



OPEN

## A multi-hazard map-based flooding, gully erosion, forest fires, and earthquakes in Iran

Soheila Pouyan<sup>1</sup>, Hamid Reza Pourghasemi<sup>1✉</sup>, Mojgan Bordbar<sup>2</sup>, Soroor Rahmanian<sup>3</sup> & John J. Clague<sup>4</sup>

We used three state-of-the-art machine learning techniques (boosted regression tree, random forest, and support vector machine) to produce a multi-hazard (MHR) map illustrating areas susceptible to flooding, gully erosion, forest fires, and earthquakes in Kohgiluyeh and Boyer-Ahmad Province, Iran. The earthquake hazard map was derived from a probabilistic seismic hazard analysis. The mean decrease Gini (MDG) method was implemented to determine the relative importance of effective factors on the spatial occurrence of each of the four hazards. Area under the curve (AUC) plots, based on a validation dataset, were created for the maps generated using the three algorithms to compare the results. The random forest model had the highest predictive accuracy, with AUC values of 0.994, 0.982, and 0.885 for gully erosion, flooding, and forest fires, respectively. Approximately 41%, 40%, 28%, and 3% of the study area are at risk of forest fires, earthquakes, floods, and gully erosion, respectively.

Hazard can be defined as a source of potential harm, a threat to natural environments, and human health. The causes of natural hazards may be geological (e.g., earthquakes, tsunamis, landslides, volcanic eruptions, etc.) and climatic (e.g., floods, windstorms, droughts, and wildfires)<sup>1</sup>. Many parts of the world are at risk from one or more natural hazards, and although many studies have focused on single hazards, there is a need for integrated assessments of multi-hazards for more efficient land management<sup>2</sup>. The concept of multi-hazards was introduced by the United Nations Environment Programme through its policies on sustainable development and its call for "comprehensive investigation of multi-hazards" to plan and manage residential areas prone to natural disasters<sup>3</sup>. Damage from natural disasters is increasing worldwide, providing an impetus to hazard researchers to develop new tools to reduce economic losses and injuries from future disasters. These tools include multi-hazard maps created using machine learning tools that show the spatial distribution of the full spectrum of hazards in a region<sup>2</sup>. Examples of areas where such multi-hazard evaluations have been performed include Greece (flooding depth, lateral erosion, earthquakes, and landslides)<sup>4</sup>, the United States (weather and climate hazards)<sup>5</sup>, Sikkim State, India (landslide and earthquake)<sup>6</sup>, the Adriatic Sea (smothering and sealing, abrasion and extraction, underwater noise, sea surface temperature variation, and sea surface salinity variation)<sup>7</sup>, and Chile (weather-related hazards, including coastal flooding, fluvial flooding, water scarcity, heat stress, and wildfire)<sup>8</sup>. The interaction of these multi-hazards could be useful for further research in this field. Previous studies have shown that landslides and flash floods are natural hazards that are frequently triggered simultaneously due to heavy or prolonged rainfall on steep mountains<sup>9</sup>. In other words, heavy rainfall causes flash floods that can lead to soil erosion and landslide events<sup>10</sup>. For instance, in April 2019, heavy rainfall had a significant effect on the whole of Iran, which caused far-reaching flooding and landslides<sup>11</sup>. Furthermore, landslides lead to seismic derangement. In addition, seismic derangement may cause landslides, leading to many victims around the world. Slides can cause catastrophic flooding, especially when landslide dams are broken, and flooding can cause slides<sup>12</sup>. Temporary soil flooding at different scales can significantly affect soil degradation. Floods on slopes in ground flow, sheet flow, return flow, groundwater ridge, etc., are connected to soil erosion and landslide occurrence. Floodwater combined with saturated status may destroy soil structure and soil organisms (<https://www.recare-hub.eu/soil-threats/floods-and-landslides>). Moreover, forest fires can significantly alter vegetation, increase soil erosion, and

<sup>1</sup>Department of Natural Resources and Environmental Engineering, College of Agriculture, Shiraz University, Shiraz 71441-65186, Iran. <sup>2</sup>Department of GIS/RS, Faculty of Natural Resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran. <sup>3</sup>Quantitative Plant Ecology and Biodiversity Research Lab, Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran. <sup>4</sup>Department of Earth Sciences, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6, Canada. ✉email: hr.pourghasemi@shirazu.ac.ir

even lead to desertification<sup>13</sup>. Admittedly, forest fires significantly impact physical, chemical, mineralogical, and biological soil features, which trigger soil erosion. Severe bushfires (i.e., wildfires) harm soil properties<sup>14</sup>. On the other hand, gully erosion influences soil structure, quality, and soil properties<sup>15</sup>.

Much of Iran is at high risk of earthquakes, floods, forest fires, and gully erosion<sup>16–18</sup>. Forest fires, for example, are a major cause of ecosystem damage and large economic losses<sup>19,20</sup>. The forests of Kohgiluyeh and Boyer-Ahmad Province in southwestern Iran are particularly vulnerable to forest fires, providing a motivation for the choice of this area for our study. Maps that show areas at high risk from wildfires help managers and planners to deal with this problem<sup>21,22</sup>.

Flooding is another serious hazard in Kohgiluyeh and Boyer-Ahmad Province and has been exacerbated in recent decades by deforestation, land use changes, poor watershed management, and climate change<sup>23</sup>. Floods have adversely affected more people than any other natural disaster type in the twenty-first century, including 127 events of differentiated natural disasters in 2018<sup>24</sup>. One such flood caused widespread damage in Iran in March and April 2019. Recurrent floods in Kohgiluyeh and Boyer-Ahmad Province inundated riverside areas, leading to soil erosion, damaged or destroyed earth and concrete dams, other watershed structures, and bridges. Identifying areas at risk<sup>25</sup> and preparing flood hazard maps are important activities in any proactive response to this hazard<sup>26–29</sup>.

Gully erosion is an important cause of soil loss and land degradation in arid and semi-arid regions in Iran and elsewhere world<sup>30,31</sup>. Maps showing where this problem exists or might develop in the future are effective tools for decision makers concerned with sustainable development. In many instances, gully erosion maps have been prepared individually based on datasets in GIS software<sup>5,32,33</sup>. In the present study, gully erosion is considered an element of the multi-hazard spectrum hazards in the Kohgiluyeh and Boyer-Ahmad Province.

The high Zagros Mountains throughout the north, east, and southeast of Kohgiluyeh and Boyer-Ahmad Province is located, experience frequent earthquakes. Although it is not possible to accurately predict future earthquakes in the province, it is possible to determine the likely locations that will suffer damaging ground motions.

Our multi-hazard study in Kohgiluyeh and Boyer-Ahmad Province involved (1) identifying and prioritizing the factors affecting the occurrence of each natural hazard in the region; (2) preparing maps of flood, gully erosion, forest fires, and earthquake hazards separately using three machine learning algorithms (boosted regression tree, support vector machine, and random forest); (3) comparing the results of the three algorithms with the ROC curve to select the best regional hazard model; and (4) preparing a multi-hazard map of the province. Our research contributes to the development and assessment of machine learning methods for mapping natural hazard zones. To the authors' knowledge, no work in the literature related to the multi-hazard spatial modeling of floods, gully erosion, forest fires, and earthquakes exists to date. Moreover, for the first time, this work was carried out mentioned hazards mapping in the Kohgiluyeh and Boyer-Ahmed province. In addition, based on the available data and sources, the SVM, BRT, and RF algorithms were used to investigate the hazards in the study area.

## Results

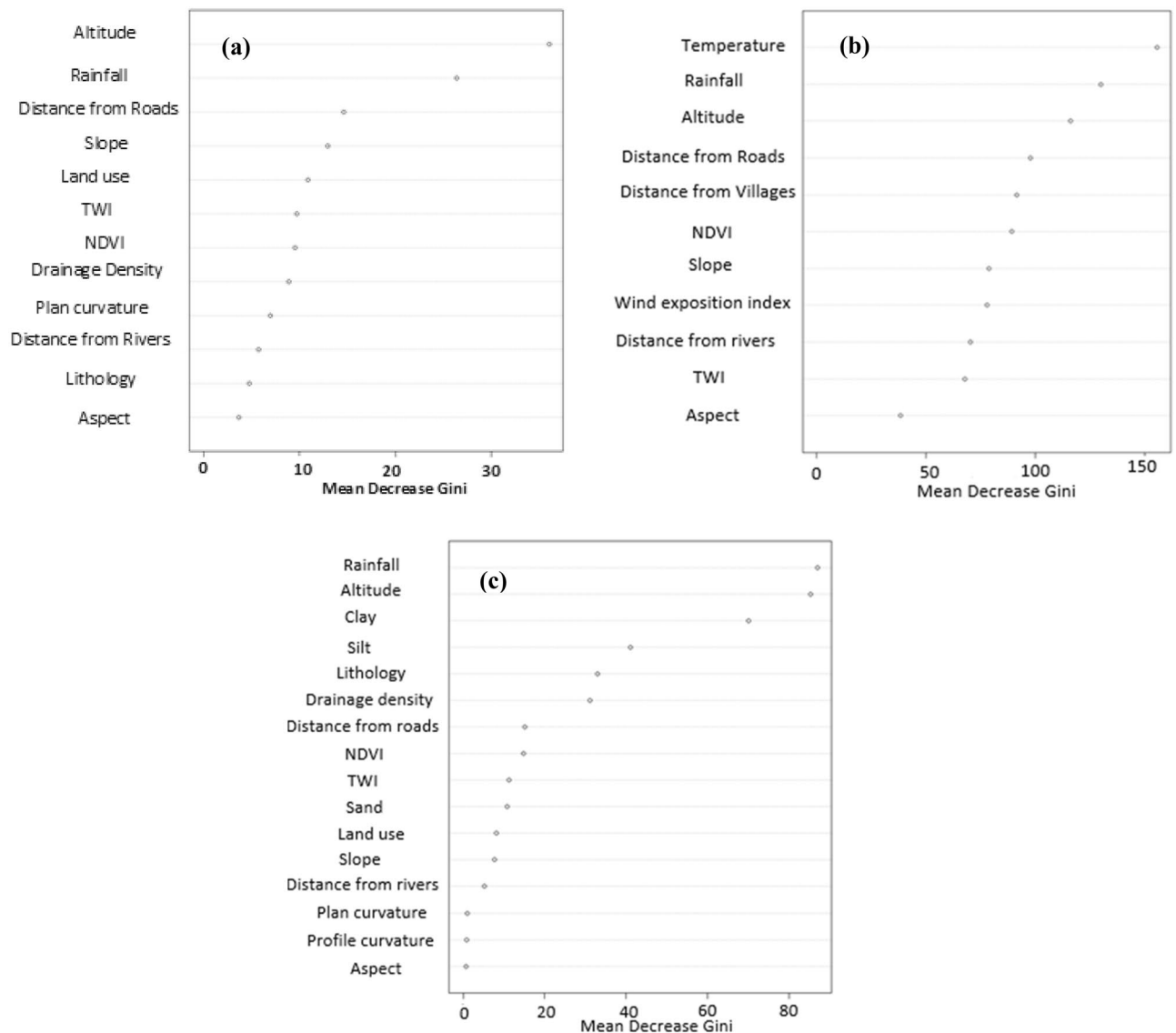
**Priority of effective factors using random forest and MDG.** Figure 1 shows the results of the prioritization of factors for the three hazards using the RF technique. The most important factor is elevation, followed by decreasing order of importance of flooding by mean annual rainfall, distance from roads, slope, land use, TWI, NDVI, drainage density, plan curvature, distance from rivers, lithology, and aspect (Fig. 1a).

Based on the MDG analysis, temperature is the most important factor for forest fire hazards, followed by mean annual rainfall, elevation, distance from roads, distance from villages, NDVI, slope, WEI, distance from rivers, TWI, and aspect (Fig. 1b). Mean annual rainfall is the most important factor for gully erosion, followed by elevation, clay, silt, lithology, drainage density, distance from roads, NDVI, TWI, sand, land use, slope, distance from rivers, plan curvature, profile curvature, and aspect (Fig. 1c).

**Natural hazard susceptibility maps.** In natural hazard studies, susceptibility is associated with the spatial features of hazards. It is defined as the tendency of a region to experience the effects of a given hazardous process (i.e., earthquakes, floods, forest fires, erosion, etc.) without considering the moment of event, fatality, and economic losses<sup>34</sup>. The RF, BRT, and SVM algorithms were applied to train the data, and susceptibility maps were prepared for flood, forest fire, and gully erosion hazards. Flood hazard (FH) maps were created by applying the RF (FHRF), SVM (FHSVM), and BRT (FHBRT) algorithms to the training dataset. Similarly, gully erosion hazard (GEH) maps were generated from the training dataset using the same three algorithms (maps GEHRF, GEHSVM, and GEHBRT), as were forest fire hazard (FFH) maps (FFHRF, FFHSVM, and FFHBRT).

Natural hazard maps prepared using the RF model (the best performance of the three models) are shown in Fig. 2. In the case of the FHRF map (Fig. 2a), approximately 50.1%, 22.1%, 15.7%, and 12.1% of the study area had low, moderate, high, and very high susceptibility to flooding, respectively. The FFHRF map (Fig. 2b) shows that approximately 17.4% and 24.0% of the study area are, respectively, in the high and very high forest fire susceptibility classes, and 30.3% and 28.3% of the area had moderate and low susceptibility, respectively. Finally, the GEHRF map (Fig. 2c) shows that 1.1% and 2.1% of the area has very high and high susceptibility to gully erosion, respectively; the moderate and low susceptibility classes cover 4.4% and 92.4% of the province, respectively.

**Peak ground acceleration map.** Figure 3 shows the PGA map of the study area. The PGA values were divided into three classes: low, moderate, and high. The PGA map shows that approximately 40%, 20.0%, and 39.6% of the area were in the low, moderate, and high classes, respectively.



**Figure 1.** Priority of the effective factors for each hazard.

**Validation and comparison of hazard maps using ROC.** We validated the hazard maps using ROC (Table 1). In the case of forest fire hazard, the random forest model performed best ( $AUC=0.885$ ), followed by the support vector machine ( $AUC=0.727$ ) and boosted regression tree ( $AUC=0.680$ ) models. For gully erosion, the AUC values for the random forest, support vector machine, and boosted regression tree maps are 0.994, 0.959, and 0.938, respectively. The AUC values for the flood hazard maps made using the random forest, support vector machine, and boosted regression tree algorithms are 0.982, 0.940, and 0.883, respectively (Table 1) (Fig. 4).

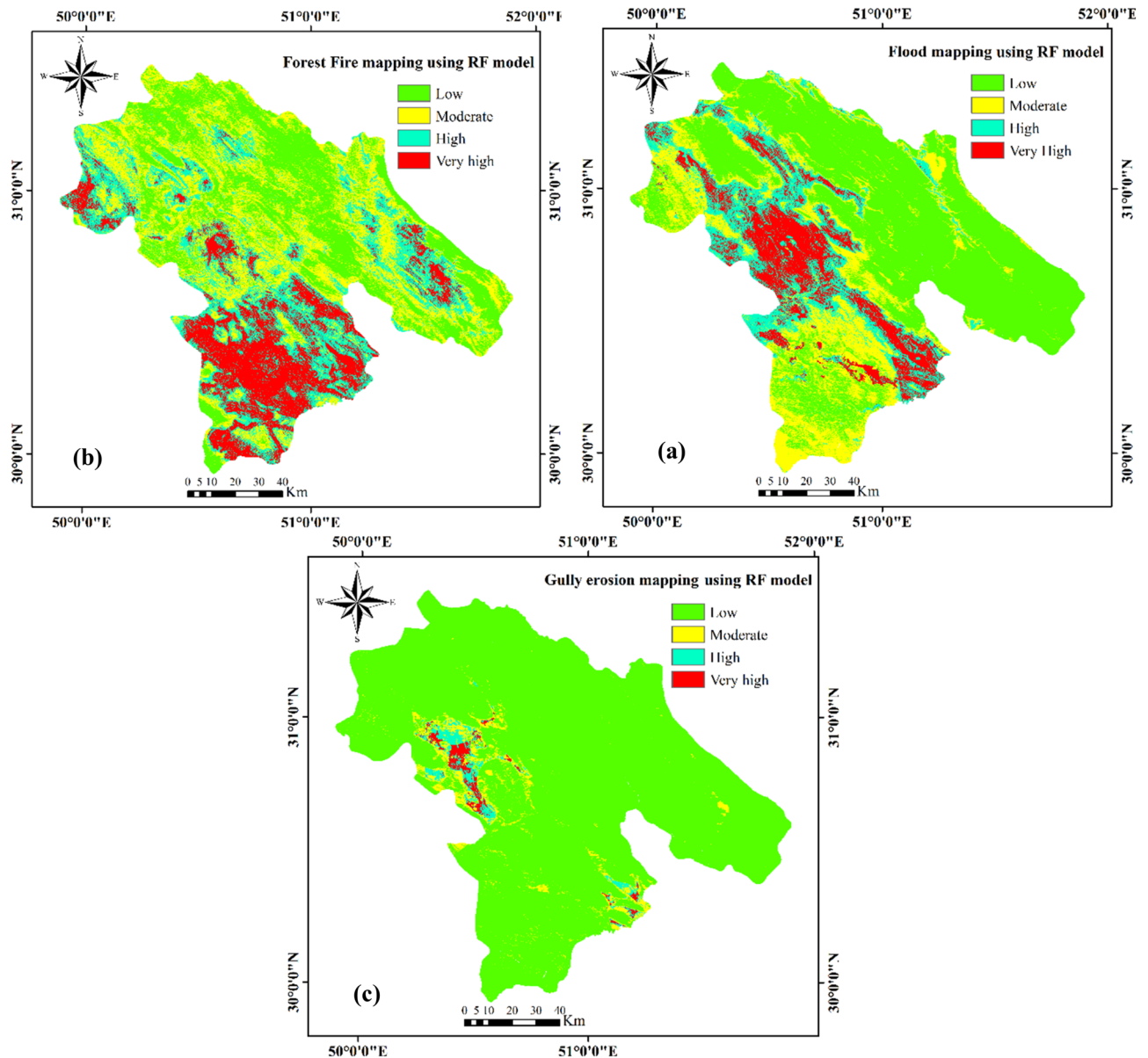
**Multi-hazard map.** The RF model was chosen to produce a multi-hazard map (MHR) of the study area (Fig. 4). We combined maps of the four hazards (flood, forest fire, gully erosion, and earthquake) in the ArcGIS 10.8 environment (<https://www.esri.com>) using the following equation:

$$MHR = FFHRF + FHRF + GEHRF + PGA \quad (1)$$

The multi-hazard map included 15 susceptibility classes (Fig. 5). Approximately 38% of the Kohgiluyeh and Boyer-Ahmad Province is safe from all-natural hazards, whereas 0.7% of the region is susceptible to all four natural hazards. Percentages of the study area with individual or combinations of hazards are shown in Figs. 6 and 7.

## Discussion

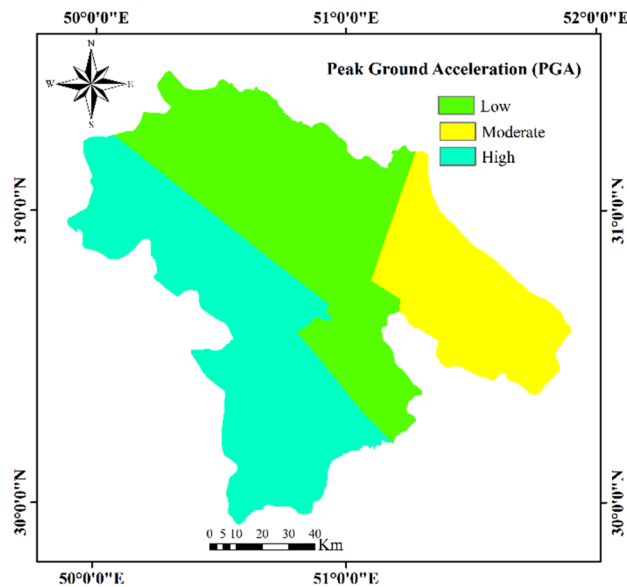
Many researchers have studied hazards individually<sup>35–39</sup>. Although such studies have been important and useful, considering multiple hazards and their interactions, their interdependencies, and possible cascading effects can be more informative and useful for reducing disaster losses, as well as for efficient land and ocean management<sup>4, 6, 40</sup>.



**Figure 2.** Hazard susceptibility maps of the study area produced using the RF model.

The motivation for our study was to provide (1) a comprehensive and integrated analysis of the full spectrum of damaging natural hazards in Kohgiluyeh and Boyer-Ahmad Province, which is a risk-prone area in Iran, and (2) a multi-hazard map useful to land-use and emergency response managers. Here, we discuss the factors that drive each hazard and identify areas within the province that are susceptible to each phenomenon. We also discuss the merits of the machine-learning algorithms used in this study. The benefits of machine learning models include the use of dichotomous dependent variables as probabilities, the ability to use different types of independent variables (i.e., binary, sequential, and continuous), and there is no need for normal distribution<sup>41</sup>. Machine learning methods can solve uncertainty factors related to the dataset modeling process. Different learning machine algorithms can also be used to solve uncertainties related to the accuracy of the models. Another source of uncertainty is the limitation of the training model, in which ML techniques such as RF do not face these issues. This algorithm uses the error rate and an indicator outside the bag indicator. One of the benefits accrued from machine learning techniques compared to traditional methods (i.e., bivariate and multivariate statistical methods) is that ML algorithms can deal with noise related to the dataset and the uncertainty of the dataset. Moreover, limited measurement errors are accurate<sup>42</sup>. However, the disadvantage of machine learning methods is their vulnerability to overfitting data, which produces unstable regression coefficients<sup>43</sup>. Therefore, it is necessary to use different techniques to improve the accuracy of prediction results.

Each of these algorithms has advantages, but the results of the AUC-based assessment showed that the RF algorithm has stronger predictive power than the BRT and SVM models. Other researchers have shown that (1) the RF algorithm performs better than conventional methods<sup>44</sup>, (2) is a powerful supervised learning method



**Figure 3.** PGA map.

Models	Forest fire	Flood area	Gully erosion
BRT	0.68	0.883	0.938
RF	0.885	0.982	0.994
SVM	0.727	0.94	0.959

**Table 1.** AUC values of hazard risk maps.

for investigating problems in the real world<sup>45</sup>, (3) is simple and fast, (4) does not require statistical hypotheses, and (5) is a reliable predictor<sup>46,47</sup>. It has been widely used in a range of environmental studies<sup>44,48–52</sup>.

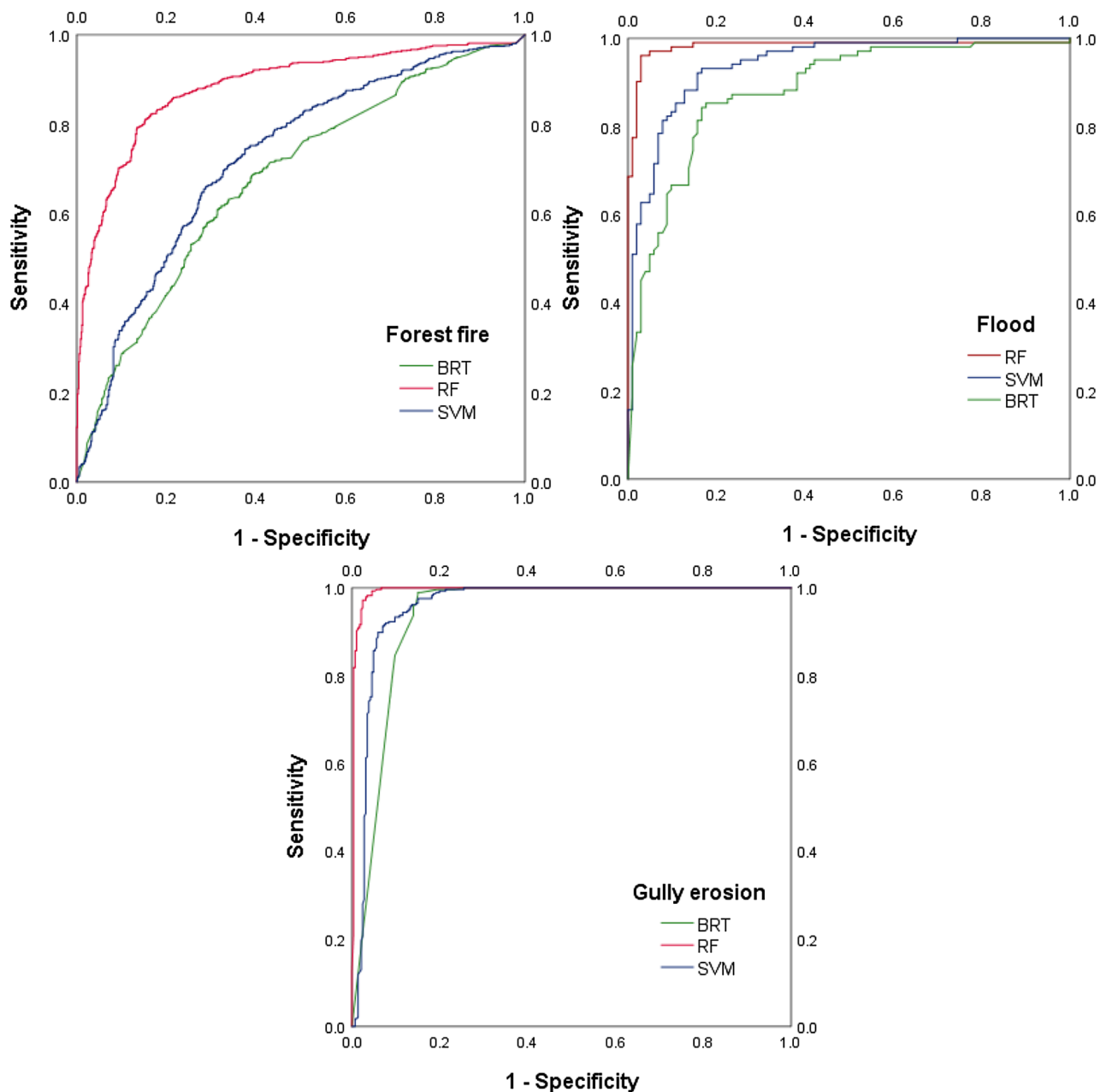
The FFHRF map shows that the southern, western, and eastern parts of the Kohgiluyeh and Boyer-Ahmad Province are most susceptible to forest fires, and the northern and central areas have low to moderate wildfire susceptibility. The FHRF map shows flood susceptibility to be high to very high in the western and southeastern parts of the province and low to moderate in the north and east. According to the GEHRF map, high to very high gully erosion susceptibility is restricted to a small area in the western part of the province. Finally, the PGA map shows that the southern and western parts of the province are highly susceptible to damaging earthquakes, whereas this hazard is low elsewhere.

In this study, we evaluated the relative importance of factors affecting flood, forest fire, and gully erosion using the random forest method. This method reliably determines the relative importance of controlling factors in hazard susceptibility applications<sup>53</sup>. Using the MDG method, we found that the annual temperature is the most important factor in predicting forest fire susceptibility, whereas elevation and mean annual precipitation are most important in the case of flood and gully erosion hazards. Aspect proved to be the least important factor for all three hazards. There is no specific aspect that favors forest fires; it occurs on all forested slopes in the study area. However, forest fires are common near residential areas and roads, where most fires are initiated. Elevation is also important because it has a strong effect on rainfall and temperature. Our map of areas susceptible to forest fires may be useful to forest managers and emergency responders who must plan and implement necessary measures to protect and preserve the remaining forest in high-risk areas.

Researchers used the RF method in a forestry study in China and, like us, concluded that vegetation type, slope, and aspect are less important than proximity to towns, temperature, and precipitation<sup>54</sup>. Another research group studied flood susceptibility along the Pearl River in China using the DMG method and concluded that the most important factors for related flooding are maximum three-day precipitation, runoff depth, typhoon frequency, elevation, and TWI<sup>55</sup>. In contrast, another group found that elevation is the most important flood risk factor in their study area<sup>56</sup>, given that flooding was greatest along trunk rivers at low elevation.

Mean annual rainfall is one of the most effective factors in the formation and spread of gullies. This metric is positively related to the volume and kinetic energy of surface runoff, which are the main factors involved in gully erosion<sup>57</sup>. Lithology is an important factor in this process. One research group working in Ilam Province, Iran, for example, found that the amount of clay in soil has a high positive correlation with gully erosion<sup>58</sup>.

In our analysis of earthquake hazards, we learned that magnitude and ground acceleration are the two most important factors for estimating seismic risk. This finding is consistent with the results of other studies that have performed similar studies<sup>40</sup>.



**Figure 4.** Validation of hazard maps using the ROC curve.

## Conclusion

We used machine learning algorithms to map areas susceptible to multiple hazards in a part of Iran with a high risk of natural hazards. Identifying these areas is an important step in making sound management decisions to reduce damage and injury from natural hazards in the future. The final product of our work is a multi-hazard map created by combining four important natural hazards (i.e., flooding, forest fires, earthquakes, and gully erosion). Factors affecting each hazard were identified and prioritized, and maps of each hazard were created using the RF, BRT, and SVM algorithms. We determined the validity of the results using ROC plots and found that the RF model had the highest AUC value and thus accuracy. We then produced a multi-hazard map for the study area by combining the maps of the four hazards. This final map shows that approximately 38% of the province is safe from all four hazards. Areas where all four hazards are a concern are restricted to the southern part of the province. Of the four hazards, forest fires affected the largest percentage of the study area and gully erosion was the least. In line with the goals of sustainable development, the results of this research can be used by managers, planners, and other stakeholders as a decision-making tool to reduce future damage from natural hazards. The critical infrastructure in Kohgiluyeh and Boyer-Ahmad Province should be examined with the aim of minimizing future losses from forest fires, floods, earthquakes, and gully erosion.

## Multi hazards mapping using RF model

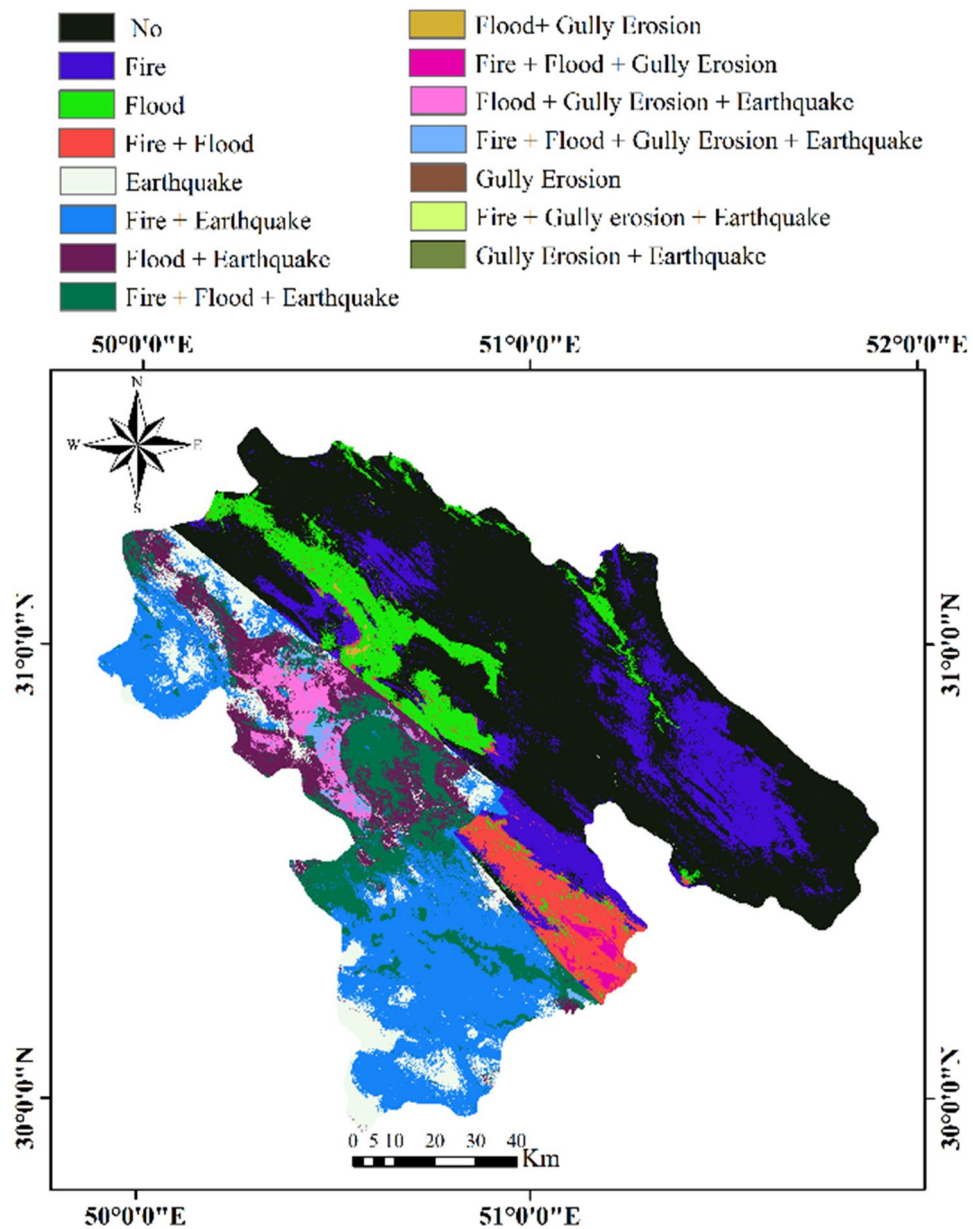


Figure 5. Multi-hazard map.

### Methodology

**Study area.** Kohgiluyeh and Boyer-Ahmad Province is a 15,500 km<sup>2</sup> area of a high mountainous region. Its borders share with Fars, Khuzestan, Isfahan, and Bushehr provinces within the Chaharmahal-Bakhtiari province, and the Zagros Mountains in southwest Iran (Fig. 8). It is located between 49° 53' 00" and 51° 53' 00" E longitudes and 29° 56' 00" and 31° 27' 00" N latitudes, and altitude from 109 to 4294 m, and has an average annual rainfall of 550.7 mm and the average temperature is between 15 and 26 °C. Plains constitute approximately 20% of province. The highest point in the province is Dena Mountain (4294 m above sea level); the lowest point is in the southwest, in the city of Gachsaran, (109 m above sea level). The difference in elevation between the more elevated northeastern part of the province and the southern and southwestern parts results in two different climatic regimes. The former region is cooler and drier than the latter<sup>59</sup>.

**Preparation of inventory maps.** Based on experience and expert opinion, we consider the four main hazards in the study area to be forest fires, flooding, earthquakes, and gully erosion. In this study, we identified the locations of these four hazard types using Google Earth images and a field survey supported by a global posi-

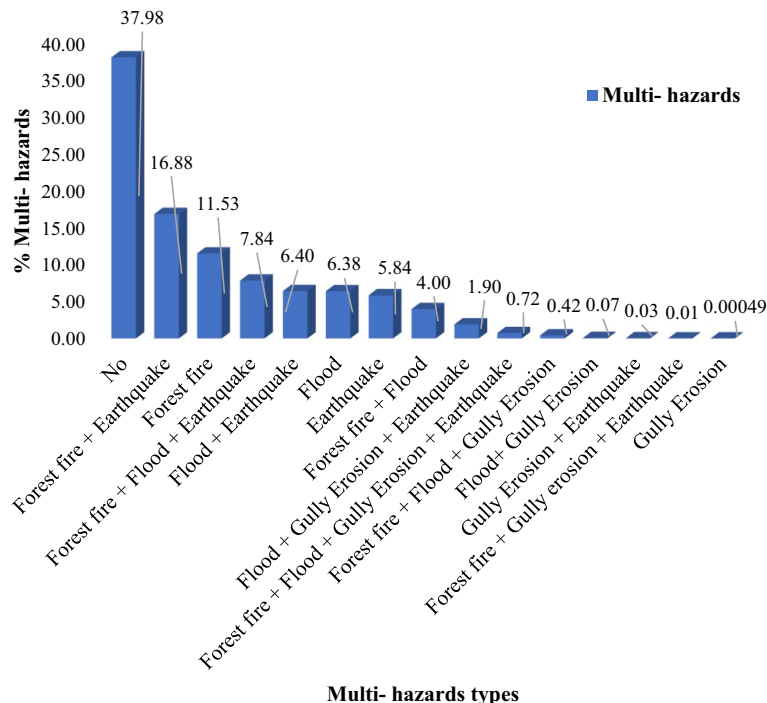


Figure 6. Percentages of areas of hazard classes.

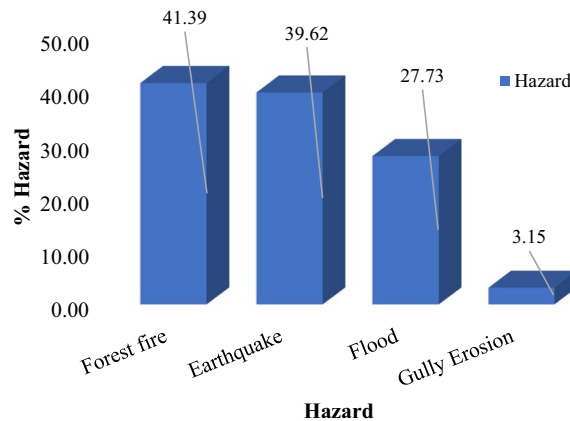
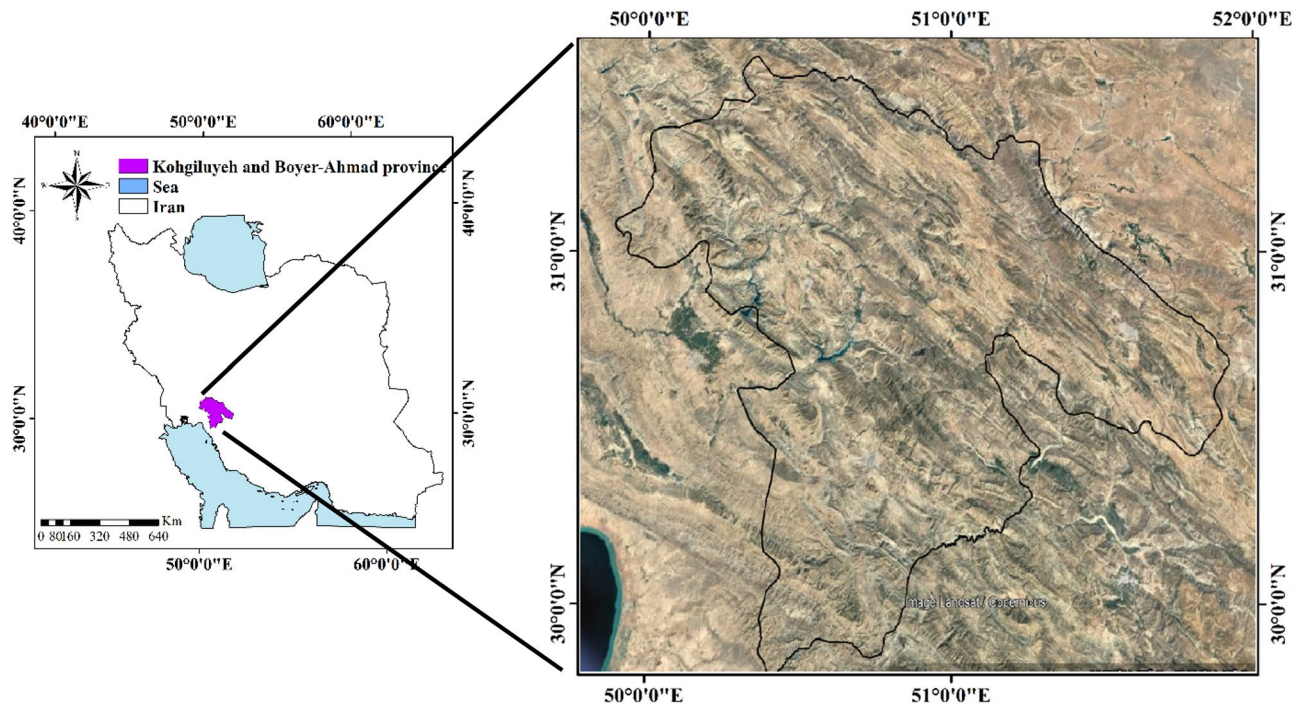


Figure 7. Percentage area of the four hazards.

tioning system (GPS). In this study, floods were investigated in 2000. Training and test data were prepared from 2000 to 2020. In addition, data from 2019 and 2020 were extracted using Sentinel satellite imagery. However, in the article, we only mentioned the data for 2019 and 2020; therefore, more details were added. Areas affected by these events were identified using the Google Earth Engine (GEE) and high-resolution Sentinel-1 and -3 images. We identified forest fire locations for the period 2015–2020 based on the GEE images. The codes were written in the GEE environment (<https://earthengine.google.com>). We randomly split the flood, gully erosion, and forest fire locations into two groups: a model training dataset with 70% of the locations, and a validation dataset with 30% of the locations<sup>60</sup>. The earthquake map was created from the epicenters and magnitudes of historic earthquakes provided by S.A.P. Consulting Engineers Co the Kohgiluyeh and Boyer-Ahmad Deputy Governor of Planning ([www.sabzandish.com](http://www.sabzandish.com)).

**Effective factors for multi-hazard assessment.** The first step in mapping and assessing natural hazards is to choose the appropriate control variables. Based on data availability and past experience<sup>61–66</sup>, we employed 19 effective factors in this study: elevation, slope, aspect, mean annual rainfall, mean annual temperature, lithology, land use, normalized difference vegetation index (NDVI), soil texture (percent clay, silt, and sand), wind exposure index, topographic wetness index (TWI), plan curvature, profile curvature, drainage density, distance from





**Figure 8.** Location of the Kohgiluyeh and Boyer-Ahmad Province in Iran.

roads, distance from rivers, and distance from villages (Table 2). These factors were rasterized with a  $30 \times 30$  m pixel size in the ArcGIS 10.8 platform and are briefly described below.

**Elevation.** Elevation is widely used as a controlling factor in natural hazard assessments<sup>27, 67</sup>. In this study, we used elevation factors in the preparation of flood, forest fire, and gully erosion maps. An elevation layer was prepared from ASTER satellite images (<https://earthexplorer.usgs.gov/>) (Fig. 9a).

**Slope angle.** The slope angle is an important contributing factor in flooding, forest fires, and gully erosion<sup>68, 69</sup>. In the case of forest fires, fire spreads more rapidly on steep slopes than on gentle ones<sup>70</sup>. In contrast, with all other things being equal, the likelihood of flooding increases with a decrease in slope, and flat and gently sloping ground in the study area is more prone to gully erosion than steep slopes<sup>58</sup>. A slope-angle layer was derived from DEM in ArcGIS 10.8. The slopes ranged from  $0^\circ$  to  $78^\circ$  (Fig. 9b).

**Slope aspect.** Sunlight, humidity, and temperature are affected by slope aspect; thus, aspect can be an important factor in the occurrence of floods, forest fires, and gully erosion. Eastern slopes in the study area receive sunlight earlier than western aspects, and therefore are more fire-prone<sup>20</sup>. The slope aspect has also been shown to be related to gully erosion through weathering mechanisms<sup>17, 71</sup>. We divided the aspect into nine classes: flat, north, northeast, northwest, southwest, west, east, south, and southeast (Fig. 9c).

**Rainfall.** Rainfall has been widely used as a factor in multi-hazard spatial modelling<sup>17, 72</sup>. Low rainfall in forested areas makes these areas susceptible to wildfire<sup>72, 73</sup>, whereas extreme rainfall triggers flooding<sup>74</sup> and gully erosion<sup>58</sup>. We prepared a rainfall layer in the ArcGIS 10.8 using data for the period 2001–2020 from nine weather stations in Kohgiluyeh and Boyer-Ahmad Province (Fig. 9d). A temperature map was obtained using the IDW interpolation<sup>75, 76</sup>.

**Temperature.** All other things being equal, high temperatures, and a lack of rainfall rapidly dry soil, making forests in environments prone to wildfire<sup>77</sup>. In the study area, temperature and humidity seem to be stronger controls for forest fires in higher parts of the province than lower ones<sup>78</sup>. We produced an average annual temperature layer based on climate data from nine stations during the period 2001–2020 (Fig. 9e). A rainfall map was created through interpolation using the IDW method<sup>79</sup>.

**Normalized difference vegetation index.** The normalized difference vegetation index (NDVI) provides a measure of vegetation health<sup>80</sup> and is thus an important factor in forest fire, gully erosion, and flood susceptibility assessments. We calculated the NDVI from a Sentinel-2 satellite image in Google Earth Engine using the following equation:

$$NDVI = (NIR - RED) / (NIR + RED) \quad (2)$$

Effective factor	Flood	Forest fire	Gully erosion
Elevation	*	*	*
Slope	*	*	*
Aspect	*	*	*
Mean annual rainfall	*	*	*
Mean annual temperature		*	
Lithology	*		*
Land use	*		*
NDVI	*	*	*
Soil texture (clay)			*
Soil texture (silt)			*
Soil texture (sand)			*
Wind exposition index		*	
TWI	*	*	*
Plane curvature	*		*
Profile curvature			*
Drainage density	*		*
Distance from roads	*	*	*
Distance from rivers	*	*	*
Distance from villages		*	

**Table 2.** The effective factors for MHA.

where NIR and RED refer to the near-infrared and red bands, respectively (Fig. 9f).

**Lithology.** Rock permeability, which is related to lithology, may affect flood susceptibility<sup>28, 81–83</sup> and gully erosion<sup>5</sup>. We obtained a geology map from the Geological Survey of Iran and produced a layer of 14 units in the ArcGIS 10.8 (Fig. 9g).

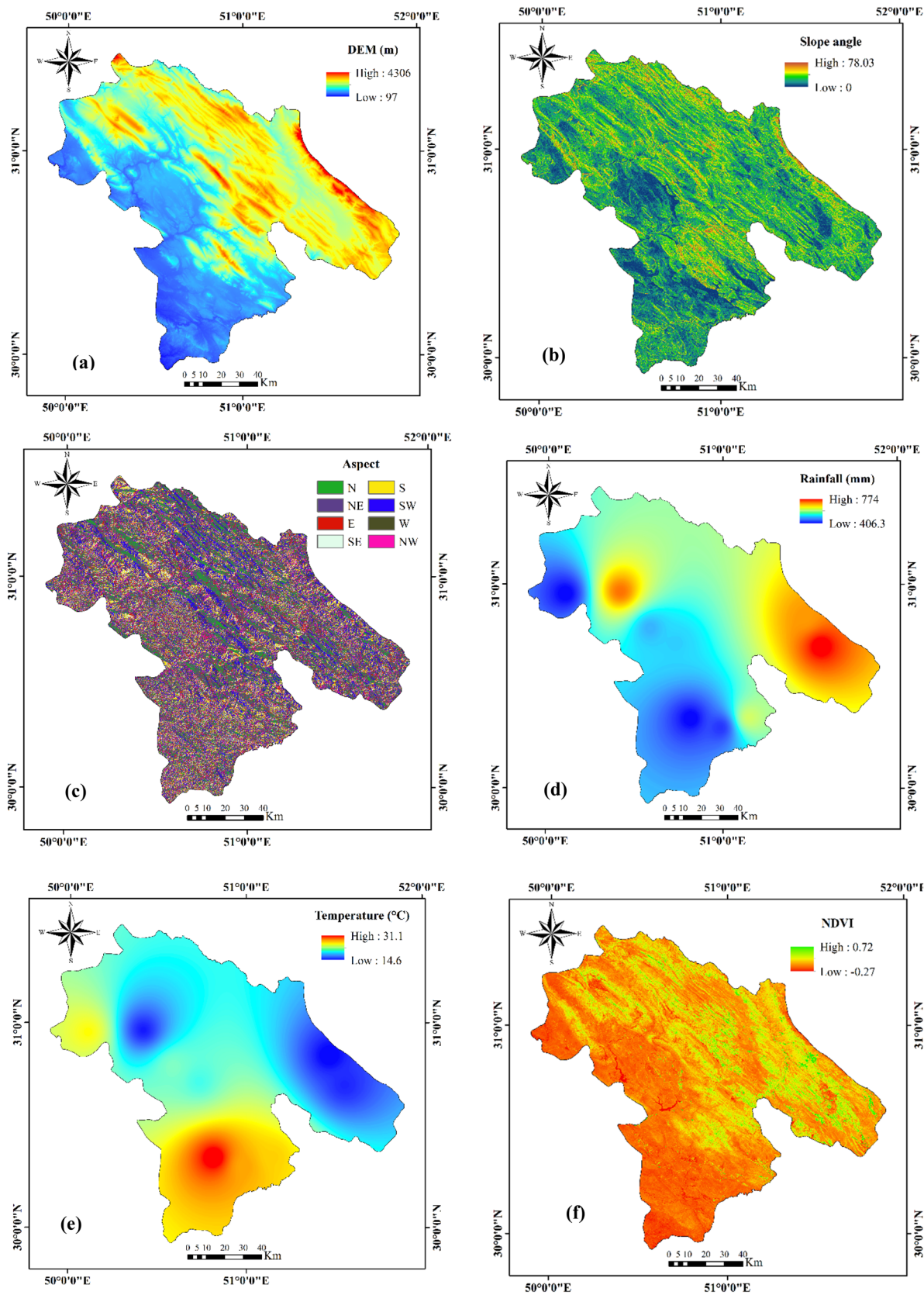
**Land use.** Land use changes, including deforestation and grazing, can initiate or exacerbate gully erosion<sup>30, 84–86</sup>, and bare lands are more susceptible to floods than naturally vegetated ones<sup>71</sup>. Therefore, we considered land use as a factor in the preparation of forest fire and flood maps. The land-use layer has 10 classes: agriculture, bare land, forest, mixture (agriculture and orchard), orchard, range land, rocky land, urban, water body, and woodlands (Fig. 9h). The land use map (with an accuracy of 91% prepared by the support vector machine (SVM) method) was obtained using the Landsat-8 (2019) satellite imagery.

**Soil texture.** A soil texture is considered one of the main factors in gully erosion<sup>87</sup>. We used data and guidelines provided by the World Soil Information Service (WoSIS, <http://soilgrids.org>) to prepare the soil layer in the ArcGIS 10.8. The SoilGrids map is a global digital soil mapping system that employs new machine learning techniques to map the spatial distribution of soil properties. The SoilGrids maps are available with a spatial resolution of 250 m. The SoilGrids maps are generated using more than 230,000 soil profile observations from the WoSIS database and over 400 environmental layers of Earth observation and other environmental data such as climate, land cover, and land morphology (<https://www.isric.org/explore/soilgrids>). The WoSIS database includes typical soil features for each soil group<sup>88</sup>. We defined soil texture in the study area using three percentage classes as proxies: clay, silt, and sand (Fig. 9i,j,k).

**Wind exposition index.** The wind exposition index (WEI) is a non-dimensional index used to quantify wind exposure at the land surface<sup>89</sup>. It takes into account the wind direction, angle of the surface earth to the horizon, and wind aspect<sup>90</sup>. Values larger than 1 indicate wind-exposed cells, and values less than 1 correspond to wind-shadowed cells<sup>91</sup>. WEI was computed from the DEM using the SAGA-GIS software (<http://www.saga-gis.org/en/index.html>) (Fig. 9l).

**Topographic wetness index.** The topographic wetness index (TWI) is a measure of the likelihood that surface water will move downslope<sup>92, 93</sup> and is used in flood, forest fire, and gully erosion assessments. TWI was calculated from ASTER imagery using the SAGA GIS software (Fig. 9m).

**Plan curvature.** Plan curvature is the curvature of an isoline constructed from the junction of a horizontal plane and the land surface<sup>94, 95</sup>. This factor was considered in flood and gully erosion evaluations. A plan curvature layer was derived from the DEM using the spatial analysis extension (curvature tool) in ArcMap (Fig. 9n).



**Figure 9.** Maps of hazard factors used in this study.

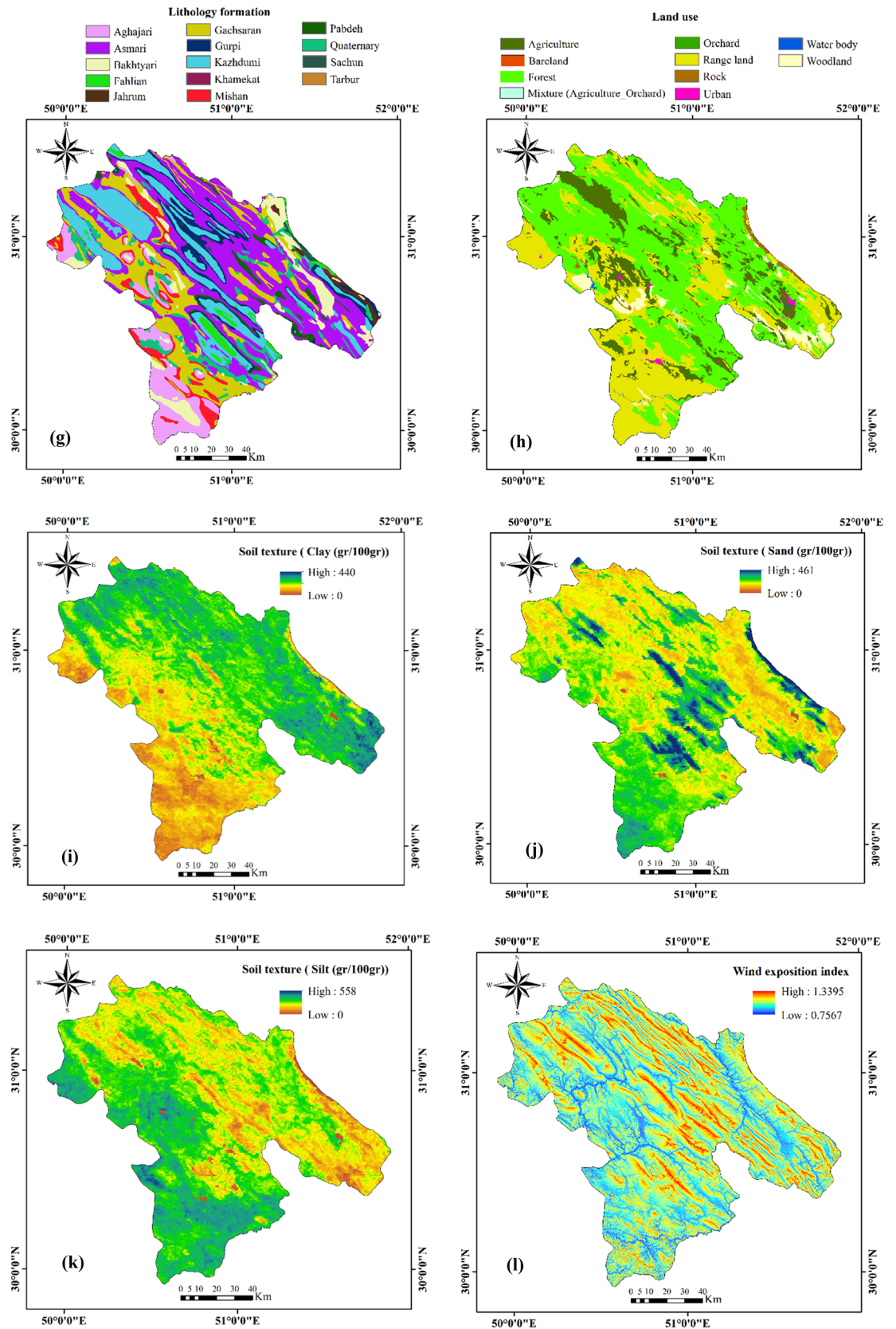


Figure 9. (continued)

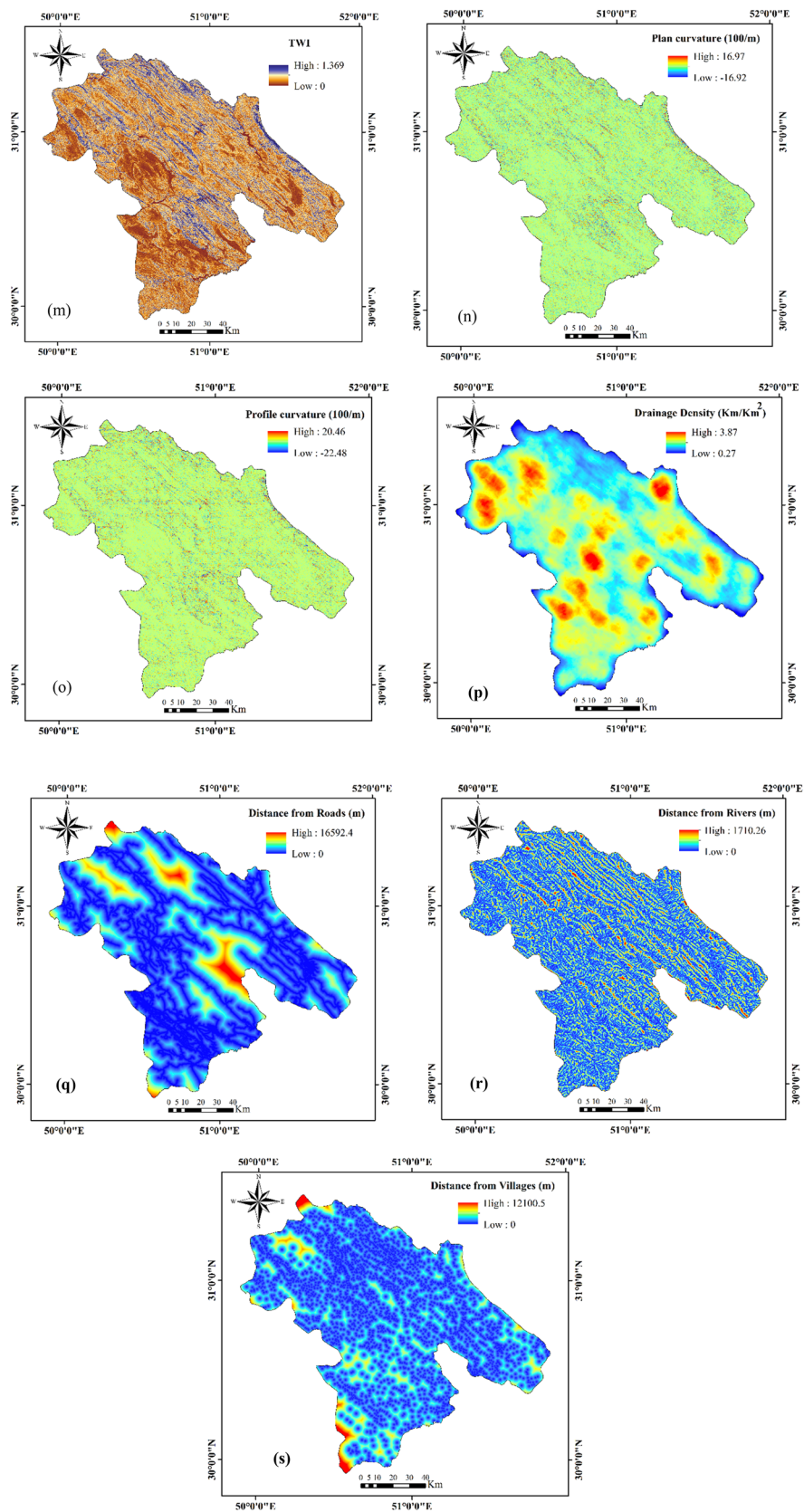


Figure 9. (continued)

**Profile curvature.** The profile curvature is a measure of the curvature of the slope surface. It controls surface and groundwater movement, and thus is a metric for flow velocity and erosion<sup>94,96</sup>. The profile curvature layer was constructed from the DEM in the ArcGIS 10.8 platform using a curvature tool (Fig. 9o). Negative and positive values indicate the concave and convex surfaces, respectively.

**Drainage density.** Drainage density (DD) is controlled by precipitation, geology, vegetation, slope, and soil, and is a factor in gully erosion and flood assessments. Drainage density scales with runoff and therefore may be an indicative factor for gully erosion<sup>97,98</sup>. The drainage density layer was created from the rivers and streams in ArcGIS 10.8 using density tools (Fig. 9p).

**Distance from roads.** Distance from roads is important in flood, forest fire, and gully erosion assessments. In forest areas, the risk of fire increases with proximity to roads<sup>21</sup>, and gully erosion, especially in bare terrain, might be related to this variable<sup>99</sup>. The distance from roads was quantified using vector line distances calculated using the Euclidean distance method in ArcGIS 10.8 (Fig. 9q).

**Distance from rivers.** The probability of flooding is related to the distance from a river, especially for rivers with low storage capacity<sup>100</sup>. In addition, most people in the study area live near rivers; thus, there is a higher likelihood of forest fires in these areas<sup>101</sup>. Gully erosion in the study area is also the greatest near rivers<sup>102</sup>. The distance from the river layer was determined using the Euclidean distance tool in ArcGIS 10.8 (Fig. 9r).

**Distance from villages.** Distance from villages, such as distance from rivers, may be related to the probability of forest fires, as most wildfire events are human-caused events<sup>103,104</sup>. Figure 9s shows the distance from the village layer produced in ArcGIS 10.8. This factor was not used in flood and gully erosion assessments.

**Multi-hazard spatial modelling. Boosted regression tree.** The boosted regression tree (BRT) algorithm is a machine learning method based on the use of classification and regression trees in combination with a boosting algorithm<sup>105</sup>. Its purpose is to improve the performance of a single model by fitting and combining a large number of models for prediction<sup>106</sup>. It can be used to predict quantitative (regression tree) or categorized (classification tree) outcomes. In this study, the regression tree model is reinforced by assessing logical conditions, rather than a linear relationship, to predict or classify landforms. When using the BRT model, there is no need to predetermine the function forms or make statistical assumptions about the data distribution. Other advantages include the ability to make multiple predictions and determine their possible nonlinear relationships with the response variable<sup>107</sup>, the high speed of the model in analyzing large volumes of data, and its high capacity to analyze and classify layers. The algorithm depends on the setting options related to the reinforced trees and tree pruning. In the case of options related to reinforced trees, an important parameter is the reduction rate as a net weight for individual and reinforced trees. Optimizing the best reduction rate is also important for preventing overfitting of the predictions. Previous studies have shown that models with a reduction rate of 0.1 or less perform best<sup>108</sup>. The best parameters are selected based on an evaluation of their results using statistical evaluation metrics (RMSE and bias). In this study, the regression trees were programmed using R 3.5.3 statistical software (<https://cran.r-project.org/bin/windows/base>) with a BRT extension (see Elith et al.<sup>109</sup> for details).

**Support vector machine.** The support vector machine (SVM) is a supervised learning method used for classification and regression. It was proposed in 1995 based on statistical learning theory using structural risk minimization<sup>110</sup>. It has been widely used in recent years as it performs better than older classification methods such as perceptron neural networks<sup>111</sup>. The SVM classifier works by linearly classifying data; in segmenting the data, it selects the most reliable boundary<sup>110</sup>. The closest samples to the decision boundary, which determine the decision boundary equation, are termed support vectors<sup>112</sup>. The principle of structural risk minimization is applied to maximize the distance between two transient hyperplanes formed from support vectors<sup>113</sup>. SVM performs better on non-modeled data than the experimental risk-minimization mode, which attempts to minimize the training error. Four types of kernel functions (linear, polynomial, sigmoid, and radial basis) can be used to prepare multi-hazard maps using the SVM algorithm. We used the radial basis function in this study because of its better performance than other functions<sup>114–116</sup>. The SVM algorithm was programmed in the R3.5.3 statistical software with the *sdm* package<sup>117</sup>.

**Random forest.** The random forest (RF) algorithm is a widely used hybrid machine learning algorithm that includes regression and classification trees<sup>118</sup>. The data used by the RF algorithm do not need to be changed, converted, or modified, and the algorithm controls the lost values automatically<sup>119</sup>. Random trees classify the input vector with each tree in the forest; the output is the class tag that receives the majority of votes. Two factors (the mean decrease in accuracy and the mean decrease in Gini) are used to determine the priority of the effect of each effective factor<sup>120</sup>. We implemented the RF model using the R3.5.3 statistical software with the random forest package<sup>121</sup>.

**Probabilistic seismic hazard analysis.** Probabilistic seismic hazard analysis (PSHA) is a method used to determine the probability that a location will experience severe ground movements during an earthquake of a particular magnitude. The numerical-analytical method for PSHA was first proposed in 1968<sup>122</sup> and has improved since then. Relevant seismic parameters include earthquake magnitude, return period, epicentral and source distances, and peak ground acceleration<sup>123</sup>. Different scales are used to quantify seismic magnitude, such

as moment, surface wave, volume wave, local, and torque magnitudes. One type of PSHA, the deterministic seismic hazard analysis (DSHA), determines the maximum kinetic parameters of the ground surface in a scenario earthquake, which is described by its magnitude and source distance<sup>124, 125</sup>. The scenario earthquake is expected to produce the greatest ground motion, such as peak ground acceleration (PGA), at the site if it were to occur<sup>126</sup>. We used this method to assess built structures whose destruction would have catastrophic consequences, such as nuclear power plants and large dams. It does not provide information on the probability of a scenario earthquake, its exact location, or the duration of severe ground motions. Uncertainties in determining the magnitude and source distance of an earthquake are typically combined to calculate the probability of exceedance of a specified PGA in a certain period of time. In this study, a peak horizontal acceleration map was created for an earthquake with a 10% probability of occurrence in 50 years (the time series 1965–2015) and an average return period of 75 years<sup>127</sup>.

**Multi-hazard mapping.** A multi-hazard map was constructed from the maps of flood hazard (FH), forest fire hazard (FFH), gully erosion hazard (GEH), and peak ground acceleration (PGA). The BRT, SVM, and RF algorithms were first used to create flood, gully erosion, and forest fire maps. Then, the results of the three algorithms were compared to select the algorithm with the best performance. Next, a PGA map was produced using PSHA. Each hazard was mapped using four groups—low, moderate, high, and very high—and subsequently reclassified into two groups: 0 and 1. Class 1 indicates high to very high hazard susceptibility, whereas class 0 indicates low and moderate susceptibility. Finally, the two-class maps for three hazards (forest fire, flood, and gully erosion) and the PGA map were integrated into a final multi-hazard map for the study area.

**Model validation.** Validation of multi-hazard models is required to evaluate their accuracy and value<sup>128</sup>. In this study, we used the receiver operating characteristic (ROC) curve to assess the model's predictive power. The ROC plot has two axes: the *x-axis* depicts the false-positive rate and the *y-axis* depicts the true-positive rate. The value of the area under the curve (AUC) ranges from 0.5 to 1.0<sup>129, 130</sup>, and the higher the AUC value, the greater the prediction accuracy<sup>128, 131</sup>. A ROC curve was created for each map using both the training and validation datasets. As mentioned previously, 70% of the hazard locations were used for model training and 30% for model validation.

**Prioritizing effective factors.** We employed the mean decrease accuracy (MDA) and mean decrease Gini (MDG) to determine the priority factors for natural hazard occurrences in the study area. The variance inflation factor (VIF) and tolerance coefficient (TOL) were first computed to check for multicollinearity. Variables with values of  $VIF \geq 5$  and  $TOL < 0.1$  indicate a multicollinearity problem<sup>132</sup>. We then determined the relative importance of these factors. According to Nicodemus<sup>133</sup>, the Gini index is more stable than the in-mean decrease for determining the priority of effective factors, especially in situations where there is a relationship between environmental factors. The mean decrease Gini (MDG) is defined as the sum of the Gini impurities decreasing from a specific variable normalized by trees<sup>120</sup>. Therefore, we also used MDG to determine the priority factors in this study.

## Data availability

The data used in this study are available upon request to the corresponding author for reasonable use in research. All figures draw by the authors using R 3.5.3 statistical software (<https://cran.r-project.org/bin/windows/base>), SAGA-GIS software (<http://www.saga-gis.org/en/index.html>), and ArcGIS 10.8 environment (<https://www.esri.com>).

Received: 27 February 2021; Accepted: 8 July 2021

Published online: 21 July 2021

## References

1. Tsakiris, G. Practical application of risk and hazard concepts in proactive planning. *Eur. Water* **19**, 47–56 (2007).
2. Kappes, M. S., Keiler, M., von Elverfeldt, K. & Glade, T. Challenges of analyzing multi-hazard risk: A review. *Nat. Hazards* **64**, 1925–1958 (2012).
3. UNEP. Agenda 21. Technical Report, United Nations Conference on Environment and Development. <https://www.un.org/en/conferences/environment/rio> (1992).
4. Van Westen, C. J., Montoya, L., Boerboom, L. & Badilla Coto, E. Multi-hazard risk assessment using GIS in urban areas: A case study for the city of Turrialba, Costa Rica. In *Proceedings of the regional workshop on best practice in disaster mitigation, Bali* 120–136 (2002).
5. Binita, K., Shepherd, J., King, A. W. & Gaither, C. J. Multi-hazard climate risk projections for the United States. *Nat. Hazards* **105**, 1963–1976 (2021).
6. Kaur, H., Gupta, S., Parkash, S. & Thapa, R. Application of geospatial technologies for multi-hazard mapping and characterization of associated risk at local scale. *Ann. GIS* **24**, 33–46 (2018).
7. Furlan, E., Torresan, S., Critto, A. & Marcomini, A. Spatially explicit risk approach for multi-hazard assessment and management in marine environment: The case study of the Adriatic Sea. *Sci. Total Environ.* **618**, 1008–1023 (2018).
8. Araya-Muñoz, D., Metzger, M. J., Stuart, N., Wilson, A. M. W. & Carvajal, D. A spatial fuzzy logic approach to urban multi-hazard impact assessment in Concepción, Chile. *Sci. Total Environ.* **576**, 508–519 (2017).
9. Larsen, M. C., Conde, M. T. V. & Clark, R. A. *Coping with Flash Floods* 259–275 (Springer, 2001).
10. Chinnasamy, P., Honap, V. U. & Maske, A. B. Impact of 2018 Kerala floods on soil erosion: Need for post-disaster soil management. *J. Indian Soc. Remote Sens.* **48**, 1373–1388 (2020).
11. Motagh, M. *et al.* Landslides triggered by 2019 extreme rainfall and flood events in Iran: Results from satellite remote sensing and field survey. In *EGU general assembly conference abstracts EGU2020-10715*, updated on 01 Jun 2021. <https://meetingorganizer.copernicus.org/EGU2020/EGU2020-10715.html>
12. Radbruch-Hall, D. & Varnes, D. Landslides—Cause and effect. *Bull. Int. Assoc. Eng. Geol.* **13**, 205–216 (1976).

13. Santín, C. & Doerr, S. H. Fire effects on soils: The human dimension. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150171 (2016).
14. Ulery, A. & Graham, R. Forest fire effects on soil color and texture. *Soil Sci. Soc. Am. J.* **57**, 135–140 (1993).
15. Shahab, H., Emami, H. & Haghnia, G. H. Effects of gully erosion on soil quality indices in northwestern Iran. *J. Agric. Sci. Technol.* **20**(6), 1317–1329 (2018).
16. Darvishi, L., Daryaei, M. & Kouchi, A. Comparison of statistical modeling and AHP methods in fire risk assessment in oak forests of Iran. *J. For. Res.* **9**, 229 (2020).
17. Lei, X. *et al.* GIS-based machine learning algorithms for gully erosion susceptibility mapping in a semi-arid region of Iran. *Remote Sens.* **12**, 2478 (2020).
18. Kamranzad, F., Memarian, H. & Zare, M. Earthquake risk assessment for Tehran, Iran. *ISPRS Int. J. Geo Inf.* **9**, 430 (2020).
19. Ozturk, M., Gucel, S., Kucuk, M. & Sakcali, S. Forest diversity, climate change and forest fires in the Mediterranean region of Turkey. *J. Environ. Biol.* **31**, 1 (2010).
20. Sari, F. Forest fire susceptibility mapping via multi-criteria decision analysis techniques for Mugla, Turkey: A comparative analysis of VIKOR and TOPSIS. *For. Ecol. Manag.* **480**, 118644 (2021).
21. Dong, X., Li-min, D., Guo-fan, S., Lei, T. & Hui, W. Forest fire risk zone mapping from satellite images and GIS for Baihe Forestry Bureau, Jilin, China. *J. For. Res.* **16**, 169–174 (2005).
22. Adab, H., Kanniah, K. D. & Solaimani, K. Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. *Nat. Hazards* **65**, 1723–1743 (2013).
23. Hirabayashi, Y. *et al.* Global flood risk under climate change. *Nat. Clim. Change* **3**, 816–821 (2013).
24. CRED. *Natural Disaster 2018* (Centre for Research on the Epidemiology of Disasters, 2020).
25. Wang, Y., Fang, Z., Hong, H. & Peng, L. Flood susceptibility mapping using convolutional neural network frameworks. *J. Hydrol.* **582**, 124482 (2020).
26. Samanta, S., Pal, D. K. & Palsamanta, B. Flood susceptibility analysis through remote sensing, GIS and frequency ratio model. *Appl. Water Sci.* **8**, 1–14 (2018).
27. Hadidi, A., Holzbecher, E. & Molenaar, R. E. Flood mapping in face of rapid urbanization: A case study of Wadi Majraf-Manumah, Muscat, Sultanate of Oman. *Urban Water J.* **17**, 407–415 (2020).
28. Souissi, D. *et al.* GIS-based MCDM–AHP modeling for flood susceptibility mapping of arid areas, southeastern Tunisia. *Geocarto Int.* **35**, 991–1017 (2020).
29. Band, S. S. *et al.* Novel ensemble approach of deep learning neural network (DLNN) model and particle swarm optimization (PSO) algorithm for prediction of gully erosion susceptibility. *Sensors* **20**, 5609 (2020).
30. Dotterweich, M. *et al.* High-resolution gully erosion and sedimentation processes, and land use changes since the Bronze Age and future trajectories in the Kazimierz Dolny area (Nałęczów Plateau, SE-Poland). *Catena* **95**, 50–62 (2012).
31. Conforti, M., Aucelli, P. P., Robustelli, G. & Scarciglia, F. Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (northern Calabria, Italy). *Nat. Hazards* **56**, 881–898 (2011).
32. 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.* **29**, 724–736 (2018).
33. Yang, Y. H. *et al.* Midcrustal thrusting and vertical deformation partitioning constraint by 2017 Mw 7.3 Sarpol Zahab earthquake in Zagros Mountain Belt, Iran. *Seismol. Res. Lett.* **89**, 2204–2213 (2018).
34. Domínguez-Cuesta, M. J. In *Encyclopedia of Natural Hazards* (ed. Bobrowsky, P. T.) 988–988 (Springer, 2013). [https://doi.org/10.1007/978-1-4020-4399-4\\_340](https://doi.org/10.1007/978-1-4020-4399-4_340).
35. Lehner, B., Döll, P., Alcamo, J., Henrichs, T. & Kaspar, F. Estimating the impact of global change on flood and drought risks in Europe: A continental, integrated analysis. *Clim. Change* **75**, 273–299 (2006).
36. Stoffel, M. & Bollschweiler, M. Tree-ring analysis in natural hazards research—An overview. *Nat. Hazards* **8**, 187–202 (2008).
37. Tofani, V., Segoni, S., Agostini, A., Catani, F. & Casagli, N. Use of remote sensing for landslide studies in Europe. *Nat. Hazards* **13**, 299–309 (2013).
38. Kundzewicz, Z. W. *et al.* Differences in flood hazard projections in Europe—Their causes and consequences for decision making. *Hydrol. Sci. J.* **62**, 1–14 (2017).
39. Paprotny, D., Kreibich, H., Morales-Nápoles, O., Terefenko, P. & Schröter, K. Estimating exposure of residential assets to natural hazards in Europe using open data. *Nat. Hazards* **20**, 323–343 (2020).
40. Bathrellos, G. D., Skilodimou, H. D., Chousianitis, K., Youssef, A. M. & Pradhan, B. Suitability estimation for urban development using multi-hazard assessment map. *Sci. Total Environ.* **575**, 119–134 (2017).
41. Lucà, F., Conforti, M. & Robustelli, G. Comparison of GIS-based gully susceptibility mapping using bivariate and multivariate statistics: Northern Calabria, South Italy. *Geomorphology* **134**, 297–308 (2011).
42. Pourghasemi, H. R., Gayen, A., Lasaponara, R. & Tiefenbacher, J. P. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. *Environ. Res.* **184**, 109321 (2020).
43. Guisan, A. & Thuiller, W. Predicting species distribution: Offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009 (2005).
44. Oliveira, S., Oehler, F., San-Miguel-Ayanz, J., Camia, A. & Pereira, J. M. Modeling spatial patterns of fire occurrence in Mediterranean Europe using multiple regression and random forest. *For. Ecol. Manag.* **275**, 117–129 (2012).
45. Kühnlein, M., Appelhans, T., Thies, B. & Nauss, T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—A random forests-based approach applied to MSG SEVIRI. *Remote Sens. Environ.* **141**, 129–143 (2014).
46. Wang, Y. *et al.* Random bits forest: A strong classifier/regressor for big data. *Sci. Rep.* **6**, 1–8 (2016).
47. Han, Q., Gui, C., Xu, J. & Laciogna, G. A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm. *Constr. Build. Mater.* **226**, 734–742 (2019).
48. Gislason, P. O., Benediktsson, J. A. & Sveinsson, J. R. Random forests for land cover classification. *Pattern Recogn. Lett.* **27**, 294–300 (2006).
49. Cutler, D. R. *et al.* Random forests for classification in ecology. *Ecology* **88**, 2783–2792 (2007).
50. Zhang, K., Wu, X., Niu, R., Yang, K. & Zhao, L. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. *Environ. Earth Sci.* **76**, 1–20 (2017).
51. Zhao, G., Pang, B., Xu, Z., Yue, J. & Tu, T. Mapping flood susceptibility in mountainous areas on a national scale in China. *Sci. Total Environ.* **615**, 1133–1142 (2018).
52. Milanović, S. *et al.* Forest fire probability mapping in eastern Serbia: Logistic regression versus random forest method. *Forests* **12**, 5 (2021).
53. Gounaridis, D., Chorianopoulos, I., Symeonakis, E. & Koukoulas, S. A random forest-cellular automata modelling approach to explore future land use/cover change in Attica (Greece), under different socio-economic realities and scales. *Sci. Total Environ.* **646**, 320–335 (2019).
54. Wu, Z., He, H. S., Yang, J., Liu, Z. & Liang, Y. Relative effects of climatic and local factors on fire occurrence in boreal forest landscapes of northeastern China. *Sci. Total Environ.* **493**, 472–480 (2014).
55. Wang, N., Shi, T., Peng, K., Zhang, W. & Jin, X. Assessment of geohazard susceptibility based on RS and GIS analysis in Jianshi County of the Three Gorges Reservoir, China. *Arab. J. Geosci.* **8**, 67–86 (2015).
56. Kanani-Sadat, Y., Arabsheibani, R., Karimipour, F. & Nasserli, M. A new approach to flood susceptibility assessment in data-scarce and ungauged regions based on GIS-based hybrid multi criteria decision-making method. *J. Hydrol.* **572**, 17–31 (2019).



57. Morgan, R. & Mngomezulu, D. Threshold conditions for initiation of valley-side gullies in the Middle Veld of Swaziland. *Catena* **50**, 401–414 (2003).
58. Azareh, A. *et al.* Modelling gully-erosion susceptibility in a semi-arid region, Iran: Investigation of applicability of certainty factor and maximum entropy models. *Sci. Total Environ.* **655**, 684–696 (2019).
59. Jahantab, E., Hatami, E., Sayadian, M. & Salahi Ardakani, A. Ethnobotanical study of medicinal plants of Boyer Ahmad and Dena regions in Kohgiluyeh and Boyer Ahmad province. *Iran. Adv. Herb. Med.* **3**, 12–22 (2018).
60. Robinson, M. C. & Glen, R. C. Validating the validation: Reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction. *J. Comput. Aided Mol. Des.* **34**, 1–14 (2020).
61. Chuvieco, E. & Congalton, R. G. Application of remote sensing and geographic information systems to forest fire hazard mapping. *Remote Sens. Environ.* **29**, 147–159 (1989).
62. Kourgialas, N. N. & Karatzas, G. P. Flood management and a GIS modelling method to assess flood-hazard areas—A case study. *Hydrol. Sci. J.* **56**, 212–225 (2011).
63. Kabenge, M., Elaru, J., Wang, H. & Li, F. Characterizing flood hazard risk in data-scarce areas, using a remote sensing and GIS-based flood hazard index. *Nat. Hazards* **89**, 1369–1387 (2017).
64. Ganteaume, A. *et al.* A review of the main driving factors of forest fire ignition over Europe. *Environ. Manag.* **51**, 651–662 (2013).
65. Eini, M., Kaboli, H. S., Rashidian, M. & Hedayat, H. Hazard and vulnerability in urban flood risk mapping: Machine learning techniques and considering the role of urban districts. *Int. J. Disaster Risk Reduct.* **50**, 101687 (2020).
66. Hembram, T. K., Paul, G. C. & Saha, S. Modelling of gully erosion risk using new ensemble of conditional probability and index of entropy in Jainti River basin of Chotanagpur Plateau fringe area, India. *Appl. Geomat.* **12**, 1–24 (2020).
67. Wang, Y., Fang, Z., Wang, M., Peng, L. & Hong, H. Comparative study of landslide susceptibility mapping with different recurrent neural networks. *Comput. Geosci.* **138**, 104445 (2020).
68. Weise, D. R. & Biging, G. S. A qualitative comparison of fire spread models incorporating wind and slope effects. *For. Sci.* **43**, 170–180 (1997).
69. Fernandez, D. & Lutz, M. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Eng. Geol.* **111**, 90–98 (2010).
70. Kushla, J. D. & Ripple, W. J. The role of terrain in a fire mosaic of a temperate coniferous forest. *For. Ecol. Manag.* **95**, 97–107 (1997).
71. Pérez-García, E. A., Sevilha, A. C., Meave, J. A. & Scariot, A. Floristic differentiation in limestone outcrops of southern Mexico and central Brazil: A beta diversity approach. *Boletín de la Sociedad Botánica de México* **84**, 45–58 (2009).
72. Moritz, M. A. *et al.* Climate change and disruptions to global fire activity. *Ecosphere* **3**, 1–22 (2012).
73. Flannigan, M. D., Stocks, B. J. & Wotton, B. M. Climate change and forest fires. *Sci. Total Environ.* **262**, 221–229 (2000).
74. Mahmoud, S. H. & Gan, T. Y. Urbanization and climate change implications in flood risk management: Developing an efficient decision support system for flood susceptibility mapping. *Sci. Total Environ.* **636**, 152–167 (2018).
75. Tewolde, M. G., Beza, T. A., Costa, A. C. & Painho, M. Comparison of different interpolation techniques to map temperature in the southern region of Eritrea. In *Geospatial Thinking, 13th AGILE international conference on geographic information science 10-14 May 2010 - Guimarães, Portugal* 1–5 (2010).
76. Xu, C., Wang, J. & Li, Q. A new method for temperature spatial interpolation based on sparse historical stations. *J. Clim.* **31**, 1757–1770 (2018).
77. Amraoui, M., Pereira, M. G., DaCamara, C. C. & Calado, T. J. Atmospheric conditions associated with extreme fire activity in the Western Mediterranean region. *Sci. Total Environ.* **524**, 32–39 (2015).
78. Hernandez-Leal, P., Arbelo, M. & Gonzalez-Calvo, A. Fire risk assessment using satellite data. *Adv. Space Res.* **37**, 741–746 (2006).
79. Ryu, S. *et al.* Spatial interpolation of gauge measured rainfall using compressed sensing. *Asia Pac. J. Atmos. Sci.* **57**, 331–345 (2021).
80. Bajocco, S. *et al.* Mapping forest fuels through vegetation phenology: The role of coarse-resolution satellite time-series. *PLoS ONE* **10**, e0119811 (2015).
81. Zeng, L. & Liu, H. The key geological factors influencing on development of low-permeability sandstone reservoirs: A case study of the Taizhao area in the Songliao Basin, China. *Energy Explor. Exploit.* **27**, 425–437 (2009).
82. Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K. & Shirzadi, A. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J. Environ. Manag.* **217**, 1–11 (2018).
83. Vojtek, M. & Vojteková, J. Flood susceptibility mapping on a national scale in Slovakia using the analytical hierarchy process. *Water* **11**, 364 (2019).
84. Dale, V. H., Efrogmson, R. A. & Kline, K. L. The land use–climate change–energy nexus. *Landsc. Ecol.* **26**, 755–773 (2011).
85. Rawat, P. K., Tiwari, P. C., Pant, C. C., Sharama, A. & Pant, P. Modelling of stream run-off and sediment output for erosion hazard assessment in Lesser Himalaya: Need for sustainable land use plan using remote sensing and GIS: A case study. *Nat. Hazards* **59**, 1277–1297 (2011).
86. Rimal, B., Baral, H., Stork, N. E., Paudyal, K. & Rijal, S. Growing city and rapid land use transition: Assessing multiple hazards and risks in the Pokhara Valley, Nepal. *Land* **4**, 957–978 (2015).
87. Auerswald, K., Fiener, P., Martin, W. & Elhaus, D. Use and misuse of the K factor equation in soil erosion modeling: An alternative equation for determining USLE nomograph soil erodibility values. *Catena* **118**, 220–225 (2014).
88. Batjes, N. H. *et al.* WoSIS: Providing standardised soil profile data for the world. *Earth Syst. Sci. Data* **9**, 1–14 (2017).
89. Conrad, O. *et al.* System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* **8**, 1991–2007 (2015).
90. Böhner, J. & Antonić, O. Land-surface parameters specific to topo-climatology. *Dev. Soil Sci.* **33**, 195–226 (2009).
91. Storey, M. A., Price, O. F., Sharples, J. J. & Bradstock, R. A. Drivers of long-distance spotting during wildfires in south-eastern Australia. *Int. J. Wildland Fire* **29**, 459–472 (2020).
92. Gokceoglu, C., Sonmez, H., Nefeslioglu, H. A., Duman, T. Y. & Can, T. The 17 March 2005 Kuzulu landslide (Sivas, Turkey) and landslide-susceptibility map of its near vicinity. *Eng. Geol.* **81**, 65–83 (2005).
93. Yanar, T., Kocaman, S. & Gokceoglu, C. Use of Mamdani fuzzy algorithm for multi-hazard susceptibility assessment in a developing urban settlement (Mamak, Ankara, Turkey). *ISPRS Int. J. Geo Inf.* **9**, 114 (2020).
94. Moore, I. D., Grayson, R. & Ladson, A. Digital terrain modelling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **5**, 3–30 (1991).
95. Minar, J. & Evans, I. S. Elementary forms for land surface segmentation: The theoretical basis of terrain analysis and geomorphological mapping. *Geomorphology* **95**, 236–259 (2008).
96. Valeriano, M. M. *et al.* Modeling small watersheds in Brazilian Amazonia with Shuttle Radar Topographic Mission 90 m data. *Comput. Geosci.* **32**, 1169–1181 (2006).
97. Liggitt, B. & Fincham, R. Gully erosion: The neglected dimension in soil erosion research. *S. Afr. J. Sci.* **85**, 18 (1989).
98. Roy, P. *et al.* Development of different machine learning ensemble classifier for gully erosion susceptibility in Gandheswari watershed of West Bengal, India. In *Machine Learning for Intelligent Decision Science*, pp 1–26 (eds Rout, J. K. *et al.*) (Springer, 2020).
99. Nyssen, J. *et al.* Impact of road building on gully erosion risk: A case study from the northern Ethiopian highlands. *Earth Surf. Proc. Land.* **27**, 1267–1283 (2002).

100. Aalto, R. *et al.* Episodic sediment accumulation on Amazonian flood plains influenced by El Niño/Southern Oscillation. *Nature* **425**, 493–497 (2003).
101. Rajabi, M., Alesheikh, A., Chehreghan, A. & Gazmeh, H. An innovative method for forest fire risk zoning map using fuzzy inference system and GIS. *Int. J. Sci. Technol. Res.* **2**, 57–64 (2013).
102. Conoscenti, C. *et al.* Gully erosion susceptibility assessment by means of GIS-based logistic regression: A case of Sicily (Italy). *Geomorphology* **204**, 399–411 (2014).
103. Zhang, Z., Zhang, H. & Zhou, D. Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires. *J. Arid Environ.* **74**, 386–393 (2010).
104. Jaiswal, R. K., Mukherjee, S., Raju, K. D. & Saxena, R. Forest fire risk zone mapping from satellite imagery and GIS. *Int. J. Appl. Earth Obs. Geoinf.* **4**, 1–10 (2002).
105. Friedman, J. H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **38**, 367–378 (2002).
106. Schapire, R. E. The boosting approach to machine learning: An overview. In *Nonlinear Estimation and Classification* (eds Denison, D. D. *et al.*) 149–171 (Springer, 2003).
107. Chung, Y.-S. Factor complexity of accident occurrence: An empirical demonstration using boosted regression trees. In *3rd International Conference on Road Safety and Simulation*, 1–23. <http://onlinepubs.trb.org/onlinepubs/conferences/2011/RSS1/Chung,Y-S.pdf> (2011).
108. Kint, V. *et al.* Forest structure and soil fertility determine internal stem morphology of Pedunculate oak: A modelling approach using boosted regression trees. *Eur. J. For. Res.* **131**, 609–622 (2012).
109. Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151 (2006).
110. Vapnik, V., Guyon, I. & Hastie, T. Support vector machines. *Mach. Learn.* **20**, 273–297 (1995).
111. Cortez, P. & Silva, A. M. G. Using data mining to predict secondary school student performance. <http://www3.dsi.uminho.pt/pcortez/student.pdf> (2008).
112. Joachims, T. Making large-scale svm learning. In *Practical Advances in Kernel Methods-Support Vector Learning*. [https://www.cs.cornell.edu/people/tj/publications/joachims\\_99a.pdf](https://www.cs.cornell.edu/people/tj/publications/joachims_99a.pdf) (1999).
113. Marjanović, M., Kovačević, M., Bajat, B. & Voženilek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **123**, 225–234 (2011).
114. Singh, V. P. & Frevert, D. K. Watershed modelling. In *World Water and Environmental Resources Congress 2003*, 1–37. [https://doi.org/10.1061/40685\(2003\)167](https://doi.org/10.1061/40685(2003)167)
115. Yu, P.-S., Chen, S.-T. & Chang, I.-F. Support vector regression for real-time flood stage forecasting. *J. Hydrol.* **328**, 704–716 (2006).
116. Sánchez, A. S., Nieto, P. G., Fernández, P. R., del Coz Díaz, J. & Iglesias-Rodríguez, F. J. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). *Math. Comput. Modell.* **54**, 1453–1466 (2011).
117. Naimi, B., Araujo, M. B., Package 'sdm'. <https://cran.r-project.org/web/packages/sdm/sdm.pdf> (2016).
118. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
119. Breiman, L., & Cutler, A. State of the art of data mining using Random forest. In *Proceedings of the Salford Data Mining Conference, San Diego, USA* 24–25 (2012).
120. Calle, M. L. & Urrea, V. Letter to the editor: Stability of random forest importance measures. *Brief. Bioinform.* **12**, 86–89 (2011).
121. RColorBrewer, S., & Liaw, M. A. Package 'randomForest' (University of California, Berkeley, CA, USA, 2018). <https://www.stat.berkeley.edu/~breiman/RandomForests/>
122. Cornell, C. Engineering seismic risk analysis. *Bull. Seismol. Soc. Am.* **58**, 1583–1606 (1968).
123. Anderson, J. G. & Trifunac, M. Uniform risk functionals for characterization of strong earthquake ground motion. *Bull. Seismol. Soc. Am.* **68**, 205–218 (1978).
124. Reiter, L. *Earthquake Hazard Analysis: Issues and Insights* Vol. 22 (Columbia University Press, 1990).
125. Krinitsky, E. L. How to obtain earthquake ground motions for engineering design. *Eng. Geol.* **65**, 1–16 (2002).
126. Wang, J. P. & Huang, D. Deterministic seismic hazard assessments for Taiwan considering non-controlling seismic sources. *Bull. Eng. Geol. Environ.* **73**, 635–641 (2014).
127. Papadopoulou-Vrynioti, K., Bathrellos, G. D., Skilodimou, H. D., Kaviris, G. & Makropoulos, K. Karst collapse susceptibility mapping considering peak ground acceleration in a rapidly growing urban area. *Eng. Geol.* **158**, 77–88 (2013).
128. Obuchowski, N. A. & Bullen, J. A. Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. *Phys. Med. Biol.* **63**, 07TR01 (2018).
129. Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn.* **30**, 1145–1159 (1997).
130. David, F. & Reiser, B. Estimation of the area under the ROC curve. *Stat. Med.* **21**, 3093–3106 (2002).
131. Fan, J., Upadhye, S. & Worster, A. Understanding receiver operating characteristic (ROC) curves. *Can. J. Emerg. Med.* **8**, 19–20 (2006).
132. O'Brien, R. M. A caution regarding rules of thumb for variance inflation factors. *Qual. Quant.* **41**, 673–690 (2007).
133. Nicodemus, K. K. Letter to the editor: On the stability and ranking of predictors from random forest variable importance measures. *Brief. Bioinform.* **12**, 369–373 (2011).

## Acknowledgements

This research was supported by the College of Agriculture at Shiraz University (Grant No. 98GRC1M271143).

## Author contributions

S.P., H.R.P., M.B., S.R., and J.J.C. designed experiments, run models, analyzed results, wrote, and reviewed manuscript. All authors reviewed the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to H.R.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021