doi: 10.17221/7/2017-JFS

Evaluation of forest fire risk using the Apriori algorithm and fuzzy c-means clustering

Ali Akbar JAFARZADEH[1], *Ali MAHDAVI*[1]*, *Heydar JAFARZADEH*[2]

[1]*Department of Forest Sciences, Faculty of Agriculture and Natural Resources, University of Ilam, Ilam, Iran*

[2]*Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Ilam, Iran*

**Corresponding author: a_amoli646@yahoo.com*

Abstract

Jafarzadeh A.A., Mahdavi A., Jafarzadeh H. (2017): Evaluation of forest fire risk using the Apriori algorithm and fuzzy c-means clustering. J. For. Sci., 63: 370–380.

In this study we evaluated forest fire risk in the west of Iran using the Apriori algorithm and fuzzy c-means (FCM) clustering. We used twelve different input parameters to model fire risk in Ilam Province. Our results with minimum support and minimum confidence show strong relationships between wildfire occurrence and eight variables (distance from settlement, population density, distance from road, slope, standing dead oak trees, temperature, land cover and distance from farm land). In this study, we defined three clusters for each variable: low, middle and high. The data regarding the factors affecting forest fire risk were distributed in these three clusters with different degrees of membership and the final map of all factors was classified by FCM clustering. Each layer was then created in a geographic information system. Finally, wildfire risks in the area obtained from overlaying these layers were classified into five categories, from very low to very high according to the degree of danger.

Keywords: wildfire; association rules; fuzzification; Ilam Province; Iran

Iran is one of the disaster-prone countries in the world. The phenomenon of forest fires is one of the most significant natural disasters that occur in the country. Even though Iranian forest fire statistics are not very reliable, some estimates indicate that more than 5,000 ha of land are affected annually (Adab et al. 2013). Forest ecosystems are important natural resources with a role in maintaining environmental balance, and their health is a good indicator of the ecological conditions prevailing in a region. Forest fires are one of the most significant factors threatening the extinction of wild animals and natural vegetation (Rajeev Kumar et al. 2002). In some cases, forest fires cause major disturbances that result in enormous physical, biological, socioeconomic and environmental losses (Bowman et al. 2009; Paz et al. 2011; Eskandari, Chuvieco 2015). Evaluation of the spatial relationship between the location of hotspot occurrence and specific geographical factors near the hotspots is essential. Therefore, the possible factors influencing forest fires can be determined to predict future hotspot occurrence. Evaluation of forest fire risk can contribute to reducing the negative impact of fire by improving the level of preparedness of forest managers, and can provide new information to guide planning. Several recent studies have presented models to predict forest fire danger on different spatial and temporal scales (Merril, Alexander 1987; Taylor, Alexander 2006; Prasad et al. 2008; Loboda 2009; Romero-Ruiz et al. 2010; Padilla, Vega-Garcia 2011; Eskandari, Chuvieco 2015), and reports have been published on the occurrence and impact factor of forest fires worldwide. Tian et al. (2012) and Hu and Jin (2002) studied the forest fire regime and conducted an analysis of the factors affecting fire distribution. Different studies have been carried out to evaluate forest fire risk in Iran. Researchers have used the analytical hierarchy process, remote sensing and geographic information systems (GIS) to assess forest fire risk on regional scales (Akbari et al. 2008; Mohammadi et al. 2011; Salamati et al. 2011; Mahdavi et al. 2012; Zarekar et al. 2013). Many studies have focused on mapping forest wildfire risk and on describing the different classes of potential risk used different methods and spatial scales. GIS, remote sensing and mathematical techniques have provided opportunities to analyse and manage forest wildfires quantitatively. These technologies represent an effective instrument to predict the area at risk of forest wildfires through modelling procedures.

In this study, for fire risk assessment, a data mining technique, namely the Apriori algorithm (association rules) was applied in the study site. Spatial data mining is a logical process used to find relevant spatial data in large datasets (Aggarwal, Rani 2013). Given the possibility of interesting associations among the data, we needed automated and efficient tools to find and to organise these associations. In spatial data mining, attributes of neighbours (factors affecting the parameter under study) of an object (hotspot) may have a significant influence on the object itself (Ester et al. 1997). Fuzzy c-means (FCM) clustering is also a useful algorithm for the determination of buffer areas, called hotspots in crime analysis, car crash analysis, disease diffusion analysis, etc. This study was aimed at determining the role of the most important factors influencing the potential risk of forest fires in Ilam Province, Iran, using the Apriori algorithm and preparing the map of risk zonation using FCM.

MATERIAL AND METHODS

Study area. The study area with a surface area of about 213,000 ha, is located in Ilam Province in western Iran, within 33°20'36" to 33°50'35"N latitude and 45°40'34" to 46°51'12"E longitude (Fig. 1). The area encompasses a diverse range of elevations, slopes, populations, land uses, etc. The climate is mostly characterised as Mediterranean arid and semi-arid. Altitude ranges from 150 to 2,750 m a.s.l. The main tree species of the forests of the area, which are part of the Zagros forests, consist of *Quercus brantii* Lindley, *Quercus libani* Olivier and *Pistacia atlantica* Desfontaines; the dominant species is *Q. brantii*. The people of Ilam Province are highly dependent on these forests for their livelihood, resulting in quantitative and qualitative reductions to the forests (Fattahi 2003).

Selection of spatial variables. Forest wildfire statistics were collected for the period from 2010 to 2015. The fire zones data were provided by the Natural Resources Management Office, Ilam, Iran

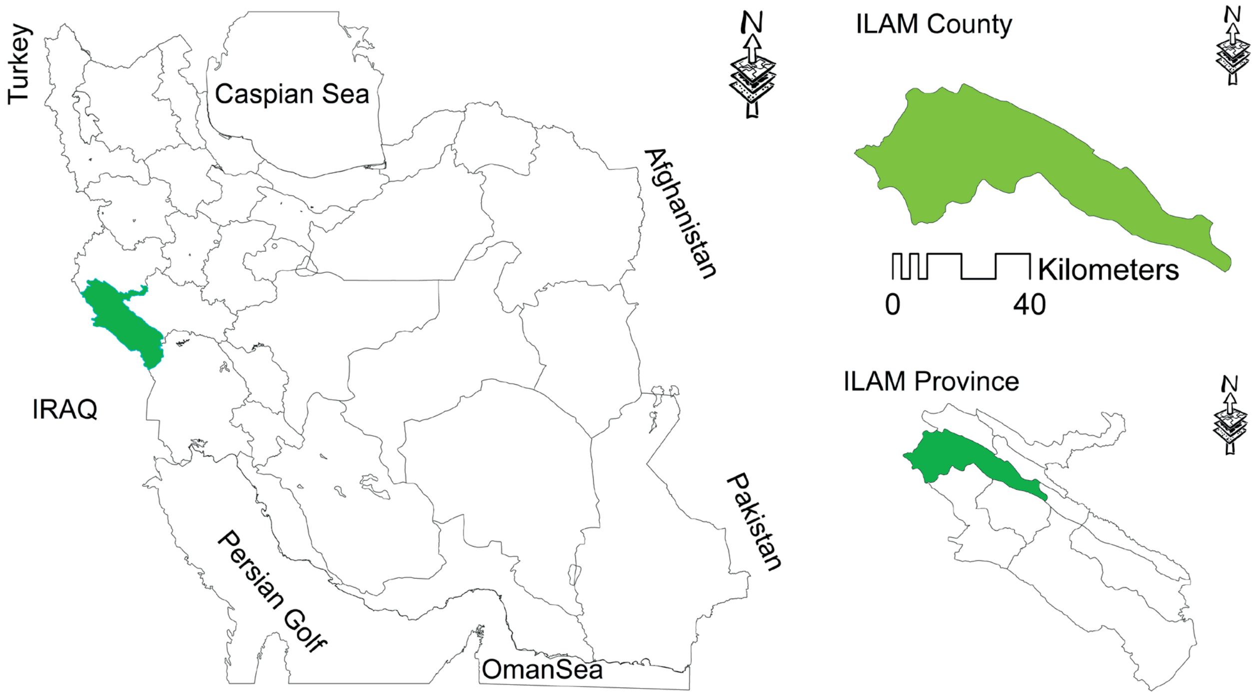

Fig. 1. Location of the study area in Ilam Province, Iran

(Forest, Range and Watershed Management Organization of Iran 2015). Fig. 2 also indicates the locations of the occurrence of wildfires in the study area. The natural and physical characteristics of the forest fire locations were determined according to the frequency of wildfire occurrence in the study area. We used twelve different input parameters (Table 1) to model fire susceptibility in Ilam Province. These parameters and their parametric effects on fire susceptibility were selected according to previous studies (Mahdavi et al. 2012; Eskandari et al. 2013).

A digital elevation model (DEM) of the study area was produced using 20-m contour lines of the region (1:25,000 scale), by an interpolation technique (Topo to Raster). Then, elevation, aspect and slope layers of the region were produced using the DEM. The cities, roads and agricultural vector maps were provided by the National Geography Organization of Iran. The layers of distance from settlement, roads, farmlands and rivers were extracted using the buffering application in ArcGIS software (Version 10.4.1, 2016). The population density layer was obtained by establishing a population database for point maps of villages and cities in the study area. Then, the population density was calculated using the territorial area of the villages and population size in each area. The annual mean precipitation

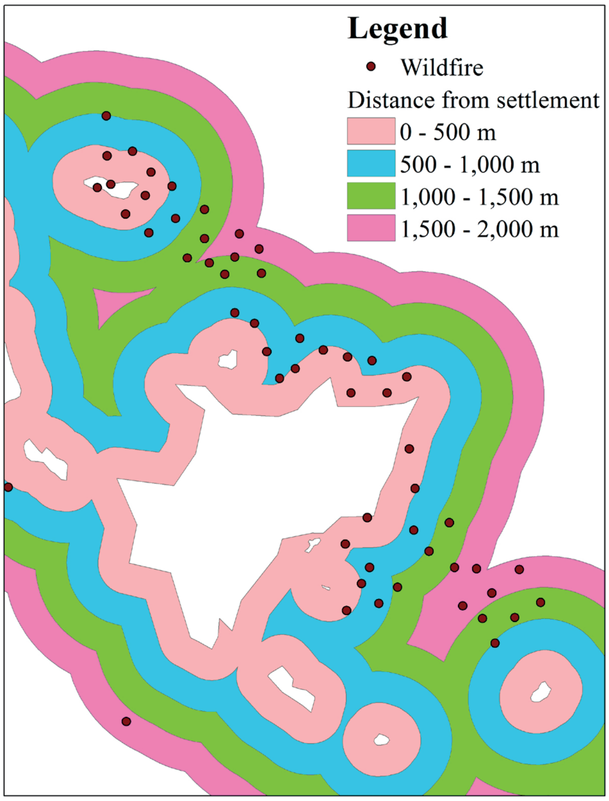

Fig. 2. Wildfire locations overlaid with the distance from settlement layer

Table 1. Input parameters used for the forest fire risk assessment system

Variable	Source
Elevation Slope Aspect	digital elevation model
Distance from settlement Distance from roads Distance from farm land Distance from rivers	National Geography Organization of Iran
Population density	population database
Temperature Precipitation	Iran Meteorological Organization
Land use/land cover	Landsat 8 OLI imagery
Standing dead oak wood	Natural Resources Management Office, Ilam, Iran

and temperature data from 1984 to 2014 were collected from the Iran Meteorological Organization. The annual mean precipitation and temperature raster values were generated using the Kriging interpolation method.

Land cover layer was derived from satellite image interpretation, using both digital image classification methods and visual interpretation techniques. The existing land use/land cover map of the region was employed as ancillary data. Landsat 8 OLI imagery (Orbital Sciences Corporation, USA), dated April 1, 2014, resized to a spatial resolution of 30 m, was the main remotely-sensed data source for this research. Remote sensing image processing was performed using ENVI (Version 4.5., 2009). The overall accuracy of the classified map was 87%, and Kappa index was 0.73. The standing dead oak wood map as a fuel load map was obtained from the General Natural Resources Office of Ilam Province, Iran. All of these spatial variables were transformed to the target resolution and geo-referenced in the Universal Transverse Mercator standard projection system (extended zone 38, using the WGS84 ellipsoid).

Data transformation. The Apriori algorithm (association rules) requires a dataset in the transaction format which contains transaction ID and item sets. Several pre-processing steps were performed to create the transaction dataset from the set of layers of influencing factors for wildfire occurrence. ArcGIS software tools were utilised in data transformation to manage the spatial database, perform spatial operations and analyse and visualise spatial data. For example, for each wildfire point, we determined whether the points were inside a population density class in which population density objects were represented by polygons (Fig. 2). This operation was also used to relate the wildfire occurrence layer to other layers. After overlaying the wildfire point with other layers, a database of the presence or absence of wildfire on the different layers was generated (Table 2). In order to discover association rules between spatial variables and wildfire occurrence using the Apriori algorithm, each layer is related to the wildfire layer.

Apriori algorithm. The Apriori algorithm was introduced by Agrawal and Srikant (1994) to discover frequent item sets and association rules in a transactional dataset. One of the most important tasks of the Apriori algorithm are association rules which can be used in different domains (Patel et al. 2011). Association rules are a general Apriori algorithm method and are used to extract useful patterns from large databases (Cakir, Aras 2012). Association rules from databases involve rules and steps based followed by all scholars in a given field to promote the methods and algorithms. A simple association rule can be represented as: bread → cheese (support = 0.1, confidence = 0.8). Simply put, this rule states that there is an association between buying bread and cheese with the support factor indicating that bread and cheese come together in 10% of the transactions and the confidence factor demonstrating that cheese has taken part in transactions where bread is also present. In this case, in 80% of the transactions including cheese, bread was also present. With this rule, we can assume that in the future, those who buy bread are most likely to also buy cheese during those transactions. Such information can help retailers explore opportunities for cross-selling (Gottwald 2006).

We represent the association rules as follows: $I = \{i_1, i_2, ..., i_m\}$ is a set of items and $T = \{t_1, t_2, ..., t_n\}$ is a set of transactions, each of which contains items from item set I. Therefore, transaction t_i contains a set of items where $t_i \subseteq I$. Association rules is a concept in the form of $X \rightarrow Y$ where $X \subset I$ and $Y \subset I$ as well as $X \cap Y = \varnothing$ where X and Y are called item sets (Liu 2007). In an association rule in the form of $X \rightarrow Y$, X is termed "antecedent" and Y "consequent". Obviously, the antecedent covers the consequent. Support and confidence are the most important measures of a rule.

The support is measured by Eq. 1:

$$\sup\left(A \Rightarrow B\right) = \frac{A \cap B}{X} \tag{1}$$

where:

A, B – different items in the database,

X – total items in the database.

Table 2. Distance from a settlement for a portion of the wildfire database

Wildfire	Distance from settlement class (m)			
	0–500	500–1,000	1,000–1,500	1,500–2,000
1	1	0	0	0
2	0	1	0	0
3	1	0	0	0
4	0	1	0	0
5	0	0	1	0

This rule mines all the transactions where A and B were present and then compares them with the minimum support specified by the user. Only those transactions in which support was equal to or bigger than the minimum support were chosen and the rest were eliminated as uninteresting rules.

Then, the confidence of the new list was calculated as follows (Eq. 2):

$$\operatorname{conf}\left(A \Rightarrow B\right) = P\left(B \middle| A\right) = \frac{\sup\left(A \rightarrow B\right)}{\sup\left(A\right)}; \sup\left(A\right) = \frac{A}{X} \tag{2}$$

where:

P – probability.

First, transactions containing item A are measured. Then, transactions containing item B as well are mined from them. The output of this equation is compared with the minimum confidence and the rules thus filtered are introduced as association rules (Srikant, Agrawal 1996). The Apriori algorithm was the first attempt to mine association rules from large sets of data (Agrawal, Srikant 1994). This algorithm mines association rules in two steps: (*i*) discovering frequent patterns, (*ii*) mining association rules.

Fuzzification. Fuzzification methods have been used in many studies (Feng 1995; Deng 1999; Mikhailov, Tsvetinov 2004; Erensal et al. 2006; Wang et al. 2008). In this study, they were used for the clustering of factors affecting forest fires (important factors that have been identified by the Apriori algorithm). From among the algorithms which lend themselves to fuzzy clustering, we chose the popular FCM clustering algorithm (Bezdek 1981; Shihab, Burger 1998). In this step, fuzzy sets were defined for all significant factors. More specifically, three fuzzy sets were defined for these variables. Three fuzzy sets, namely low, medium and high were defined in a qualitative 0–58 scale as shown in Fig. 3. Membership function defines a fuzzy set by mapping crisp values from its domain to the set's associated degree of membership. The degree to which a crisp value is compatible with a membership function

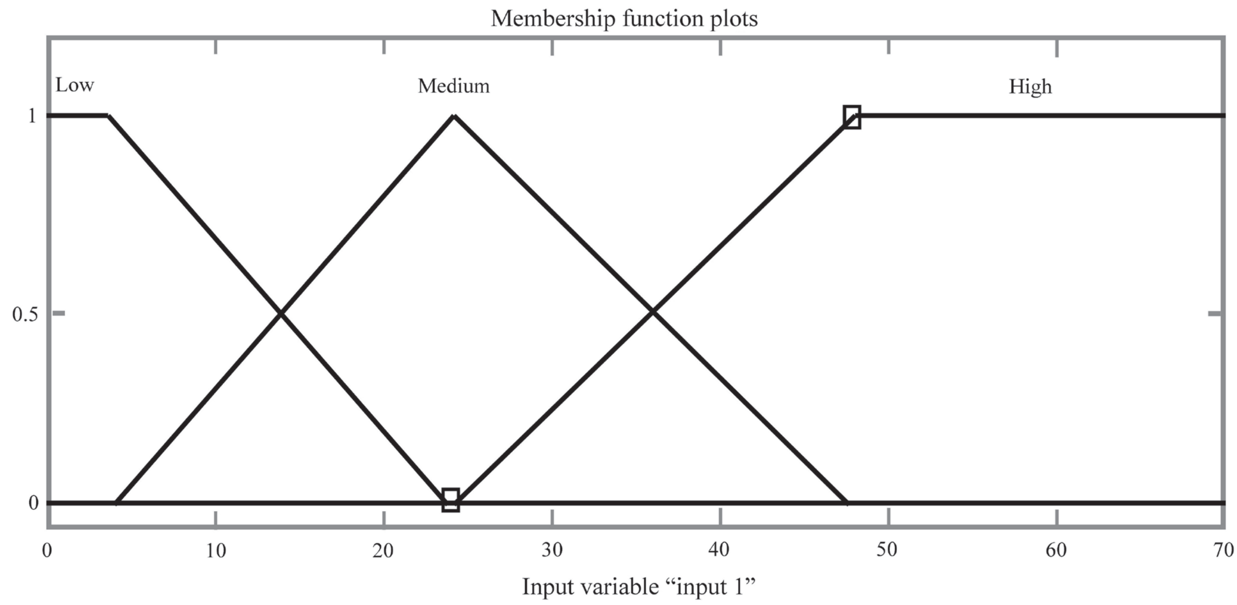

Fig. 3. Membership functions defined for significant factors

is referred to as the degree of membership, and the value ranges from 0 to 1. This is also known as the truth value or fuzzy input. A label is the descriptive name used to identify a membership function. The number of labels corresponds to the number of regions that the universe should be divided into, so that each label describes a region of behaviour. A scope must be assigned to each membership function that numerically identifies the range of input values that corresponds to a label. The type of representation of the membership function depends on the base set. A membership function can be simply viewed as a curve that defines how each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. The input space is sometimes referred to as the universe of discourse. The only condition a membership function must ultimately satisfy is that it must range between 0 and 1. There are a number of ways in which membership function can be represented, including: (*i*) triangular membership function, (*ii*) trapezoidal membership function, (*iii*) Gaussian function, (*iv*) generalized Bell membership function and (*v*) sigmoidal membership function.

In this study, we used a triangular membership function. This is specified by three parameters (a, b, c) with ($a < b < c$) determining the x coordinates of the three angles. Variable x is the crisp value whose membership function is to be determined within the universe of discourse. The graphical representation of the triangular membership function shown in Fig. 3 can be represented mathematically by either of the two mathematical models in Eq. 3. To state the membership function degree of the desired attribute (significant factors), we used Eqs. 4–6:

$$\text{triangle}(x{:}\ a,b,c) = \max\left\{\min\left[\frac{x-a}{b-a},\frac{c-x}{c-b}\right],0\right\}$$

$$\text{triangle}(x{:}\ a,b,c) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ 0, & c \le x \end{cases} \tag{3}$$

$$\mu_{\text{low}}(\text{attribute}) = \begin{cases} 1 & \text{attribute} \le 3.7293 \\ \dfrac{24.0745-\text{attribute}}{20.3452} & 3.7293 \le \text{attribute} \le 24.0745 \end{cases} \tag{4}$$

$$\mu_{\text{medium}}(\text{attribute}) = \begin{cases} \dfrac{\text{attribute}-3.7293}{20.3452} & 3.7293 \le \text{attribute} \le 24.0745 \\ \dfrac{47.8889-\text{attribute}}{23.8144} & 24.0745 \le \text{attribute} \le 47.8889 \end{cases} \tag{5}$$

$$\mu_{\text{high}}(\text{attribute}) = \begin{cases} \dfrac{\text{attribute}-24.0745}{23.8144} & 24.0745 \le \text{attribute} \le 47.8889 \\ 1 & \text{attribute} \ge 47.8889 \end{cases} \tag{6}$$

where:
μ – membership degree.

Wildfire map and its validation. After clustering significant factors in three levels using the FCM approach, we assigned them to each data layer in GIS. Finally, forest fire risk in the study area was obtained from overlaying these layer maps which are classified into five categories depending on the risk of danger, from very low to very high. Unfortunately, there are no up-to-date long-term national forest fire statistics for the study area. Therefore, validation of our forest fire risk analysis was based on a comparison of estimated risk values with actual fire occurrence.

RESULTS

The GIS layers of all investigated variables and their corresponding classes in this study are shown in Fig. 4. The Apriori algorithm which is available in the statistical computing Weka software (Version 3.8, 2016) was executed on the dataset and it generated 6,581 association rules. The purpose of this study was to find factors that strongly influence wildfire occurrence. Therefore, all further analysis, used only those association rules that included a wildfire occurrence. As a result, there were 14 association rules containing wildfire occurrence generated from the dataset with the minimum support of 10% and the minimum confidence of 80% (Table 3). Our results with the minimum support of 80% and the minimum confidence of 100% show strong relationships among wildfire occurrence and eight variables (distance from settlement, population density, distance from road, slope, standing dead

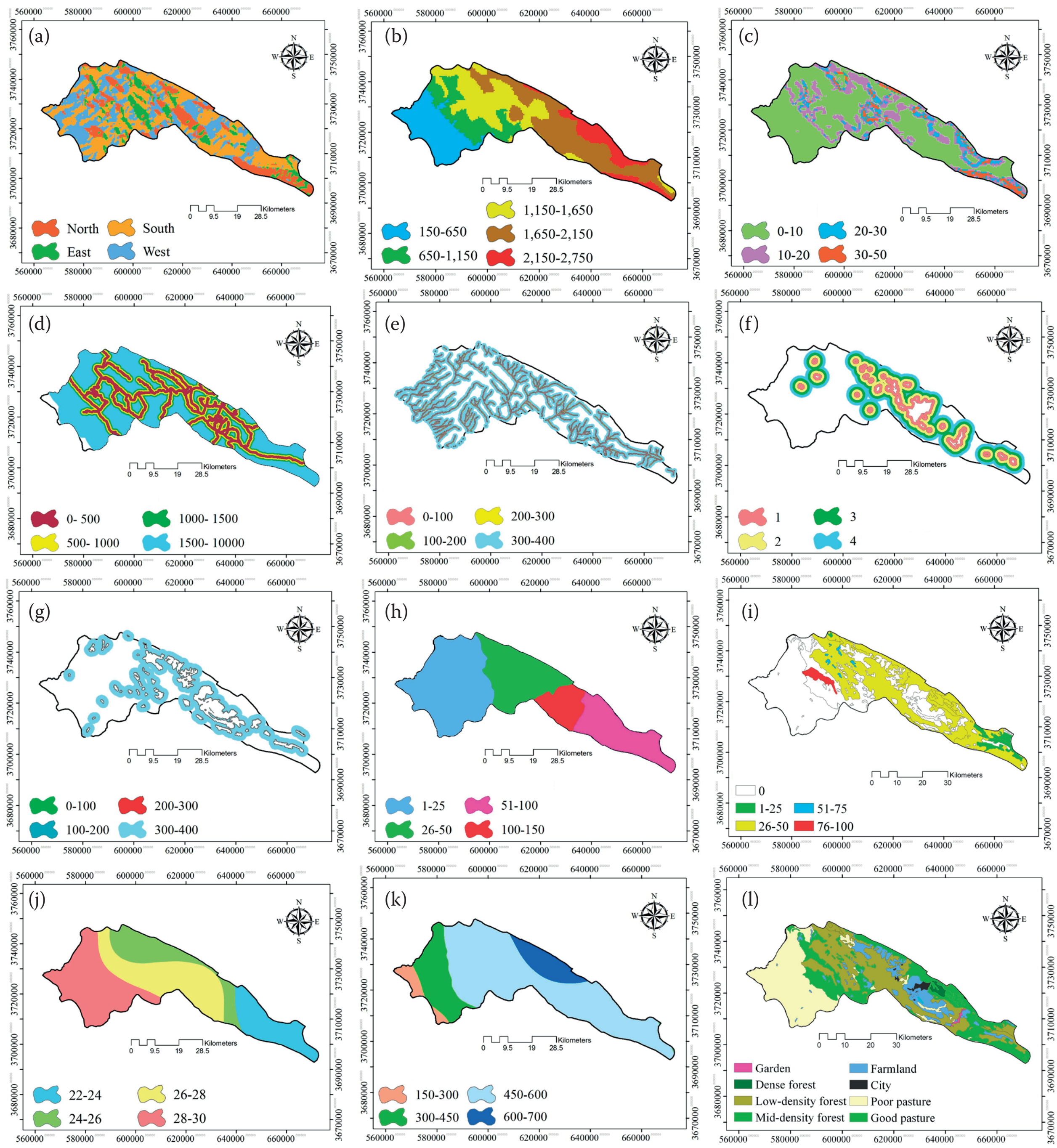

Fig. 4. Maps of input parameters: aspect (a), elevation (m) (b), slope (%) (c), distance from road (m) (d), distance from river (m) (e), distance from settlement (m) (f), distance from farmland (m) (g), population density per hectare (h), oak decline (%) (i), temperature (°C) (j), precipitation (mm) (k), land use (l)

Table 3. The most important association rules

Association rules	Minimum support	Minimum confidence
[DS = 1]: 56 ==> [PD = 1, DR = 1, S = 1, OD = 1, T = 1, LU = 1, DF = 1]	**0.8**	**1**
[LU = 1, PD = 1, E = 1]: 9 ==> [DRI = 1, DS = 1, DF = 1, S = 1]	0.1	0.89
[DR = 1, DS = 1, LU = 1, OD = 1]: 17 ==> [DR = 1, DF = 1	0.3	0.88
[DR = 1, T = 1, S = 1]: 8 ==> [DRI = 1, PD = 1]	0.6	0.88
[OD = 1, DR = 1]: 40 ==> [DS = 1]	0.8	0.81
[PD = 1, DF = 1, T = 1]: 22 ==> [DS = 1]	0.72	1
[LU = 1, LU = 1, DR = 1]: 13 ==> [DS = 1, DF = 1]	0.65	1
[OD = 1, DR = 1, T = 1, E = 1, FR = 1]: 10 ==> [DS = 1]	0.3	1
[DF = 1, DR = 1, E = 1, P = 1]: 7 ==> [OD = 1, DS = 1, T = 1]	0.42	1
[E = 1, DR = 1, DS = 1, LU = 1, S = 1]: 7 ==> [T = 1, DF = 1, DRI = 1]	0.1	1
[PD = 1]: 58 ==> [DS = 1, DF = 1, LU = 1, DR = 1, E = 1, P = 1, OD = 1]	0.1	0.90
[DRI = 1]: 58 ==> [DF = 1, T = 1, DR = 1, E = 1, S = 1, PD = 1]	0.5	0.86
[DS = 1]: 56 ==> [PD = 1, DF = 1, T = 1, E = 1, LU = 1, OD = 1, A = 1]	0.1	1

the bold values represent the number and percentage of association rules, DS – distance from the settlement, PD – population density, DR – distance from the road, S – slope, OD – standing dead oak wood, T – temperature, LU – land use/land cover, DF – distance from farm land, E – elevation, DRI – distance from river, FR – forest road, P – precipitation, A – aspect

oak wood, temperature, land use/land cover and distance from farm land).

Table 4 contains statistical information of the minimum, maximum, and midpoints of the input crisp data (factors significantly influencing wildfire occurrence). In this case study, we defined three clusters, low, middle and high, for each field and distributed all the items or classes into them. Using the FCM algorithm, the data describing the factors were distributed in these three clusters with different degrees of membership. Table 5 shows each field with its corresponding clusters and membership function. The final map of each factor classified by FCM is included in Fig. 5. Spatial distribution of this map shows higher values in the northeast and north of the study area (mostly covered by dense forests) where most fires occur. This is related to a higher density of settlement, roads, agricultural lands and other human factors, and higher presence of fuels. Finally, the forest fire risk map (using a weighted overlay of eight affective factor maps) was developed for five classes (Fig. 6).

Table 4. The resulting clusters centre for significant factors

Minimum	0
Maximum	58
Centre 1 (high)	47.8889
Centre 2 (medium)	24.0745
Centre 3 (low)	3.7293

Table 5. Distribution of data between fuzzy clusters using fuzzy c-means

Significant factor	Cluster		
	high	medium	low
Distance from settlement	0.9186	0.0593	0.0221
Population density	0.8900	0.0791	0.0309
Distance from road	0.8837	0.0954	0.0209
Slope	0.6885	0.2474	0.0640
Standing dead oak wood	0.7534	0.1874	0.0592
Temperature	0.6680	0.2892	0.0428
Land use	0.6299	0.3176	0.0524
Distance from farm land	0.6180	0.3156	0.0664

DISCUSSION

In this study, a forest fire susceptibility map was produced by combining the capabilities of the Apriori algorithm, FCM and GIS. We will now discuss the application of the association rule algorithm to discover strong relationships among forest wildfire occurrence and other geographical factors. Apriori algorithm analysis resulting in a minimum support of 0.8% and a minimum confidence of 100% showed strong relationships among wildfire occurrence, distance from settlement, population density, distance from road, slope, standing dead wood, temperature, land cover and distance from farmland. Other variables were tested in the Apriori algorithm, but they did not have strong association rules with wildfire occurrence and so were not considered in the forest fire risk map.

The results show that those areas near settlement centers and roads and with high population densities are at highest risk of forest fire occurrence. Forest wildfires mostly occur in highly-populated areas with population densities of 50 or more persons per hectare, which is why the highest forest wildfire risk was around the main city of the region,

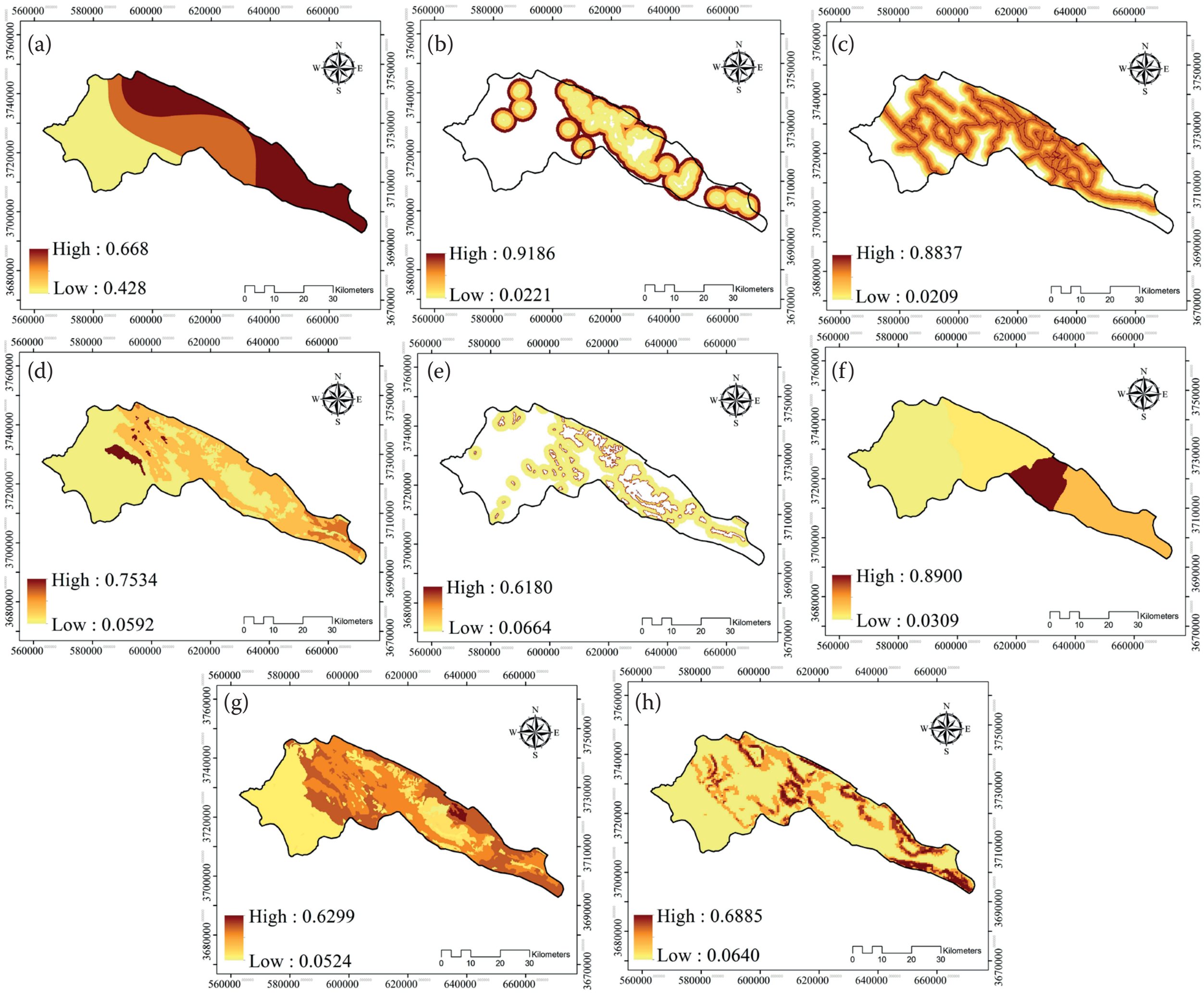

Fig. 5. Factors classified by fuzzy c-means: temperature (a), settlement (b), road (c), oak decline (d), farm land (e), population density (f), land use (g) and legend slope (h)

Ilam. Distance to settlements and distance to roads with high population densities have more importance than other variables, and are always important to explain regional patterns. These observations have also been reported by other researchers in their case studies (Erten et al. 2005; Eskandari et al. 2013; Zarekar et al. 2013). Thus, the impact of population density, local trails and roads (and activities such as grazing, hunting and agricultural activities in the understory of these forests) might be more relevant for forest fire occurrence than the impact of less populated centres and the national road network. The people of Ilam city have been forced to be highly dependent on these forests for recreation, so the forests have been reduced quantitatively and qualitatively (Arekhi, Jafarzadeh 2014). The results showed that the role of natural factors in forest wildfires is less significant than that of human activities. This is related to economic and social issues including the population's dependence on the forest and the lack of awareness regarding principles of forest protection and the use of forests. Man-made wildfires can result from campfires being left unattended, the burning of debris, negligently discarded cigarettes and unintentional acts of arson (because of carelessness).

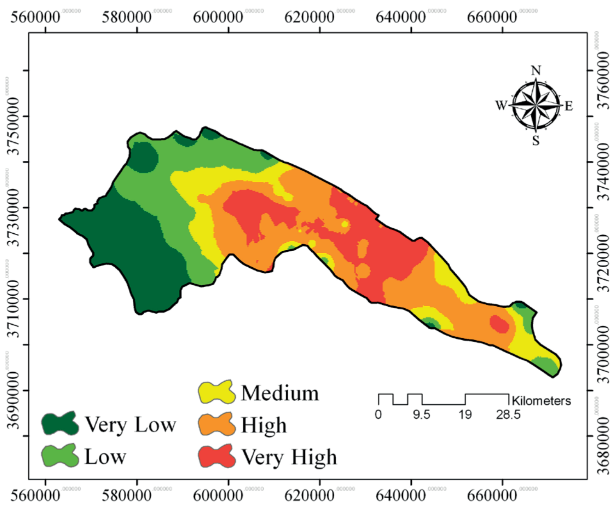

Fig. 6. Wildfire risk map

Distance from farmland was one of the most important variables to explain fire ignition patterns in Iran. In this study, fire occurrence was found to increase with proximity to farmland. This role of agricultural burnings has also been identified by several other researchers (Koutsias et al. 2010; Eskandari et al. 2013; Rodrigues et al. 2014; Eskandari, Chuvieco 2015). The starting of fires in the forestry/agriculture interface may be caused by negligence (sometimes farmers are not cautious enough when burning agricultural debris in the western forests of Iran (Forest, Range and Watershed Management Organization of Iran 2015)) or may be intentional for agricultural purposes (to clean and prepare the farmlands), which includes border regions (especially in the northern regions of the case study due to the higher value of the agricultural land). Mahdavi et al. (2012) stated that orchards and farms are supervised by their owners to prevent the initiation of fire on their properties. This increases the probability of extending unwanted fires, used for agricultural activities, towards forests and rangelands in their neighborhood. The majority of wildfires happen far from riversides, because orchards and farms owned by local communities are located there.

The only topographic factor affecting the occurrence of forest fires in the region was slope. In general, our study area is not characterised by steep slopes. The most susceptible areas to fires are the 30–50% classes of slope, and the research carried out by Rajeev Kumar et al. (2002), Xu et al. (2005) and Keane et al. (2009) conform with these results.

In the recent decades, the Mediterranean and semi-Mediterranean climatic regions have faced global warming (Nogués Bravo et al. 2008; Parry et al. 2008), and in this study, temperature was one of the most important factors explaining wildfire ignition patterns, as also found by other researchers (Stolle et al. 2003; Eskandari, Chuvieco 2015). The increasing temperatures and drought due to climatic changes in the dry seasons have been recognised as one of the most important causes of wildfires in some enclosure pastures and forests of western Iran (Mahdavi et al. 2012; Garavand et al. 2013). In addition, areas with higher oak decline have shown more potential for the starting of fires, as fuel is more abundant. The decline and mortality of oak forests has been noted across the forests of western Iran forests since 2000 and standing dead oak wood is one component of the wider issue of forest sustainability (Mirabolfathy 2013; Ahmadi et al. 2014). These dead trees as have been abandoned as flammable fuels in these areas.

Validation of the forest fire risk map was based on known fire occurrence. Cross tabulation between estimated and actual fire occurrence found a global accuracy of our forest fire risk map of 85%, with significantly higher probability values for cells that had active fires. About 40% of wildfires occurred in very high- and high-risk classes and 45% were located in risky classes (very high-, high- and medium-risk classes). So, it seems the map of forest fire risk produced in this research predicts more than 85% of forest wildfires occurring in the study area, and these data would be helpful in designing better annual wildfire management plans at national and regional levels. Regional fire forest management plans can benefit from a spatial knowledge of fire risks and conditions, by promoting the establishment of firebreaks or look-out towers in the most susceptible forest areas (very high-, high- and medium-classes), as well as introducing restrictions on access to forest areas prone to forest fires, particularly in natural forest reserves. Eventually, we suggest that forest fire risk management can be incorporated into forestry plans and in future experimental studies we will focus more attention on economic and social issues and use other data mining algorithms to analyse forest fires.

CONCLUSIONS

Mapping forest fire risk in different regions requires an analysis of the involved factors and the generation of a consistent model in each area. In this study, the Apriori algorithm (association rules) was applied for analysis of the forests in western Iran. The purpose of this study was to discover relationships between wildfire occurrence using association rules and the characteristics of factors determining wildfire hotspots as well as to map forest fire risk using FCM. The Apriori algorithm which is available in statistical computing tools was executed on the dataset and it generated 6,581 association rules. There were 14 association rules containing forest wildfire occurrence generated from the dataset with a minimum support of 10% and a minimum confidence of 80%. Our results show strong relationships between forest wildfire occurrence and eight variables (distance from settlement, population density, distance from road, slope, standing dead oak wood, temperature, land cover and distance from farm land). Finally, forest fire risk in the study area was obtained from overlaying these layer maps, and, depending on level of danger, risk was classified into five categories, ranging from very low to very high. This study has described the development of a forest fire risk assessment system for a given study area using different spatial databases. The produced maps can answer important questions concerning the causative factors of forest fires in the spatial domain and will be useful to forest managers when undertaking necessary protective measures at the regional level.

References

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

Aggarwal S., Rani B. (2013): Optimization of association rule mining process using Apriori and Ant Colony Optimization algorithm. International Journal of Current Engineering and Technology, 3: 620–623.

Agrawal R., Srikant R. (1994): Fast algorithms for mining association rules. In: Bocca J.B., Jarke M., Zaniolo C. (eds): Proceedings of 20th International Conference on Very Large Databases, Santiago de Chile, Sept 12–15, 1994: 487–499.

Ahmadi R., Kiadaliri H., Mataji A., Kafaki S. (2014): Oak forest decline zonation using AHP model and GIS technique in Zagros Forests of Ilam Province. Journal of Biodiversity and Environmental Sciences (JBES), 4: 141–150.

Akbari D., Amini J., Saadatseresht M. (2008): Developing a simple and fast model to produce wildfire map for forest. In: 2nd Symposium on Facing Natural Hazard, Tehran, Dec 25–26, 2008: 1–7.

Arekhi S., Jafarzadeh A.A. (2014): Forecasting areas vulnerable to forest conversion using artificial neural network and GIS (case study: northern Ilam forests, Ilam province, Iran). Arabian Journal of Geosciences, 7: 1073–1085.

Bezdek J.C. (1981): Pattern Recognition with Fuzzy Objective Function Algorithms. Norwell, Kluwer Academic Publishers: 272.

Bowman D., Balch J.K., Artaxo P., Bond W.J., Carlson J.M., Cochrane M.A., D'Antonio C.M., DeFries R.S., Doyle J., Harrison S.P., Johnston F.H., Keeley J.E., Krawchuk M.A., Kull C.A., Marston J.B., Moritz M.A., Prentice I.C., Roos C.I., Scott A.C., Swetnam T.W., van der Werf G.R., Pyne S.J. (2009): Fire in the Earth system. Science, 324: 481–484.

Cakir O., Aras M.E. (2012): A recommendation engine by using association rules. Procedia – Social and Behavioral Sciences, 62: 452–456.

Deng H. (1999): Multicriteria analysis with fuzzy pairwise comparisons. International Journal of Approximate Reasoning, 21: 215–231.

Erensal Y.C., Öncan T., Demircan M.L. (2006): Determining key capabilities in technology management using fuzzy analytic hierarchy process: A case study of Turkey. Information Sciences, 176: 2755–2770.

Erten E., Kurgun V., Musaoğlu N. (2005): Forest Fire Risk Zone Mapping from Satellite Imagery and GIS: A Case Study. Fire Risk Report. Istanbul, Civil Engineering Faculty, Remote Sensing Division: 7.

Eskandari S., Chuvieco E. (2015): Fire danger assessment in Iran based on geospatial information. International Journal of Applied Earth Observation and Geoinformation, 42: 57–64.

Eskandari S., Ghadikolaei J.O., Jalilvand H., Saradjian M.R. (2013): Fire risk modeling and prediction in district three of Neka-Zalemroud forest using geographical information system. Iranian Journal of Forest and Poplar Research, 21: 203–217.

Ester M., Kriegel H.P., Sander J. (1997): Spatial data mining: A database approach. In: Scholl M., Voisard A. (eds): Advances in Spatial Databases. Proceedings of the 5th International Symposium on Spatial Databases, Berlin, July 15–18, 1997: 47–66.

Fattahi M. (2003): Zagros Oak Forests and the Most Important Degradation Factors. Tehran, Research Institute of Forests and Rangelands: 324.

Feng C. (1995): Fuzzy multicriteria decision-making in distribution of factories: An application of approximate reasoning. Fuzzy Sets and Systems, 71: 197–205.

Forest, Range and Watershed Management Organization of Iran (2015): Annual Report on Forest Fires. Tehran, Forest & Rangelands Publications: 30.

Garavand S., Yaralli N., Sadeghi H. (2013): Spatial pattern and mapping fire risk occurrence at natural lands of Lorestan province. Iranian Journal of Forest and Poplar Research, 21: 231–242.

Gottwald S. (2006): Universes of fuzzy sets and axiomatizations of fuzzy set theory. Part I: Model-based and axiomatic approaches. Studia Logica, 82: 211–244.

Hu H., Jin S. (2002): Study on forest fire regime of Heilongjiang province II. Analysis on factors affecting fire dynamics and distributions. Scientia Silvae Sinicae, 38: 98–102.

Keane R.E., Drury S.A., Karau E.C., Hessburg P.F., Reynolds K.M. (2009): A method for mapping fire hazard and risk across multiple scales and its application in fire management. Ecological Modelling, 221: 2–18.

Koutsias N., Martínez-Fernández J., Allgöwer B. (2010): Do factors causing wildfires vary in space? Evidence from geographically weighted regression. GIScience & Remote Sensing, 47: 221–240.

Liu B. (2007): Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data. Berlin, Heidelberg, New York, Springer-Verlag: 532.

Loboda T.V. (2009): Modeling fire danger in data-poor regions: A case study from the Russian Far East. International Journal of Wildland Fire, 18: 19–35.

Mahdavi A., Fallah Shamsi S., Nazari R. (2012): Forests and rangelands' wildfire risk zoning using GIS and AHP techniques. Caspian Journal of Environmental Sciences, 10: 43–52.

Merril D.F., Alexander M.E. (eds) (1987): Glossary of Forest Fire Management Terms. 4th Ed. Ottawa, Canadian Committee on Forest Fire Management, National Research Council of Canada: 91.

Mikhailov L., Tsvetinov P. (2004): Evaluation of services using a fuzzy analytic hierarchy process. Applied Soft Computing, 5: 23–33.

Mirabolfathy M. (2013): Outbreak of charcoal disease on Quercus spp. and Zelkova carpinifolia trees in forests of Zagros and Alborz Mountains in Iran. Iranian Journal of Plant Pathology, 49: 257–263.

Mohammadi F., Shabanian N., Pourhashemi M., Fatehi P. (2011): Risk zone mapping of forest fire using GIS and AHP in a part of Paveh forests. Iranian Journal of Forest and Poplar Research, 18: 569–586. (in Persian)

Nogués Bravo D., Araújo M.B., Lasanta T., López Moreno J.I. (2008): Climate change in Mediterranean mountains during the 21st century. Ambio: A Journal of the Human Environment, 37: 280–285.

Padilla M., Vega-Garcia C. (2011): On the comparative importance of fire danger rating indices and their integration with spatial and temporal variables for predicting daily human-caused fire occurrences in Spain. International Journal of Wildland Fire, 20: 46–58.

Parry M., Palutikof J., Hanson C., Lowe J. (2008): Squaring up to reality. Nature Reports Climate Change, 2: 68–70.

Patel B., Chaudhari V.K., Karan R.K., Rana Y.K. (2011): Optimization of association rule mining Apriori algorithm using ACO. International Journal of Soft Computing and Engineering, 1: 24–26.

Paz S., Carmel Y., Jahshan F., Shoshany M. (2011): Post-fire analysis of pre-fire mapping of fire-risk: A recent case study from Mt. Carmel (Israel). Forest Ecology and Management, 262: 1184–1188.

Prasad V.K., Badarinath K.V.S., Eaturu A. (2008): Biophysical and anthropogenic controls of forest fires in the Deccan Plateau, India. Journal of Environmental Management, 86: 1–13.

Rajeev Kumar J., Saumitra M., Kumaran D.R., Rajesh S. (2002): Forest fire risk zone mapping from satellite imagery and GIS. International Journal of Applied Earth Observation and Geoinformation, 4: 1–10.

Rodrigues M., de la Riva J., Fotheringham S. (2014): Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. Applied Geography, 48: 52–63.

Romero-Ruiz M., Etter A., Sarmiento A., Tansey K. (2010): Spatial and temporal variability of fires in relation to ecosystems, land tenure and rainfall in savannas of northern South America. Global Change Biology, 16: 2013–2023.

Salamati H., Mostafalou H., Mastoori A., Honardoost F. (2011): Evaluation and provision of forest fire risk map using GIS in Golestan forests. In: Proceedings of the 1st International Conference on Wildfire in Natural Resources Lands, Gorgan, Oct 26–28, 2011: 56–65. (in Persian)

Shihab A.I., Burger P. (1998): The analysis of cardiac velocity MR images using fuzzy clustering. In: Proceedings of SPIE Medical Imaging. Physiology and Function from Multidimensional Images, San Diego, Feb 21, 1998: 176–183.

Srikant R., Agrawal R. (1996): Mining quantitative association rules in large relational tables. In: Widom J. (ed.): Proceedings of the ACM SIGMOD International Conference on Management of Data, Montreal, June 4–6, 1996: 1–12.

Stolle F., Chomitz K.M., Lambin E.F., Tomich T.P. (2003): Land use and vegetation fires in Jambi Province, Sumatra, Indonesia. Forest Ecology and Management, 179: 277–292.

Taylor S.W, Alexander M.E. (2006): Science, technology, and human factors in fire danger rating: The Canadian experience. International Journal of Wildland Fire, 15: 121–135.

Tian X., Shu L., Zhao F., Wang M. (2012): Analysis of the conditions for lightning fire occurrence in Daxing' anling region. Scientia Silvae Sinicae, 48: 98–103.

Wang Y.M., Luo Y., Hua Z. (2008): On the extent analysis method for fuzzy AHP and its applications. European Journal of Operational Research, 186: 735–747.

Xu D., Dai L.M., Shao G.F., Tang L., Wang H. (2005): Forest fire risk zone mapping from satellite images and GIS for Baihe Forestry Bureau, Jilin, China. Journal of Forestry Research, 16: 169–174.

Zarekar A., Kazemi Zamani B., Ghorbani S., Ashegh Moalla M., Jafari H. (2013): Mapping spatial distribution of forest fire using MCDM and GIS (case study: Three forest zones in Guilan province). Iranian Journal of Forest and Poplar Research, 21: 218–230.

Received for publication January 10, 2017
Accepted after corrections July 27, 2017