Natural Hazards (2022) 114:457–473
https://doi.org/10.1007/s11069-022-05397-6

ORIGINAL PAPER

Anthropogenic activities amplify wildfire occurrence in the Zagros eco-region of western Iran

Abolfazl Jaafari[1] · Omid Rahmati[2] · Eric K. Zenner[3] · Davood Mafi-Gholami[4]

Received: 30 September 2021 / Accepted: 9 May 2022 / Published online: 2 June 2022

Abstract

The aim of this study was to improve our understanding of factors that affect the spatial distribution of wildfire occurrences at the regional scale. We employed the random forest, boosted regression tree, and genetic algorithm rule-set production models to assess the spatial interplay between fire events and climate, topography, and anthropogenic factors in order to characterize wildfire occurrence in the Zagros eco-region of western Iran. We constructed a geospatial database using the historical fires from the period 2007–2020 and topography, climate, and human related factors. The results demonstrated that human activities (i.e., land use and distance from the settlements and roads) contributed 45% to the probability model of wildfire occurrence in the study region. The models ranked the climate factors (rainfall, temperature, and wind effect) as the second most influential drivers of fire occurrences, whereas topographic features (slope, elevation, and aspect) did not significantly influence fire probability in the landscape. Overall model performance was assessed with the area under the receiver operating characteristic (AUROC) method that showed the superior performance of the RF model in the training phase (AUROC = 0.92) and in its ability to predict upcoming fires (AUROC = 0.90). The insights obtained from this research can bring into focus both the locations and the types of suppression policies that are required to alleviate the effects of the upcoming wildfires in the early twenty-first century.

Keywords Human activity · Geographic information systems (GIS) · Machine learning · Probability mapping

✉ Abolfazl Jaafari
jaafari@rifr-ac.ir

1 Forest Research Division, Research Institute of Forests and Rangelands, Agricultural Research, Education, and Extension Organization (AREEO), Tehran 1496813111, Iran

2 Soil Conservation and Watershed Management Research Department, Kurdistan Agricultural and Natural Resources Research and Education Center, AREEO, Sanandaj, Iran

3 Department of Ecosystem Science and Management, The Pennsylvania State University, Forest Resources Building, University Park, PA 16802, USA

4 Department of Forest Sciences, Faculty of Natural Resources and Earth Sciences, Shahrekord University, Shahrekord, Iran

1 Introduction

The potential occurrences of wildland fires threaten terrestrial ecosystems in many parts of the globe (Pausas and Keeley 2009; Swain 2021). Fire damage can be reduced or even be avoided by exploring wildfire-ecosystem relationships. Estimating the probability that a fire may occur provides the basis for probability mapping, land classification, firefighting, resource allocation, and landscape management (Hyde et al. 2013; Zhang et al. 2019a).

Presently, wildfire research has been increasingly fueled by advances in geographic information systems (GIS) and remote sensing (Ghorbanzadeh et al. 2019), along with the improvement in data availability and computation power (Jaafari et al. 2019a; Elia et al. 2020). Despite these advances, there is a shortage of knowledge of the connectivity between wildfires and different geo-environmental factors in many rugged landscapes of large extents around the world (Pourtaghi et al. 2016; Spano et al. 2021; Tavakkoli Piralilou et al. 2022). Further, previous works have rarely focused on additional complexities (e.g., human–environment interactions) associated with the predictive modeling of wildfires (Hong et al. 2019; Jaafari et al. 2019a, b). Developing and collecting the input data, i.e., factors that directly or indirectly cause wildfire occurrence in different fire-prone landscapes, remains a difficult task that discourages and limits spatially explicit mapping of wildfire susceptibilities.

Causative factors of wildfire have been well researched and formalized and fall into the four main groups of climate, land cover, topography, and anthropogenic factors (Adab et al. 2018; Rodrigues et al. 2018; Hong et al. 2019; Spano et al. 2021; Tavakkoli Piralilou et al. 2022). Whereas climate-related factors (e.g., wind, drought, rainfall, lightning, and evapotranspiration) largely affect the frequency and intensity of wildfire occurrence (Dennison et al. 2014; Rasooli et al. 2021), topographic features (e.g., altitude, slope, and aspect) affect fire ignitions mostly indirectly (Hong et al. 2019; Naderpour et al. 2021) through changes to the vegetation, local climate, and accessibility to people (Adab et al. 2013, 2018). Land cover (i.e., vegetation) strongly affects fire occurrence and fire spread via fuel type, fuel load, and moisture content (Chuvieco et al. 2010; Adab et al. 2018; Sari 2021). Finally, humans can alter natural local conditions (Li et al. 2021; Lan et al. 2021; Liu et al. 2022) in ways that may either intensify or suppress wildfire occurrences (Gralewicz et al. 2012; Viedma et al. 2018; Naderpour et al. 2021).

Being able to quantify the relative importance of different causative factors and their associations with historical fire events is central to an improved understanding of underlying patterns of wildfire susceptibilities that facilitates the development and adaptation of more efficient fire management strategies. The current approach to predictive modeling of wildfire, however, is an extremely subjective approach that requires a significant time investment, typically refers to expert knowledge (Goleiji et al. 2017; Sari 2021), applies multiple methods (Pourtaghi et al. 2016; Jaafari and Pourghasemi 2019; Jaafari et al. 2019a, b), and often uses trial-and error processes. These disadvantages have provided the impetus for new studies to develop new approaches for more efficient field inspections and the incorporation of advances made possible by big data to enhance our understanding of the connectivity between landscape characteristics and wildfire probabilities.

Several factor selection techniques have been suggested in the wildfire literature for the identification of explanatory factors that are closely associated with wildfire occurrences. The inclusion of too many conditioning factors, however, may introduce inherent noise that decreases the prediction accuracy. Equally, the use of a narrow range of explanatory factors will likely fail to provide sufficient information to build predictive models and make accurate predictions. Several researchers suggested the use of factor selection methods as a crucial pre-processing phase for every wildfire prediction modeling project (Pourtaghi et al. 2016; Jaafari et al. 2018), because factor selection practices safeguard the predictive system from the inclusion of those factors that are not relevant and might even decrease the accuracy of model outputs (Jaafari et al. 2018). Factor selection techniques are generally separated into three groups: (1) factor ranking methods such as random forest (RF) and boosted regression tree (BRT) that are used for quantifying the importance of each factor (Pourtaghi et al. 2016; Jaafari and Pourghasemi 2019), (2) correlation analysis methods such as evidential belief function, weight of evidence, frequency ratio, and step-wise assessment ratio analysis (SWARA) (Jaafari et al. 2017, 2019a, b), and (3) factor subset selection methods such as Gain Ratio, and Relief-F (Jaafari et al. 2018).

Here, we present a comparative analysis of three predictive models derived from machine learning techniques, i.e., RF, BRT, and genetic algorithm rule-set production (GARP), to rank different geo-environmental factors known to influence fire occurrence and to predict the probability of fire occurrence in a mountainous eco-region in western Iran where wildfires are recurrent. The main contributions and novelties of the current study are as follows:

- *General* by applying the state-of-the-art machine learning methods to a typical fire-prone landscape, the study enhances our understanding of wildfires and the likely factors that determine its occurrence.
- *Methodical* this is the first application of the GARP model in wildfire modeling and factor ranking, enabling direct comparisons with results from RF and BFT models.
- *Regional* improved knowledge of underlying patterns of wildfire occurrences in the Zagros eco-region of Iran. Our findings are useful for adopting fire management strategies adjusted to the local conditions, and bring into focus both the locations and the types of suppression policies that are necessary to mitigate the impacts of the future wildfires.
- *Global* Our study extends the scope of fire behavior inferences beyond America, Europe, and Oceania/Australia to an Asian country, providing the basis for government authorities and scientists to enhance preparedness of human communities toward fire safety techniques.

2 Study area

Our study area is located in the central highlands of the Zagros Mountains of Iran and extends between latitudes 31° 9′ N to 32° 48′ N and longitudes 49° 28′ E to 51° 25′ E between 783 and 4178 m a.s.l. elevation and covers an area of approximately 16,532 km^2 (Fig. 1). This area is a typical fire-prone portion of the Zagros eco-region where many wildfires occur. The area enjoys a Mediterranean climate with warm and dry summers, below-freezing winter temperatures, and mean annual temperature between 5 and 16 °C. Mean annual precipitation ranges between 250 and 1400 mm that mostly falls during the autumn and winter months between November and March. Land use/cover patterns are strongly shaped by elevation and within each vegetation zone, aspect-induced micro-climate permits different vegetation communities and land uses at small scales. Generally, the area exhibits a range of vegetation conditions that includes grasslands and scattered forests dominated by *Quercus brantii* with a mixture of several broadleaved

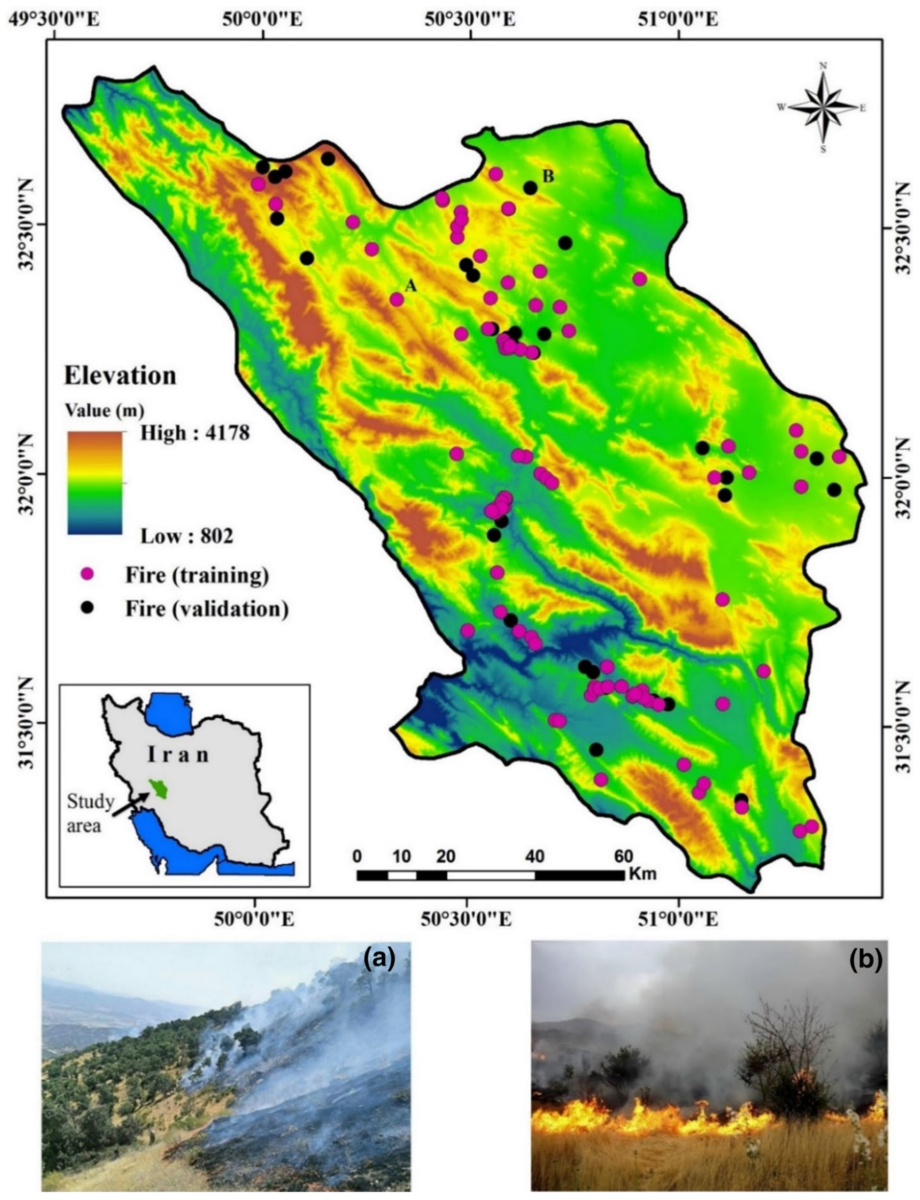


Fig. 1 Location of the study area with locations of observed wildfire sites that occurred between 2007 and 2016 (used for model training) and between 2017 and 2020 (used for model validation) (**A** and **B** are field photographs of recent wildfires in the study area)

species (e.g., *Amygdalus scoparia*, *Pistacia vera*, *Acer monspessulanum*, *Pyrus glabra*, *Crataegus microphylla*, *Fraxinus rotundifolia*, *Lonicera nummulariifolia*). Most of the past wildfires that have occurred in the Zagros eco-region were either due to a combination of low rainfall, drought occurrences or anthropogenic phenomena (Jaafari et al. 2017, 2019a). Although wildfires are a recurrent feature of this eco-region, few studies have been conducted to predict the probability of wildfire occurrences and to quantify the main drivers of wildfire.

3 Data used

While fire data recorded over eight years between 2007 and 2014 have been used previously by Jaafari et al. (2017) and (2019b), these data were supplemented with recent field survey data (2015–2020) for this study. These data include the spatial location of wildfires and a set of different geo-environmental and anthropogenic factors that serve as wildfire predictors and are described in more detail below.

3.1 Historical fires

Identifying and mapping the historical locations of wildfires is crucial for exploring the association between fire probability and the geo-environmental factors. Current approaches typically rely on remotely sensed satellite image data. The Moderate Resolution Imaging Spectrometer (MODIS) active fire products have been widely used for detecting fire events (Giglio et al. 2003; Ghorbanzadeh et al. 2019). MODIS has a 1 km pixel resolution and detects active fires through uses thermal data anomalies at the time the satellite passes over the ground. The MODIS data are often validated in terms of the records of fire occurrence and area burned using the finer resolution sensors such as Landsat and ASTER. For this study, we compiled a fire inventory map based on the chronological documents of the administrative office of natural resources of the Chaharmahal and Bakhtiari Province related to wildfires and multiple field surveys. For further verification of the time and locations of fire events, we used data from the MODIS active fire products obtained from the NASA's Terra MODIS and Aqua MODIS satellites (https://modis.gsfc.nasa.gov) and Landsat images (https://landsat.visibleearth.nasa.gov) (Zhang et al. 2019b; Zhao et al. 2021). Since many of the fires were small, we omitted any fires < 0.3 ha in extent. The value of < 0.3 ha was adopted after considering minimum (0.07 ha), maximum (12.61 ha), and average (5.39 ha) areas of historical fire events across the study area. We thus used 164 confirmed fire events that occurred during the period between 2007 and 2020 as the basis for the inventory map (Fig. 1). Records for each of the confirmed fire events include the geographical coordinates, date of the fire, and the size of the burned area.

3.2 Predictor factors

To select the predictor factors (i.e., wildfire influencing factors), an initial study of local wildfire properties and their spatial distribution was conducted. Then, relying on other peer-reviewed wildfire modeling research (Chuvieco et al. 2010; Carmo et al. 2011; Oliveira et al. 2012; Satir et al. 2016; Adab et al. 2018; Viedma et al. 2018; Zhang et al. 2019a; Ghorbanzadeh et al. 2019; Sari 2021) and data availability, the final set of predictor factors was selected: Slope degree, aspect, elevation (m), land use, mean annual rainfall (mm) and temperature (°C), normalized difference vegetation index (NDVI), wind effect, and distance (m) from human settlements, streams, and roads. These factors were generated using the 10-year data from 2007 to 2020 (Fig. 2). Topographic features (i.e., slope, aspect, elevation) were derived from a 30-m resolution digital elevation model (DEM) collected from ASTER Global DEM Explorer tool (http://earthexplorer.usgs.gov) (Zhang

Fig. 2 Predictive factors used in this study: **a** elevation, **b** aspect, **c** slope, **d** rainfall, **e** temperature, **f** wind ▸ effect, **g** land use (1: Forest, 2: Good range, 3: Poor range, 4: Farmland, 5: Wetland, and 6: Irrigated farming), **h** NDVI, **i** distance from settlements, **j** distance from rivers, and **k** distance from roads

et al. 2019c). Using the Landsat satellite images, we provided the land use and mean annual NDVI maps for the study area (Chen et al. 2021; Quan et al. 2022). The climate factors (i.e., temperature, rainfall, wind speed, and wind direction) were computed using the 20-year data collected from the national meteorological organization of Iran (Chen et al. 2022; Yin et al. 2022). Wind effect that is a non-dimensional factor combining wind speed and wind direction was produced within the SAGA GIS software. The distance maps were generated by buffering streams, roads, and settlements areas within the study area using the Euclidean distance tool available in the ArcGIS software that computes the distance of each fire cell to the closest cell depicting road, river or settlement. The information related to all factors was integrated and manipulated in a GIS environment and then transformed to raster format at 30 m resolution. A detailed description of the significance of each predictor factor on wildfire occurrences is available in the corresponding literature (e.g., Adab et al. 2013, 2018; Pourtaghi et al. 2016; Zhang et al. 2019a; Hong et al. 2019; Elia et al. 2020; Liu et al. 2020, and references therein).

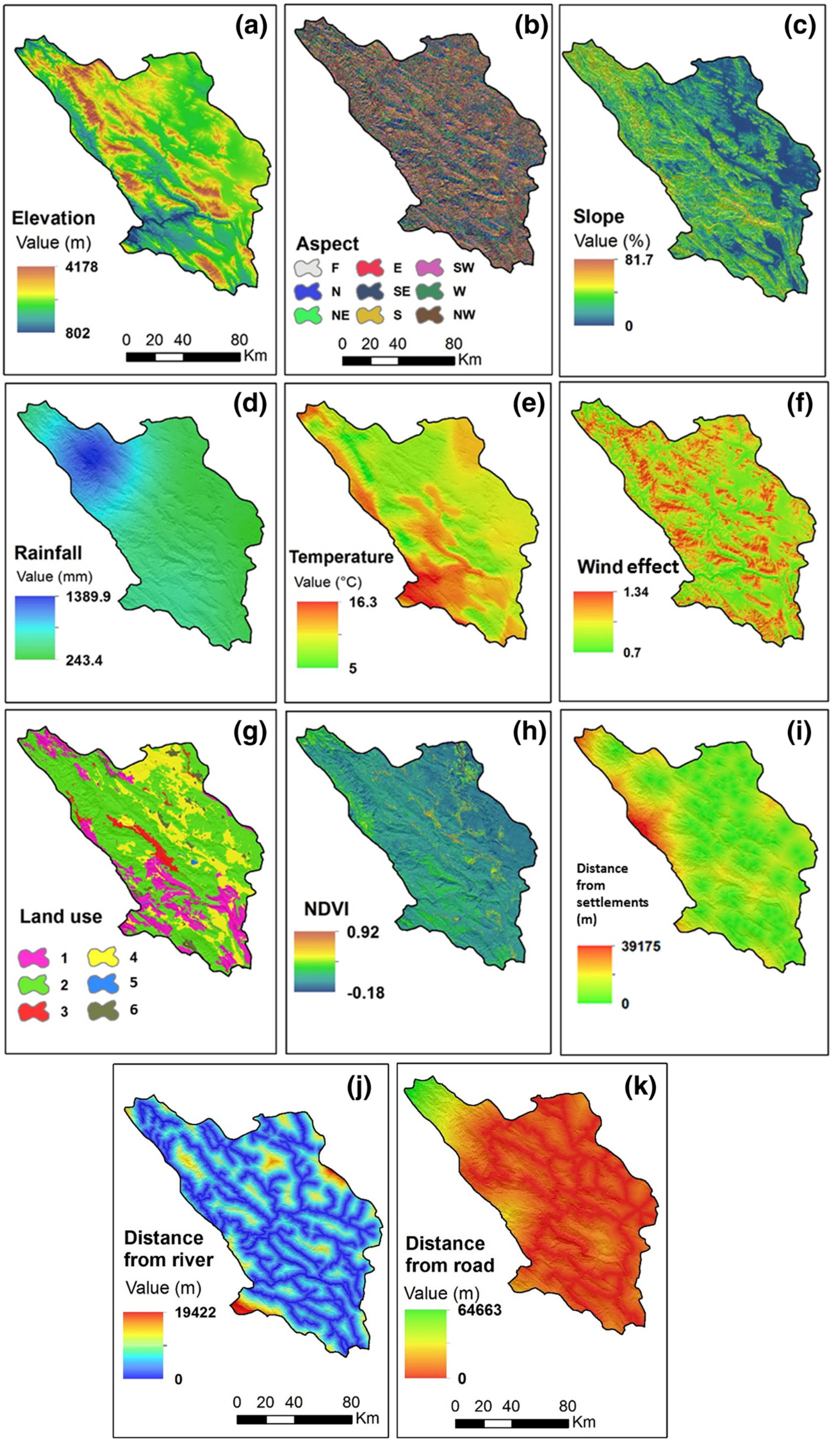
4 Probability modeling

Modeling of wildfire probability deals with estimating the likelihood of wildfire occurrence and results in a probability value between 0 and 1 (i.e., probability values), with greater probability values indicating a greater probability for a fire to occur. Probability modeling is typically based on exploring the spatial association between wildfire events that have occurred in a given landscape and different geo-environmental and anthropogenic factors using a data-driven method (Pourtaghi et al. 2016). The general work flow of the modeling process proposed in this study for modeling of wildfire probability and identifying the most influential factors is shown in Fig. 3 and described in the following subsections.

4.1 Background of the methods used

Based on an initial analysis of the many methods used in the literature, we selected three machine learning methods that can efficiently elucidate relationships among predictor factors and would be most appropriate for future applications. Here, we briefly describe these methods and refer to the corresponding literature (Breiman 2001; Stockwell 1999) for a more detailed description of each method.

RF uses Breiman's "bagging" idea to integrate a collection of separately trained binary decision trees with controlled variance to perform a classification task. RF can handle both continuous and categorical datasets and is not very sensitive to over-fitting, yielding promising results in many fields of science (Shabani et al. 2021). In a wildfire modeling using the RF model, the training dataset D is split into m subsets of the samples D_1, D_2, …, D_k using the bootstrap resampling method. Then, the m decision trees are generated corresponding to the m subsets and a random vector k of the predictor factors. Finally, the probability of wildfire occurrence is calculated by estimating the proportion of the decision trees that predict the wildfire occurrence among all the decision trees within the RF algorithm. In this study, we used the random-forest package (Breiman 2001) for the


(a)
Elevation
Value (m)
4178
802
0 20 40 80
Km
(b)
Aspect
F
E
SW
N
SE
W
NE
S
NW
0 20 40 80
Km
(c)
Slope
Value (%)
81.7
0
(d)
Rainfall
Value (mm)
1389.9
243.4
(e)
Temperature
Value (°C)
16.3
5
(f)
Wind effect
1.34
0.7
(g)
Land use
1
2
3
4
5
6
(h)
NDVI
0.92
-0.18
(i)
Distance from settlements (m)
39175
0
(j)
Distance from river
Value (m)
19422
0
0 20 40 80
Km
(k)
Distance from road
Value (m)
64663
0
0 20 40 80
Km

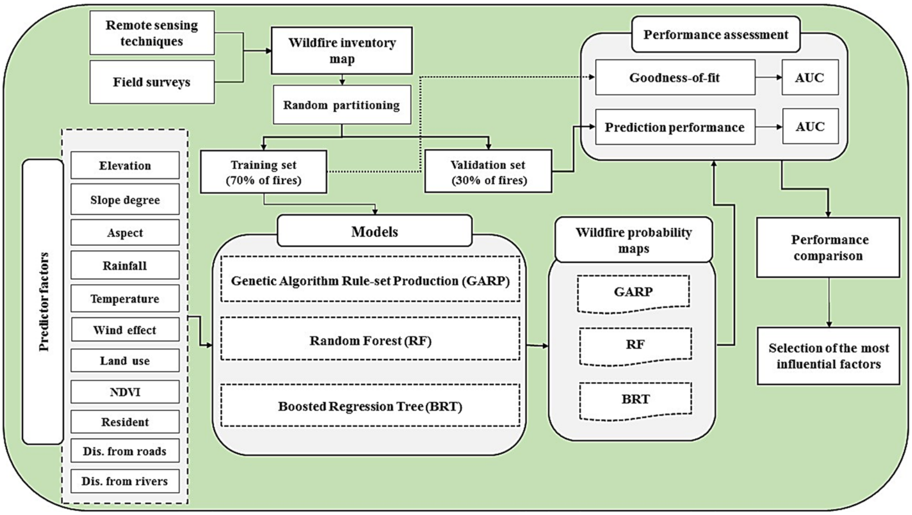


Fig. 3 Flowchart detailing the methodology employed in this study

implementation of RF model; the computational process was carried out in the R statistical software. For the minimum out of-bag (OOB) error of the RF model, we tuned the number of decision trees and factors tried at each split on 1000 and 3.

BRT is a combination of regression and boosting techniques intended to reduce the weakness of the single regression tree methods. No prior data transformation or omission of outliers is required for BRT, which can easily handle different types of environmental factors and spatial data. Further, this method can fit complex nonlinear relationships such as natural hazards, and efficiently explore the interaction effects among predictors. Here, we performed the BRT model using the gbm package within the R 3.0.2 statistical software (Ridgeway 2007).

GARP is a machine learning method based on a genetic algorithm that classifies data using an iterative procedure of rule (if–then) selection, assessment, testing, and integration or rejection (Feria and Peterson 2002). Originally developed for species distribution modeling (Stockwell 1999), this method was adopted in recent years for natural hazards modeling (Darabi et al. 2019; Rahmati et al. 2019). GARP predicts the probability of incidence of an event using the historic presence-only occurrence records in relation to continuous environmental factors, providing an estimate of future probability of occurrence. For this study, we used the GARPTools package within the R 3.0.2 statistical software for the implementation of the GARP model. Table 1 lists the optimized parameters that we used to perform the GARP model.

4.2 Model training and validation

To apply the RF, BRT, and GARP models for wildfire modeling and mapping, we split the historical fires detected within the study landscape into two groups such that the 115 fires (70% of all fires) that occurred in the period between 2007 and 2016 were used for training the models and 49 fires (30% of all fires) that occurred between 2017 and 2020 were used for validation. Along with these datasets, 164 non-fire locations were randomly sampled

Table 1 Optimization of the GARP parameters

Parameter	Value
Runs	50
Convergence limit	0.01
Maximum iterations	500
Best subset selection parameters	
Omission measure type	Extrinsic-soft
Omission threshold	25
Commission threshold	55

from the unburned/unburnable portions of the study area and were used for producing the final training and validation datasets (Xie et al. 2021a). This process produced a total of 230 records for the training and 98 records for the validation datasets.

The receiver operating characteristic (ROC) curve was used to evaluate the goodness-of-fit (i.e., training performance) and predictive capability (i.e., validation performance) of the models. The x-axis of ROC curve signifies the false positive rate (1-specificity) and the y-axis specifies the false negative rate (sensitivity). An area under curve (AUC) = 1 represents 100% sensitivity and 100% specificity and is the best possible outcome. AUC values of > 0.9 indicate an outstanding performance of the predictive model (Xie et al. 2021b).

4.3 Probability mapping

After successful training and validation, the resulting predictive models were applied to the entire research landscape to generate the landscape-level probability values. To generate the final wildfire probability maps, we classified the probability values derived from each model into different probability levels by means of the geometrical interval classification method. Five classes (i.e., very low, low, moderate, high, and very high) were adopted to the probability maps after Pourtaghi et al. (2016) and Hong et al. (2019). Finally, the probability maps were analyzed and the land area allocated to each probability class was summed.

5 Results

5.1 Factor importance

The three models tested in this study provided a rank of relative importance for each wildfire influencing factor (Table 2). All three models consistently identified distance to human settlements, NDVI, and distance to roads as the most important factors, corresponding to more than 50% of the total contribution to wildfire probability. Conversely, distance from streams and slope were considered least influential (less than 3% each) affecting the probability of wildfire occurrence on the landscape. In the case of other factors, the models did not agree on their relative importance. For example, while the BRT model ranked aspect as the fifth influential factors, the other two models gave a much lower influence/weight to this factor.

Table 2 Relative importance (%) of influencing factors extracted using GARP, BRT, and RF

Factor	GARP	BRT	RF
Distance from settlements	**26.5**	**25.4**	**27.1**
NDVI	**16.3**	**14.2**	**17.3**
Distance from roads	**11.2**	**12.3**	**12.5**
Rainfall	9.8	9.4	7.6
Elevation	8.4	3.7	5.5
Land use	7	5.7	7.4
Temperature	6.3	7.2	9.6
Wind effect	5.5	8.3	6.4
Aspect	3.9	8.5	2.7
Distance from streams	2.6	2.4	2.5
Slope	2.5	2.9	1.4

The most influential factors are shown in bold

5.2 Model performance

The best training performance of the three predictive models of wildfires was achieved by the RF model (ROC-AUC = 0.924; Fig. 4a), followed by GARP (ROC-AUC = 0.881) and BRT (ROC-AUC = 0.821). The results of the validation performance resulted in the same ranking of the models. RF was the most powerful model for the prediction of wildfires (AUC = 0.904), followed by GARP (AUC = 0.849) and BRT (AUC = 0.748) (Fig. 4b). The variance and covariance between the fire and non-fire grid cells estimated using the OOB error index (Fig. 5) showed an accuracy rate of 90.82%, indicating a near perfect performance of the RF model for classifying non-fire and fire grid cells over the landscape.

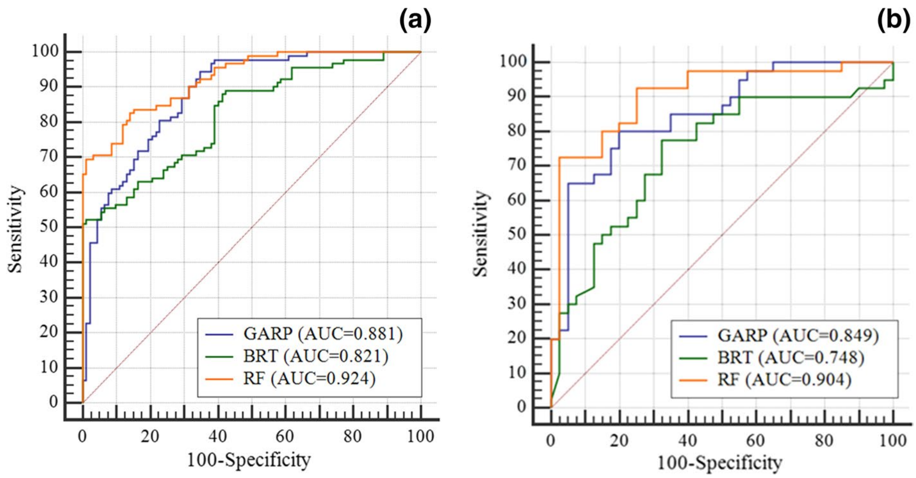


Fig. 4 ROC curves and AUC values of the models in the training phase (**a**) and validation phase (**b**)

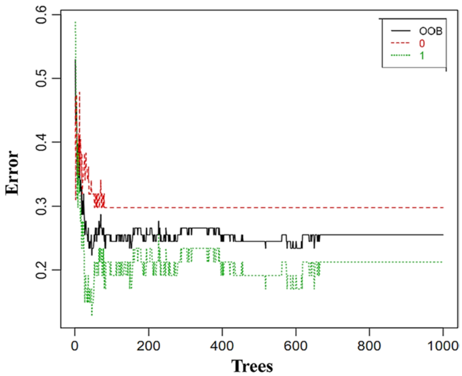


Fig. 5 The error rate of the RF model (OOB: out of bag, 0: absence fire, and 1: presence of fire). OOB estimate of error rate: 9.18%; model accuracy: 90.82%

5.3 Probability maps

The probability values derived from the GARP model that ranged from 0.02 to 0.98 were grouped into five probability classes (Fig. 6a). This model delineated approximately 45% and 40% the land area into the very low to low and high to very high probability classes, respectively (Fig. 7). The application of the BRT model led to probability values ranging from 0.2 to 0.97 with an overestimation of portion of the landscape (~60% of land area) classified into the high and very high probability to fire occurrences compared to other two models (Figs. 6b and 7). The probability values derived from the RF model were between 0.07 and 1 and were more evenly distributed across all probability classes over the landscape (Figs. 6c and 7).

6 Discussion

Our earlier research in the Zagros eco-region (Jaafari et al. 2017, 2018, 2019b; Jaafari and Pourghasemi 2019) led us to specifically focus on quantifying the effects of different geo-environmental and anthropogenic factors on the probability of a fire to occur and how the inclusion/exclusion of different variables would change the results. In this study, we tested the performances of three machine learning methods for developing wildfire predictive models. Whereas the RF and BRT models have a history of widespread use in wildfire modeling (Pourtaghi et al. 2016), this study applied the GARP model for the first time in the context of wildfire modeling. GARP utilizes a stochastic approach based on a genetic algorithm that allows different outputs to be run simultaneously to obtain a near optimum result (Stockwell 1999). While this advantage has enabled the GARP model to successfully outperform several other machine learning methods in other modeling projects (Darabi et al. 2019; Rahmati et al. 2019), our comparative study revealed that the RF model provided slightly better training and validation performances than the GARP model. The RF model inherently benefits from several important features (Ließ et al. 2012) that make this model a suitable tool for the prediction of wildfires: (1) RF is a straightforward learning machine that can easily be coupled with a GIS; (2) RF has the ability to handle very high-dimensional datasets with discrete and continuous factors, (3) during the training phase of the RF model, interactions among input factors can be detected, (4) RF is capable of

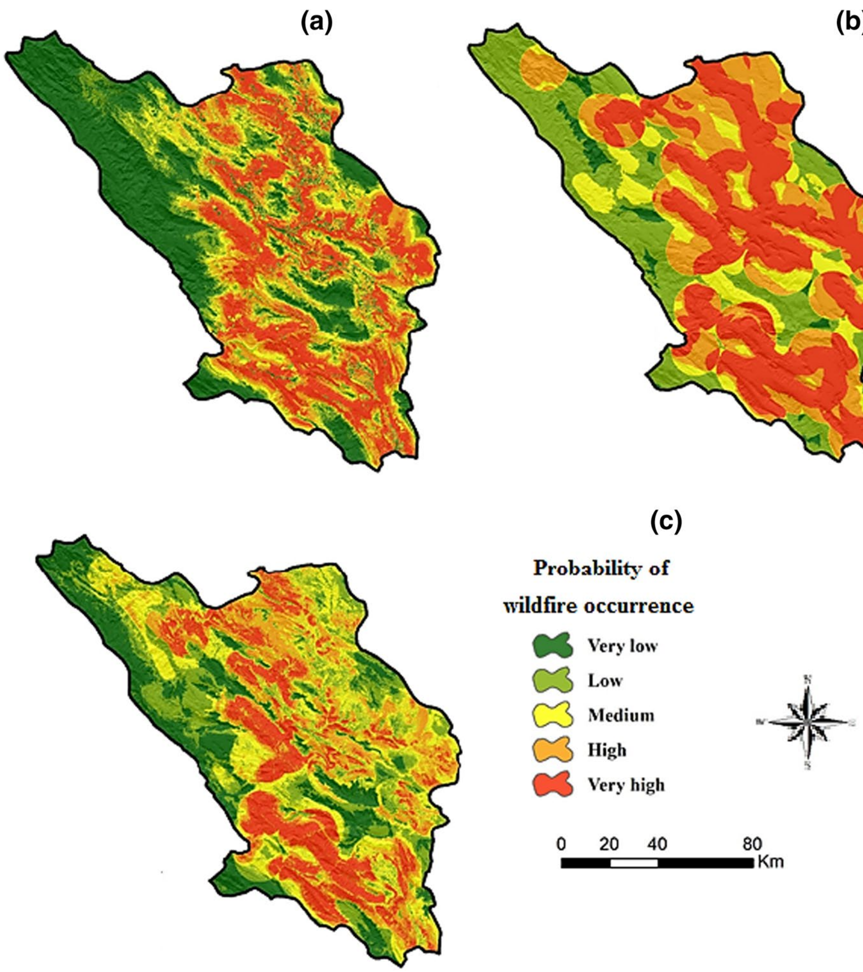
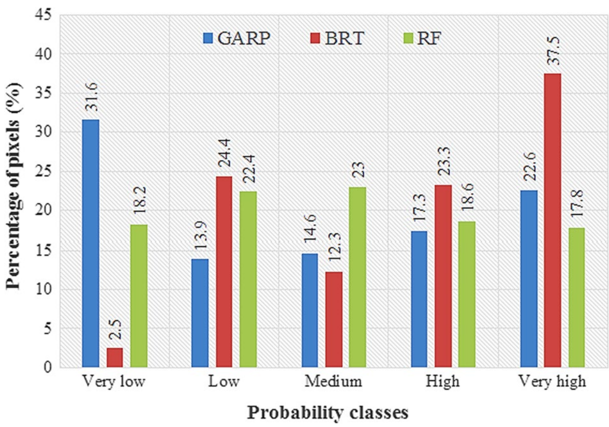


Fig. 6 Probability maps of wildfire occurrence produced using **a** GARP, **b** BRT, and **c** RF

Fig. 7 Distribution of probability classes in the four wildfire probability maps



capturing nonlinear relationships among the input factors, (5) the computational cost is low and the training speed is fast, and (6) RF achieves a high level of quality performance while being less prone to overfitting.

Although the superiority of the RF model over other models has been acknowledged by different authors (Oliveira et al. 2012; Valdez et al. 2017), Pourtaghi et al. (2016) reported that BRT out-performed RF predicting upcoming fires. Pourtaghi et al. (2016) attributed these results to the nature of BRT that integrates the strengths of regression trees and boosting to provide enhanced predictive performance. From these divergent results, we are inclined to conclude that the spatially explicit modeling using machine learning techniques are perhaps site-specific and likely depend on the factors included in the geospatial database that the models are built upon.

Although the probability of wildfire occurrences is strongly dependent on the local climate (Stocks et al. 1998; Gillett et al. 2004; Flannigan et al. 2009; Wu et al. 2015), anthropogenic influences have been increasingly identified as the main driver of wildland fires across the world (Syphard et al. 2008; Vilar et al. 2010; Oliveira et al. 2012; Collins et al. 2015). As a consequence, climate factors (e.g., rainfall, temperature, and wind effect) contribute less to empirical models predicting fire probability across the landscape than climatic variables, despite ongoing climate change over the past years. Similarly, our results show that the contribution of individual human-related factors exceeded the total contribution of climate- and topographic-related factors in the models of wildfire probability. Human-related factors in this study that accounted for human infrastructure (e.g., settlements, roads) and human-created land-use types (e.g., agriculture) contributed approximately 45% to the model of wildfire probability in this study. In contrast to previous works that relied on data for the period from 2007 to 2014 and limited the influence of human-related factors on wildfire occurrence to the analysis to distance to national roads, which failed to document a significant association between proximity to roads and intensity of fire occurrence works (Jaafari et al. 2017, 2018, 2019b; Jaafari and Pourghasemi 2019), the expanded current database included the a layer of local roads and clearly showed a significant effect of human activities along these roads that resulted in increased probability of fire occurrence. Specifically, distance from human settlements and roads jointly contributed approximately 38% to the model of wildfire probability in this study area, lending support to previous studies that showed that human activities are a strong driver of increasing fire occurrence (Vilar et al. 2010; Syphard et al. 2008; Collins et al. 2015).

The strength of the relationship between wildfires and roads seems to be dependent upon the region, the scale of analysis, and human population densities and activities (Collins et al. 2015). For example, in the densely poulated Santa Monica Mountains in California, distance to roads was the strongest factor of fire occurrence (Syphard et al. 2008) whereas this was not the case in Chinese boreal forests where climate-related factors were more strongly associated with fire occurrence than roads (Wu et al. 2014). Increased human infrastructure and local road density does not necessarily result in greater wildfire occurrence, however; In fact, the probability of fire occurrence may even be lower close to human settlements and roads than at greater distances to human infrastructure if fire suppression policies exist and are enforced and strengthened by well-equipped firefighting services (Gralewicz et al. 2012). Thus, in cases where prevention measures that targeted wildfire-related human activities were successfully implemented in Portugal, climate-related factors once again replaced human-related factors as the most important factors that controlled the probability of wildfire occurrence (Rodrigues et al. 2018). We interpret the strong association of human-related factors with past wildfire occurrences in our study area with Mediterranean climate as reflecting mostly unintended fire initiations following increased human activities or landscape modifications. If so, increased human influences on wildland landscapes should be accompanied by appropriate safety measures (e.g., monitoring, fire stations) that can be readily called upon to address such wildfires.

Although the topographic factor elevation was more strongly associated with wildfire occurrence than the topographic factors slope and aspect, which has also been seen in Canada (Gralewicz et al. 2012), the additional contribution of topographic factors to wildfire occurrence in the Zagros eco-region was limited after accounting for human-related and climatic factors. This weak additional contribution of topographic factors has also been reported for the Hyrcanian eco-region of northern Iran (Adab et al. 2018) and several European countries (Oliveira et al. 2012; Viedma et al. 2018) whereas other authors have found a strong correlation of topography with fire occurrence (Carmo et al. 2011; Satir et al. 2016). In general, however, because topography typically has a strong association with human-related and climatic factors (e.g., strong relation of elevation with temperature or land use), accounting for the effects of human-related factors on fire occurrence likely diminishes the additional contribution of topography in empirical models of fire probability that incorporate many correlated factors.

Overall, our findings may be useful for adopting fire management strategies that are adjusted to local conditions. These strategies may include differential prevention measurements with respect to the main drivers in different portions of the landscape, regulation of human activities in areas that are more fire-prone, fuel reduction measures adjacent to the human settlements and along rivers, and an informed allocation of public resources and support systems in hazardous areas prior to the start of the main fire season. Finally, it worth noting that due to significant anthropogenic activities in the Zagros eco-region, the zones of probability to wildfire occurrence delimited by these three models are only reflective of the time period of the study and should not be understood to persist unchanged for a long time. Thus, wildfire probability maps should be periodically updated for more informed and timely fire prevention.

7 Conclusions

Using three machine learning methods, we showed how different geo-environmental and anthropogenic factors interact to determine wildfire probability in a mountainous landscape in western Iran. Anthropogenic influences such as land use and distance from the settlements and roads had a relative contribution of 45% in models predicting the probability of wildfire occurrence in the study area. The models ranked climate factors as the second most influential drivers of fire occurrences, whereas topographic features only contributed an additional 10–15% to the fire probability model after accounting for anthropocentric and climatic factors in the model. We conclude that wildfire dynamics in the Zagros landscape are now strongly influenced by human activities and infrastructure and suggest that future fire management strategies should be directed to enhance preparedness of human communities through the application of fire safety techniques. Our approach to develop spatially explicit models of the probability of wildfire occurrence throughout large landscapes can identify high-risk areas to which scarce resources can be deployed for efficacious suppression activities to prevent fire ignitions outright and to develop suppression policies (e.g., monitoring protocols) to reduce the probability of future wildfire occurrence where fires are most likely to occur.

Funding This study was funded by Research Institute of Forests and Rangelands (RIFR) as part of the National Research Project No. 0-09-09-002-000095.

Declarations

Conflict of interest The authors declared that they have no conflict of interest.

References

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

Adab H, Atabati A, Oliveira S, Gheshlagh AM (2018) Assessing fire hazard potential and its main drivers in Mazandaran province, Iran: a data-driven approach. Environ Monit Assess 190:670

Breiman L (2001) Random forests. Mach Learn 45:5–32

Carmo M, Moreira F, Casimiro P, Vaz P (2011) Land use and topography influences on wildfire occurrence in northern Portugal. Landsc Urban Plan 100:169–176

Chen X, Quan Q, Zhang K, Wei J (2021) Spatiotemporal characteristics and attribution of dry/wet conditions in the Weihe River Basin within a typical monsoon transition zone of East Asia over the recent 547 years. Environ Model Softw 143:105116

Chen Z, Liu Z, Yin L, Zheng W (2022) Statistical analysis of regional air temperature characteristics before and after dam construction. Urban Clim 41:101085

Chuvieco E et al (2010) Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecol Model 221:46–58

Collins KM, Price OF, Penman TD (2015) Spatial patterns of wildfire ignitions in south-eastern Australia. Int J Wildland Fire 24:1098–1108

Darabi H et al (2019) Urban flood risk mapping using the GARP and QUEST models: a comparative study of machine learning techniques. J Hydrol 569:142–154

Dennison PE, Brewer SC, Arnold JD, Moritz MA (2014) Large wildfire trends in the western United States, 1984–2011. Geophys Res Lett 41:2928–2933

Elia M, D'Este M, Ascoli D, Giannico V, Spano G, Ganga A, Colangelo G, Lafortezza R, Sanesi G (2020) Estimating the probability of wildfire occurrence in Mediterranean landscapes using artificial neural networks. Environ Impact Assess Rev 85:106474

Feria A, Peterson T (2002) Prediction of bird community composition based on point-occurrence data and inferential algorithms: a valuable tool in biodiversity assessments. Divers Distrib 8:49–56

Flannigan MD, Krawchuk MA, de Groot WJ, Wotton BM, Gowman LM (2009) Implications of changing climate for global wildland fire. Int J Wildland Fire 18:483–507

Ghorbanzadeh O et al (2019) Spatial prediction of wildfire susceptibility using field survey GPS data and machine learning approaches. Fire 2(3):43

Giglio L, Descloitres J, Justice CO, Kaufman YJ (2003) An enhanced contextual fire detection algorithm for MODIS. Remote Sens Environ 87:273–282

Gillett N, Weaver A, Zwiers F, Flannigan M (2004) Detecting the effect of climate change on Canadian forest fires. Geophys Res Lett 31:18

Goleiji E, Hosseini SM, Khorasani N, Monavari SM (2017) Forest fire risk assessment-an integrated approach based on multicriteria evaluation. Environ Monit Assess 189:612

Gralewicz NJ, Nelson TA, Wulder MA (2012) Factors influencing national scale wildfire susceptibility in Canada. For Ecol Manag 265:20–29

Hong H, Jaafari A, Zenner EK (2019) Predicting spatial patterns of wildfire susceptibility in the Huichang County, China: an integrated model to analysis of landscape indicators. Ecol Ind 101:878–891

Hyde K, Dickinson MB, Bohrer G, Calkin D, Evers L, Gilbertson-Day J, Nicolet T, Ryan K, Tague C (2013) Research and development supporting risk-based wildfire effects prediction for fuels and fire management: status and needs. Int J Wildland Fire 22(1):37–50

Jaafari A, Pourghasemi HR (2019) Factors influencing regional-scale wildfire probability in Iran: an application of random forest and support vector machine. In: Pourghasemi HR, Gokceoglu C (eds) Spatial modeling in GIS and R for earth and environmental sciences. Elsevier, pp 607–619

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

Jaafari A, Zenner EK, Pham BT (2018) Wildfire spatial pattern analysis in the Zagros Mountains, Iran: a comparative study of decision tree based classifiers. Ecol Inform 43:200–211

Jaafari A, Mafi-Gholami D, Pham BT, Tien Bui D (2019a) Wildfire probability mapping: bivariate vs. multivariate statistics. Remote Sens 11:618

Jaafari A, Razavi Termeh SV, Bui DT (2019b) Genetic and firefly metaheuristic algorithms for an optimized neuro-fuzzy prediction modeling of wildfire probability. J Environ Manag 243:358–369

Lan Z, Zhao Y, Zhang J, Jiao R, Khan MN, Sial TA, Si B (2021) Long-term vegetation restoration increases deep soil carbon storage in the Northern Loess Plateau. Sci Rep 11(1):1–11

Li J, Zhao Y, Zhang A, Song B, Hill RL (2021) Effect of grazing exclusion on nitrous oxide emissions during freeze–thaw cycles in a typical steppe of Inner Mongolia. Agric Ecosyst Environ 307:107217

Ließ M, Glaser B, Huwe B (2012) Uncertainty in the spatial prediction of soil texture: comparison of regression tree and Random Forest models. Geoderma 170:70–79

Liu Y, Zhang K, Li Z, Liu Z, Wang J, Huang P (2020) A hybrid runoff generation modelling framework based on spatial combination of three runoff generation schemes for semi-humid and semi-arid watersheds. J Hydrol 590:125440

Liu B, Spiekermann R, Zhao C, Püttmann W, Sun Y, Jasper A, Uhl D (2022) Evidence for the repeated occurrence of wildfires in an upper Pliocene lignite deposit from Yunnan, SW China. Int J Coal Geol 250:103924

Naderpour M, Rizeei HM, Ramezani F (2021) Forest fire risk prediction: a spatial deep neural network-based framework. Remote Sens 13(13):2513

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

Pausas JG, Keeley JE (2009) A burning story: the role of fire in the history of life. Bioscience 59:593–601

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 Ind 64:72–84

Quan Q, Liang W, Yan D, Lei J (2022) Influences of joint action of natural and social factors on atmospheric process of hydrological cycle in Inner Mongolia. China Urban Clim 41:101043

Rahmati O, Golkarian A, Biggs T, Keesstra S, Mohammadi F, Daliakopoulos IN (2019) Land subsidence hazard modeling: machine learning to identify predictors and the role of human activities. J Environ Manag 236:466–480

Rasooli SB, Rashidi S, Shabanian N (2021) Fire sensitivity of broadleaf tree species in plantations of Kurdistan, Iran. J for Res 32(3):1167–1176

Ridgeway G (2007) Generalized boosted models: a guide to the gbm package. Update 1(1):1–15

Rodrigues M, Jiménez-Ruano A, Peña-Angulo D, de la Riva J (2018) A comprehensive spatial-temporal analysis of driving factors of human-caused wildfires in Spain using geographically weighted logistic regression. J Environ Manag 225:177–192

Sari F (2021) Forest fire susceptibility mapping via multi-criteria decision analysis techniques for Mugla, Turkey: a comparative analysis of VIKOR and TOPSIS. For Ecol Manag 480:118644

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

Shabani S, Jaafari A, Bettinger P (2021) Spatial modeling of forest stand susceptibility to logging operations. Environ Impact Assess Rev 89:106601

Spano G, Elia M, Cappelluti O, Colangelo G, Giannico V, D'Este M, Lafortezza R, Sanesi G (2021) Is experience the best teacher? Knowledge, perceptions, and awareness of wildfire risk. Int J Environ Res Public Health 18(16):8385

Stocks BJ et al (1998) Climate change and forest fire potential in Russian and Canadian boreal forests. Clim Change 38:1–13

Stockwell D (1999) The GARP modelling system: problems and solutions to automated spatial prediction. Int J Geogr Inf Sci 13:143–158

Swain DL (2021) A shorter, sharper rainy season amplifies California wildfire risk. Geophys Res Lett 48(5):e2021GL092843

Syphard AD et al (2008) Predicting spatial patterns of fire on a southern California landscape. Int J Wildland Fire 17:602–613

Tavakkoli Piralilou S, Einali G, Ghorbanzadeh O, Nachappa TG, Gholamnia K, Blaschke T, Ghamisi P (2022) A Google Earth Engine approach for wildfire susceptibility prediction fusion with remote sensing data of different spatial resolutions. Remote Sens 14(3):672

Valdez MC, Chang KT, Chen CF, Chiang SH, Santos JL (2017) Modelling the spatial variability of wildfire susceptibility in Honduras using remote sensing and geographical information systems. Geomat Nat Haz Risk 8(2):876–892

Viedma O, Urbieta I, Moreno J (2018) Wildfires and the role of their drivers are changing over time in a large rural area of west-central Spain. Sci Rep 8:17797

Vilar L, Woolford DG, Martell DL, Martín MP (2010) A model for predicting human-caused wildfire occurrence in the region of Madrid, Spain. Int J Wildland Fire 19:325–337

Wu Z, He HS, Yang J, Liu Z, Liang Y (2014) Relative effects of climatic and local factors on fire occurrence in boreal forest landscapes of northeastern China. Sci Total Environ 493:472–480

Wu Z, He HS, Yang J, Liang Y (2015) Defining fire environment zones in the boreal forests of northeastern China. Sci Total Environ 518:106–116

Xie W, Li X, Jian W, Yang Y, Liu H, Robledo LF, Nie W (2021a) A novel hybrid method for landslide susceptibility mapping-based geodetector and machine learning cluster: a case of Xiaojin county, China. ISPRS Int J Geo-Inf 10(2):93

Xie W, Nie W, Saffari P, Robledo LF, Descote PY, Jian W (2021b) Landslide hazard assessment based on Bayesian optimization–support vector machine in Nanping City, China. Nat Hazards 109(1):931–948

Yin L, Wang L, Zheng W, Ge L, Tian J, Liu Y, Yang B, Liu S (2022) Evaluation of empirical atmospheric models using Swarm-C satellite data. Atmosphere 13(2):294

Zhang G, Wang M, Liu K (2019a) Forest Fire susceptibility modeling using a convolutional neural network for Yunnan Province of China. Int J Disaster Risk Sci 10(3):386–403

Zhang K, Ali A, Antonarakis A, Moghaddam M, Saatchi S, Tabatabaeenejad A, Chen R, Jaruwatanadilok S, Cuenca R, Crow WT, Moorcroft P (2019b) The sensitivity of North American terrestrial carbon fluxes to spatial and temporal variation in soil moisture: an analysis using radar-derived estimates of root-zone soil moisture. J Geophys Res Biogeosci 124(11):3208–3231

Zhang K, Wang S, Bao H, Zhao X (2019c) Characteristics and influencing factors of rainfall-induced landslide and debris flow hazards in Shaanxi Province, China. Nat Hazard 19(1):93–105

Zhao X, Xia H, Pan L, Song H, Niu W, Wang R, Li R, Bian X, Guo Y, Qin Y (2021) Drought monitoring over Yellow River basin from 2003–2019 using reconstructed MODIS land surface temperature in Google Earth Engine. Remote Sens 13(18):3748