to reduce the connectivity of the flows [7].

5. Conclusions

Recognition of a precise and calibrated model to alleviate errors in modelling gully erosion susceptibility and determining gully erosion susceptible areas is of great importance. The present study enriches the systematic assessment of the multivariate adaptive regression splines model (MARS) to model gully erosion susceptibility among others. The major concluding remarks might be written as below: the aforementioned MARS model not just confirmed superiority for either based on limits-independent and limits-driven approaches, at the same time led to precise forecasting while changing the sample dataset. Hence, it can be inferred that this model is superior to evaluate gully erosion susceptibility while research aims to generate an exact gully erosion susceptibility map (GESM) and to pave the way to offering information on the outstanding potential of predictors. Reconnaissance and gully alleviation controls are not cost-effective and at the same time are not time-effective, and hence, to develop comprehensive susceptibility forecasting system as per modelling seems to be a remarkable possibility. In the current research based on past studies [34] and multicollinearity tests, digital elevation model, aspect map, slope percent, curvature of profile, curvature of plan, land use (LU), soil texture, TWI, distance to streams, distance to roads, drainage density, annual rainfall, stream power index, relative slope position, lithological formation, K factor, melton ruggedness number, and topographic position index are significant factors that influence gully erosion in the study area. Additionally, in this study we tried to investigate gully erosion susceptibility through the MARS model and analyzed the performance and accuracy of this technique for zoning gully erosion. While changing training and validation spots, the precision of the forecasting model was determined as the consistency of outputs of the model in respect to model precision. Here, predictive performance and goodness of fit of the model was subjected to assessment using both threshold-driven and non-threshold-dependent methods. This validates the robustness as well as the effectiveness of this model. Zonation of the gully erosion susceptibility for that section of Gorganrood display regions with high hazard susceptibility as well as demonstrates a valid map, in which the result is useful to managers and stakeholders to recognize the most prone area gully erosion and dedicate inputs for soil protection actions in the best manner.

Author Contributions: N.J., A.K., Z.J., H.R.P. and C.C. designed the experiments, performed the experiments, analyzed the data, and contributed reagents/materials/analysis tools. N.J. wrote the paper, Writing—review and editing was done by A.K. and H.R.P.

Funding: This research received no external funding.

Acknowledgments: We are grateful to the Sari Agricultural Sciences and Natural Resources University for financial support (SANRU) and Department of Earth and Sea Sciences (DISTEM), University of Palermo.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lal, R. Offsetting global CO₂ emissions by restoration of degraded soils and intensification of world agriculture and forestry. *Land Degrad. Dev.* **2003**, *14*, 309–322. [[CrossRef](#)]
2. Kosmas, C.; Danalatos, N.; Cammeraat, L.H.; Chabart, M.; Diamantopoulos, J.; Farand, R.; Gutierrez, L.; Jacob, A.; Marques, H.; Martinez-Fernandez, J.; et al. The effect of land use on runoff and soil erosion rates under Mediterranean conditions. *Catena* **1997**, *29*, 45–59. [[CrossRef](#)]
3. Vandekerckhove, L.; Poesen, J.; Oostwoud Wijdenes, D.; Gyssels, G.; Beuselinck, L.; de Luna, E. Characteristics and controlling factors of bank gullies in two semi-arid mediterranean environments. *Geomorphology* **2000**, *33*, 37–58. [[CrossRef](#)]
4. Vanwalleghe, T.; Poesen, J.; Nachtergaele, J.; Verstraeten, G. Characteristics, controlling factors and importance of deep gullies under cropland on loess-derived soils. *Geomorphology* **2005**, *69*, 76–91. [[CrossRef](#)]
5. Chaplot, V.; Giboire, G.; Marchand, P.; Valentin, C. Dynamic modelling for linear erosion initiation and development under climate and land-use changes in northern Laos. *Catena* **2005**, *63*, 318–328. [[CrossRef](#)]
6. Li, Y. Gully erosion: Impacts, factors and control. *Catena* **2005**, *63*, 132–153.
7. Imeson, A.C.; Kwaad, F. Gully types and gully prediction. *Geogr. Tydschr.* **1980**, *14*, 430–441.
8. Angileri, S.E.; Conoscenti, C.; Hochschild, V.; Märker, M.; Rotigliano, E.; Agnesi, V. *Water Erosion Susceptibility Mapping by Applying Stochastic Gradient Treeboost to the Imera Meridionale River Basin (Sicily, Italy)*; Elsevier: Amsterdam, The Netherlands, 2016; Volume 262, ISBN 3909123864643.
9. Conforti, M.; Aucelli, P.P.C.; Robustelli, G.; Scarciglia, F. Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). *Nat. Hazards* **2011**, *56*, 881–898. [[CrossRef](#)]
10. Govers, G.; Giménez, R.; Van Oost, K. Rill erosion: Exploring the relationship between experiments, modelling and field observations. *Earth Sci. Rev.* **2007**, *84*, 87–102. [[CrossRef](#)]
11. Guzzetti, F.; Reichenbach, P.; Ardizzone, F.; Cardinali, M.; Galli, M. Estimating the quality of landslide susceptibility models. *Geomorphology* **2006**, *81*, 166–184. [[CrossRef](#)]
12. Carrara, A.; Pike, R.J. GIS technology and models for assessing landslide hazard and risk. *Geomorphology* **2008**, *3*, 257–260. [[CrossRef](#)]
13. Vergari, F.; Della Seta, M.; Del Monte, M.; Fredi, P.; Palmieri, E.L. Landslide susceptibility assessment in the Upper Orcia Valley (Southern Tuscany, Italy) through conditional analysis: A contribution to the unbiased selection of causal factors. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 1475. [[CrossRef](#)]
14. Pozdnoukhov, A.; Matasci, G.; Kanevski, M.; Purves, R.S. Spatio-temporal avalanche forecasting with Support Vector Machines. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 367–382. [[CrossRef](#)]
15. Costanzo, D.; Cappadonia, C.; Conoscenti, C.; Rotigliano, E. Exporting a Google Earth™ aided earth-flow susceptibility model: A test in central Sicily. *Nat. Hazards* **2012**, *61*, 103–114. [[CrossRef](#)]
16. Ballabio, C.; Sterlacchini, S. Support vector machines for landslide susceptibility mapping: The Staffora River Basin case study, Italy. *Math. Geosci.* **2012**, *44*, 47–70. [[CrossRef](#)]
17. Conoscenti, C.; Ciaccio, M.; Caraballo-Arias, N.A.; Gómez-Gutiérrez, Á.; Rotigliano, E.; Agnesi, V. Assessment of susceptibility to earth-flow landslide using logistic regression and multivariate adaptive regression splines: A case of the Belice River basin (western Sicily, Italy). *Geomorphology* **2015**, *242*, 49–64. [[CrossRef](#)]
18. Magliulo, P. Soil erosion susceptibility maps of the Janare Torrent Basin (Southern Italy). *J. Maps* **2010**, *6*, 435–447. [[CrossRef](#)]
19. Magliulo, P. Assessing the susceptibility to water-induced soil erosion using a geomorphological, bivariate statistics-based approach. *Environ. Earth Sci.* **2012**, *67*, 1801–1820. [[CrossRef](#)]

20. Conoscenti, C.; Agnesi, V.; Angileri, S.; Cappadonia, C.; Rotigliano, E.; Märker, M. A GIS-based approach for gully erosion susceptibility modelling: A test in Sicily, Italy. *Environ. Earth Sci.* **2013**, *70*, 1179–1195. [[CrossRef](#)]
21. Lucà, F.; Conforti, M.; Robustelli, G. Comparison of GIS-based gullying susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy. *Geomorphology* **2011**, *134*, 297–308. [[CrossRef](#)]
22. Conoscenti, C.; Angileri, S.; Cappadonia, C.; Rotigliano, E.; Agnesi, V.; Märker, M. Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). *Geomorphology* **2014**, *204*, 399–411. [[CrossRef](#)]
23. Geissen, V.; Kampichler, C.; López-de Llergo-Juárez, J.J.; Galindo-Acántara, A. Superficial and subterranean soil erosion in Tabasco, tropical Mexico: Development of a decision tree modeling approach. *Geoderma* **2007**, *139*, 277–287. [[CrossRef](#)]
24. Märker, M.; Pelacani, S.; Schröder, B. A functional entity approach to predict soil erosion processes in a small Plio-Pleistocene Mediterranean catchment in Northern Chianti, Italy. *Geomorphology* **2011**, *125*, 530–540. [[CrossRef](#)]
25. Gutiérrez, Á.G.; Schnabel, S.; Contador, J.F.L. Using and comparing two nonparametric methods (CART and MARS) to model the potential distribution of gullies. *Ecol. Model.* **2009**, *220*, 3630–3637. [[CrossRef](#)]
26. Gomez Gutierrez, A.; Schnabel, S.; Felicísimo, Á.M. Modelling the occurrence of gullies in rangelands of southwest Spain. *Earth Surf. Process. Landf. J. Br. Geomorphol. Res. Gr.* **2009**, *34*, 1894–1902. [[CrossRef](#)]
27. Gómez-Gutiérrez, Á.; Conoscenti, C.; Angileri, S.E.; Rotigliano, E.; Schnabel, S. Using topographical attributes to evaluate gully erosion proneness (susceptibility) in two mediterranean basins: Advantages and limitations. *Nat. Hazards* **2015**, *79*, 291–314. [[CrossRef](#)]
28. Zabihi, M.; Pourghasemi, H.R.; Motevalli, A.; Zakeri, M.A. Gully erosion modeling using GIS-based data mining techniques in Northern Iran: A comparison Between Boosted Regression Tree and Multivariate Adaptive Regression Spline. In *Natural Hazards GIS-Based Spatial Modeling Using Data Mining Techniques*; Springer: Cham, Switzerland, 2019; pp. 1–26.
29. Rahmati, O.; Tahmasebipour, N.; Haghizadeh, A.; Pourghasemi, H.R.; Feizizadeh, B. Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion. *Geomorphology* **2017**, *298*, 118–137. [[CrossRef](#)]
30. Nazari Samani, A.; Ahmadi, H.; Mohammadi, A.; Ghoddousi, J.; Salajegheh, A.; Boggs, G.; Pishyar, R. Factors controlling gully advancement and models evaluation (Hableh Rood Basin, Iran). *Water Resour. Manag.* **2010**, *24*, 1531–1549. [[CrossRef](#)]
31. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67. [[CrossRef](#)]
32. Water Resources Company of Golestan (WRCG). Precipitation and Temperature Reports. Available online: <http://www.gsrw.ir/Default.aspx> (accessed on 15 May 2013).
33. Lee, S.; Ryu, J.-H.; Kim, I.-S. Landslide susceptibility analysis and its verification using likelihood ratio, logistic regression, and artificial neural network models: Case study of Youngin, Korea. *Landslides* **2007**, *4*, 327–338. [[CrossRef](#)]
34. Pradhan, B. Flood Susceptible Mapping and Risk Area Delineation Using Logistic Regression, GIS and Remote Sensing. *J. Spat. Hydrol.* **2009**, *9*, 1–18.
35. Cama, M.; Lombardo, L.; Conoscenti, C.; Rotigliano, E. Improving transferability strategies for debris flow susceptibility assessment: Application to the Saponara and Itala catchments (Messina, Italy). *Geomorphology* **2017**, *288*, 52–65. [[CrossRef](#)]
36. Rotigliano, E.; Martinello, C.; Agnesi, V.; Conoscenti, C. Evaluation of debris flow susceptibility in El Salvador (CA): A comparison between Multivariate Adaptive Regression Splines (MARS) and Binary Logistic Regression (BLR). *Hung. Geogr. Bull.* **2018**, *67*, 361–373. [[CrossRef](#)]
37. Conoscenti, C.; Rotigliano, E.; Cama, M.; Caraballo-Arias, N.A.; Lombardo, L.; Agnesi, V. *Exploring the Effect of Absence Selection on Landslide Susceptibility Models: A Case Study in Sicily, Italy*; Elsevier: Amsterdam, The Netherlands, 2016; Volume 261, ISBN 3909123864.
38. Kia, M.B.; Pirasteh, S.; Pradhan, B.; Mahmud, A.R.; Sulaiman, W.N.A.; Moradi, A. An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environ. Earth Sci.* **2012**, *67*, 251–264. [[CrossRef](#)]

39. Lee, M.J.; Kang, J.; Jeon, S. Application of Frequency Ratio Model and Validation for Predictive Flooded Area Susceptibility Mapping Using GIS. In Proceedings of the 2012 IEEE international Geoscience and Remote Sensing Symposium; IEEE: Piscataway, NJ, USA, 2012; pp. 895–898.
40. Youssef, A.M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Al-Katheeri, M.M. Erratum to: Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia (Landslides, 10.10. *Landslides* **2016**, *13*, 1315–1318. [[CrossRef](#)]
41. Jaafari, A.; Najafi, A.; Pourghasemi, H.R.; Rezaeian, J.; Sattarian, A. GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. *Int. J. Environ. Sci. Technol.* **2014**, *11*, 909–926. [[CrossRef](#)]
42. Jiménez-Perálvarez, J.D.; Irigaray, C.; El Hamdouni, R.; Chacón, J. Landslide-susceptibility mapping in a semi-arid mountain environment: An example from the southern slopes of Sierra Nevada (Granada, Spain). *Bull. Eng. Geol. Environ.* **2011**, *70*, 265–277. [[CrossRef](#)]
43. Nagarajan, R.; Roy, A.; Kumar, R.V.; Mukherjee, A.; Khire, M. V Landslide hazard susceptibility mapping based on terrain and climatic factors for tropical monsoon regions. *Bull. Eng. Geol. Environ.* **2000**, *58*, 275–287. [[CrossRef](#)]
44. Gallardo-Cruz, J.A.; Pérez-García, E.A.; Meave, J.A. β -Diversity and vegetation structure as influenced by slope aspect and altitude in a seasonally dry tropical landscape. *Landsc. Ecol.* **2009**, *24*, 473–482. [[CrossRef](#)]
45. Geroy, I.J.; Gribb, M.M.; Marshall, H.-P.; Chandler, D.G.; Benner, S.G.; McNamara, J.P. Aspect influences on soil water retention and storage. *Hydrol. Process.* **2011**, *25*, 3836–3842. [[CrossRef](#)]
46. Kornejady, A.; Ownegh, M.; Bahremand, A. Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. *Catena* **2017**, *152*, 144–162. [[CrossRef](#)]
47. Kornejady, A.; Ownegh, M.; Rahmati, O.; Bahremand, A. Landslide susceptibility assessment using three bivariate models considering the new topo-hydrological factor: HAND. *Geocarto Int.* **2018**, *33*, 1155–1185. [[CrossRef](#)]
48. Ercanoglu, M.; Gokceoglu, C. Assessment of landslide susceptibility for a landslide-prone area (north of Yenice, NW Turkey) by fuzzy approach. *Environ. Geol.* **2002**, *41*, 720–730.
49. Sidle, R.C.; Ochiai, H. Landslides: Processes, Prediction, and Land Use. In *Water Resources Monograph*; American Geophysical Union: Washington, DC, USA, 2006; Volume 18, p. 307. [[CrossRef](#)]
50. Yalcin, A. GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): Comparisons of results and confirmations. *Catena* **2008**, *72*, 1–12. [[CrossRef](#)]
51. Vahidnia, M.H.; Alesheikh, A.A.; Alimohammadi, A.; Hosseinali, F. A GIS-based neuro-fuzzy procedure for integrating knowledge and data in landslide susceptibility mapping. *Comput. Geosci.* **2010**, *36*, 1101–1114. [[CrossRef](#)]
52. Meinhardt, M.; Fink, M.; Tünschel, H. Landslide susceptibility analysis in central Vietnam based on an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. *Geomorphology* **2015**, *234*, 80–97. [[CrossRef](#)]
53. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J. Hydrol.* **2014**, *512*, 332–343. [[CrossRef](#)]
54. Tehrany, M.S.; Lee, M.-J.; Pradhan, B.; Jebur, M.N.; Lee, S. Flood susceptibility mapping using integrated bivariate and multivariate statistical models. *Environ. Earth Sci.* **2014**, *72*, 4001–4015. [[CrossRef](#)]
55. Khosravi, K.; Nohani, E.; Maroufinia, E.; Pourghasemi, H.R. A GIS-based flood susceptibility assessment and its mapping in Iran: A comparison between frequency ratio and weights-of-evidence bivariate statistical models with multi-criteria decision-making technique. *Nat. Hazards* **2016**, *83*, 947–987. [[CrossRef](#)]
56. Jenness, J. *DEM Surface Tools*; Jenness Enterp.: Flagstaff, AZ, USA, 2013. Available online: http://www.jennessent.com/arcgis/surface_area.htm (accessed on 20 May 2013).
57. Maestre, F.T.; Cortina, J. Spatial patterns of surface soil properties and vegetation in a Mediterranean semi-arid steppe. *Plant Soil* **2002**, *241*, 279–291. [[CrossRef](#)]
58. Zucca, C.; Canu, A.; Della Peruta, R. Effects of land use and landscape on spatial distribution and morphological features of gullies in an agropastoral area in Sardinia (Italy). *Catena* **2006**, *68*, 87–95. [[CrossRef](#)]
59. Cosby, B.J.; Hornberger, G.M.; Clapp, R.B.; Ginn, T. A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.* **1984**, *20*, 682–690. [[CrossRef](#)]

60. Gyssels, G.; Poesen, J.; Nachtergaele, J.; Govers, G. The impact of sowing density of small grains on rill and ephemeral gully erosion in concentrated flow zones. *Soil Tillage Res.* **2002**, *64*, 189–201. [[CrossRef](#)]
61. Vandekerckhove, L.; Poesen, J.; Govers, G. Medium-term gully headcut retreat rates in Southeast Spain determined from aerial photographs and ground measurements. *Catena* **2003**, *50*, 329–352. [[CrossRef](#)]
62. Moore, I.D.; Grayson, R.B.; Ladson, A.R. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 3–30. [[CrossRef](#)]
63. Grabs, T.; Seibert, J.; Bishop, K.; Laudon, H. Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model. *J. Hydrol.* **2009**, *373*, 15–23. [[CrossRef](#)]
64. Jungerius, P.D.; Matundura, J.; Van De Ancker, J.A.M. Road construction and gully erosion in West Pokot, Kenya. *Earth Surf. Process. Landf.* **2002**, *27*, 1237–1247. [[CrossRef](#)]
65. Pourghasemi, H.R.; Moradi, H.R.; Fatemi Aghda, S.M. Landslide susceptibility mapping by binary logistic regression, analytical hierarchy process, and statistical index models and assessment of their performances. *Nat. Hazards* **2013**, *69*, 749–779. [[CrossRef](#)]
66. Pourtaghi, Z.S.; Pourghasemi, H.R. GIS-based groundwater spring potential assessment and mapping in the Birjand Township, southern Khorasan Province, Iran. *Hydrogeol. J.* **2014**, *22*, 643–662. [[CrossRef](#)]
67. Bui, D.T.; Pradhan, B.; Lofman, O.; Revhaug, I.; Dick, O.B. Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): A comparative assessment of the efficacy of evidential belief functions and fuzzy logic models. *Catena* **2012**, *96*, 28–40.
68. Nefeslioglu, H.A.; Gokceoglu, C.; Sonmez, H. An assessment on the use of logistic regression and artificial neural networks with different sampling strategies for the preparation of landslide susceptibility maps. *Eng. Geol.* **2008**, *97*, 171–191. [[CrossRef](#)]
69. Kakembo, V.; Xanga, W.W.; Rowntree, K. Topographic thresholds in gully development on the hillslopes of communal areas in Ngqushwa Local Municipality, Eastern Cape, South Africa. *Geomorphology* **2009**, *110*, 188–194. [[CrossRef](#)]
70. Böhner, J.; Selige, T. Spatial Prediction of Soil Attributes Using Terrain Analysis and Climate Regionalisation. In *SAGA—Analysis and Modelling Applications*; Verlag Erich Goltze GmbH: Göttingen, Germany, 2006; Volume 115.
71. Song, Y.; Gong, J.; Gao, S.; Wang, D.; Cui, T.; Li, Y.; Wei, B. Susceptibility assessment of earthquake-induced landslides using Bayesian network: A case study in Beichuan, China. *Comput. Geosci.* **2012**, *42*, 189–199. [[CrossRef](#)]
72. Zhu, A.-X.; Wang, R.; Qiao, J.; Qin, C.-Z.; Chen, Y.; Liu, J.; Du, F.; Lin, Y.; Zhu, T. An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic. *Geomorphology* **2014**, *214*, 128–138. [[CrossRef](#)]
73. Vaezi, A.R.; Sadeghi, S.H.R.; Bahrami, H.A.; Mahdian, M.H. Modeling the USLE K-factor for calcareous soils in northwestern Iran. *Geomorphology* **2008**, *97*, 414–423. [[CrossRef](#)]
74. O’Callaghan, J.F.; Mark, D.M. The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **1984**, *28*, 323–344. [[CrossRef](#)]
75. Melton, M.A. The geomorphic and paleoclimatic significance of alluvial deposits in southern Arizona. *J. Geol.* **1965**, *73*, 1–38. [[CrossRef](#)]
76. Marchi, L.; Dalla Fontana, G. GIS morphometric indicators for the analysis of sediment dynamics in mountain basins. *Environ. Geol.* **2005**, *48*, 218–228. [[CrossRef](#)]
77. De Reu, J.; Bourgeois, J.; Bats, M.; Zwertvaegher, A.; Gelorini, V.; De Smedt, P.; Chu, W.; Antrop, M.; De Maeyer, P.; Finke, P. Application of the topographic position index to heterogeneous landscapes. *Geomorphology* **2013**, *186*, 39–49. [[CrossRef](#)]
78. Daoud, J.I. Multicollinearity and Regression Analysis. *J. Phys. Confer. Ser.* **2017**, *949*, 12009. [[CrossRef](#)]
79. O’Brien, R.M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **2007**, *41*, 673–690. [[CrossRef](#)]
80. Ozdemir, A. Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey). *J. Hydrol.* **2011**, *405*, 123–136. [[CrossRef](#)]
81. Leathwick, J.R.; Rowe, D.; Richardson, J.; Elith, J.; Hastie, T. Using multivariate adaptive regression splines to predict the distributions of New Zealand’s freshwater diadromous fish. *Freshw. Biol.* **2005**, *50*, 2034–2052. [[CrossRef](#)]

82. Felicísimo, Á.M.; Cuartero, A.; Remondo, J.; Quirós, E. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. *Landslides* **2013**, *10*, 175–189. [[CrossRef](#)]
83. Milborrow, S.; Hastie, T.; Tibshirani, R. Earth: Multivariate Adaptive Regression Spline Models; R Software Package. 2019. Available online: <http://www.milbo.users.sonic.net/earth> (accessed on 12 April 2019).
84. Demšar, J.; Curk, T.; Erjavec, A.; Gorup, Č.; Hočevar, T.; Milutinovič, M.; Možina, M.; Polajnar, M.; Toplak, M.; Starič, A. Orange: Data mining toolbox in Python. *J. Mach. Learn. Res.* **2013**, *14*, 2349–2353.
85. Galili, T. dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics* **2015**, *31*, 3718–3720. [[CrossRef](#)]
86. Lombardo, L.; Cama, M.; Conoscenti, C.; Märker, M.; Rotigliano, E. Binary logistic regression versus stochastic gradient boosted decision trees in assessing landslide susceptibility for multiple-occurring landslide events: Application to the 2009 storm event in Messina (Sicily, southern Italy). *Nat. Hazards* **2015**, *79*, 1621–1648. [[CrossRef](#)]
87. Umar, Z.; Pradhan, B.; Ahmad, A.; Jebur, M.N.; Tehrany, M.S. Earthquake induced landslide susceptibility mapping using an integrated ensemble frequency ratio and logistic regression models in West Sumatera Province, Indonesia. *Catena* **2014**, *118*, 124–135. [[CrossRef](#)]
88. Oh, H.-J.; Pradhan, B. Application of a neuro-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput. Geosci.* **2011**, *37*, 1264–1276. [[CrossRef](#)]
89. Mouton, A.M.; De Baets, B.; Goethals, P.L.M. Ecological relevance of performance criteria for species distribution models. *Ecol. Model.* **2010**, *221*, 1995–2002. [[CrossRef](#)]
90. Shirzadi, A.; Soliamani, K.; Habibnejhad, M.; Kavian, A.; Chapi, K.; Shahabi, H.; Chen, W.; Khosravi, K.; Thai Pham, B.; Pradhan, B. Novel GIS based machine learning algorithms for shallow landslide susceptibility mapping. *Sensors* **2018**, *18*, 3777. [[CrossRef](#)] [[PubMed](#)]
91. Saito, H.; Nakayama, D.; Matsuyama, H. Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. *Geomorphology* **2009**, *109*, 108–121. [[CrossRef](#)]
92. Kontijevskis, A.; Wikberg, J.E.S.; Komorowski, J. Computational proteomics analysis of HIV-1 protease interactome. *Proteins Struct. Funct. Bioinform.* **2007**, *68*, 305–312. [[CrossRef](#)] [[PubMed](#)]
93. Dai, Q.; Yang, Y.; Wang, T. Markov model plus k-word distributions: A synergy that produces novel statistical measures for sequence comparison. *Bioinformatics* **2008**, *24*, 2296–2302. [[CrossRef](#)] [[PubMed](#)]
94. Park, S.; Choi, C.; Kim, B.; Kim, J. Landslide susceptibility mapping using frequency ratio, analytic hierarchy process, logistic regression, and artificial neural network methods at the Inje area, Korea. *Environ. Earth Sci.* **2013**, *68*, 1443–1464. [[CrossRef](#)]
95. Naghibi, S.A.; Dashtpajardi, M.M. Evaluation of four supervised learning methods for groundwater spring potential mapping in Khalkhal region (Iran) using GIS-based features. *Hydrogeol. J.* **2017**, *25*, 169–189. [[CrossRef](#)]
96. Rahmati, O.; Naghibi, S.A.; Shahabi, H.; Bui, D.T.; Pradhan, B.; Azareh, A.; Rafiei-Sardooi, E.; Samani, A.N.; Melesse, A.M. Groundwater spring potential modelling: Comprising the capability and robustness of three different modeling approaches. *J. Hydrol.* **2018**, *565*, 248–261. [[CrossRef](#)]
97. Hosmer, D.W.; Lemeshow, S. Wiley Series in Probability and Statistics. Applied Logistic Regression. 1989. Available online: <https://www.worldcat.org/title/applied-logistic-regression/oclc/19514573> (accessed on 8 June 1989).
98. Pourghasemi, H.R.; Yousefi, S.; Kornejady, A.; Cerdà, A. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci. Total Environ.* **2017**, *609*, 764–775. [[CrossRef](#)]
99. Youden, W.J. Index for rating diagnostic tests. *Cancer* **1950**, *3*, 32–35. [[CrossRef](#)]
100. Ayalew, L.; Yamagishi, H. The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* **2005**, *65*, 15–31. [[CrossRef](#)]
101. Akgun, A.; Dag, S.; Bulut, F. Landslide susceptibility mapping for a landslide-prone area (Findikli, NE of Turkey) by likelihood-frequency ratio and weighted linear combination models. *Environ. Geol.* **2008**, *54*, 1127–1143. [[CrossRef](#)]
102. Conoscenti, C.; Agnesi, V.; Cama, M.; Caraballo-Arias, N.A.; Rotigliano, E. Assessment of Gully Erosion Susceptibility Using Multivariate Adaptive Regression Splines and Accounting for Terrain Connectivity. *Land Degrad. Dev.* **2018**, *29*, 724–736. [[CrossRef](#)]

103. Poesen, J.; Nachtergaele, J.; Verstraeten, G.; Valentin, C. Gully erosion and environmental change: Importance and research needs. *Catena* **2003**, *50*, 91–133. [[CrossRef](#)]
104. Dewitte, O.; Daoudi, M.; Bosco, C.; Van Den Eeckhaut, M. Predicting the susceptibility to gully initiation in data-poor regions. *Geomorphology* **2015**, *228*, 101–115. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).