

Article

Comparisons of Diverse Machine Learning Approaches for Wildfire Susceptibility Mapping

Khalil Gholamnia ¹, Thimmaiah Gudiyangada Nachappa ² , Omid Ghorbanzadeh ^{2,*}  and Thomas Blaschke ² 

¹ Department of Remote Sensing and GIS, University of Tabriz, Tabriz 5166616471, Iran; khalil.gh3@gmail.com

² Department of Geoinformatics–Z_GIS, University of Salzburg, 5020 Salzburg, Austria; thimmaiah.gudiyangadanachappa@sbg.ac.at (T.G.N.); Thomas.Blaschke@sbg.ac.at (T.B.)

* Correspondence: omid.ghorbanzadeh@stud.sbg.ac.at

Received: 7 March 2020; Accepted: 31 March 2020; Published: 10 April 2020



Abstract: Climate change has increased the probability of the occurrence of catastrophes like wildfires, floods, and storms across the globe in recent years. Weather conditions continue to grow more extreme, and wildfires are occurring quite frequently and are spreading with greater intensity. Wildfires ravage forest areas, as recently seen in the Amazon, the United States, and more recently in Australia. The availability of remotely sensed data has vastly improved, and enables us to precisely locate wildfires for monitoring purposes. Wildfire inventory data was created by integrating the polygons collected through field surveys using global positioning systems (GPS) and the data collected from the moderate resolution imaging spectrometer (MODIS) thermal anomalies product between 2012 and 2017 for the study area. The inventory data, along with sixteen conditioning factors selected for the study area, was used to appraise the potential of various machine learning (ML) methods for wildfire susceptibility mapping in Amol County. The ML methods chosen for this study are artificial neural network (ANN), dmne regression (DR), DM neural, least angle regression (LARS), multi-layer perceptron (MLP), random forest (RF), radial basis function (RBF), self-organizing maps (SOM), support vector machine (SVM), and decision tree (DT), along with the statistical approach of logistic regression (LR), which is very apt for wildfire susceptibility studies. The wildfire inventory data was categorized as three-fold, with 66% being used for training the models and 33% being used for accuracy assessment within three-fold cross-validation (CV). Receiver operating characteristics (ROC) was used to assess the accuracy of the ML approaches. RF had the highest accuracy of 88%, followed by SVM with an accuracy of almost 79%, and LR had the lowest accuracy of 65%. This shows that RF is better suited for wildfire susceptibility assessments in our case study area.

Keywords: Bushfire; artificial neural network (ANN); dmne regression (DR); DM neural; least angle regression (LARS); multi-layer perceptron (MLP); random forest (RF); radial basis function (RBF); logistic regression (LR); self-organizing maps (SOM); support vector machine (SVM); decision tree (DT); MODIS; k-fold cross-validation

1. Introduction

Forests are considered crucial natural resources that play an integral function in preserving the ecological equilibrium of the environment and shielding one-third of the earth. According to FAO, forests across the world inhabit an area of about 4000 million ha, which is approximately 30% of the earth's total surface area [1]. Forests play a vital role in the production of oxygen and purifying the environment [2]. Ecological health is measured by the state and well-being of the forest which are true signs of the condition of the region. In addition, forests have economic and social importance and play important roles in the existence of all living things on planet Earth. Forests also regulate the climate

and carbon cycle, which includes weather alterations [3]. Forest ecosystems account for around 66% of total terrestrial carbon and are exceedingly important to the carbon budget globally [4]. However, in recent years, forest ecosystems have been gradually becoming endangered by fires across the world, which have predominantly been triggered by anthropogenic factors [5].

Fires are a devastating ecological feature that have several adverse consequences in diverse ways, including with respect to life, health, the economy, and demolition of natural habitats [6]. In recent years, global warming and industrialization coupled with human interference have considerably increased the occurrence and severity of forest fires in many parts of the world. Not all wildfires are disasters; rather, when a wildfire affects the community or ecosystem it becomes a disaster, and these effects can be rather substantial [7]. Forest fires embody the most life-threatening form of devastation and widespread destruction of forests, infrastructures, and natural habitats. Severe fire seasons have resulted in considerable bearings on ecosystems, infrastructures, and the lives of humans for Greece in 2007 and 2018, for Australia in 2009, 2013 and 2019, for the USA in 2007, 2013, 2016 and 2018, for Siberia in 2019, for Canada in 2018 and 2019, and for the Amazon rainforest in 2019 [8]. Forest fires can be a natural or manmade hazard, and are categorized within the cluster of climatological disasters. With reference to the destruction and concerns, in addition to the magnitude of destruction each year, forest fires indicate a global issue rather than a local problem. To tackle wildfires, the involvement of all government and civic organizations, along with institutions and all units of the community, have to be coordinated in order to prevent and extinguish forest fires. Forest fires have become multifaceted both locally and globally, greatly impacting ecosystems, landscapes, people and economies [9]. In Iran, forests are valued as a substantial resource that supports the local economy, in the sense that products are harvested from forests, and supports local recreation activities. The majority of forests in northern Iran are located in the mountainous regions, and these act as barriers against natural hazards such as rockfalls and erosion [10]. In northern Iran, the frequency of wildfires has increased, along with the intensities. Northern Iran, in particular, has approximately 1.2 million ha, of which more than 300 ha are burnt annually by forest fires [11]. This mainly affects the younger trees at the ground level, rather than the older ones, and the regeneration of trees takes a long time, which has adverse impacts like deforestation and desertification. Despite the frequent forest fires in the region, there have been no comprehensive studies in the region, which could help in mitigating the wildfires. Thus, this study regarding forest fires in Northern Iran provides a comprehensive and detailed assessment of the methodologies and susceptibility of wildfires.

Remote sensing and the geographic information system (GIS) are significant tools for wildfire analysis and susceptibility mapping. Natural hazards like landslides, floods, earthquakes, and wildfires have been assessed using diverse methodologies like analytical hierarchical process (AHP) [12], frequency ratio (FR) [13,14], support vector machines (SVM) [15], analytical network process (ANP) [16], random forest (RF) [17], Dempster-Shafer [14,18,19], artificial neural networks (ANN) [20], decision tree (DT) [21], logistic regression (LR) [22–24] convolution neural network (CNN) [20,25] and evidence belief function (EBF) [26]. The results of some of these methodologies and models were evaluated and compared to each other. For example, in a study done on tallgrass prairie in the US, the results from both ANN and stepwise multiple linear regression methods were compared in estimating fuel moisture content for wildfire estimation [27]. Their accuracy assessment compared the results of ANN with that of a stepwise linear regression (SLR) for estimating fuel moisture content for wildfire estimation. Their accuracy assessment metrics revealed that the ANN presented higher accuracy than the SLR.

Thus, in recent years, ML methods have proved to be better methods for natural hazard assessments [28–31]. The wide use of ML methods is based on two factors: (1) the availability of inventory datasets for both training and test datasets, and (2) advancements in computing platforms. Each model has its own merits and demerits and depends on the availability and amount of the training data. However, there is no evidence that a particular model is best for certain hazards, as this also depends on the study area and availability of data for that particular region [32].

The main goal of this study is to assess various machine learning approaches like RF, ANN, DT, least angle regression (LARS), SVM, multi-layer perceptron (MLP), self-organizing maps (SOM), DM neural, and radial basis function (RBF) along with a statistical approach of logistic regression (LR) for wildfire susceptibility mapping for Amol County in Northern Iran. This study is quite relevant to the present scenario and bodes well for the research of assessing the best approach for northern Iran. Some of the mentioned ML approaches were used for wildfire susceptibility mapping in our study area. In this study, we compare the most common ML approaches.

2. Study Area

Amol County is part of the Mazandaran province in Iran, with a total area of 4374 km². The total population is around 343,747 [6]. The study area is located on the southern coast of the Caspian Sea. The province of Mazandaran is at a strategic location and has borders with Russia, Azerbaijan, Turkmenistan, and Kazakhstan. The province is rich in natural resources, including vast areas of forests and large reservoirs of oil and natural gas. The study area selected was the forest area in Amol County, located in the central of Mazandaran Province, as shown in Figure 1. The total area covered is around 646 km² and is crucial for the growth of the local economy, which is mainly focused on animal husbandry and harvesting of forest products. The study area is quite popular with tourists due to the recreational activities available, mostly during the summer months. The elevation in the study area varies from 100 m in the valleys to 2500 m in the mountainous regions. The study area experienced several harsh wildfires in recent years, which has impacted more than twenty settlements and villages in the region. In recent years, the occurrences of wildfires have increased due to the impact of climate change in the region due to the increase in temperatures and droughts.

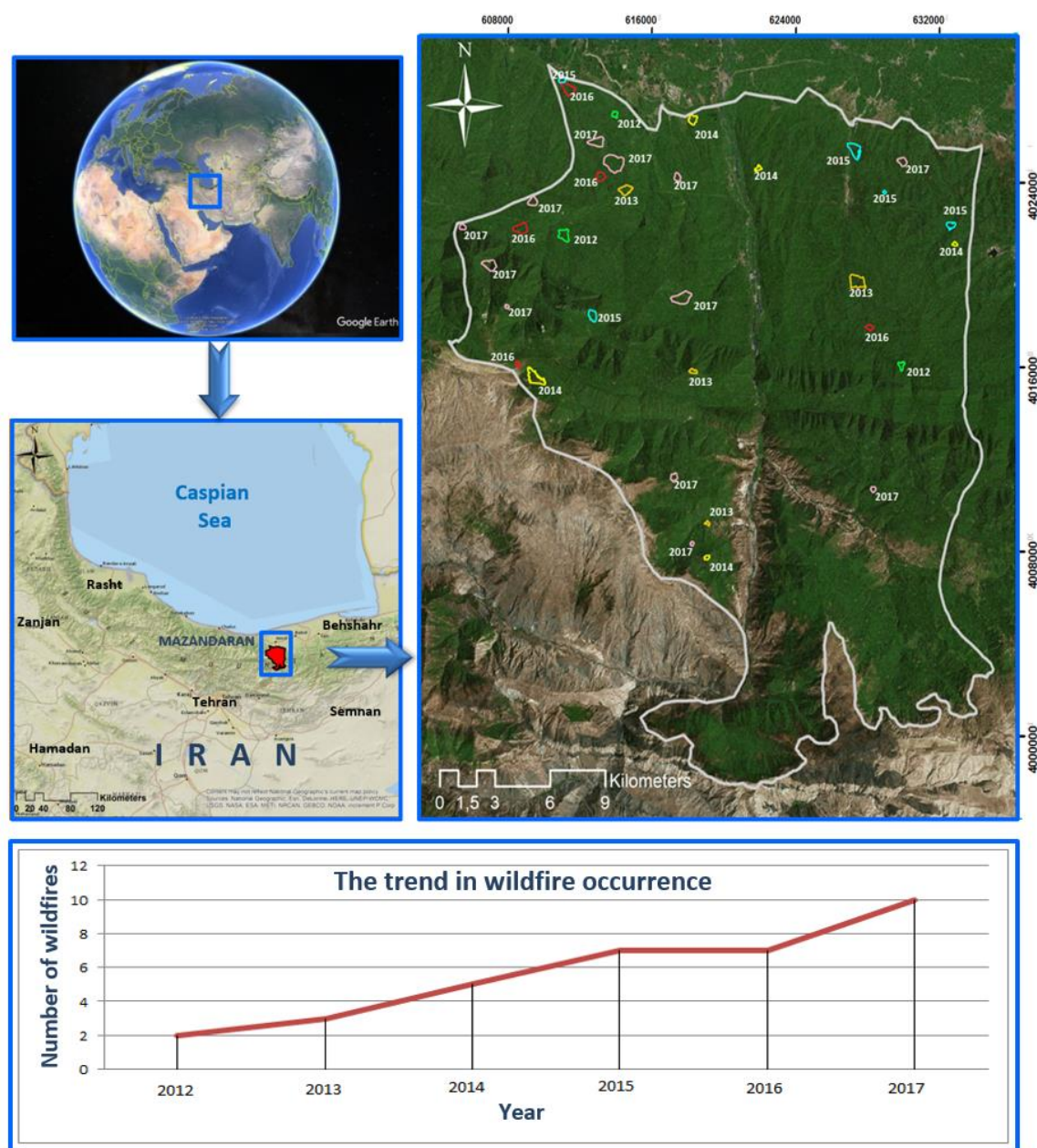


Figure 1. The location of the study area and the wildfire inventory data from 2012 to 2017 that was created from moderate-resolution imaging Spectroradiometer (MODIS) and field surveys.

3. Materials and Methods

3.1. Wildfire Inventory Data

Inventory data is crucial for training the model, and adequate and reliable inventory data is a prerequisite for machine learning approaches in particular. **Moderate-resolution imaging spectrometer (MODIS) is an instrument that is in operation on the Terra as well as Aqua spacecraft.** This instrument has electromagnetic spectral bands that can differ from visible bands to thermal infrared bands [33]. To generate wildfire inventory data, we used the GPS-derived data from the state wildlife organization and MODIS fire event data, which is available without restriction from 2012 to 2017 to precisely map the affected areas as polygons. The MODIS fire data includes information about the spatial distribution of fires and also their timestamp. The MODIS fire data is characterized by the exclusion of vegetation in the charred areas, deposits of charcoal and ash, and alteration in the vegetation structure. MODIS fire

products or thermal anomalies are primarily derived from MODIS 4 and 11 micrometer radiances, and we used MOD14, which has a temporal resolution of 5 min with level 2 processing. MODIS data on Terra and Aqua are attained from each platform twice every day at mid-latitudes. The fire recognition method is centered on the absolute recognition of the fire when the strength of the fire is sufficient to be detected. A total of 34 fire hotspots (17,420 pixels) were identified for the study area for the given period of time with a resolution of 1km which is appropriate to identify the burned area. The hotspots identified were enriched by means of the GPS data collected through field surveys for credibility. The wildfire inventory data was thus created with a combination of GPS data and data from MODIS. All the corrections were carried out using ArcGIS software. The dataset for the state wildlife collected through GPS surveys may not be complete and may have missed smaller fires in the region. This was completed based on the MODIS dataset, while the larger wildfires were incorporated from the state department, resulting in comprehensive inventory data for wildfires for the study area.

3.2. Conditioning Factors

The spatial probability of wildfire occurrence is the probability that the given region will be affected by wildfires based on the environmental conditions. These conditions can be geomorphological, topographical, meteorological, hydrological, and anthropological. These factors are known as conditioning factors and can highly influence the final susceptibility mapping; they are selected based on their relevance to the study area. It is pivotal to standardize the causative factors for the hazard in preparing for any of the natural hazard susceptibility mapping [13]. For this study, sixteen conditioning factors were selected for Amol County based on their relevance to the study area and the availability of the data. The conditioning factors were categorized based on topographic, hydrological, meteorological and anthropological factors, along with normalized difference vegetation index (NDVI) as the vegetation factor, as seen in Table 1.

Table 1. Categorization of conditioning factors selected for the study area.

Topographical	Hydrological	Meteorological	Anthropological	Vegetation
Slope (%)	Distance to Stream (m)	Annual Temperature (°C)	Land use	Normalized difference vegetation index (NDVI)
Slope Aspect	Annual Rainfall (mm)	Wind Effect	Distance to Village (m)	
Altitude (m)		Potential Solar Radiation	Distance to Road (m)	
Topographic Wetness Index (TWI)			Recreation Area (m)	
Landforms				
Plan Curvature (100/m)				

The ranges for the input conditioning factors were classified based on their relevance and importance to forest fires, the literature review, and expert experiences. See Table 2. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) (NASA, California Institute of Technology, USA) on board the Terra spacecraft of the National Aeronautics and Space Administration (NASA) delivers the Global Digital Elevation Model (GDEM) with a resolution of 30 m. This was used to derive topographic conditioning factors like distance to stream, landform, topographic wetness index (TWI), plan curvature, slope, slope aspect, and altitude.

The data for annual temperature, annual rainfall, wind effect, distance to road, recreation area, potential solar radiation and distance to village were provided by the state wildlife organization of Amol County (SWOAC) and the state meteorological organization of Amol County (SMOAC). Landsat-8 was used for generating the NDVI conditioning factor for the highly vegetative period and Land Use from the United States Geological Survey (USGS) archive (<http://earthexplorer.usgs.gov>) with a resolution of 30 m.

All the input conditioning factors were resampled using ArcGIS software with a resolution of 30 m and were classified based on expert opinion, literature review, and relevance to the study area.

Table 2. Factor classification by corresponding areas and percent of the burned areas in each class, along with the source for each conditioning factor.

Factors	Class	# of Pixels in Domain	Area (ha)	% of Domain	Area of Forest Fires (ha)	% of Forest Fires	Source
Slope aspect	(1) Flat	413	32.66	0.05	0.23	0.04	ASTER DEM
	(2) North	163,477	12,929.5	20.017	78.87	15.74	
	(3) Northeast	157,185	12,431.86	16.25	74.90	14.95	
	(4) East	111,057	8783.56	13.60	59.27	11.83	
	(5) Southeast	64,513	5102.32	7.9	35.55	7.09	
	(6) South	59,425	4699.96	7.27	53.12	10.6	
	(7) Southwest	69,288	5480.03	8.48	65.06	12.98	
	(8) West	89,748	7098	10.98	83.87	16.71	
	(9) Northwest	101,549	8031.57	12.43	50.21	10.02	
Slope (%)							
	(1) 0–5	52,438	4147.35	6.42	49.13	9.8	ASTER DEM
	(2) 5–10	131,189	10,375.82	16.06	129.72	25.89	
	(3) 10–15	165,158	13,062.45	20.22	160.07	31.95	
	(4) 15–20	132,343	10,467.09	16.20	68.49	13.67	
	(5) 20–30	172,740	13,662.11	21.15	55.58	11.09	
	(6) 30<	162,787	12,874.92	19.93	37.93	7.57	
Altitude (m)							
	(1) 500>	267,103	20,609.83	31.76	272.50	54.39	ASTER DEM
	(2) 500–1000	221,070	17,057.90	26.28	139.98	27.93	
	(3) 1000–1500	175,496	13,541.38	20.86	33.66	6.72	
	(4) 1500–2000	131,112	10,116.68	15.59	51.22	10.23	
	(5) 2000–2500	44,074	3400.77	5.59	3.57	0.71	
	(6) 2500<	2064	159.25	0.24	0		
Annual temperature (°C)							
	(1) 10>	30,663	2425.1	3.75	0	0	SMOAC
	(2) 10–12	190,487	15,065.7	23.29	3.61	0.72	
	(3) 12–14	213,835	16,912.3	26.15	92.93	18.55	
	(4) 14–16	234,441	18,542.0	28.67	162.79	32.48	
	(5) 16<	148,230	11,723.6	18.1	241.12	48.25	
Annual rainfall (mm)							
	(1) 400–450	40,288	3186.40	4.92725	0	0	SMOAC
	(2) 450–500	129,427	10,236.4	15.8290	0	0	
	(3) 500–550	138,521	10,955.7	16.9412	30.56	6.10	
	(4) 550–600	311,886	24,667.2	38.1439	146.55	29.25	
	(5) 600<	197,534	15,623.0	24.1585	323.83	64.64	
Wind effect							
	(1) 0.73–0.93	203,575	16,100.8	24.9279	161.16	32.25	ASTER DEM & SMOAC
	(2) 0.93–1.09	204,281	16,156.7	25.0143	143.42	28.62	
	(3) 1.09–1.25	204,979	16,211.9	25.0998	123.72	24.69	
	(4) 1.25–1.35	203,820	16,120.2	24.9579	72.25	14.42	
Plan curvature (100/m)							
	(1) Concave	153,099	12,108.7	18.73	62.9	12.55	ASTER DEM
	(2) Flat	499,095	39,473.7	61.05	351.45	70.15	
	(3) Convex	165,204	13,066	20.21	86.59	17.28	
Topographic wetness index (TWI)							
	(1) 5–10	89,647	7090.23	10.97	61.82	12.34	ASTER DEM
	(2) 10–15	186,858	14,778.7	22.8	117.62	23.48	
	(3) 15–20	113,587	8983.66	13.9	61.22	12.22	
	(4) 20 <	259,476	20,522.1	31.7	174.21	34.72	
		167,087	13,215.	20.45	86.07	17.18	

Table 2. Cont.

Factors	Class	# of Pixels in Domain	Area (ha)	% of Domain	Area of Forest Fires (ha)	% of Forest Fires	Source
Landform							
	(1) canyon	39,975	3161.64	4.8	16.10	3.21	ASTER DEM
	(2) Gentle slopes	159,331	12,601.5	19.48	63.23	12.62	
	(3) steep slope	513,481	40,611.5	62.79	375.23	75.02	
	(4) ridges	104,869	8294.15	12.825	45.75	9.13	
Land use							
	(1) Forest	748,822	59,224.8	91.4729	491.8	98.03	LANDSAT satellite image
	(2) Non-forest	56,744	4487.91	6.93160	9.87	1.97	
	(3) Farm	10,619	839.863	1.29717	0	0	
	(4) village	2442	193.139	0.29830	0	0	
NDVI							
	(1) −0.08–0.1	162,431	12,846.7	19.86	38.03	7.59	LANDSAT 8
	(2) 0.1–0.36	153,261	12,121.5	18.74	72.30	14.44	
	(3) 0.36–0.41	161,025	12,735.5	19.69	103.78	20.73	
	(4) 0.41–0.43	176,758	13,979.9	21.617	160.03	31.94	
	(5) 0.43<	164,181	12,985.1	20.07	121.70	25.29	
Distance to stream (m)							
	(1) 200>	78,797	6232.1	9.636	22.56	4.5	ASTER DEM
	(2) 200–500	106,507	8423.7	13.02	83.04	16.57	
	(3) 500–800	106,173	8397.2	12.985	97.99	19.57	
	(4) 800–1200	131,936	10,434.9	16.135	67.93	13.56	
	(5) 1200<	394,243	31,180.93	48.216	229.43	45.79	
Distance to road (m)							
	(1) 0–300	141,880	11,221.3	17.352	115.99	23.15	SWOAC
	(2) 300–600	116,931	9248.14	14.30	107.178	21.49	
	(3) 600–1200	172,493	13,642.5	21.096	99.06	19.77	
	(4) 1200–1800	129,926	10,275.9	15.890	88.82	17.73	
	(5) 1800<	256,426	20,280.9	31.36	89.40	17.78	
Recreation area (m)							
	(1) 0–300	32,430	2689.05	3.881	13.87	2.77	SWOAC
	(2) 300–700	72,251	5985.99	9.006	0.098	0.019	
	(3) 700<	751,341	59,830.23	87.021	468.21	97.20	
Potential solar radiation							
	(1) 282.943–983.084	64,516	5102.61	7.89	98.04	3.9	SWOAC
	(2) 983.084–1.189.376	21,641	1711.60	2.646	1.26	0.25	
	(3) 1.189.376–1.339.406	54,780	4332.58	6.699	2.47	0.49	
	(4) 1.339.406–1.501.939	113,723	8994.4	13.90	59.65	11.9	
	(5) 1.501.939–1.877.015	562,996	44,527.71	68.85	339.51	67.71	
Distance to village (m)							
	(1) 0–300	33,175	2623.83	4.05	0.094	0.018	SWOAC
	(2) 300–600	33,140	2621.06	4.053	13.85	2.76	
	(3) 600–1200	82,832	6551.23	10.13	16.99	3.39	
	(4) 1200–2400	203,181	16,069.71	24.84	73.72	14.71	
	(5) 2400>	465,328	36,803.0	56.90	396.28	79.1	

3.3. Methods

Machine learning approaches have contributed significantly during recent years in the evolution of prediction systems providing enhanced performance and efficient results. The persistent advancement of ML approaches over the last few years recognized their suitability for various natural hazard predictions with an adequate degree of surpassing traditional approaches. ML is independent of expert knowledge, but it exclusively hinges on the inventory data. For this study, we have used eleven ML approaches for wildfire susceptibility assessment like ANN, DR, DM neural, LARS, MLP, RF, RBF, LR, SOM, SVM, and DT.

3.3.1. Artificial Neural Network (ANN)

Artificial neural network impersonates the performance of the human brain through a set of nodes that are interconnected [34]. The ANN imitates the human brain in two main respects: firstly, it obtains knowledge through a learning procedure; and secondly, the knowledge gained is stored through synaptic weights [35]. The ANN approach is trained to differentiate and simplify the association between input and output. In general, there is a link in multiple-input non-linear procedures among minor single interrelated and interconnected neural networks and weighted interconnections. Spatial prediction of wildfire susceptibility is a complex and non-linear problem, where an optimal solution can be found through ANN by determining the patterns between the conditioning factors and responses. There are various neural networks for different purposes, and we used the most widely used MLP architecture and used the backpropagation algorithm (BPA) in order to train the model. Neurons that are present in the same hidden layer are not connected in the MLP architecture, but there are connections neurons belonging to one layer and the neurons belonging to the next layer. The amount of concealed layers in the neural network depends on the complexity of the problem. The size and number of the hidden layers in the ANN model are generally established based on the application [36]. To minimize the error from the random selection of initial weightings, the repeated backward process is used to update the initial weightings. For this study, the input layer was made up of 16 neurons (based on the number of input conditioning factors), one concealed layer (20 neurons), and one output layer that acted as the network structure. The learning rate was fixed at 0.09, and the number of epochs was set to 500 for our model.

3.3.2. Dmine Regression (DR)

The Dmine Regression (DR) technique executes a regression analysis on data sets that have a binary or interval level target variable. The DR technique calculates a forward stepwise least squares regression. An independent variable is selected at each step, which contributes greatly to the model r-square value. The DR is able to calculate all two-way interactions of classification variables and uses the AOV16 variables to recognize non-linear relations between the interval and target variables. Another advantage is that the DR is able to use the group variables to decrease the number of levels in the classification of variables.

3.3.3. Dmneural

The Dmneural technique is a ML modelling tool that can be used for fitting a non-linear model. The nonlinear model uses transformed principal components as inputs to forecast a binary or interval target variable. The Dmneural technique is intended to offer a flexible target prediction by means of an algorithm that has similarities to a neural network. Using the principle components approach, the problem of selecting useful inputs is avoided. The complexity of the model is controlled by selecting the quantity of stages in the multi-stage prediction formula.

3.3.4. Least Angle Regression (LARS)

LARS is a classic model selection approach well known as forwarding selection or forward stepwise regression [37]. LARS is the formalized version of the stage-wise technique that uses a simple mathematical formula to fast-track the calculations. Forward stepwise regression develops a model in sequence, adding a single variable at a time. At each step, it recognizes the finest variable to include in the active set and then updates the least squares fit to comprise all the active variables. LARS uses a similar approach but only enters predictors as it needs. As a first step, it recognizes the variable best correlated with the response. Rather than fitting this variable entirely, LAR transfers the coefficient of this variable continuously toward its least square value. As soon as another variable catches up, the process is stopped. LARS has a simple structure and lends itself to inferential analysis [38].

3.3.5. Multi-Layer Perceptron (MLP)

The MLP is a neural network that has several hidden layers and the neurons are connected between the neighboring layers. MLP is normally used as a feed-forward supervised neural network and is widely used owing to its clean architecture, fast operation, ease of implementation, and competency when resolving intricate problems arising through classification [39,40]. The MLP system consists of three main layers: an input layer, a hidden layer, and an output layer. These three main layers are used for input of data, transmission of data, and data output, respectively. The function of the hidden layer is to transfer the results to the output layer [41]. Each neuron output can be described scientifically as shown in Equation (1):

$$y_i = f\left(\sum w_{ij}x_j\right) \quad (1)$$

where y_i denotes the input received by a single node j . The function f is a function that can be a threshold, sigmoid or hyperbolic tangent. The weights among the nodes i and j are represented by w_{ij} , and x_i signifies the output from node i .

3.3.6. Random Forest (RF)

The RF algorithm was first created using the random subspace method [42]. Random forest (RF) is a machine learning technique in which the input dataset is classified based on an ensemble of multiple decision trees. RF has received increasing attention in recent years due to its ability to produce excellent classification results with a rapid processing speed [43]. Additionally, the feature set is selected randomly at each stage at which the output is forecasted, and then each of the outputs is given a weighting with a value that is based on the votes obtained. Based on the outputs of decision tree assessments, the majority vote converges on a single decision tree for the final classification [44]. In order to overcome the uncertainty problem, a single decision tree can be used, and this will result in higher prediction accuracy [45]. The crucial step in the RF classification is to derive high variance from different decision trees. Obtaining a high degree of variance from diverse decision trees is vital for this classification method. RF is regarded as being one of the best operative non-parametric ensemble learning approaches for the purpose of wildfire susceptibility mapping and modeling. The main training options in the RF model are the use of the maximum number of trees, the variable number required for the split search, and the variant for the sampling process [46]. The first and the second training options are options considered in the split searching of the RF. For the purposes of this analysis, the maximum number of trees used was set to 200. The selection process category was measured as a fraction that recognizes the percentage of observations applied for each tree. For the final forest model, an out-of-the-bag (OOB) sample statistic was used. This OOB sample statistic identifies how a model will be executed when using new inputs. The inputs used in the training samples are called bagged observations, and the input data is referred to as bagged data for the decision tree in the RF approach.

3.3.7. Radial Basis Function (RBF)

The RBF is a neural network that contains an input layer, a hidden layer and an output layer with a feed-forward structure [39,47]. The hidden layer in the network assembles the data from the input layer and passes the data to the Gaussian transfer function, which transforms the data and regulates the data nonlinearly. The Gaussian function reactions are then linearly fused to generate the data of the output layer. The RBF is broadly applied in numerous applications like data classification, time series prediction, system control, and dynamic system problems due to the capability of estimating the conduct directly from the input and output data [41]. RBF systems attempt to curtail the training error, and can be described as shown in Equation (2):

$$E = \frac{1}{2} \sum_{t=1}^k \sum_{j=1}^p e_j^2(t) \quad (2)$$

where $e_{j(t)}$ is the error of each output unit.

3.3.8. Logistic Regression (LR)

LR is classified with the statistical models known as generalized linear models. Although LR is not considered to be a ML model, it was used for wildfire susceptibility mapping in this study due to its popularity in this field. LR is able to investigate a series of glitches where the results are impacted by one or more factors. The factors influencing the results are referred to as independent variables, which can be discrete or continuous, or a combination of discrete and continuous [48]. Logistic regression allows the forecasting of discrete outcomes, such as group membership, from a set of variables that may be continuous, discrete, or a mixture of any of these types. LR is intended to build regressions that are able to fit fundamental associations among several explanatory variables and dependent variables [49]. The dependent variable is typically binary, with values of either 0 or 1. Logistic regression also offers knowledge on the relationships and strengths among the variables. The main advantage of LR is that, with the addition of a fitting link function to the typical linear regression model, the variables may be either continuous or discrete, or a combination of both types. They usually do not have normal distributions [50].

3.3.9. Self-Organizing Maps (SOM)

Self-Organizing Map (SOM) is an unsupervised learning technique which is neutrally motivated and used in various data analysis tasks. When the classifications of a test set are not known, unsupervised neural networks are used to carry out classifications. Initially, Kohonen generally defined the initial unsupervised neural network as an approach for repetitively segregating the classification space [51]. This approach was designated as a self-organizing map (SOM). SOMs comprise a non-parametric analysis approach built on neural networks that derive general arrangements from the user-defined gridded data over a region [52]. Multivariate and multi-dimensional inputs can be processed by SOMs on a user-defined grid by making a spatially arranged set of general patterns from the input data. User-defined grids can be 4×3 , i.e., twelve, nodes [53].

3.3.10. Support Vector Machines (SVM)

A support vector machine is a data mining ML approach universally used with a set of linear indicator functions that are used to issue function estimations [54]. SVM is also known as the maximum-margin method, which provides better performance and superior results with an inadequate number of data points. SVM is grounded in statistical knowledge theory, which maps the datasets into a high dimensional feature space through nonlinear transformers to generate the best hyperplane [55]. When there is maximal separations between the margins among the defined classes of the problem, the best hyperplane can be achieved. SVMs have two layers that can implement diverse functions like linear, radial, polynomial or sigmoid; hence, they are unidirectional. The performance of the SVM model is greatly influenced by kernel functions like linear, polynomial, sigmoid and radial basis function (RBF) [56].

3.3.11. Decision Tree (DT)

The DT can be described as a non-supervised non-parametric learning method for prediction and classification [57]. DT has the distinct advantage of easy interpretation and comprehension for the purposes of comparison and validation of options by decision-makers. DTs are easy to build and interpret, and their predictions are efficient. The core idea of the decision tree is to divide the data recursively into subsets to make sure that each of the subsets consists of more or less homogeneous states of the target variable (predictable attribute). All input attributes are assessed for their impact on the predictable attribute at each split in the tree, and when this recursive process is finished, a decision tree has been designed [58]. The decision tree is called a classification tree if the predictable target

attribute consists of discrete data, and if it is a continuous variable, then it is called a regression tree. The whole process of decision tree building is known as decision tree induction. Many approaches have been developed for performing decision tree induction. However, the general approach to decision tree induction is similar for all types of decision tree approach. Each approach employs a different learning algorithm to determine a model that fits the relationship between the attribute set and class label of input data best. A model generated by a learning algorithm should both fit the input data well and correctly predict the class labels of records it has never seen before [59].

4. Results

Each of the eleven selected ML approaches was applied using the sixteen wildfire conditioning factors in order to derive the susceptibility mapping for each ML approach for Amol County in Iran. The resulting susceptibility map for each ML approach is shown below in Figure 2. The susceptibility map indicates the probability of wildfire occurrences based on the relevant conditioning factors for a given region. There is no specific classification method for classifying the resulting susceptibility map. However, the natural break classification method, which is widely used, was used to classify the resulting susceptibility maps derived from the eleven ML approaches. The natural break classification method is useful for interpreting values that lie close to each class boundary. Each of the eleven ML approaches used for attaining the susceptibility maps was classified based on the natural break classification method and into one of five classes (very high, high, moderate, low and very low), which served as a uniform classification system for the purpose of comparing the results. The natural break classification method is based on the data itself, whereby the groupings are done naturally and the classification interval is chosen based on where the groups themselves suggest optimal groupings or similar values [60]. This also helps in the accuracy assessments for the machine learning approaches, which were used to ascertain the best machine learning approach to wildfire susceptibility mapping. Based on the differences between the resulting susceptibility maps, it can be seen that each model was influenced by one or more conditioning factors that distinguished it from the others. However, the role of the slope aspect factor was much more significant.

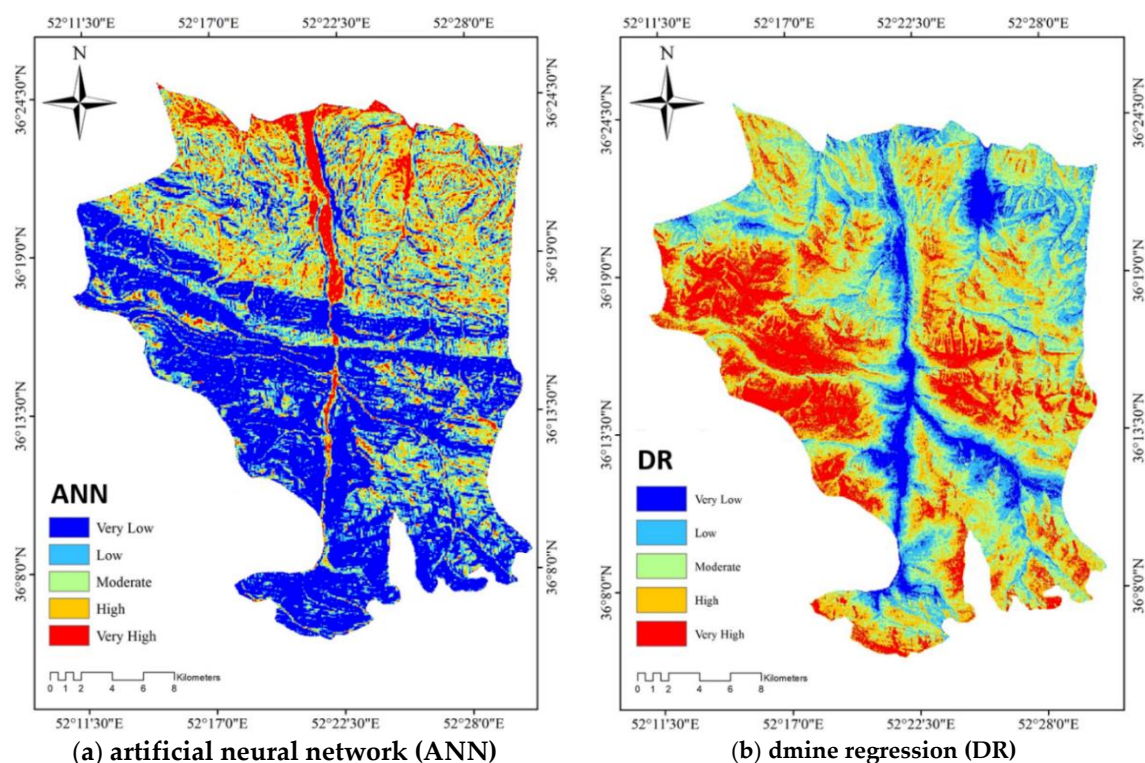


Figure 2. Cont.

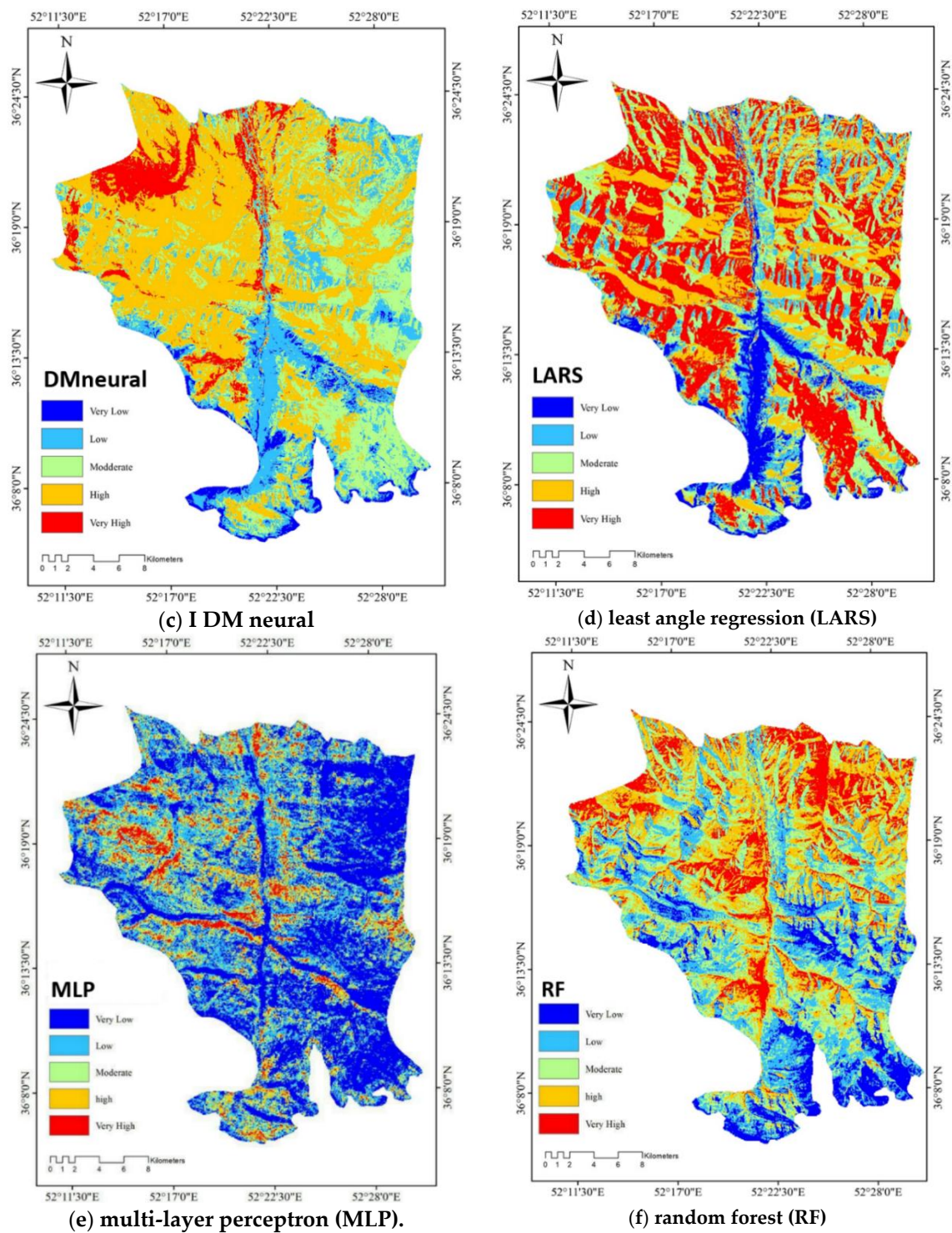


Figure 2. Cont.

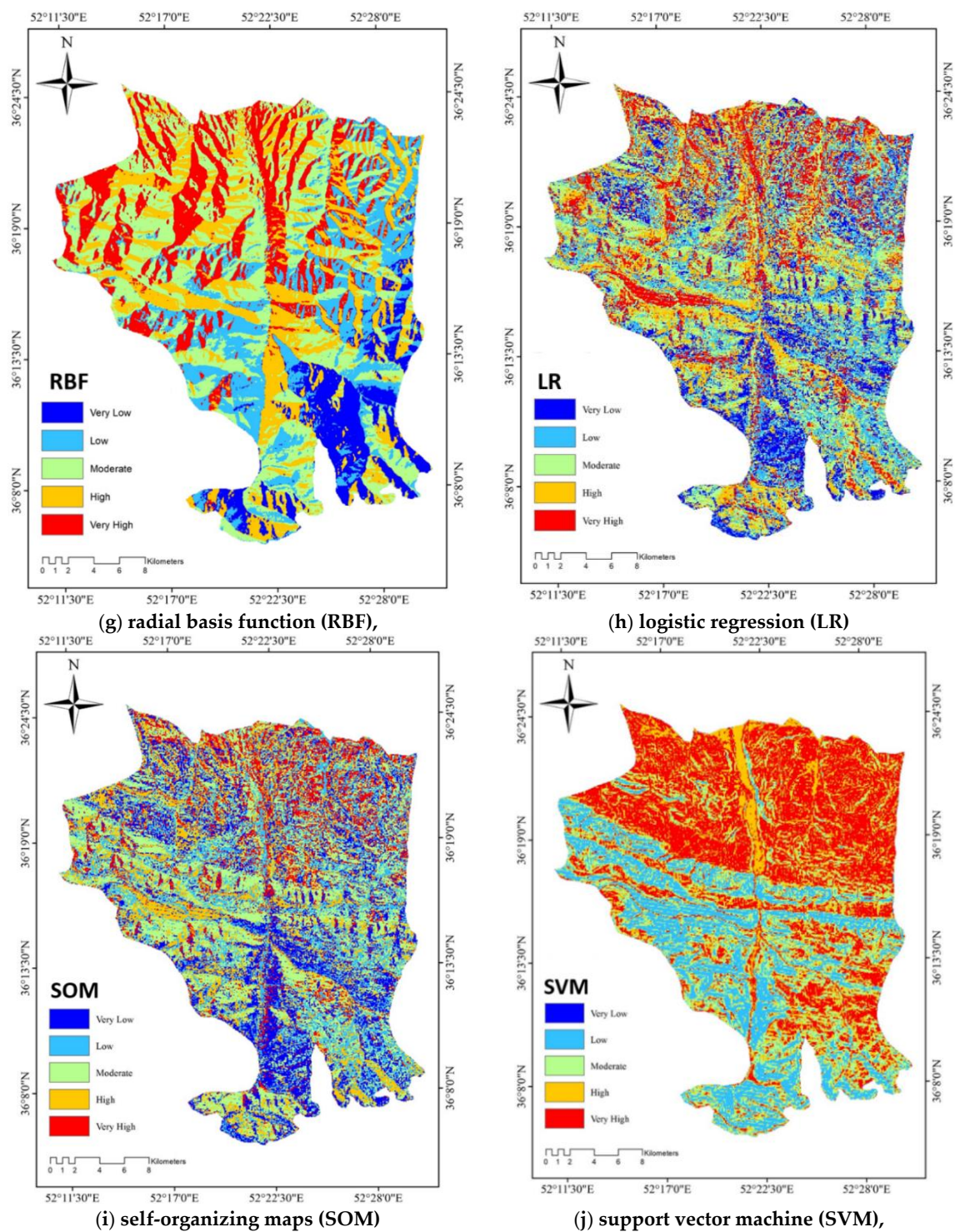


Figure 2. Cont.

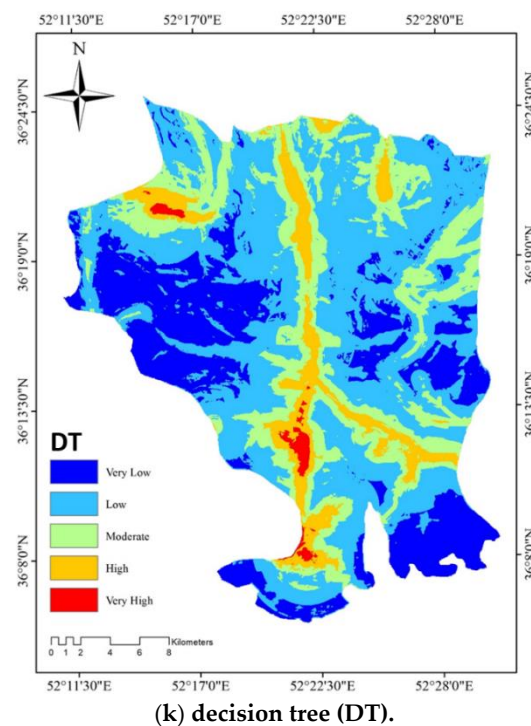


Figure 2. Susceptibility maps derived using each machine learning approach: (a) artificial neural network (ANN), (b) dm mine regression (DR), (c) DM neural, (d) least angle regression (LARS), (e) multi-layer perceptron (MLP), (f) random forest (RF), (g) radial basis function (RBF), (h) logistic regression (LR), (i) self-organizing maps (SOM), (j) support vector machine (SVM), and (k) decision tree (DT).

5. Accuracy Assessment

The accuracy assessment is a crucial step in understanding the accuracy of the wildfire susceptibility maps. For the validation, we used three-fold CV along with the receiver operating characteristics (ROC) curve approach in order to determine the accuracy of each model.

5.1. Receiver Operating Characteristics (ROC)

The ROC method is frequently used to illustrate the performance of a model [61]. The ROC curves plotted present a comparison between the true positive rate and the false positive rate on the y -axis and x -axis. In order to assess the accuracy of each approach, the area under the curve (AUC) was used, for which values closer to 1 indicate higher accuracy and values closer to zero indicate lower accuracy of the susceptibility map. The ROC curves were calculated for all of the wildfire susceptibility maps derived from the eleven ML approaches (see Figure 3). The corresponding AUC values for each wildfire susceptibility map based on each ML and each fold are presented in Table 3.

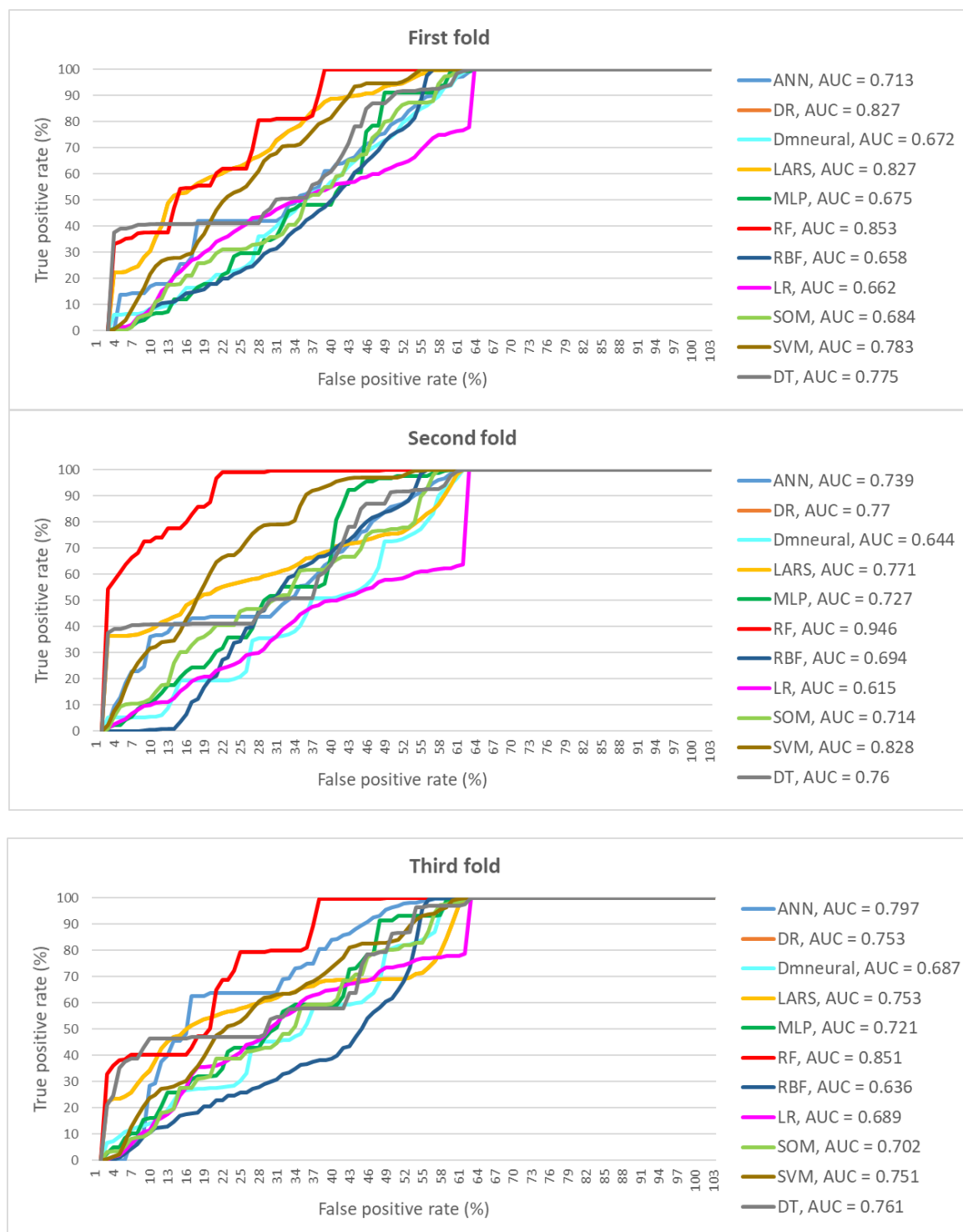


Figure 3. Values for all the machine learning approaches. The susceptibility maps based on each of the machine learning approaches: artificial neural network (ANN), dmne regression (DR), DM neural, least angle regression (LARS), multi-layer perceptron (MLP), random forest (RF), radial basis function (RBF), logistic regression (LR), self-organizing maps (SOM), support vector machine (SVM), and decision tree (DT).

5.2. Cross-Validation (CV)

The three-fold CV was applied in order to prepare the inventory dataset of wildfire pixels for the training and testing processes of each applied ML. For the three-fold CV validation method, the

wildfire inventory data pixels are arbitrarily distributed into three reciprocally exclusive folds: D_1 , D_2 , and D_3 . Each ML is run using two folds for training the model, and the resulting wildfire susceptibility map is tested with the third fold, which was not used for training. For example, when an ML is trained the two folds D_2 and D_3 , the resulting maps will be tested with the D_1 fold. As our inventory dataset of wildfire pixels was randomly divided into three folds, each time, 66% (11,614 pixels) were used for training the applied ML approaches, while 33% (5806 pixels) of the wildfire inventory data was reserved for the accuracy assessment. The three folds were used to study the volume of the wildfire inventory data pixels along with eleven ML approaches. Since we used several ML approaches, the CV method was helpful for dealing with the negative effects of randomness on the resulting maps based on each ML approach. The application of a k-fold CV for using MLs for hazard susceptibility mapping has been described in detail by Ghorbanzadeh et al. (2018) [62]. The average accuracy assessment of all applied folds is considered as the CV value for each wildfire susceptibility map based on each ML, as presented in Table 3. Thus, the whole process was implemented three times, using different folds of the inventory dataset of wildfire pixels each time. For the accuracy assessment using the three-fold CV along with ROC-AUC, the values ranged from 0 to 1 (0–100%) when determining the performance of each ML approach. If the ROC values ranged between 0.5 and 0.6 (<60%) then the performance of the model was considered to be poor or bad, ROC values between 0.6 and 0.7 (<70%) were considered to be moderate, ROC values between 0.7 and 0.8 (<80%) were considered to be good, ROC values between 0.8 and 0.9 (<90%) were considered to be very good, and ROC values of more than 0.9 and ranging up to 1 (>90%) were considered to indicate exceptional performance of the model.

Table 3. Factor classification.

Model	AUC _{fold1}	AUC _{fold2}	AUC _{fold3}	CV
ANN	0.713	0.739	0.797	0.749
DR	0.827	0.77	0.753	0.783
DMneural	0.672	0.644	0.687	0.667
LARS	0.827	0.771	0.753	0.783
MLP	0.675	0.727	0.721	0.707
RF	0.853	0.946	0.851	0.883
RBF	0.658	0.694	0.636	0.662
LR	0.662	0.615	0.689	0.655
SOM	0.684	0.714	0.702	0.7
SVM	0.783	0.828	0.751	0.787
DT	0.775	0.76	0.761	0.765

6. Discussion

Susceptibility mapping is a crucial component of addressing wildfire risk, which poses a high risk to lives and the ecosystem. It is a fundamental module in emergency management and in the planning of mitigation measures aimed at decreasing adverse impacts [63]. The ML approaches fare better than statistical approaches, as shown in previous studies [64]. Moreover, compared to the other wildfire susceptibility studies in the same area and neighboring forestry areas [4,6,9,65], we used and compared many more of the available and commonly used ML approaches in order to show their capabilities with respect to wildfire susceptibility mapping.

Of all the ML approaches used in this study, RF showed the highest accuracy in any fold of the inventory data set. The highest accuracy for this ML approach was obtained within the second fold, with an AUC value of more than 94%. However, the third fold had a lower AUC value of 85%. Two ML approaches—DR and LARS—presented almost the same accuracies in each fold and subsequently in the resulting CV values. Among all of the applied MLs, RBF had the lowest CV accuracy.

7. Conclusions

Wildfires have become a frequent hazard across the globe, ravaging forests, burning acres of habitat and causing a loss of lives. The prediction of wildfires is a significant component of emergency management, and mapping susceptible wildfire areas will help in mitigating the impact of fires. Susceptibility maps are widely and increasingly frequently used in order to prioritize locations with respect to managing hazards. However, each susceptibility map might differ based on the input parameters and the methodology used for producing them, which may have varying accuracies. For this study, we used eleven machine learning approaches that were developed and trained based on the historic wildfire events from 2012 to 2017, along with sixteen relevant conditioning factors for the study area. The performance of each ML approach was assessed with respect to accuracy using the ROC curve. Most of the ML approaches showed an accuracy above 70%, except for LR, Dmneural, and RBF, which had the lowest accuracies, indicating that those approaches were not suitable for wildfire susceptibility mapping in our study area. However, this might be different for other study areas, in which these ML approaches might have higher accuracies depending on the available conditioning factors and training data sets.

We possessed sufficient wildfire inventory data for training the ML approaches and also for testing the ML models. However, the manner in which the quantity of training data influences the performance of each ML approach is still not clear, and this might be a limitation of this study that could be studied in more detail in future work. However, we used three-fold CV to address this limitation, as well as the randomness among the resulting wildfire susceptibility maps. Thus, the application of a k-fold CV is highly recommended for similar studies. Furthermore, for future work, we would like to consider the vulnerable areas in the study area, as susceptibility maps can demonstrate the locations of elements at risk and can be incorporated in order to derive wildfire risk maps that incorporate a vulnerability analysis. Obtaining information on the communities within the study area with limited capacities and capabilities to prevent wildfires will be crucial for deriving risk maps and will help in mitigation and planning. In addition, seasonal aspects such as seasonal climate data will be considered for the wildfire susceptibility mappings of our future work.

Author Contributions: Conceptualization, K.G., O.G., and T.B.; Data curation, K.G. and O.G.; Funding acquisition, T.B.; Investigation, K.G., T.G.N., and O.G.; Methodology, K.G. and O.G.; Supervision, T.B.; Validation, K.G. and O.G.; Visualization, T.G.N., and O.G.; Writing—original draft, K.G., T.G.N., and O.G.; Writing—review & editing, T.G.N., O.G. and T.B. All authors read and approved the final manuscript.

Funding: This research was partly funded by the Austrian Science Fund (FWF) through the GIScience Doctoral College (DK W 1237-N23).

Acknowledgments: Open Access Funding by the Austrian Science Fund (FWF).

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

References

1. MacDicken, K.G. Global Forest Resources Assessment 2015: What, why and how? *For. Ecol. Manag.* **2015**, *352*, 3–8. [[CrossRef](#)]
2. Pourtaghi, Z.S.; Pourghasemi, H.R.; Aretano, R.; Semeraro, T. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. *Ecol. Indic.* **2016**, *64*, 72–84. [[CrossRef](#)]
3. Moayedi, H.; Mehrabi, M.; Bui, D.T.; Pradhan, B.; Foong, L.K. Fuzzy-metaheuristic ensembles for spatial assessment of forest fire susceptibility. *J. Environ. Manag.* **2020**. [[CrossRef](#)] [[PubMed](#)]
4. Ghorbanzadeh, O.; Valizadeh Kamran, K.; Blaschke, T.; Aryal, J.; Naboureh, A.; Einali, J.; Bian, J. Spatial Prediction of Wildfire Susceptibility Using Field Survey GPS Data and Machine Learning Approaches. *Fire* **2019**, *2*, 43. [[CrossRef](#)]
5. Ghomi Motazeh, A.; Farahi Ashtiani, E.; Baniasadi, R.; Masoumpoor Choobar, F. Rating and mapping fire hazard in the hardwood Hyrcanian forests using GIS and expert choice software. *For. Ideas* **2013**, *19*, 141–150.

6. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Aryal, J. Forest fire susceptibility and risk mapping using social/infrastructural vulnerability and environmental variables. *Fire* **2019**, *2*, 50. [\[CrossRef\]](#)
7. Tymstra, C.; Stocks, B.J.; Cai, X.; Flannigan, M.D. Wildfire management in Canada: Review, challenges and opportunities. *Prog. Disaster Sci.* **2020**, *5*. [\[CrossRef\]](#)
8. Hantson, S.; Pueyo, S.; Chuvieco, E. Global fire size distribution is driven by human impact and climate. *Glob. Ecol. Biogeogr.* **2015**, *24*, 77–86. [\[CrossRef\]](#)
9. Jaafari, A.; Pourghasemi, H.R. Factors Influencing Regional-Scale Wildfire Probability in Iran: An Application of Random Forest and Support Vector Machine. In *Spatial Modeling in GIS and R for Earth and Environmental Sciences*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 607–619.
10. Berger, F.; Rey, F. Mountain Protection Forests against Natural Hazards and Risks: New French Developments by Integrating Forests in Risk Zoning. *Nat. Hazards* **2004**, *33*, 395–404. [\[CrossRef\]](#)
11. Jahdi, R.; Darvishsefat, A. Wind Effect on Wildfire and Simulation of its Spread (Case Study: Siahkal Forest in Northern Iran). *J. Agric. Sci. Technol.* **2014**, *16*, 1109–1121.
12. Rahmati, O.; Zeinivand, H.; Besharat, M. Flood hazard zoning in Yasooj region, Iran, using GIS and multi-criteria decision analysis. *Geomat. Nat. Hazards Risk* **2015**, *7*, 1000–1017. [\[CrossRef\]](#)
13. Rahmati, O.; Pourghasemi, H.R.; Zeinivand, H. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golestan Province, Iran. *Geocarto Int.* **2015**, *31*, 42–70. [\[CrossRef\]](#)
14. Gudiyangada Nachappa, T.; Tavakkoli Piralilou, S.; Ghorbanzadeh, O.; Shahabi, H.; Blaschke, T. Landslide Susceptibility Mapping for Austria Using Geons and Optimization with the Dempster-Shafer Theory. *Appl. Sci.* **2019**, *9*, 5393. [\[CrossRef\]](#)
15. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 1149–1165. [\[CrossRef\]](#)
16. Nazmfar, H.; Saredeh, A.; Eshgi, A.; Feizizadeh, B. Vulnerability evaluation of urban buildings to various earthquake intensities: A case study of the municipal zone 9 of Tehran. *Hum. Ecol. Risk Assess. Int. J.* **2019**, *25*, 455–474. [\[CrossRef\]](#)
17. Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [\[CrossRef\]](#)
18. Mohammady, M.; Pourghasemi, H.R.; Pradhan, B. Landslide susceptibility mapping at Golestan Province, Iran: A comparison between frequency ratio, Dempster-Shafer, and weights-of-evidence models. *J. Asian Earth Sci.* **2012**, *61*, 221–236. [\[CrossRef\]](#)
19. Pourghasemi, H.; Pradhan, B.; Gokceoglu, C.; Moezzi, K.D. A comparative assessment of prediction capabilities of Dempster-Shafer and Weights-of-evidence models in landslide susceptibility mapping using GIS. *Geomat. Nat. Hazards Risk* **2013**, *4*, 93–118. [\[CrossRef\]](#)
20. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sens.* **2019**, *11*, 196. [\[CrossRef\]](#)
21. Tehrany, M.S.; Pradhan, B.; Jebur, M.N. Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS. *J. Hydrol.* **2013**, *504*, 69–79. [\[CrossRef\]](#)
22. Pradhan, B.; Lee, S.; Mansor, S.; Buchroithner, M.; Jamaluddin, N.; Khujaimah, Z. Utilization of optical remote sensing data and geographic information system tools for regional landslide hazard analysis by using binomial logistic regression model. *J. Appl. Remote Sens.* **2008**, *2*, 023542.
23. Felicísimo, Á.M.; Cuartero, A.; Remondo, J.; Quirós, E. Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: A comparative study. *Landslides* **2013**, *10*, 175–189. [\[CrossRef\]](#)
24. Sharma, S.; Dhakal, K.; Wagle, P.; Kilic, A. Retrospective tillage differentiation using the Landsat-5 TM archive with discriminant analysis. *Agrosyst. Geosci. Environ.* **2020**, *3*, e20000. [\[CrossRef\]](#)
25. Zhang, G.; Wang, M.; Liu, K. Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China. *Int. J. Disaster Risk Sci.* **2019**, *10*, 386–403. [\[CrossRef\]](#)
26. Nampak, H.; Pradhan, B.; Manap, M.A. Application of GIS based data driven evidential belief function model to predict groundwater potential zonation. *J. Hydrol.* **2014**, *513*, 283–300. [\[CrossRef\]](#)

27. Sharma, S.; Ochsner, T.E.; Twidwell, D.; Carlson, J.; Krueger, E.S.; Engle, D.M.; Fuhlendorf, S.D. Nondestructive estimation of standing crop and fuel moisture content in tallgrass prairie. *Rangel. Ecol. Manag.* **2018**, *71*, 356–362. [\[CrossRef\]](#)
28. Tien Bui, D.; Khosravi, K.; Shahabi, H.; Daggupati, P.; Adamowski, J.F.; M.Melesse, A.; Thai Pham, B.; Pourghasemi, H.R.; Mahmoudi, M.; Bahrami, S.; et al. Flood Spatial Modeling in Northern Iran Using Remote Sensing and GIS: A Comparison between Evidential Belief Functions and Its Ensemble with a Multivariate Logistic Regression Model. *Remote Sens.* **2019**, *11*, 1589. [\[CrossRef\]](#)
29. Jaafari, A.; Zenner, E.K.; Panahi, M.; Shahabi, H. Hybrid artificial intelligence models based on a neuro-fuzzy system and metaheuristic optimization algorithms for spatial prediction of wildfire probability. *Agric. Forest Meteorol.* **2019**, 266–267, 198–207. [\[CrossRef\]](#)
30. Watson, G.L.; Telesca, D.; Reid, C.E.; Pfister, G.G.; Jerrett, M. Machine learning models accurately predict ozone exposure during wildfire events. *Environ. Pollut.* **2019**, *254*, 112792. [\[CrossRef\]](#)
31. Sayad, Y.O.; Mousannif, H.; Al Moatassime, H. Predictive modeling of wildfires: A new dataset and machine learning approach. *Fire Saf. J.* **2019**, *104*, 130–146. [\[CrossRef\]](#)
32. Khosravi, K.; Panahi, M.; Bui, D.T. Spatial prediction of groundwater spring potential mapping based on an adaptive neuro-fuzzy inference system and metaheuristic optimization. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 4771–4792. [\[CrossRef\]](#)
33. Junpen, A.; Garivait, S.; Bonnet, S. Estimating emissions from forest fires in Thailand using MODIS active fire product and country specific data. *Asia-Pac. J. Atmos. Sci.* **2013**, *49*, 389–400. [\[CrossRef\]](#)
34. Ghorbanzadeh, O.; Blaschke, T.; Aryal, J.; Gholaminia, K. A new GIS-based technique using an adaptive neuro-fuzzy inference system for land subsidence susceptibility mapping. *J. Spat. Sci.* **2018**, *63*, 1–17. [\[CrossRef\]](#)
35. Haykin, S. Neural Network—A comprehensive foundation. *Neural Netw.* **2004**, *2*, 41.
36. Safi, Y.; Bouroumi, A. Prediction of Forest Fires Using Artificial Neural Networks. *Appl. Math. Sci.* **2013**, *7*, 271–286. [\[CrossRef\]](#)
37. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R. *Least Angle Regression*; Statistics Department, Stanford University: Stanford, CA, USA, 2003.
38. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Science & Business Media: New York, NY, USA, 2009.
39. Al_Janabi, S.; Al-Shourbaji, I.; Salman, M.A. Assessing the suitability of soft computing approaches for forest fires prediction. *Appl. Comput. Inform.* **2018**, *14*, 214–224. [\[CrossRef\]](#)
40. Kavzoglu, T.; Mather, P.M. The use of backpropagating artificial neural networks in land cover classification. *Int. J. Remote Sens.* **2010**, *24*, 4907–4938. [\[CrossRef\]](#)
41. Zare, M.; Pourghasemi, H.R.; Vafakhah, M.; Pradhan, B. Landslide susceptibility mapping at Vaz Watershed (Iran) using an artificial neural network model: A comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. *Arab. J. Geosci.* **2013**, *6*, 2873–2888. [\[CrossRef\]](#)
42. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, Quebec, Canada, 14–16 August 1995; pp. 278–282.
43. Pourghasemi, H.R.; Rahmati, O. Prediction of the landslide susceptibility: Which algorithm, which precision? *Catena* **2018**, *162*, 177–192. [\[CrossRef\]](#)
44. Xu, R.; Lin, H.; Lü, Y.; Luo, Y.; Ren, Y.; Comber, A. A Modified Change Vector Approach for Quantifying Land Cover Change. *Remote Sens.* **2018**, *10*, 1578. [\[CrossRef\]](#)
45. Valdez, M.C.; Chang, K.-T.; Chen, C.-F.; Chiang, S.-H.; Santos, J.L. Modelling the spatial variability of wildfire susceptibility in Honduras using remote sensing and geographical information systems. *Geomat. Nat. Hazards Risk* **2017**, *8*, 876–892. [\[CrossRef\]](#)
46. Tavakkoli Pirailou, S.; Shahabi, H.; Jarihani, B.; Ghorbanzadeh, O.; Blaschke, T.; Gholaminia, K.; Meena, S.R.; Aryal, J. Landslide Detection Using Multi-Scale Image Segmentation and Different Machine Learning Models in the Higher Himalayas. *Remote Sens.* **2019**, *11*, 2575. [\[CrossRef\]](#)
47. Broomhead, D.S.; Lowe, D. *Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks*; Royal Signals and Radar Establishment Malvern (United Kingdom): Malvern, UK, 1988.
48. Li, Y.; Chen, W. Landslide Susceptibility Evaluation Using Hybrid Integration of Evidential Belief Function and Machine Learning Techniques. *Water* **2019**, *12*, 113. [\[CrossRef\]](#)

49. Rafique, W.; Zheng, D.; Barras, J.; Joglekar, S.; Kosmas, P. Predictive Analysis of Landmine Risk. *IEEE Access* **2019**, *7*, 107259–107269. [[CrossRef](#)]
50. Lee, S.; Pradhan, B. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. *Landslides* **2006**, *4*, 33–41. [[CrossRef](#)]
51. Oja, M.; Kaski, S.; Kohonen, T. Bibliography of Self Organizing Maps (SOM) Papers: 1998–2001 Addendum. *Neural Comput. Surv.* **2002**, *3*, 1–156.
52. Nauslar, N.J.; Hatchett, B.J.; Brown, T.J.; Kaplan, M.L.; Mejia, J.F. Impact of the North American monsoon on wildfire activity in the southwest United States. *Int. J. Climatol.* **2019**, *39*, 1539–1554. [[CrossRef](#)]
53. Reusch, D.B.; Alley, R.B.; Hewitson, B.C. Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. *Polar Geogr.* **2005**, *29*, 188–212. [[CrossRef](#)]
54. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin, Germany, 2013.
55. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2013**, *11*, 425–439. [[CrossRef](#)]
56. 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* **2016**, *13*, 361–378.
57. Nefeslioglu, H.A.; Sezer, E.; Gokceoglu, C.; Bozkir, A.S.; Duman, T.Y. Assessment of Landslide Susceptibility by Decision Trees in the Metropolitan Area of Istanbul, Turkey. *Math. Probl. Eng.* **2010**, *2010*, 1–15. [[CrossRef](#)]
58. Tang, Z.; Maclellan, J. *Data Mining with SQL Server 2005*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
59. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [[CrossRef](#)] [[PubMed](#)]
60. Chen, W.; Pourghasemi, H.R.; Naghibi, S.A. A comparative study of landslide susceptibility maps produced using support vector machine with different kernel functions and entropy data mining models in China. *Bull. Eng. Geol. Environ.* **2017**, *77*, 647–664. [[CrossRef](#)]
61. Maser, B.; Söllinger, D.; Uhl, A. PRNU-based Finger Vein Sensor Identification in the Presence of Presentation Attack Data. In Proceedings of the Joint ARW/OAGM Workshop 2019 (ARW/OAGM'19), Steyr, Austria, 9–10 May 2019.
62. Ghorbanzadeh, O.; Rostamzadeh, H.; Blaschke, T.; Gholaminia, K.; Aryal, J. A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping. *Nat. Hazards* **2018**, *94*, 497–517. [[CrossRef](#)]
63. Haas, J.R.; Calkin, D.E.; Thompson, M.P. A national approach for integrating wildfire simulation modeling into Wildland Urban Interface risk assessments within the United States. *Landsc. Urban Plan.* **2013**, *119*, 44–53. [[CrossRef](#)]
64. Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [[CrossRef](#)]
65. Pourtaghi, Z.S.; Pourghasemi, H.R.; Rossi, M. Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran. *Environ. Earth Sci.* **2015**, *73*, 1515–1533. [[CrossRef](#)]

