Using Machine Learning Techniques In Multi - Hazards Assessment of Golestan National Park, Iran

Hassan Faramarzi

Tarbiat Modares University

Seyed Mohsen Hosseini (✉ hosseini@modares.ac.ir)

Tarbiat Modares University

Hamid Reza Pourghasemi

Shiraz University

Mahdi Farnaghi

University of Twente

Research Article

Keywords: Flood, Forest Fire, Landslide, Environmental hazards, Optimization

Posted Date: August 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-814390/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Golestan National Park is one of the oldest biosphere reserves exposed to environmental hazards due to growing demand, geographical location of the park, mountainous conditions, and developments in the last five decades. The purpose of this study was to evaluate potential environmental hazards using machine-learning techniques. In this study, maximum entropy, random forest, boosted regression tree, generalized additive model, and support vector machine methods were applied to model environmental hazards and evaluate the impact of affecting agents, their area of influence, and interactions. After data collection and preprocessing, the models were implemented, tuned, and trained, and their accuracies were determined using the receiver operating characteristic curve. The results indicate the high importance of climatic and human variables, including rainfall, temperature, presence of shepherds, and villagers for fire hazards, elevation, transit roads, temperature, and rainfall for the formation of floodplains, and elevation, transit roads, rainfall, and topographic wetness index in the occurrence of landslides in the national park. The boosted regression tree model with a ROC value of 0.98 for flooding, 0.97 for fire, and 0.93 for landslide hazards, had the best performance. The modeling estimated that, on average, 16.2% of the area of Golestan National Park has a high potential for landslides, 14% has a high potential for fire, and 7.2% has a high potential for flooding.

Introduction

Golestan National Park (GNP) is the first national park in Iran ¹ and one of the oldest biosphere reserves and biomass of mixed mountain systems. It is considered the only representative example of a temperate forest biome worldwide ². Because of its particular vegetation cover and its location in the wind corridor caused by the collision of wet and dry air fronts ³, GNPs are prone to various environmental hazards. Every year, several wildfires occur in hundreds of hectares of parks. Additionally, improper use of land, construction of inappropriate structures, the high slope of the basin, and heavy rains result in the occurrence of flood and landslide events in the area ⁴, where the floods and landslides in 2001, 2019, and 2020 were the most destructive events.

Climate change and global warming have resulted in irregularities in the pattern and intensity of rainfall, resulting in an increase in environmental hazards, including landslides, floods, and wildfires ⁵⁻⁹. Precipitation anomalies are increasing rapidly worldwide ¹⁰⁻¹², causing significant damage to natural resources, agricultural communities, and human structures every year ^{13,14}. The average annual economic loss due to floods and landslides in Iran is estimated to be \$1705,000 ¹⁵. Recent floods have suffered more severe damage due to rapid urbanization and population growth, especially in developing countries ¹⁶. In parallel, landslides cause human casualties, significant financial losses, and severe damage to natural ecosystems and man-made structures ^{17,18}. Additionally, the number and severity of forest fires are increasing ¹⁹. These fires result in damage to the soil, seriously endangering ecosystem services, and threatening sustainability ²⁰. Forest fires also adversely affect the hydrological response at the watershed level ²¹.

These environmental hazards are typically studied independently. However, in order to advance knowledge, improve management, and reduce the resulting financial and human losses, we need to study them simultaneously as a comprehensive set of interrelated phenomena^{22–24}. "Multi-hazard" is a term used by the United Nations in the context of the Sustainable Development Agenda 21 as the global program to reduce natural and man-made disasters and achieve sustainable management²⁵. A coherent solution to control these hazards is to develop models that identify areas that are prone to different possible risks simultaneously^{26–28}. Determining the severity, recurrence, and extent of hazards is of great importance for planning the response and making well-informed management decisions²⁹. In particular, identifying areas with high potential for landslides, flooding, and fire hazards can help managers in the GNP plan better and reduce losses.

This study contributes to the assessment of GNP, as the most important national park in Iran, through multi-hazard modeling of flooding, landslides, and wildfires using machine learning (ML) algorithms. The modeling considers the effects of climatic factors, environmental characteristics, and human activities to determine the most vulnerable areas in the park. High-risk areas of environmental hazards and their extent of overlap were also determined. Finally, the results of the optimal model, selected based on the accuracy measure, were presented.

Materials And Methods

Study area

GNP, a UNESCO World Heritage Natural Site, is located at the intersection of two biogeographical areas: Iran-Turani and Hyrcania. The yearly rainfall varies between 150 mm and 750 mm in different regions of the park^{30–32}. This national park is one of the inscribed components of Hyrcanian forests which with an area of approximately 91859 hectares in the geographical range of 37°16'34" to 37°31'00" northern latitude and 55°43'00" to 56°17'45" eastern longitude (Fig. 1). The national park's minimum and maximum elevations were 450 and 2,411 m, respectively. It is located between two temperate climates: humid temperate and cold and dry. The average annual temperature in this national park is between 11.5 to 17.5 degrees Celsius. In winter, the relative humidity is saturated (100%), and in summer, it decreases to 18%. In addition, there is a potential for strong winds in this area.

Methods

In this study, first, the agents that influence the environmental hazards of flooding, landslides, and wildfires and their effects were determined by reviewing the literature. The data layers related to each factor were then prepared using the Geospatial Information System (GIS) software. Next, the locations of historic hazards in the study area were determined using global positioning system (GPS) sensors. Having access to the influencing factors along with the locations and types of events, we trained ML models that could predict the occurrence of hazards based on the input spatial and temporal data. Finally, the trained models were used to produce a multi-hazard map for the entire study area.

Influencing factors

To assess flood, landslide, and fire hazards, we need to have access to spatial and temporal data to properly model the static and dynamic factors³³. In this respect, the influential factors that have a static nature include topographical aspects (elevation, slope, and aspect), geological factors (rock type, soil depth, and bedrock characteristics), hydrological factors (distance from rivers, drainage density, drainage, and soil moisture), geomorphological aspects (physiographic units and Earth morphologies), and environmental factors (roads, structures, and vegetation characteristics)^{34,35}. Dynamic factors also include rainfall, temperature, and humidity, which vary over time and space³⁶. We used ArcGIS 10.4 and SAGA GIS 2 to prepare the data layers related to each affecting factor.

Topographical factors

Topography affects environmental hazards by altering the composition and distribution of vegetation, climate, and soil^{37–39}, modifying the temperature and microclimate^{40,41}. In this regard, slope, aspect, elevation, curvature, stream power index (SPI), and topographic wetness index (TWI) play essential roles in creating and intensifying these hazards. These variables were calculated using a 30 × 30 m digital elevation model (DEM) prepared by the Iran National Cartographic Center.

TWI is widely used to determine the spatial distribution of moisture status in a particular area⁴². This factor was obtained using Eq. 1.

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

The SPI index is also one of the most widely used factors for determining the erosive power of water flow, and in areas with high values of these indicators, there is a high probability of floods and landslides⁶. This factor was calculated using Eq. 2.

$$SPI = As \tan \beta \quad (2)$$

In Eqs. 1 and 2, As is the specific contributing area of the basin in terms of the cumulative upstream area (upstream area per unit length of alignment line), and β is the degree of slope^{43,44}.

Plan curvature and stream variables control drainage and water retention by changing the movement of water, thus affecting the formation of floodplains and landslide hazards^{45,46}.

Human factors

Human factors include accidents, traffic, traditional activities in the surrounding village areas, and factors that can lead to unrest⁴⁷. In this regard, the presence of villagers, their population, the role of tourists, and the presence of shepherds and farmers affect environmental hazards^{48,49}. Roads in any ecosystem

influence the stimulation of the domains and the frequency of natural disturbances^{28,50}. To include these factors in the modeling, we created distance maps from villages, roads, and areas where the Shepherds were present (identified by the Central National Park Administration) with a 1/25,000 scale.

Hydrological factors

Streams significantly affect brae stability through erosion and saturation of the lower parts⁵¹. In other words, water on the surface of braes (runoff) and water that penetrates between soil particles causes irritation. In areas where slopes are steep and slippery, rivers play an essential role in creating all types of mass movements, especially landslides, by washing the brae and destroying the support of the brae material⁵². Therefore, the landslide potential decreases with an increase in the distance from waterways. Rivers flood adjacent lands⁵³, and because they attract passers-by and tourists, they also increase the risk of fire⁵⁴. To include these factors in the modeling, the distance layers from the primary and secondary rivers, as well as the drainage density layer, were prepared with a 1/25,000 scale.

Climate factors

Climatic factors, including temperature, humidity, and rainfall, directly affect the occurrence of environmental hazards⁵⁵. On the other hand, these factors implicitly influence the occurrence of environmental hazards because other factors, such as human factors, are a function of climatic factors.

Rainfall is an influential factor in the creation of slope instability. The highest number of slope ruptures occurred after heavy rains or snowmelt in spring and due to the infiltration of water in the crevices⁵⁶. Rainfall is one of the factors controlling wildfires because increasing rainfall in an area increases humidity. The higher the amount of rainfall in the spring in an area, the higher the density of grass cover in summer and, consequently, the greater the likelihood of wildfires occurring and spreading in warm seasons⁵⁷. Climatic factor maps were generated from the average data of 15 meteorological stations around the national park from 2003 to 2018 using kriging interpolation.

Temperature reduces humidity and increases evaporation, resulting in the drying of flammable materials. Wildfires are expected to occur more frequently in areas with higher temperatures and lower rainfall⁵⁸. The 15-year maximum, minimum, and average temperature variables were generated using a simple kriging interpolation.

Humidity is an important atmospheric parameter for determining the potential of wildfires. In previous studies, relative humidity had the highest correlation with the occurrence and extent of wildfires⁵⁹. Therefore, this layer was obtained using the 15-year average of data from synoptic meteorological stations around the park and simple kriging interpolation.

Geological factors

The distance from the fault plays a vital role in the landslide, so that the fault movement can initiate a landslide on nearby slopes. This factor is an effective landslide occurrence by creating and reactivating,

fragmentation and shear, water infiltration into the domains, discontinuities around the fault, and erosion differences in the slopes ⁶⁰. To include the effect of faults in the models, the distance from the fault layer was created using the fault map of the region (published by the Geological Survey of Iran with a scale of 1: 100,000).

Different geological units have different degrees of sensitivity to landslide occurrence ⁶¹. The presence of loose and sensitive structures such as marl, clay, and deep aerated materials increases the occurrence of landslides in a particular area ⁶². The geological unit layer, which distinguishes five classes of lime, loess, marl, sandstone, and alluvium (from the Geological Survey of Iran with a scale of 1/100,000), was used in the modeling to account for geological units.

Soil types with different moisture contents, specific gravities, grain sizes, and structures have different sensitivities to hazards. Soil types have different effects on vegetation, soil moisture, and soil depth, which all together affect the occurrence of environmental hazards ⁶³. The study area is covered by four soil types: Alfisol, Entisols, Inceptisols, and Mollisols, with the largest area belonging to Entisols. Therefore, the soil category layer (from the Geological Survey of Iran) with a scale of 1:250,000 was used in this study. The soil types were coded as numbers 1 to 4.

Vegetation and landuse factors

Normalized differential vegetation index (NDVI) layer from the 2017 Landsat satellite imagery with a 30 × 30 meter resolution and a 1:25,000 land use map was prepared by the environmental protection organization were used in the modeling. The NDVI is one of the most well-known vegetation indices used to characterize the health of plants using satellite imagery data ⁶⁴. This index shows the surface cover of plants and their density in the images, the value of which is between - 1 and + 1, and was prepared using Eq. 3 ⁶⁵.

$$\text{NDVI} = (\text{NIR}-\text{VIS}) / (\text{NIR} + \text{VIS}) \quad (3)$$

In Eq. 3, VIS and NIR represent the spectral reflectance measurements acquired in the red (visible) and near-infrared regions, respectively.

Land use reflects the characteristics of the earth's surface and is closely related to the occurrence of environmental hazards ⁶⁶. Land use in national parks is vegetation with different densities. This type of land use has different functions in terms of slope instability. Vegetation reduces landslides in certain areas by regulating the permeability of surfaces and water flow. The type and density of vegetation also affect the onset and intensity of wildfires due to changes in the amount of fuel ⁶⁷.

Hazards maps

The locations of previous environmental hazards were available as descriptive reports in the central office of the national park. We digitized the locations of the events using GPS sensors while being helped by national park officers. In this regard, 76 fire hazards, 69 landslide hazards, and 44 flood hazard areas

were recorded. The hazard layers were converted into Boolean maps with a pixel size of 30×30 m, where the true value of each pixel represented the occurrence of a particular hazard type.

Modelin using Machine Learning

Machine learning (ML) refers to a branch of algorithms and statistical models that enable computer systems to learn from the data and perform a specific task without explicit instructions from the user. In this study, we used random forest (RF) ⁶⁸, boosted regression tree (BRT) ⁶⁹, support vector machine (SVM) ⁷⁰, and generalized additive model (GAM) ⁷¹ for multi-hazard modeling.

After determining the influencing factors, the importance of each of the three hazards of flooding, landslides, and wildfires was examined using the maximum entropy (ME) method in MaxEnt software, and less important factors were removed from the analysis. This model uses real data, discrete and continuous data, interaction of factors, and correlation of hazard probability to locations of occurrence, variables that are sufficient input to maximize the accuracy of the model, prediction and participation percentage ⁷².

To implement the ML models based on RF, BRT, SVM, and GAM, hazard maps were converted into pixels of 30×30 m, in which 48,574 pixels of wildfire, 1,999 pixels of flood, and 141 pixels of landslides were obtained. Of these pixels, 70%, including 34,001 of wildfire, 1,399 pixels of flood, and 99 pixels of landslide were used for training, and the remaining pixels were left aside to evaluate the accuracy of the models. These models were developed in R programming language using "randomForest," "rminer," "gbm," "dismo," and "GRASP" packages.

After training these models in R for observational data and generalization of the modeling process to the whole region in the form of black box modeling, the output was prepared as a csv file, which has two columns, where the first column indicates the probability of a hazard event and pixels, and the second column indicates the values of hazard values for each pixel of the entire study area. Hazard maps were generated by importing the csv file in GIS software.

Evaluation

After preparing environmental hazard maps using the proposed models, their accuracy was assessed to determine the optimal model. For evaluation, we used 30% of the data that were not used in the training. Receiver operating characteristic (ROC) was calculated as the evaluation metric in the IDRISI Taiga software. The ROC is a relative factor used to evaluate the accuracy of a model by comparing a continuous image with Boolean results that show exactly where it happened and determine its accuracy in the class ⁷³. The ROC statistic is the area under the curve calculated using Eq. 4 ⁷⁴:

$$AreaUnderCurve = \sum_{i=1}^n ([x_{i+1} + 1 - x_i] * [y_i + (y_{i+1} - y_i)/2]) \quad (4)$$

In Eq. 4, X_i is the percentage of incorrect pixels for i , y_i is the percentage of correct pixels for i , and n is the number of pixels.

After preparing the hazard maps for the wildfire, flood, and landslide hazards, these maps were converted to Boolean in the GIS software and integrated with each other, and a multi-hazard map was obtained for each model with combined and reclass maps, where each map had eight classes from the location of hazards. Finally, the area of each hazard class was calculated, the risk overlap for the models was determined, and the average was obtained.

Results

In the present study, ME was used to determine the importance of factors in the occurrence of environmental hazards. The results of the Jack Knife test, highlighting the essential variables in the occurrence of environmental hazards in GNP, are presented in Figs. 2–4. After determining the important factors affecting the hazards using this algorithm, the variables that were of low importance were removed, and then other machine learning models were implemented. Therefore, the slope and aspect factors in the wildfire model and geology factor in the landslide model were eliminated.

Figure 5 shows the values obtained from the accuracy of the hazard maps in the form of ROC, obtained from the 30% test data, which were not entered into the model. The results of the evaluation in almost all cases showed ROC values higher than 0.9. The results of the BRT are presented below because of their relatively better performance.

Figures 6 to 8 present the results of the effect of influencing factors on the outputs of the BRT model for wildfire, flood, and landslide hazards, respectively. These figures show that the most significant impact on fire hazard is related to the distance from the village (22.2%) and at distances less than 8 km from them. The flood hazard is associated with the elevation factor (40%) for elevations less than 700 m. For landslide hazards, the distance from the transit road (26.7%) at distances of less than 1 km has the highest importance.

Figure 9 shows the environmental hazard maps derived from the BRT model. In the fire hazard map (fig 9a), the distances between the transit road and the northern areas of the national park were at high risk. The flood risk map (fig 9b) also shows the areas of the Maderso River in the center of the national park and the villages in the south, southeast, and northeast, with a high risk of flooding. The landslide map (fig 9c) also identified areas of the one-kilometer buffer around the transit road and the northern areas as high-risk areas. A multi-hazard map (fig 9d) was obtained from the collection of hazard maps, which also showed high-risk areas in most of the central and northwestern regions of the national park.

The fuzzy hazard maps were classified into five classes: very high, high, medium, low, and very low. Figure 10 shows the percentage of the area occupied by each class. According to this figure, 7% of the area has a high risk of wildfire, 2.5% flooding, and 15% landslides. The multi-hazard map also marks 2.3% of GNP as areas with a very high or high probability of all three hazards together.

The output hazard maps of wildfires, floods, and landslides were combined, and a multi-hazard map was generated (Fig. 11). This map consists of eight classes, including the areas with the probability of one

hazard and the areas where there are possibilities for two or three hazards. The areas related to each class are presented in Table 1.

Table 1
Area (ha) of environmental hazards resulting from machine learning modeling

Hazards	ME	SVM	GAM	BRT	RF	Mean
Landslide	2212	17182	8681	12932	15868	9560
Fire	17550	4854	12073	5268	3938	8145
Flood	2852	1537	5827	1402	1331	2777
Flood-Landslide	531	2051	3915	2242	2864	1957
Flood-Fire	1270	156	884	90	62	470
Fire-Landslide	1999	4506	5456	2191	1115	2563
Flood-Landslide-Fire	807	182	3231	479	247	827
No Hazards	61678	60070	50469	65933	65112	63709

Figure 12 shows AUC in ROC graph in the models used in environmental hazards modeling compared to the reality of hazards, so that all models have high accuracy and in general, the BRT model has the highest accuracy.

Discussion

Hazard management requires intermittent and frequent monitoring of hazard factors ⁷⁵; therefore, classification of wildfires, floods, and landslide hazard indicators is necessary for spatial hazard modeling ⁴⁴. The Jack Knife test in the ME model is a method of determining the importance of variables as an alternative method in predictive models that works by deleting independent variables, which is executed by the number of variables (n), in which an independent variable is deleted (n -1), and model evaluation is performed to determine the importance and position of each variable in the model ⁷⁶. The results of Jack Knife ME indicate the high importance of climatic and human factors in the occurrence of fire hazards in GNP; therefore, it can be said that these factors determine the occurrence of fire. Among these, the rainfall factor had the highest impact, and this study revealed that areas with rainfall less than 260 mm and more than 460 mm per year are most likely to have fires. Rain can affect the occurrence of fire in two ways. First, increasing the humidity in the area reduces the chance of fire. Therefore, the probability of wildfire in low rain areas is higher, which describes the increase in the risk of wildfire in areas with less than 260 mm yearly rain. Second, rainfall can affect the growth of annual and eventually perennial grass cover, which influences wildfire; in this case, the high probability of wildfire in areas with more than 460 mm annual rainfall can be attributed to the dense vegetation cover resulting from the abundance of water.

The graph of the average annual temperature shows the effect of increasing temperature and a significant increase in the probability of wildfire at temperatures above 17.5°C. Such a high temperature is observed in the northern and northeastern parts of the national park near the Qaraqum Deserts in the north of the national park. ⁷⁷ analyzed the synchronous patterns of wildfire hazards in the Golestan province forests and pointed to the adverse effects of hot desert winds from Qarahqum on wildfire events in the region. Results from the humidity variable also show that in areas with low humidity, and hence poor rangeland cover, the risk of wildfire increases ⁷⁸.

Vegetation factors are important variables that affect wildfires. The NDVI variable increases from west to east of the national park. So since the cover ranges from west to the center of the national park, which is the beginning of the Hyrcanian forests, there is an ecotonic cover with a variety of plant species and grass cover, the graph of this variable shows an increasing risk of wildfire in this range. In fact, the diversity and dense cover of annual and perennial forest in this area has increased the risk of wildfires. The most important land cover variable is related to shrub cover, which has been completely burned, and a low canopy forest is the least important in the occurrence of national park wildfires. This cover, despite being located at high elevation, is less vulnerable to human and therefore less exposed to wildfire.

Regarding the high importance of human agents in the possibility of wildfire, it should be said that despite the transit of the road inside the national park and the traffic of passengers and passers-by in this section, the probability of wildfire is very high within 100 m of the road so that with the slightest spark, including cigarette butts thrown by drivers around the road, which provides access to the side roads inside the national park and nature walkers, tourists, and even hunters have access to a depth of about 14 km from the national park, the probability of wildfire despite this road is still high up to this distance.

The 500-meter area of villages has a high probability of wildfire, which can be attributed to the agricultural lands of these villages, which are bordered by national parks. Farmers burn agricultural waste to replant their land. In this case, the control of this wildfire is out of reach and penetrates the national park ⁷⁸. However, at longer distances can be attributed to the interaction of other variables and the livelihoods of these villagers are in turn tied to the forests and pastures, and in this regard, the managerial and legal problems for managing the national park. In fact, in these villages, there is a conflict between human activities and nature protection, such as turning national park areas into farmland, poachers, and grazing livestock ³². Therefore, in some cases, these people take wildfire in natural areas in retaliatory actions due to legal problems, which can continue up to 15 km in the villages.

The results of the studies of ⁴⁷ and ⁷⁹ also confirm the high impact of climatic and human agents on the occurrence of wildfire, which is the same as the results of the present study.

The two variables of elevation above sea and distance from the transit road are the most important variables of flood and landslide hazards according to the ME results. One of the reasons for the importance of these variables in flood hazards is that the main river of the national park, Maderso, where most of the runoff created in the national park collects and exits the park, is located at the lowest

elevation of the area, which is also prone to flooding ⁴. Also, the transit road is located exactly on the same river, so that due to the twisted state of the river, several bridges have been created on this river along this road. This situation has caused losses in the event of rising river water and flooding damage to the passengers who travel on this route. The graph of how these variables affect also shows a significant decrease in this hazard at a distance of more than 300 m from the transit road. The elevation graph also shows a greater hazard of flooding at an elevation of less than 500 m above sea level, which is actually produced at high elevation if runoff conditions are provided and finally reach each other at the lowest point of the basin, causing flooding areas.

Because of the mountainous nature of the national park, the possibility of flooding in the flat areas located on the slopes and upstream is very small, because in the first place, the volume of runoff in these areas is low and there is no possibility of flooding and the gravity of the earth; in the upstream areas, the runoff speed is high in these areas, which causes water to pass through the flat areas and reduce the risk of flooding ^{80,81}.

The graph of how elevation affects shows the occurrence probability of landslide hazard at low elevations, such that elevations 700–800 m above sea level are the most likely to occur. Meanwhile, the graph of the road distances shows that approximately 500 m is the most hazardous area in terms of landslides. The point that can be added to these two important variables is that this road has almost exceeded the elevation range affecting this hazard. In fact, high elevations are less prone to landslides because they are rocky and inaccessible, and low elevations are more likely to occur because of man-made structures and road traffic located in these places. In fact, excavations and embankments that have taken place around roads have also caused instability of the slopes and have increased this hazard.

Because of the high impact of rainfall on the occurrence of flood hazards, it can be said that the amount and manner of rainfall play an important role in the formation of these areas; therefore, heavy rains can cause runoff production in a short time and the occurrence of this risk. In fact, production runoffs do not have the opportunity to penetrate the soil and increase flooding hazards.

Faults play an effective role in increasing the instability of slopes; therefore, seismic events are more intense near faults, and these faults in crushing, surrounding rocks, and water penetration into the crushed masses ⁸² Therefore, by moving away from them, the probability of slope instability decreases.

The history of landslide hazard studies also shows the importance of road distance, elevation, slope, and distance from faults in the occurrence of this hazard ^{6,83,84}, which is consistent with the results of this study.

The largest area related to the overlap of hazards with 3% of the area is related to the two wildfire and landslide hazards, which the study of ⁶ confirmed the simultaneous occurrence of these two hazards.

The highest hazard area, accounting for 17.2% of the area, is related to landslide hazards. This hazard occurred because of its location on different faults and the mountainous nature of the national park and

its high slope, as well as the existence of roads and waterways. Meanwhile, the fire hazard was more than 14%, and floods accounted for 7.1% of the national park area. In fact, the high slope and mountainous nature of the national park have caused production runoff to exit and reduce the flood risk area. Since more than 47% of the national park wildfire area is related to the impact of the transit road ⁸⁵, it can be said that approximately 14% about 6.6% of the national park wildfire area is the route of this road. In addition to the excavations and embankments that have taken place around it, have caused instability of the slopes, which has led to landslide hazards, and this has caused an increased overlap of the hazards of landslides and wildfires. With the loss of vegetation by wildfire, it can be considered an agent in creating a landslide hazard, as ⁸⁶ also pointed out the effect of forest fires in increasing the risk of landslides in mountainous areas.

ROC is a common method for evaluating the accuracy of models; therefore, in this method, the area under curve (AUC) analysis from the value 0.5 to 1 is used to evaluate the model. The value of this characteristic in the BRT model for the fire hazard map was 0.97, flood risk was 0.98, and landslide hazard was 0.93, indicating the excellent prediction of this model in identifying hazards. In fact, the high stability of this model against data turbulence has led to a reduction in its forecasting error, as many studies have shown that it performs well in data mining, classification, forecasting, and cluster analysis⁸⁷. The results of evaluating the accuracy of this model using 30% of the data that were not included in the model and the use of value range analysis under the curve in the detection of relative performance indicate the excellent performance of this model in identifying potential hazards so that the results Studies ^{88, 89, 90, 86, 88, 91 6, 92}. In a study conducted ⁹³ on landslide hazard assessment using five models of machine learning techniques, the highest accuracy was obtained for the BRT model.

If the AUC value is between 0.9 to 1 excellent prediction, 0.8 to 0.9 very good, 0.7 to 0.8 good, 0.6 to 0.7 moderate and less than 0.6 ⁸⁶. Therefore, other machine learning techniques also exhibit excellent performance. The SVM model has the complex advantages of nonlinear and insensitive relationships to parasites; thus, the study records also show the high accuracy of this model in assessing landslide and flood risks ^{44,94}. One of the most important features of the GAM is its similarity to machine learning methods and the possibility of easy interpretation of this algorithm⁹⁵ and which has a great ability to analyze data and determine nonlinear relationships. Different variables ⁹⁶. The results of the present study show the excellent accuracy of this model in risk assessment, which is the same as the study by ⁹⁷. In the RF model, data fragmentation with one variable is performed randomly, and this process continues with other predictor variables in a suitable set to grow classification trees ⁷⁹. In fact, the high stability of this model against data turbulence has led to a reduction in its forecasting error, as many studies have shown that it performs well in data mining, classification, clusters, forecasting, and analysis ⁸⁷.

Finally, we compared the hazard map obtained in the present study and the distribution of important mammals in the national park, which indicates that the habitat of these mammals is endangered. Thus, the Mirzabaylo area, which is the habitat of deer, is at risk of flooding and the habitat of marals and leopards. There is a possibility that all three environmental hazards and the habitat of ewes are at risk of

fire and bears are at risk of landslides, so the hazard maps from the present study can be used to manage and protect the habitat of these important species. In addition, because of the location of the bridges built on the Madersu River and Cheshmeh-e-Zav, it has increased the vulnerability of these areas, and the path of these rivers should be laminated in these sections as much as possible.

According to the results of this study, the implementation of the transit road transfer policy outside the national park due to its high impact on all hazards is felt more than before, and it also suggests controlling these hazards by creating hazard monitoring systems and preventing the construction of human structures in high-risk areas, especially the Zaw area, which is being developed for tourism programs.

Conclusion

Every year, environmental hazards cause serious damage to GNP, which is one of the most important biosphere reserves in Iran; therefore, mapping the potential of these hazards and describing them simultaneously can be an essential tool for reducing and managing them. Therefore, in the present study, machine learning techniques were used to determine an accurate method for identifying susceptible areas, while the extent and effect of each variables were determined. The results show the importance of variables such as rainfall, temperature, presence of farmers and villagers, elevation, transit roads, and topographic humidity in the occurrence of hazards. Therefore, the elevation controls the hazards by controlling the type of erosion, concentration of human activities in the lower and middle elevations, and sediments in the lower part and hard structures in the highlands. The transit road at lower elevations and its growing traffic have increased the probability of fire and landslides in the surrounding areas. The results of accuracy assessment with ROC values higher than 0.9 show excellent performance of machine learning techniques. The BRT model showed the best result among other models. Environmental hazard maps estimated that, on average, 16.2% of the area had a high landslide potential, 14% had a high fire potential, and 7.2% had a high flood potential. Fire and landslide hazards had the highest overlap in terms of hazard. The outcomes shows that the existence of the transit road in the national park has strongly increased the probability of fire and landslide hazards.

References

- 1 Khorozyan, I., Soofi, M., Ghoddousi, A., Hamidi, A. K. & Waltert, M. The relationship between climate, diseases of domestic animals and human-carnivore conflicts. *Basic and Applied Ecology* **16**, 703-713 (2015).
- 2 Sarikhani, N. & Majnonian, B. Guideline for production of forest roads project. *J. Plan. Bud. Org. Pub* **148**, 178-192 (1999).
- 3 Shokri, M., Safaian, N. & Atrakchali, A. Investigation of the Effects of Fire on Vegetation Variations in Takhti-Golestan National Park. *Journal of Natural Environment (In persian)* **55**, 273-281 (2002).

- 4 Faramarzi, H., Hosseini, S. M., pourghasemi, H. R. & Farnaghi, M. Assessment and Zoning of Flood Risk in Golestan National Park. *Eco Hydrology (In persian)* **6**, 1055-1068 (2020).
- 5 Lan, H. *et al.* A web-based GIS for managing and assessing landslide data for the town of Peace River, Canada. *Natural Hazards and Earth System Sciences* **9**, 1433-1443 (2009).
- 6 Pourghasemi, H. R., Gayen, A., Panahi, M., Rezaie, F. & Blaschke, T. Multi-hazard probability assessment and mapping in Iran. *Science of the total environment* **692**, 556-571 (2019).
- 7 Mioc, D. *et al.* Web-GIS application for flood prediction and monitoring. *WIT Transactions on Ecology and the Environment* **118**, 145-154 (2008).
- 8 Guhathakurta, P., Sreejith, O. & Menon, P. Impact of climate change on extreme rainfall events and flood risk in India. *Journal of earth system science* **120**, 359-373 (2011).
- 9 Hennequin, T., Sørup, H. J. D., Dong, Y. & Arnbjerg-Nielsen, K. A framework for performing comparative LCA between repairing flooded houses and construction of dikes in non-stationary climate with changing risk of flooding. *Science of the total environment* **642**, 473-484 (2018).
- 10 Hao, W., Hao, Z., Yuan, F., Ju, Q. & Hao, J. Regional frequency analysis of precipitation extremes and its spatio-temporal patterns in the Hanjiang River Basin, China. *Atmosphere* **10**, 130 (2019).
- 11 Li, Y. *et al.* In situ reconstruction of long-term extreme flooding magnitudes and frequencies based on geological archives. *Science of the total environment* **670**, 8-17 (2019).
- 12 Lyubchich, V., Newlands, N. K., Ghahari, A., Mahdi, T. & Gel, Y. R. Insurance risk assessment in the face of climate change: Integrating data science and statistics. *Wiley Interdisciplinary Reviews: Computational Statistics* **11**, e1462 (2019).
- 13 Toda, L. L., Yokingco, J. C. E., Paringit, E. C. & Lasco, R. D. A LiDAR-based flood modelling approach for mapping rice cultivation areas in Apalit, Pampanga. *Applied Geography* **80**, 34-47 (2017).
- 14 Pham, B. T. *et al.* Flood risk assessment using deep learning integrated with multi-criteria decision analysis. *Knowledge-Based Systems* **219**, 106899 (2021).
- 15 Norouzi, G. & Taslimi, M. The impact of flood damages on production of Iran's Agricultural Sector. *Middle East J Sci Res* **12**, 921-926 (2012).
- 16 Ahmadalipour, A., Moradkhani, H., Castelletti, A. & Magliocca, N. Future drought risk in Africa: Integrating vulnerability, climate change, and population growth. *Science of the Total Environment* **662**, 672-686 (2019).
- 17 Dai, F. & Lee, C. Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology* **42**, 213-228 (2002).

- 18 Guzzetti, F., Reichenbach, P., Cardinali, M., Galli, M. & Ardizzone, F. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology* **72**, 272-299 (2005).
- 19 Turetsky, M. R. *et al.* Recent acceleration of biomass burning and carbon losses in Alaskan forests and peatlands. *Nature Geoscience* **4**, 27-31 (2011).
- 20 Zaitsev, A. S., Gongalsky, K. B., Malmström, A., Persson, T. & Bengtsson, J. Why are forest fires generally neglected in soil fauna research? A mini-review. *Applied soil ecology* **98**, 261-271 (2016).
- 21 Shearman, T. M., Varner, J. M., Hood, S. M., Cansler, C. A. & Hiers, J. K. Modelling post-fire tree mortality: Can random forest improve discrimination of imbalanced data? *Ecological Modelling* **414**, 108855 (2019).
- 22 Gayen, A., Pourghasemi, H. R., Saha, S., Keesstra, S. & Bai, S. Gully erosion susceptibility assessment and management of hazard-prone areas in India using different machine learning algorithms. *Science of the total environment* **668**, 124-138 (2019).
- 23 Karimi, A., Abdollahi, S., Ostad-Ali-Askari, K., Eslamian, S. & Singh, V. P. Predicting fire hazard areas using vegetation indexes, case study: forests of Golestan province, Iran. *Journal of Geography and Cartography* **2** (2019).
- 24 Wang, C. *et al.* Machine learning-based regional scale intelligent modeling of building information for natural hazard risk management. *Automation in Construction* **122**, 103474 (2021).
- 25 Keesstra, S. *et al.* Soil-related sustainable development goals: Four concepts to make land degradation neutrality and restoration work. *Land* **7**, 133 (2018).
- 26 Ding, Q., Chen, W. & Hong, H. Application of frequency ratio, weights of evidence and evidential belief function models in landslide susceptibility mapping. *Geocarto international* **32**, 619-639 (2017).
- 27 Allard-Duchêne, A., Pothier, D., Dupuch, A. & Fortin, D. Temporal changes in habitat use by snowshoe hares and red squirrels during post-fire and post-logging forest succession. *Forest Ecology and Management* **313**, 17-25 (2014).
- 28 Jaiswal, R. K., Mukherjee, S., Raju, K. D. & Saxena, R. Forest fire risk zone mapping from satellite imagery and GIS. *International Journal of Applied Earth Observation and Geoinformation* **4**, 1-10 (2002).
- 29 Whittier, T. R. & Gray, A. N. Tree mortality based fire severity classification for forest inventories: a Pacific Northwest national forests example. *Forest Ecology and Management* **359**, 199-209 (2016).
- 30 Djamali, M. *et al.* Modern pollen rain-vegetation relationships along a forest-steppe transect in the Golestan National Park, NE Iran. *Review of Palaeobotany and Palynology* **153**, 272-281 (2009).

- 31 Rad, E. B., Mesdaghi, M., Ahmad, N. & Abdullah, M. Nutritional quality and quantity of available forages relative to demand: a case study of the goitered gazelles of the Golestan National Park, Iran. *Rangelands* **37**, 68-80 (2015).
- 32 Ghoddousi, S., Pintassilgo, P., Mendes, J., Ghoddousi, A. & Sequeira, B. Tourism and nature conservation: A case study in Golestan National Park, Iran. *Tourism management perspectives* **26**, 20-27 (2018).
- 33 Van Westen, C., Van Asch, T. W. & Soeters, R. Landslide hazard and risk zonation—why is it still so difficult? *Bulletin of Engineering geology and the Environment* **65**, 167-184 (2006).
- 34 Schlögl, M. & Matulla, C. Potential future exposure of European land transport infrastructure to rainfall-induced landslides throughout the 21st century. *Natural Hazards and Earth System Sciences* **18**, 1121-1132 (2018).
- 35 Hu, Q., Zhou, Y., Wang, S. & Wang, F. Machine learning and fractal theory models for landslide susceptibility mapping: Case study from the Jinsha River Basin. *Geomorphology* **351**, 106975 (2020).
- 36 Zhu, A.-X. *et al.* An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic. *Geomorphology* **214**, 128-138 (2014).
- 37 Flannigan, M. D., Stocks, B. J. & Wotton, B. M. Climate change and forest fires. *Science of the total environment* **262**, 221-229 (2000).
- 38 Syphard, A. D. *et al.* Predicting spatial patterns of fire on a southern California landscape. *International Journal of Wildland Fire* **17**, 602-613 (2008).
- 39 Chen, L. *et al.* Integrating expert opinion with modelling for quantitative multi-hazard risk assessment in the Eastern Italian Alps. *Geomorphology* **273**, 150-167 (2016).
- 40 Chandra, S. in *Geo-information for Disaster management* 1239-1254 (Springer, 2005).
- 41 Fernandez, D. & Lutz, M. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology* **111**, 90-98 (2010).
- 42 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. *Journal of Hydrology* **373**, 15-23 (2009).
- 43 Moore, I. D., Grayson, R. & Ladson, A. Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological processes* **5**, 3-30 (1991).
- 44 Tehrany, M. S., Jones, S. & Shabani, F. Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques. *Catena* **175**, 174-192 (2019).

- 45 Thach, N. N. *et al.* Spatial pattern assessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: A comparative study. *Ecological informatics* **46**, 74-85 (2018).
- 46 Hong, H., Naghibi, S. A., Dashtpajardi, M. M., Pourghasemi, H. R. & Chen, W. A comparative assessment between linear and quadratic discriminant analyses (LDA-QDA) with frequency ratio and weights-of-evidence models for forest fire susceptibility mapping in China. *Arabian Journal of Geosciences* **10**, 167 (2017).
- 47 Sevinc, V., Kucuk, O. & Goltas, M. A Bayesian network model for prediction and analysis of possible forest fire causes. *Forest Ecology and Management* **457**, 117723 (2020).
- 48 Nourani, V., Ejlali, R. G. & Alami, M. T. Spatiotemporal groundwater level forecasting in coastal aquifers by hybrid artificial neural network-geostatistics model: a case study. *Environmental Engineering Science* **28**, 217-228 (2011).
- 49 Ghobadi, G. J., Gholizadeh, B. & Dashliburun, O. M. Forest fire risk zone mapping from geographic information system in Northern Forests of Iran (Case study, Golestan province). *International Journal of Agriculture and Crop Sciences* **4**, 818-824 (2012).
- 50 Schicker, R. & Moon, V. Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. *Geomorphology* **161**, 40-57 (2012).
- 51 Dai, F., Lee, C., Li, J. & Xu, Z. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environmental Geology* **40**, 381-391 (2001).
- 52 Yerro, A., Soga, K. & Bray, J. Runout evaluation of Oso landslide with the material point method. *Canadian Geotechnical Journal* **56**, 1304-1317 (2019).
- 53 Sofia, G. & Nikolopoulos, E. Floods and rivers: a circular causality perspective. *Scientific reports* **10**, 1-17 (2020).
- 54 Agee, J. K. *et al.* The use of shaded fuelbreaks in landscape fire management. *Forest ecology and management* **127**, 55-66 (2000).
- 55 Sullivan-Wiley, K. A. & Gianotti, A. G. S. Risk perception in a multi-hazard environment. *World Development* **97**, 138-152 (2017).
- 56 Gariano, S. L. & Guzzetti, F. Landslides in a changing climate. *Earth-Science Reviews* **162**, 227-252 (2016).
- 57 Cortez, P. & Morais, A. d. J. R. A data mining approach to predict forest fires using meteorological data. (2007).

- 58 Hoffmann, W. A., Orthen, B. & Nascimento, P. K. V. d. Comparative fire ecology of tropical savanna and forest trees. *Functional Ecology* **17**, 720-726 (2003).
- 59 Jaafari, A., Gholami, D. M. & Zenner, E. K. A Bayesian modeling of wildfire probability in the Zagros Mountains, Iran. *Ecological informatics* **39**, 32-44 (2017).
- 60 Kamranzad, F., MohselAfshar, E., Mojarab, M. & Memarian, H. Landslide Hazard Zonation in Tehran Province Using Data-Driven and AHP Methods. *Geoscience (In persian)* **25**, 101-114 (2016).
- 61 Duman, T., Can, T., Gokceoglu, C., Nefeslioglu, H. & Sonmez, H. Application of logistic regression for landslide susceptibility zoning of Cekmece Area, Istanbul, Turkey. *Environmental Geology* **51**, 241-256 (2006).
- 62 Cerri, R. I., Reis, F. A., Gramani, M. F., Giordano, L. C. & Zaine, J. E. Landslides Zonation Hazard: relation between geological structures and landslides occurrence in hilly tropical regions of Brazil. *Anais da Academia Brasileira de Ciências* **89**, 2609-2623 (2017).
- 63 Zhuo, L. *et al.* Evaluation of remotely sensed soil moisture for landslide hazard assessment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12**, 162-173 (2019).
- 64 Chuvieco, E. *et al.* Combining NDVI and surface temperature for the estimation of live fuel moisture content in forest fire danger rating. *Remote Sensing of Environment* **92**, 322-331 (2004).
- 65 Nampak, H., Pradhan, B. & Abd Manap, M. Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *Journal of Hydrology* **513**, 283-300 (2014).
- 66 Vilar, L., Woolford, D. G., Martell, D. L. & Martín, M. P. A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. *International Journal of Wildland Fire* **19**, 325-337 (2010).
- 67 Cantarello, E. *et al.* Simulating the potential for ecological restoration of dryland forests in Mexico under different disturbance regimes. *Ecological Modelling* **222**, 1112-1128 (2011).
- 68 Breiman, L. Random forests. *Machine learning* **45**, 5-32 (2001).
- 69 Elith, J., Leathwick, J. R. & Hastie, T. A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802-813 (2008).
- 70 Pouteau, R., Meyer, J.-Y. & Stoll, B. A SVM-based model for predicting distribution of the invasive tree *Miconia calvescens* in tropical rainforests. *Ecological modelling* **222**, 2631-2641 (2011).
- 71 Guisan, A., Edwards Jr, T. C. & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling* **157**, 89-100 (2002).
- 72 Convertino, M., Troccoli, A. & Catani, F. Detecting fingerprints of landslide drivers: a MaxEnt model. *Journal of Geophysical Research: Earth Surface* **118**, 1367-1386 (2013).

- 73 Rossiter, D. & Loza, A. Analyzing land cover change with logistic regression in R. *University of Twente, Faculty of Geo-Information Science & Earth Observation (ITC), Enschede (NL)* (2012).
- 74 Schneider, L. C. & Pontius Jr, R. G. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment* **85**, 83-94 (2001).
- 75 Tingsanchali, T. & Karim, F. Flood-hazard assessment and risk-based zoning of a tropical flood plain: case study of the Yom River, Thailand. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques* **55**, 145-161 (2010).
- 76 Baldwin, R. A. Use of maximum entropy modeling in wildlife research. *Entropy* **11**, 854-866 (2009).
- 77 Rahimi, D. & Khadem, S. Analysis Synoptic Patterns for Forest Fires Risk in Northern of Iran. *Natural Environmental Hazards (in persian)* **7**, 19-36 (2018).
- 78 Faramarzi, H., Hosseini, S. M., Pourghasemi, H. R. & Farnaghi, M. Forest fire spatial modelling using ordered weighted averaging multi-criteria evaluation. *Journal of Forest Science* **67**, 87-100 (2021).
- 79 Collins, L., Griffioen, P., Newell, G. & Mellor, A. The utility of Random Forests for wildfire severity mapping. *Remote Sensing of Environment* **216**, 374-384 (2018).
- 80 Termeh, S. V. R., Kornejady, A., Pourghasemi, H. R. & Keesstra, S. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Science of the Total Environment* **615**, 438-451 (2018).
- 81 Khosravi, K., Pourghasemi, H. R., Chapi, K. & Bahri, M. Flash flood susceptibility analysis and its mapping using different bivariate models in Iran: a comparison between Shannon's entropy, statistical index, and weighting factor models. *Environmental monitoring and assessment* **188**, 1-21 (2016).
- 82 Abedini, M. & Tulabi, S. Assessing LNRF, FR, and AHP models in landslide susceptibility mapping index: a comparative study of Nojian watershed in Lorestan province, Iran. *Environmental Earth Sciences* **77**, 1-13 (2018).
- 83 Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J. D. & Chacón, J. Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: application to the river Beiro basin (Spain). *Natural Hazards and Earth System Sciences* **12**, 327-340 (2012).
- 84 Pradhan, B. & Lee, S. Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environmental Modelling & Software* **25**, 747-759 (2010).
- 85 Faramarzi, H., Hosseini, S. M., Pourghasemi, H. R. & Farnaghi, M. Evaluation of the Asian Highway Role on Fire Golestan National Park in GIS Environment. *Wood and Forest science and technology (In*

persian) **25**, 33-48 (2019).

86 Pourghasemi, H. R. & Kerle, N. Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. *Environmental earth sciences* **75**, 185 (2016).

87 Cutler, D. R. *et al.* Random forests for classification in ecology. *Ecology* **88**, 2783-2792 (2007).

88 Rahmati, O. *et al.* Multi-hazard exposure mapping using machine learning techniques: A case study from Iran. *Remote Sensing* **11**, 1943 (2019).

89 Thüring, T., Schoch, M., van Herwijnen, A. & Schweizer, J. Robust snow avalanche detection using supervised machine learning with infrasonic sensor arrays. *Cold Regions Science and Technology* **111**, 60-66 (2015).

90 Pozdnoukhov, A., Purves, R. S. & Kanevski, M. Applying machine learning methods to avalanche forecasting. *Annals of Glaciology* **49**, 107-113 (2008).

91 Arabameri, A. *et al.* Novel ensembles of COPRAS multi-criteria decision-making with logistic regression, boosted regression tree, and random forest for spatial prediction of gully erosion susceptibility. *Science of the Total Environment* **688**, 903-916 (2019).

92 Shabani, S., Pourghasemi, H. R. & Blaschke, T. Forest stand susceptibility mapping during harvesting using logistic regression and boosted regression tree machine learning models. *Global Ecology and Conservation* **22**, e00974 (2020).

93 Bui, D. T., Tuan, T. A., Klempe, H., Pradhan, B. & Revhaug, I. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **13**, 361-378 (2016).

94 Pourghasemi, H. R., Jirandeh, A. G., Pradhan, B., Xu, C. & Gokceoglu, C. Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. *Journal of Earth System Science* **122**, 349-369 (2013).

95 Goetz, J. N., Guthrie, R. H. & Brenning, A. Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology* **129**, 376-386 (2011).

96 Hanspach, J., Kühn, I., Pompe, S. & Klotz, S. Predictive performance of plant species distribution models depends on species traits. *Perspectives in Plant Ecology, Evolution and Systematics* **12**, 219-225 (2010).

97 Abeare, S. Comparisons of boosted regression tree, GLM and GAM performance in the standardization of yellowfin tuna catch-rate data from the Gulf of Mexico lonline [sic] fishery. (2009).

Figures

Figure 1

Map of GNP in the northeast of Iran with the location of sampling sites and nearby villages

Figure 2

Jackknife of regularized training gain for fire

Figure 3

Jackknife of regularized training gain for flood

Figure 4

Jack Knife of regularized training gain for Landslide

Figure 5

Graph the ROC value obtained by evaluating the accuracy of machine learning models

Figure 6

The importance and form of the effect of factors on fire hazard using BRT model

Figure 7

The importance and form of the effect of factors on flood hazard using BRT model

Figure 8

The importance and form of the effect of factors on landslide hazard using BRT model

Figure 9

Maps of environmental hazards resulting from BRT; a. wildfire model, b.flood model, c.landslide model and d.multihazards model.

Figure 10

Percentage diagram of the classification area of hazards from the BRT

Figure 11

Reclassification multi hazards map of BRT model

Figure 12

ROC graphs, a.Fire models; b.Flood models; c.Landslide Models