CrossMark

ORIGINAL PAPER

Spatial prediction of wildfire probability in the Hyrcanian ecoregion using evidential belief function model and GIS

M. H. Nami[1] · A. Jaafari[2] · M. Fallah[1] · S. Nabiuni[1]

Abstract Accurate estimates of wildfire probability and production of distribution maps are the first important steps in wildfire management and risk assessment. In this study, geographical information system (GIS)-automated techniques were integrated with the quantitative data-driven evidential belief function (EBF) model to predict spatial pattern of wildfire probability in a part of the Hyrcanian ecoregion, northern Iran. The historical fire events were identified using earlier reports and MODIS hot spot product as well as by carrying out multiple field surveys. Using the GIS-based EBF model, the relationships among existing fire events and various predictor variables predisposing fire ignition were analyzed. Model results were used to produce a distribution map of wildfire probability. The derived probability map revealed that zones of moderate, high, and very high probability covered nearly 60% of the landscape. Further, the probability map clearly demonstrated that the probability of a fire was strongly dependent upon human infrastructure and associated activities. By comparing the probability map and the historical fire events, a satisfactory spatial agreement between the five probability levels and fire density was observed. The probability map was further validated by receiver operating characteristic using both success rate and prediction rate curves. The validation results confirmed the effectiveness of the GIS-based EBF model that achieved AUC values of 84.14 and 81.03% for success and prediction rates, respectively.

Keywords Natural hazard · Forest fire · Fire ignition · Probability mapping

Editorial responsibility: Tan Yigitcanlar.

✉ A. Jaafari
a.jaafari@modares.ac.ir

[1] Islamic Azad University, Science and Research Branch, Tehran, Iran

[2] Young Researchers and Elite Club, Karaj Branch, Islamic Azad University, Karaj, Iran

Introduction

Wildfires play an important role as a natural disturbance across the world. Wildfires threaten human life and property, affect ecological processes and functions, and can alter the structure of ecosystems (Stephens et al. 2013). Over the past decade, substantial global increases in number and severity of wildfires have been reported (Parisien et al. 2016; Jaafari et al. 2017a). The relative likelihood of wildfire occurrence and its spread are strongly affected by various variables that can be grouped under four main categories: climate, vegetation, topography, and human activities. Climate variables (e.g., temperature, rainfall, wind, and evapotranspiration) exert both direct and indirect influences on fire ignitability (Parisien et al. 2012). Vegetation (i.e., land cover) effects on fire ignition and spread through fuel characteristics such as type, load, and moisture content (Adab et al. 2013, 2015). The effect of topographic variables (e.g., slope, aspect, and elevation) on fire activity is largely indirect (Parisien et al. 2012). Topography exerts its effect on fire mainly by influencing patterns in ignitions, vegetation, local climate, and human accessibility. Humans affect the spatial pattern and frequency of fire by providing ignition sources and altering natural vegetation in ways that may either promote or limit fire (Parisien et al. 2016).

🖄 Springer

The probability that a fire occurs is mostly estimated by investigating the spatial relationship between the predictor variables and historical fire locations of a given landscape (Syphard et al. 2008; Oliveira et al. 2012; Adab et al. 2013; Tien Bui et al. 2016; Abdi et al. 2016). Chuvieco and Congalton (1989) were probably the first authors to clearly introduce and articulate the use of geo-environmental data and a statistical model (logistic regression) for fire probability mapping. This approach inspired a growing stream of research on wildfire modeling that the variety of works published by different modelers demonstrates its vitality (e.g., Syphard et al. 2008; Martínez et al. 2009; Oliveira et al. 2012; Pourghasemi 2016; Eskandari and Chuvieco 2015; Abdi et al. 2016). Other popular methods with large number of applications are classification and regression tree models (e.g., Lozano et al. 2008), MaxEnt algorithm (e.g., Parisien et al. 2012; Arpaci et al. 2014; Chen et al. 2015), neural networks (e.g., Satir et al. 2016), frequency ratio (Pourtaghi et al. 2015); Shannon's entropy (Pourtaghi et al. 2015), weights of evidence approach (e.g., Dickson et al. 2006; Jaafari et al. 2017a), random forests (e.g., Oliveira et al. 2012; Arpaci et al. 2014; Guo et al. 2016; Pourtaghi et al. 2016), fuzzy systems (e.g., Pourghasemi et al. 2016; Tien Bui et al. 2017), support vector machines (Tien Bui et al. 2016), Bayesian modeling (e.g., Silva et al. 2015), and evidential belief function (EBF) (Pourghasemi 2016). These methods that have taken advantage of Remote Sensing and Geographical Information System (GIS) techniques enable the delineation of locations prone to fire ignition and the development of spatially explicit hazard mitigation plans. However, each method has its inherent advantages and limitations. For instance, while input process, calculations, and output process are very simple and easy to understand in the frequency ratio, Shannon's entropy, and EBF models (Pourtaghi et al. 2015; Pourghasemi 2016), other models (e.g., fuzzy systems, neural networks, and support vector machines) require the conversion of data to ASCII or other formats that is too time-consuming for large amount of data. Further, these models (i.e., fuzzy systems, neural networks, and support vector machines) need an appropriate calibration and optimization of the parameters to achieve the maximum level of predictive ability, but computation complexity and the lack of available computing power usually prevent the modelers from proper exploring this issue (Rodrigues and de la Riva 2014; Pourghasemi et al. 2016; Tien Bui et al. 2016; Satir et al. 2016). On the other hand, the random forests model that was shown to be superior to MaxEnt algorithm (Arpaci et al. 2014), support vector machines (Rodrigues and de la Riva 2014), and logistic regression (Rodrigues and de la Riva 2014; Guo et al. 2016) achieved a lower predictive ability than boosted regression tree, generalized additive model in the comparative study conducted by Pourtaghi et al. (2016). From the literature reviewed, it can be concluded that, despite much effort to identify the best modeling approach, it is still unclear which approach should be employed for wildfire modeling.

This study was aimed to evaluate the predictive ability of a GIS-EBF model for wildfire probability modeling. The model exploits information obtained from an inventory map of historical fire locations and a wide range of predictor variables to predict where wildfires may occur in the future. Density graph and receiver operating characteristics (ROC) methods were utilized for the assessment and validation of the wildfire probability map. The proposed scheme is illustrated via a case study that was conducted during 2015–2016 in the eastern part of the Hyrcanian ecoregion, Iran. Since this is a prototype study performed in one of the characteristic fire susceptible regions of northern Iran, the findings can be utilized for regions that show similar environmental characteristics.

Characteristics of the study area

The study area encompasses 10,552 km^2 of the Hyrcanian ecoregion in northern Iran (Fig. 1). The area is characterized by several land cover types, i.e., forest, rangeland, farmland, and orchard that exhibit quite a range of vegetation. This area generally contains gently sloping areas, but the southeastern zone which is covered by deciduous forests is steep. The elevation gradients of the study area vary between −117 and 2455 m, whereas about 60% of the area is less than 300 m. The area has a relatively semi-desert climate due to the topographic features (the Alborz mountain ranges), distance from the Caspian Sea, and proximity to desert areas in south of Turkmenistan. The mean annual precipitation and temperature are recorded to be 210 mm and 16 °C, while the average annual evapotranspiration is about 1980 mm. The area has experienced extensive fire activity both historically and in recent years (Fig. 2). Although fire activity has changed through time according to climate, human population, and land use changes, it typically extends from June until December with a single modal seasonal distribution that peaks in July and August.

Materials and methods

Geospatial database

Wildfire inventory map

An inventory map of historical fire events is the main component of every geospatial database constructed for wildfire probability modeling (Adab et al. 2013; Arpaci et al. 2014; Chen et al. 2015; Pourtaghi et al. 2016; Jaafari et al. 2017a). The wildfire inventory map for the study area was compiled using documentations from the administrative office of
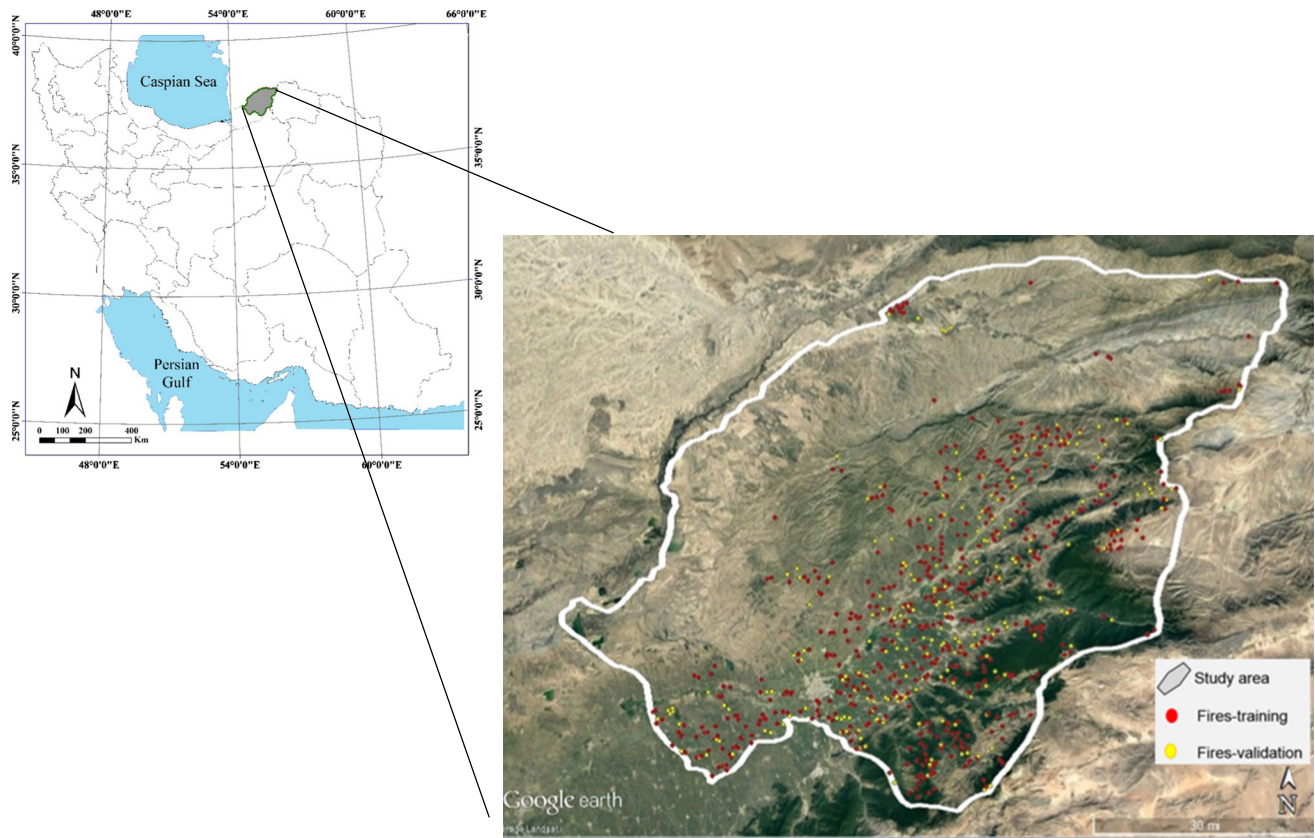
Fig. 1 Location of study area with fire event locations

Fig. 2 Examples of wildfire events in the study area (photographs by: Aboutaleb Nadri and Mahmoud Hazini)

376

Int. J. Environ. Sci. Technol. (2018) 15:373–384

natural resources of the Golestan Province, national reports, and the MODIS hot spot product (http://earthdata.nasa.gov/firms). Multiple field surveys and screening processes were conducted to remove records with inaccurate locations (e.g., records with zero as the fire detection time or a crown fire type without any associated burn area in forests) and blank records that did not contain any information. Finally, a total of 772 fire events (1162 fire pixels) from the period of 2002–2014 were verified for modeling purpose (Fig. 1). The fire locations were randomly divided into two subsets: Seventy percent of the locations (540 fires comprising 813 fire pixels) were retained for training the probability model, and the remaining locations (232 fires comprising 349 fire pixels) were set aside for model validation (Pourghasemi 2016; Pourtaghi et al. 2016; Jaafari et al. 2017a).

Predictor variables

The second component of the geospatial database for wildfire probability modeling is the set of predictor variables that have been directly or indirectly related to fire occurrence in previous studies as reported in the literature (e.g., Rollins et al. 2004; Syphard et al. 2008; Martínez et al. 2009; Oliveira et al. 2012; Adab et al. 2013; Tien Bui et al. 2016; Abdi et al. 2016; Nunes et al. 2016; Jaafari et al. 2017a). These variables are: slope degree, aspect, elevation, plan curvature, topographic wetness index (TWI), topographic roughness index (TRI), temperature, rainfall, evapotranspiration, land use/cover (LULC), soil type, and proximity to roads, rivers, and settlements. To assess possible multicollinearity among the variables, variance inflation factors (VIF) and tolerance (Hair et al. 2006; Liao and Valliant 2012) were computed. Since threshold values of VIF and tolerance (VIF < 5 and tolerance >0.2) indicated that slope degree and TWI had a significant linear relationship with other variables (Table 1), these variables were excluded for further analysis and model building.

Maps of the variables were constructed using available data from the study area. Specifically, topographic variables (i.e., aspect, elevation, plan curvature, and TRI) were extracted from a digital elevation model (DEM) with 30 × 30 m pixel size, which were obtained from ASTER Global DEM Explorer tool of United States Geological Survey (http://earthexplorer.usgs.gov). The climate variables (i.e., temperature, rainfall, and evapotranspiration) were calculated as the 20-year mean for the months from June through December using data obtained from the meteorological stations for the study area. The maps of soil type and LULC were extracted from state maps on a 1:100,000 scale. Road networks and settlements areas were digitalized from Google Earth images and co-registered with other maps, and then the proximity maps were produced by buffering settlements areas and road segments slopes in the study area. River networks extracted from the DEM, and then the map of proximity to rivers was produced by buffering river sections. All

Table 1 Multicollinearity diagnosis statistics for variables

Variable	Collinearity statistics	
	Tolerance	VIF
Slope	**0.182**	**5.494**
Aspect	0.991	1.009
Altitude	0.279	4.014
Plan curvature	0.859	1.164
TWI	**0.188**	**5.305**
TRI	0.455	2.198
Temperature	0.201	4.981
Rainfall	0.816	1.225
Evapotranspiration	0.534	1.872
LULC	0.550	1.819
Soil type	0.684	1.462
Proximity to rivers	0.951	1.052
Proximity to roads	0.849	1.178
Proximity to settlements	0.582	1.719

Excluded variables are shown in bold

calculations and generations of predictor variables were performed in ArcGIS 9.3 and SAGA GIS 2.1 for a 30 × 30 m pixel size.

In the final step, all thematic maps were classified into different classes based on suggestions found in the literature (e.g., Pourghasemi 2016; Pourtaghi et al. 2016) that were further informed by field surveys and studying the study area (Jaafari et al. 2017a) (Fig. 3).

Evidential belief function (EBF)

EBF, also known as Dempster–Shafer theory, is a mathematical-based model for reasoning with uncertainty, with understood connections to other frameworks such as probability, possibility, and imprecise probability theories. Originally introduced by Dempster (1967a, b) in the context of statistical inference, the method was later developed by Shafer (1976) into a general framework for modeling epistemic uncertainty. Over the last years, the EBF model has been frequently used in scientific studies, most often in the domain of environmental studies (e.g., Rahmati and Melesse 2016) and hazard assessment (e.g., Althuwaynee et al. 2014; Pourghasemi 2016; Regmi and Poudel 2016). This model has the capability to handle heterogeneous and incomplete datasets and represents a flexible framework to accept uncertainty and combine beliefs from multiple sources of evidence (Thiam 2005). The EBF model includes degree of belief (Bel), degree of disbelief (Dis), degree of uncertainty (Unc), and degree of plausibility (Pls) that are confined in range of 0–1 (Carranza et al. 2008; Naghibi et al. 2016). In this model, Bel and Pls represent the lower and upper probability of the generalized Bayesian theory. The difference between Pls and Bel is Unc, which represents the doubt that the
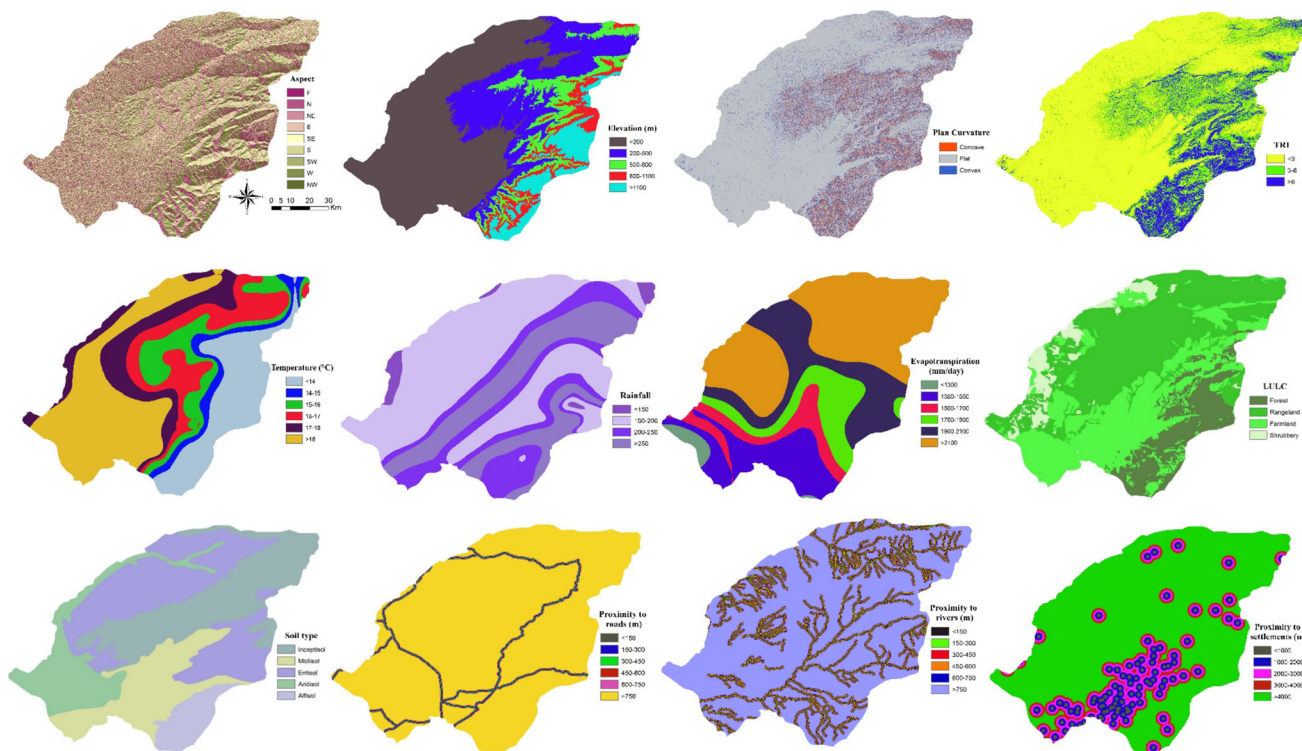
Fig. 3 Predictor variables used in this study

evidence supports a proposition (Carranza et al. 2005). The Unc values are always positive because the minimum possible value for Pls is equal to Bel. Dis is a degree of disbelief in evidence with respect to the proposition (Gorum and Carranza 2015), which is equal to $1 -$ Pls (or $1 -$ Unc $-$ Bel). Therefore, Bel $+$ Unc $+$ Dis for evidence with respect to any proposition is always equal to 1 (i.e., maximum probability) (Gorum and Carranza 2015; Naghibi et al. 2016). The schematic representation of these combinations is shown in Fig. 4.

In spatial pattern analysis of wildfire probability based on the EBF model, a structure of discernment can be defined as follows:

$$m : 2^{\Theta} = \{\phi, F_{\mathrm{P}}, \bar{F}_{\mathrm{P}}, \Theta\} \quad \text{With } \Theta = \{F_{\mathrm{P}}, \bar{F}_{\mathrm{P}}\} \tag{1}$$

where F_{P} class pixels affected by fire ignition, $\bar{F}_{\mathrm{P}}$ class pixels not affected by fire ignition, ϕ is empty set.

The function $m: 2\Theta \rightarrow [0,1]$ is called a basic probability assignment when:

$$m(\phi) = 0 \tag{2}$$

and

$$\sum_{A \subset \Theta} m(L) = 1 \tag{3}$$

where L is a subset of Θ. The function m is considered as a measure of belief committed to each possibility (Walley 1987). Based on mass function, belief (B) function can be calculated as follows:
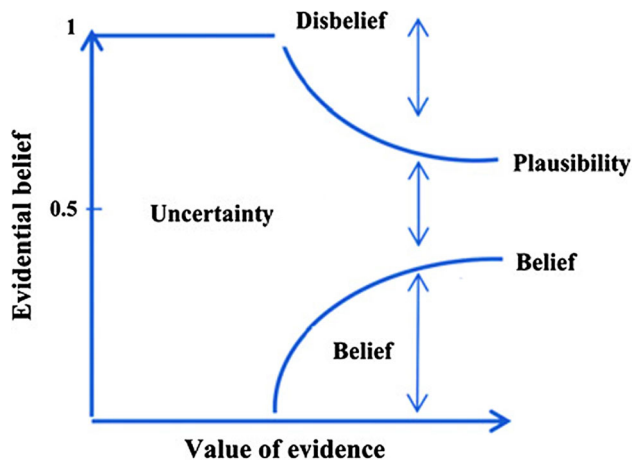

Fig. 4 Schematic relationships of evidential belief functions (modified after Carranza et al. 2005)

$$\lambda(F_{\mathrm{P}})_{B_{ij}} = \frac{\frac{N\{A \cap B_{ij}\}}{N(A)}}{\frac{N(B_{ij}) - N\{A \cap B_{ij}\}}{N(C) - N(A_{ij})}} \tag{4}$$

$$m(F_{\mathrm{P}})_{B_{ij}} = \frac{\lambda(F_{\mathrm{P}})_{B_{ij}}}{\sum \lambda(F_{\mathrm{P}})_{B_{ij}}} \tag{5}$$

where $N\{A \cap B_{ij}\}$ density of fire pixels in the predictor category B_{ij}, $N(A)$ density of fires pixels in the study area, $N(A_{ij})$ density of pixels in the predictor category B_{ij}, $N(C)$ density of pixels in the study area, $m(F_{\mathrm{P}})_{B_{ij}}$ belief function.

378

Int. J. Environ. Sci. Technol. (2018) 15:373–384

Further, the disbelief $(m(\bar{F}_{\mathrm{P}})_{B_{ij}})$, uncertainty $(m(\Theta))$, and plausibility (Pls) functions are given as:

$$\lambda(F_{\bar{\mathrm{P}}})_{B_{ij}} = \frac{\frac{N(A)-N(A\cap B_{ij})}{N(A)}}{\frac{N(C)-N(A)-N(B_{ij})+N(A\cap B_{ij})}{N(C)-N(A)}} \tag{6}$$

$$m(\bar{F}_{\mathrm{P}})_{B_{ij}} = \frac{\lambda(F_{\bar{\mathrm{P}}})_{B_{ij}}}{\sum \lambda(\bar{F}_{\bar{\mathrm{P}}})_{B_{ij}}} \tag{7}$$

$$m(\Theta) = 1 - m(F_{\mathrm{P}})_{B_{ij}} - m(\bar{F}_{\mathrm{P}})_{B_{ij}} \tag{8}$$

$$\mathrm{Pls} = 1 - m(\bar{F}_{\mathrm{P}})_{B_{ij}} \tag{9}$$

The application of the EBF model characterizes each predictor variable category with the four mentioned mass functions that indicate the strength of the correlation between the predictor category and the probability of fire occurrence. These functions were used to produce multi-category weighted maps for all predictor variables, which were overlaid and numerically added using the following equation to produce four fire probability index maps:

$$\text{Index map} = \sum_{i=1}^{12} \text{function}_{\text{Bel,Dis},Unc,Pls} \tag{10}$$

where 12 refers to the number of predictor variables in the model.

The final integrated map of wildfire probability was produced using a GIS-based raster calculator technique by overlaying the four index maps. Finally, to facilitate visual interpretation of the integrated map, the data were classified using the Natural Breaks (Jenks) method and grouped into relative levels that depict very low, low, moderate, high, and very high probabilities of fire occurrence.

Model evaluation

Density graph

The density graph is a proper approach to show how the historical fire events are distributed in different probability levels of a wildfire probability map. To plot a wildfire density graph, the density of fire events (the ratio of fire pixels over the ratio of total pixels) per each probability level is plotted on a diagram. In accordance with a theoretical base, the density of historical fire events should increase from the very low to the very high probability levels with an increasing rate.

ROC method

As an evaluation method that validates the quality of the modeling process, the receiver operating characteristic

(ROC) (Hanley and McNeil 1982) with success rate and prediction rate curves was used. The ROC curve is a useful method that has already enjoyed great success as a measure of performance for real case modeling efforts (Jaafari et al. 2014; Pourtaghi et al. 2015; Regmi and Poudel 2016; Tien Bui et al. 2016; Jaafari et al. 2017a, b). This method is a graphical representation of 100-specificity versus sensitivity rates for every modeling approach (Hanley and McNeil 1982). The plot shows the 100-specificity on the X-axis and the sensitivity on the Y-axis:

$$X = 100 - \text{Specificity} = 100 - \left(\frac{\text{TN}}{\text{TN} + \text{FP}}\right) \tag{11}$$

$$Y = \text{Sensitivity} = \left(\frac{\text{TP}}{\text{TP} + \text{FN}}\right) \tag{12}$$

where TN (true negative) and TP (true positive) are the number of pixels that are properly assigned as fire occurrences and FP (false positive) and FN (false negative) are the numbers of pixels erroneously assigned (Jaafari et al. 2017b). The best possible ROC curve passes through the point of (0, 1), where the area under curve (AUC) = 1 and represents 100% specificity (no false positives; the proportion of non-fire correctly predicted) and 100% sensitivity (no false negatives; the proportion of fire correctly predicted). Excellent models have AUC values greater than 0.9, and good models have AUC values greater than 0.7 (Hosmer et al. 2013). The success rate of a ROC curve that uses the training dataset reflects how well the model fits the training dataset, whereas the prediction rate uses the validation dataset and explains how well the model predicts the general probability of wildfire occurrence across the study area.

Results and discussion

Model results

All the twelve predictor variables were characterized by four mass functions of belief, disbelief, uncertainty, and plausibility that reveal the correlation between the fire ignitions and the variables across the study area (Table 2). Given to the similar proportion of fire pixels in the different classes of aspect, it appears that landscape-level difference in fire ignition among the different classes of this variable was modest (cf. Jaafari et al. 2017a). Although aspect significantly effects on local conditions such as exposure to sunshine, prevailing direction of winds, amount of rainfall, drying winds, and the morphologic structure that have been associated with fire incidents (de Vasconcelos et al. 2001; Adab et al. 2013; Chen et al. 2015; Jaafari et al. 2017a), values of mass functions failed to indicate a strong association between fire

Int. J. Environ. Sci. Technol. (2018) 15:373–384

379

Table 2 Spatial relationship between each predictor variable and wildfires extracted by using EBF model

Variable	Class	Number of pixels in domain	Number of fire pixels	Belief functions components			
				Bel	Dis	Unc	Pls
Aspect	F	1,293,545	104	0.128	0.109	0.763	0.891
	N	1,125,592	68	0.096	0.113	0.791	0.887
	NE	1,080,931	88	0.130	0.109	0.761	0.891
	E	1,149,299	82	0.114	0.111	0.776	0.889
	SE	1,351,149	91	0.107	0.112	0.781	0.888
	S	1,273,339	89	0.111	0.111	0.778	0.889
	SW	1,509,005	113	0.119	0.110	0.771	0.890
	W	1,503,418	107	0.113	0.111	0.776	0.889
	NW	1,399,058	71	0.081	0.115	0.804	0.885
Altitude (m)	<200	5,718,412	444	0.229	0.179	0.592	0.821
	200–500	2,944,720	163	0.163	0.215	0.622	0.785
	500–800	1,240,373	82	0.195	0.203	0.603	0.797
	800–1100	830,649	65	0.231	0.200	0.570	0.800
	>1100	951,182	59	0.183	0.203	0.614	0.797
Plan curvature (100/m)	Concave	844,087	51	0.326	0.356	0.318	0.644
	Flat	9,224,888	679	0.397	0.276	0.327	0.724
	Convex	1,616,361	83	0.277	0.368	0.355	0.632
TRI	<3	7,761,777	591	0.396	0.281	0.323	0.719
	3–6	2,344,592	118	0.262	0.370	0.368	0.630
	>6	1,578,967	104	0.342	0.349	0.309	0.651
Temperature (°C)	<14	2,310,677	183	0.201	0.161	0.637	0.839
	14–15	617,880	30	0.123	0.170	0.707	0.830
	15–16	1,397,226	70	0.127	0.173	0.699	0.827
	16–17	1,831,376	118	0.164	0.169	0.667	0.831
	17–18	1,745,637	137	0.199	0.163	0.637	0.837
	>18	3,782,540	275	0.185	0.163	0.652	0.837
Rainfall (mm)	<150	306,212	8	0.097	0.247	0.656	0.753
	150–200	5,861,249	266	0.169	0.328	0.503	0.672
	200–250	2,672,594	311	0.434	0.194	0.371	0.806
	>250	2,845,281	228	0.299	0.231	0.470	0.769
Evapotranspiration (mm/day)	<1300	227,151	7	0.057	0.163	0.780	0.837
	1300–1500	1,739,279	301	0.319	0.120	0.561	0.880
	1500–1700	984,916	185	0.346	0.136	0.517	0.864
	1700–1900	1,417,211	136	0.177	0.153	0.670	0.847
	1900–2100	2,387,681	81	0.063	0.183	0.754	0.817
	>2100	4,929,098	103	0.039	0.244	0.717	0.756
LULC	Forest	1,705,745	136	0.325	0.240	0.434	0.760
	Rangeland	5,104,500	74	0.059	0.398	0.543	0.602
	Farmland	4,158,113	598	0.587	0.101	0.312	0.899
	Shrubbery	717,178	5	0.028	0.261	0.711	0.739
Soil type	Inceptisol	3,157,383	145	0.110	0.222	0.668	0.778
	Mollisol	2,063,469	417	0.485	0.117	0.398	0.883
	Entisol	3,950,429	138	0.084	0.248	0.669	0.752
	Aridisol	1,763,401	22	0.030	0.226	0.744	0.774
	Alfisol	750,654	91	0.291	0.187	0.522	0.813
Proximity to roads (m)	<150	175,923	43	0.218	0.121	0.661	0.879
	150–300	154,972	30	0.173	0.123	0.704	0.877

Table 2 continued

| Variable	Class	Number of pixels in domain	Number of fire pixels	Belief functions components			
				Bel	Dis	Unc	Pls
	300–450	153,767	40	0.232	0.121	0.646	0.879
	450–600	152,271	22	0.129	0.124	0.747	0.876
	600–750	150,666	33	0.195	0.123	0.682	0.877
	>750	10,897,737	645	0.053	0.387	0.560	0.613
Proximity to rivers (m)	<150	806,703	64	0.178	0.163	0.658	0.837
	150–300	704,907	65	0.207	0.161	0.631	0.839
	300–450	685,411	61	0.200	0.162	0.638	0.838
	450–600	656,600	42	0.144	0.166	0.691	0.834
	600–750	617,301	33	0.120	0.167	0.713	0.833
	>750	8,214,414	548	0.150	0.181	0.669	0.819
Proximity to settlements (m)	<1000	320,996	46	0.205	0.160	0.635	0.840
	1000–2000	812,958	159	0.280	0.143	0.578	0.857
	2000–3000	929,602	186	0.286	0.138	0.575	0.862
	3000–4000	827,374	102	0.176	0.155	0.668	0.845
	>4000	8,794,406	320	0.052	0.404	0.544	0.596

ignition and this variable (Table 2). The spatial association between fire events and elevation shows that when the elevation increases, the probability of wildfire decreases. This result is supported by previous findings that low-elevation areas are more vulnerable to fire occurrence due to intensive human activities (Syphard et al. 2008; Oliveira et al. 2012; Adab et al. 2015; Guo et al. 2016), and favorable weather conditions, soil moisture, and vegetation cover (Adab et al. 2013). Although 83.5% of all fires recorded on the flat class of plan curvature, 79% of the study area was classified as flat terrain, which shows why plan curvature was identified as being marginally associated with fire occurrence. In contrast, in some studies (Pourtaghi et al. 2015; Pourghasemi 2016), the probability of fire occurrence was lower on flat terrain and higher on concave slopes. This is likely attributable to the dominant effect of other topographic features in the study area that swamped the influence of plan curvature on fire occurrence (cf. Parisien et al. 2012). For TRI, the class of <3 shows the highest wildfire susceptibility with almost 73% of recorded fire ignitions and a Bel value of 0.396. The spatial association between fire occurrence and annual temperature shows that up to 65.2% of fire events are related to the areas with >16 °C. For rainfall, while the classes of >200 mm covered 47.2% of the landscape, 66.3% of all fires occurred here; in contrast, the classes of <200 mm covered 52.8% of the landscape, but only 33.7% of all fires occurred here. Portions of the landscape with average evapotranspiration of 1500–1700 mm that encompassed 8.4% of the landscape and experienced 22.8% of all fires were the predestined places for wildfires. Climate variables have been widely cited as critical parameters affecting relative likelihood of wildfire occurrence (Syphard et al. 2008; Arpaci et al. 2014; Adab et al. 2015;

Pourtaghi et al. 2016; Jaafari et al. 2017a), presumably because fuels moisture content is largely a function of precipitation and temperature (Oliveira et al. 2014). Further, higher temperatures are expected to increase the amount of moisture that evaporates from land and water (Christian-Smith et al. 2012), which can lead to increase the probability of fire occurrence (Mhawej et al. 2016). The attribution of fire occurrences to the different categories of LULC shows that most fires regularly occur at those portions where land use has favored farmlands (73.5% of all fires). Here, fire events are often deliberately set to remove weeds, shrubs, and stubbles from the farmlands. Further, many of these fires are caused by arson due to the conflicts among people, or to expand the farmlands into the forests. Among different categories of soil type, fire probability was more likely on Mollisol soils, which mostly permit farmlands and forests and contain 51.3% of recorded fire events, but cover 17.7% of the landscape. In the case of the effect of proximity to rivers, the results showed that none of the distance classes deviated from an expectation of random fire occurrence that reveals a weak association between proximity to rivers and fire ignition in the study area. For the effect of proximity to roads, the results showed that there was a clear contrast between areas within 750 m and those >750 m from road networks. Although only 6.7% of the land area was within 750 m of road networks, 20.7% of all fires occurred there; 79.3% of fires occurred in places >750 m from road networks that comprised 93.3% of the land area. These results are supported by previous findings that demonstrated a positive relationship between the spatial patterns of ignitions and landscape accessibility (e.g., de Vasconcelos et al. 2001; Martínez-Fernández et al. 2013; Jaafari et al. 2017a). Similarly, proximity to settlements showed a

Fig. 5 Function maps generated from EBF model results
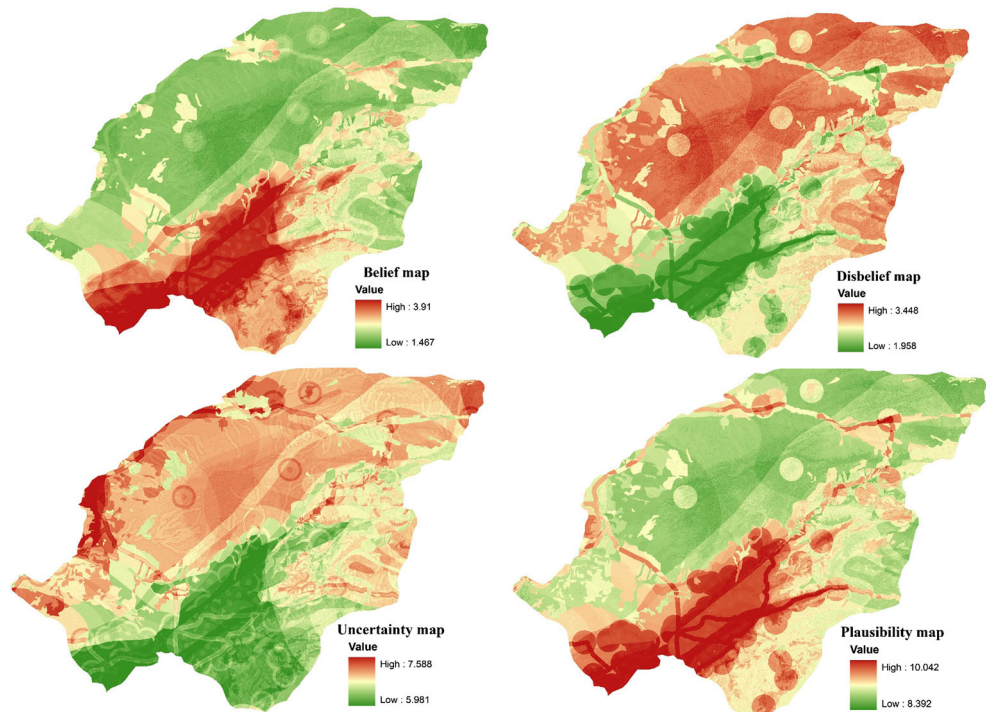
clear contrast in fire probability: landscape portions within 4000 m to a human settlement were positively associated with fires (60.6% of all fires on 24.7% of the landscape area), whereas portions >4 km from settlements were negatively associated with fire occurrence (39.4% of all fires on 75.3% of the landscape area). The proximity variables directly (proximity to roads and settlements) or indirectly (proximity to rivers) provide significant opportunities for people to change fire regimes in different ways, such as through changing fuel types, modifying fuel structure and continuity, and igniting the forest and other vegetated areas in different seasons and under different weather conditions than would occur naturally (Bowman et al. 2011; Adab et al. 2015; Jaafari et al. 2017a).

The final results of the EBF model are shown in Fig. 5. The spatial distribution of the mass functions components can be interpreted in terms of physiography (Rahmati and Melesse 2016) and the characteristics of landscapes. Here, higher degrees of Bel and Pls are near to roads and human settlements and correlated with flat areas, while lower degrees are correlated with those portions of the landscape where land use has favored farmlands. Further, higher probability of fire occurrence is associated with those portions of the landscape that are characterized with higher degrees of Bel and lower degree of Dis values. The Unc map offers insights into the degree of uncertainty associated with the predictions results on those portions of the landscape that geospatial data failed to provide enough evidence for confirmation (Althuwaynee et al. 2014; Mogaji et al. 2016).

The integrated map of wildfire probability (Fig. 6) depicts five probability classes that range from very low to very high across the landscape. The map provides a relative ranking of the likelihood of wildfire occurrence that conforms somewhat to human infrastructure and associated activities, i.e., farmlands, settlement areas, and road networks. Farmlands and the areas surrounding human settlements and along road networks that are visible as circular rings (Fig. 6) have mostly been classified into high to very high level of probability to fire occurrence. These results are in close agreement with those who found a positive relationship between intensive human activities and increased probability of fire occurrence (e.g., Catry et al. 2009; Adab et al. 2015; Guo et al. 2016).

A comparison of the spatial distribution of the probability levels indicates that the zones of high and very high probability cover 38.7% of the entire study area, zone of moderate probability covers 20.4% of the area and zones of low and very low probability cover 40.9% of the study area (Fig. 7). Thus, resources and infrastructure, mitigations strategies, and suppression efforts should be allocated to the roughly 60% of the landscape where fires are most likely to occur.

Validation results

Figure 7 shows the fire density in five classes of the wildfire probability map. Fire density that indicates to a ratio between the number of fire pixels and the number of class pixels in each probability-level class ranges from 0.18 to 2.94. The highest value is for very high probability levels which is followed by

382

Int. J. Environ. Sci. Technol. (2018) 15:373–384

**Fig. 6** Wildfire probability distribution map
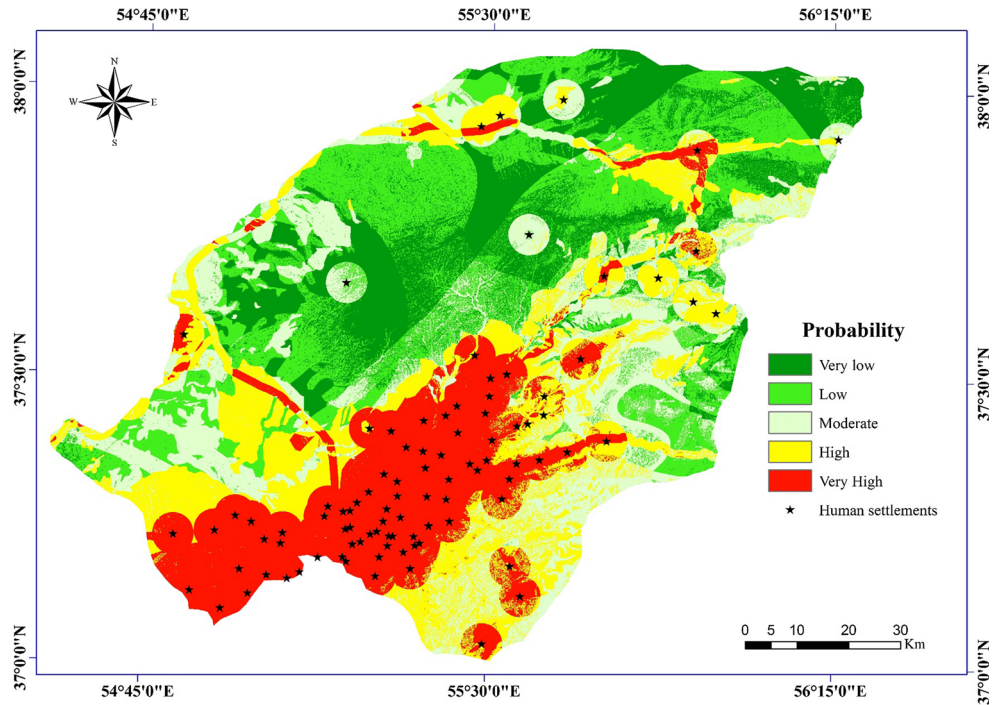
high probability, moderate probability, low probability, and very low probability, respectively. The quantized increase in the fire density from very low probability level to very high probability level that was achieved in this study indicates to the capability of the EBF model to delineate the general probability levels of fire occurrence with respect to the previous ignitions.

In addition, ROC-AUC method with success rate and prediction rate curves was used to measure the overall performance of the EBF model. Figure 8 shows the success rate and prediction rate curves that were plotted using ''sensitivity'' and ''100-specificity'' values. Since the steep slope of the first part of the curves and their further distance from the diagonal trend indicate a more reliable model performance (Jaafari et al. 2015), the results of this

study show the good performance of the GIS-based EBF model and therefore the good reliability of the modeling process, including the selection of the predictor variables. More precisely, the success rate curve shows that nearly 78% of the fire ignitions (of the training subset) were captured within 10% of the probability map, indicating high model efficiency. The prediction rate curve also showed that 70% of the fire events (of the validation subset) can predict within 10% of the probability map. Furthermore, the validation results show that the AUC for success rate and prediction rate curves is 84.14 and 81.03%, respectively, indicating a very good reliability of the EBF model to estimate relative wildfire likelihood (cf. Hosmer et al. 2013). However, some drawbacks exist due to its statistical assumptions (Althuwaynee et al. 2014). The main weak point of this model is to neglect the relationship between the predictor variables that can actually lead to assume that all the variables carry equal weight in determining the level of vulnerability of the pixels across the landscape. In this situation, some of the variables might be ignored during the calculations, which can decrease the prediction quality and enhance the model uncertainty (Jaafari et al. 2014). Nonetheless, despite this limitation, the GIS-based EBF model adopted in this study had the ability to deal with generalized and manifold information from multiple sources. The model was quite flexible and easy to apply within the GIS environment to explore spatial association between each category of the predictor variables and historical fire locations to predict where wildfires may occur in the future. Overall, the model represented a
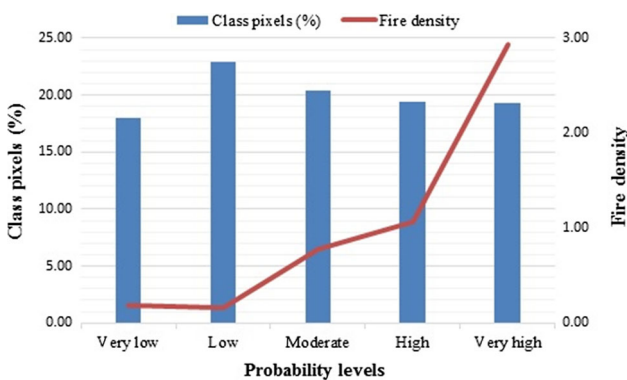
Fig. 7 Distribution of probability levels and fire density in wildfire probability map
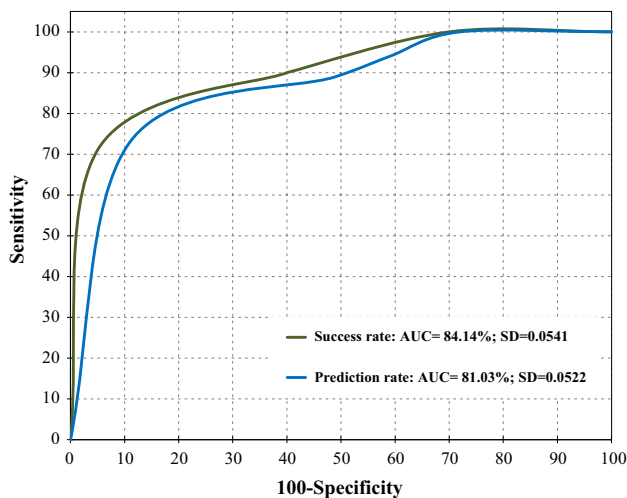
Fig. 8 ROC curves and AUC for EBF model

robust and analytic framework for studying wildland fires and for supporting spatial planning. This assertion is supported by previous applications of the EBF model in different contexts of hazard assessment such as forest fire (Pourghasemi 2016), landslide (e.g., Althuwaynee et al. 2014; Gorum and Carranza 2015; Regmi and Poudel 2016), and environmental studies (e.g., Thiam 2005; Carranza et al. 2005, 2008; Rahmati and Melesse 2016; Naghibi et al. 2016; Mogaji et al. 2016).

Conclusion

In this paper, GIS-automated techniques were integrated with the quantitative data-driven EBF model to estimate wildfire probability in a part of the Hyrcanian ecoregion. A wide array of predictor variables including aspect, elevation, plan curvature, TRI, annual temperature and rainfall, evapotranspiration, LULC, soil type, and proximity to roads, rivers, and settlements were used as the inputs to the EBF model. The application of model was divided into training and validation steps. 70% of historical fire events were used for training the model, and the remainder was used for validation purpose. The results showed that the GIS-based EBF framework is successful for identifying fire-prone areas. The derived probability map revealed that zones of moderate, high, and very high probability covered nearly 60% of the land area and the probability of fire occurrence was strongly dependent upon human infrastructure and associated activities. As a whole, the results revealed a good reliability of the adopted approach for investigating the spatial relationship between each category of the predictor variables and historical fire locations throughout the study area that enabled us to study in detail the effect of each variable on the probability of fire occurrence.

Due to the continuous human activity throughout the landscape, the probability of fire occurrence is subject to change. Thus, a periodic updating of the results is desirable. Recommendations for future research include evaluating the capability of different bivariate, multivariate, and knowledge-based models, as well as assessing and employing a wider range of predictor variables, especially those related to human actions, to produce more reliable estimates of wildfire probability.

References

Abdi O, Kamkar B, Shirvani Z, Teixeira da Silva JA, Buchroithner MF (2016) Spatial-statistical analysis of factors determining forest fires: a case study from Golestan, Northeast Iran. Geomat Nat Hazards Risk. doi:10.1080/19475705.2016.1206629

Adab H, Kanniah KD, Solaimani K (2013) Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. Nat Hazards 65(3):1723–1743

Adab H, Kanniah KD, Solaimani K, Sallehuddin R (2015) Modelling static fire hazard in a semi-arid region using frequency analysis. Int J Wildland Fire 24(6):763–777

Althuwaynee OF, Pradhan B, Park HJ, Lee JH (2014) A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. CATENA 114:21–36

Arpaci A, Malowerschnig B, Sass O, Vacik H (2014) Using multivariate data mining techniques for estimating fire susceptibility of Tyrolean forests. Appl Geogr 53:258–270

Bowman DM, Balch J, Artaxo P, Bond WJ, Cochrane MA, D'antonio CM, Kull CA (2011) The human dimension of fire regimes on Earth. J Biogeogr 38(12):2223–2236

Carranza EJM, Woldai T, Chikambwe EM (2005) Application of data-driven evidential belief functions to prospectivity mapping for aquamarine-bearing pegmatites, Lundazi district, Zambia. Nat Resour Res 14(1):47–63

Carranza EJM, Van Ruitenbeek F, Hecker C et al (2008) Knowledge-guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE Spain. Int J Appl Earth Obs 10:374–387

Catry FX, Rego FC, Bação FL, Moreira F (2009) Modeling and mapping wildfire ignition risk in Portugal. Int J Wildland Fire 18(8):921–931

Chen F, Du Y, Niu S, Zhao J (2015) Modeling forest lightning fire occurrence in the Daxinganling mountains of Northeastern China with MAXENT. Forests 6(5):1422–1438

Christian-Smith J, Gleick PH, Cooley H, Allen L, Vanderwarker A, Berry KA (2012) A twenty-first century US water policy. Oxford University Press, Oxford

Chuvieco E, Congalton RG (1989) Application of remote sensing and geographic information systems to forest fire hazard mapping. Rem Sens Environ 29(2):147–159

de Vasconcelos MP, Silva S, Tome M, Alvim M, Pereira JC (2001) Spatial prediction of fire ignition probabilities: comparing logistic regression and neural networks. Photogramm Eng Rem Sens 67(1):73–81

Dempster AP (1967a) Upper and lower probabilities induced by a multivalued mapping. Ann Math Stat 38:325–339

Dempster AP (1967b) Upper and lower probability inferences based on a sample from a finite univariate population. Biometrika 54:515–528

Dickson BG, Prather JW, Xu Y, Hampton HM, Aumack EN, Sisk TD (2006) Mapping the probability of large fire occurrence in northern Arizona, USA. Landsc Ecol 21:747–761

Eskandari S, Chuvieco E (2015) Fire danger assessment in Iran based on geospatial information. Int J Appl Earth Obs Geoinf 42:57–64

Gorum T, Carranza EJM (2015) Control of style-of-faulting on spatial pattern of earthquake-triggered landslides. Int J Environ Sci Technol 12(10):3189–3212

Guo F, Wang G, Su Z, Liang H, Wang W, Lin F, Liu A (2016) What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. Int J Wildland Fire 25(5):505–519

Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL (2006) Multivariate data analysis, vol 6. Pearson Prentice Hall, Upper Saddle River, NJ

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36

Hosmer DW Jr, Lemeshow S, Sturdivant RX (2013) Applied logistic regression, vol 398. Wiley, London

Jaafari A, Najafi A, Pourghasemi HR, Rezaeian J, Sattarian A (2014) GIS-based frequency ratio and index of entropy models for landslide susceptibility assessment in the Caspian forest, northern Iran. Int J Environ Sci Technol 11(4):909–926

Jaafari A, Najafi A, Rezaeian J, Sattarian A, Ghajar I (2015) Planning road networks in landslide-prone areas: a case study from the northern forests of Iran. Land Use Pol 47:198–208

Jaafari A, Gholami DM, Zenner EK (2017a) A Bayesian modeling of wildfire probability in the Zagros Mountains, Iran. Ecol Informat 39:32–44

Jaafari A, Rezaeian J, Omrani MSO (2017b) Spatial prediction of slope failures in support of forestry operations safety. Croat J Forest Eng 38(1):107–118

Liao D, Valliant R (2012) Variance inflation factors in the analysis of complex survey data. Surv Method 38(1):53–62

Lozano FJ, Suárez-Seoane S, Kelly M, Luis E (2008) A multi-scale approach for modeling fire occurrence probability using satellite data and classification trees: a case study in a mountainous Mediterranean region. Rem Sens Environ 112(3):708–719

Martínez J, Vega-Garcia C, Chuvieco E (2009) Human-caused wildfire risk rating for prevention planning in Spain. J Environ Manag 90(2):1241–1252

Martínez-Fernández J, Chuvieco E, Koutsias N (2013) Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. Nat Hazards Earth Syst Sci 13(2):311–327

Mhawej M, Faour Gh, Abdallah Ch, Adjizian-Gerard J (2016) Towards an establishment of a wildfire risk system in a Mediterranean country. Ecol Inf 32:167–184

Mogaji KA, Omosuyi GO, Adelusi AO, Lim HS (2016) Application of GIS-based evidential belief function model to regional groundwater recharge potential zones mapping in hardrock geologic Terrain. Environ Process 3(1):93–123

Naghibi SA, Pourghasemi HR, Dixon B (2016) GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environ Monit Assess 188(1):1–27

Nunes AN, Lourenço L, Meira AC (2016) Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). Sci Total Environ 573:1190–1202

Oliveira S, Oehler F, San-Miguel-Ayanz J, Camia A, Pereira JM (2012) Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. For Ecol Manag 275:117–129

Parisien MA, Snetsinger S, Greenberg JA, Nelson CR, Schoennagel T, Dobrowski SZ, Moritz MA (2012) Spatial variability in wildfire probability across the western United States. Int J Wildland Fire 21(4):313–327

Parisien MA, Miller C, Parks SA, DeLancey ER, Robinne FN, Flannigan MD (2016) The spatially varying influence of humans on fire probability in North America. Environ Res Lett 11(7):075005

Pourghasemi HR (2016) GIS-based forest fire susceptibility mapping in Iran: a comparison between evidential belief function and binary logistic regression models. Scand J For Res 31(1):80–98

Pourghasemi HR, Beheshtirad M, Pradhan B (2016) A comparative assessment of prediction capabilities of modified analytical hierarchy process (M-AHP) and Mamdani fuzzy logic models using Netcad-GIS for forest fire susceptibility mapping. Geomat Nat Hazards Risk 7(2):861–885

Pourtaghi ZS, Pourghasemi HR, Rossi M (2015) Forest fire susceptibility mapping in the Minudasht forests, Golestan province, Iran. Environ Earth Sci 73(4):1515–1533

Pourtaghi ZS, Pourghasemi HR, Aretano R, Semeraro T (2016) Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. Ecol Indic 64:72–84

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

Regmi AD, Poudel K (2016) Assessment of landslide susceptibility using GIS-based evidential belief function in Patu Khola watershed, Dang, Nepal. Environ Earth Sci 75(9):1–20

Rodrigues M, de la Riva J (2014) An insight into machine-learning algorithms to model human-caused wildfire occurrence. Environ Model Softw 57:192–201

Rollins MG, Keane RE, Parsons RA (2004) Mapping fuels and fire regimes using remote sensing, ecosystem simulation, and gradient modeling. Ecol Appl 14(1):75–95

Satir O, Berberoglu S, Donmez C (2016) Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. Geomat Nat Hazards Risk 7(5):1645–1658

Shafer G (1976) A mathematical theory of evidence. Princeton University Press, Princeton, p 297

Silva GL, Soares P, Marques S, Dias MI, Oliveira MM, Borges JG (2015) A Bayesian modelling of wildfires in Portugal. In: Bourguignon J-P, Jeltsch R, Pinto AA, Viana M (eds) Dynamics, games and science. Springer, Berlin, pp 723–733

Stephens SL, Agee JK, Fulé PZ et al (2013) Managing forests and fire in changing climates. Science 342:41–42

Syphard AD, Radeloff VC, Keuler NS, Taylor RS, Hawbaker TJ, Stewart SI, Clayton MK (2008) Predicting spatial patterns of fire on a southern California landscape. Int J Wildland Fire 17(5):602–613

Thiam AK (2005) An evidential reasoning approach to land degradation evaluation: Dempster–Shafer theory of evidence. Trans GIS 9:507–520

Tien Bui D, Le KTT, Nguyen VC, Le HD, Revhaug I (2016) Tropical forest fire susceptibility mapping at the cat Ba National Park area, Hai Phong City, Vietnam, using GIS-based kernel logistic regression. Rem Sens 8(4):347

Tien Bui D, Bui QT, Nguyen QP, Pradhan B, Nampak H, Trinh PT (2017) A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. Agr For Meteorol 233:32–44

Walley P (1987) Belief function representations of statistical evidence. Ann Stat 15:1439–1465