ISSN: 2538-2152; ECOPERSIA 2025;13(1): 13-36.

Enhancing Fire Susceptibility Mapping in Semnan Province: Integrating Machine Learning and Geospatial Analysis

ARTICLEINFO

Article Type
Original Research

Authors
Ali Asghar Zolfaghari, *Ph.D.*[1*]
Maryam Raeesi, *Ph.D.*[2]
Zahra Sheikh, *Ph.D.*[3]
Azadeh Soltani, *M.Sc.*[4]
Soghra Poodineh, *M.Sc.*[5]
Mojtaba Amiri, *Ph.D.*[6]

How to cite this article
Zolfaghari A.A., Raeesi M., Sheikh Z., Soltani A., Poodineh S., Amiri M. Enhancing Fire Susceptibility Mapping in Semnan Province: Integrating Machine Learning and Geospatial Analysis. ECOPERSIA 2025;13(1): 13-36.

DOI:
10.22034/ECOPERSIA.13.1.13

[1] Associate Professor, Faculty of Desert Studies, Semnan University, Semnan, Iran.
[2] Ph.D., Faculty of Desert Studies, Semnan University, Semnan, Iran.
[3] Ph.D., Faculty of Desert Studies, Semnan University, Semnan, Iran. Soil Conservation and Watershed Management Research Institute (SCWMRI), Tehran, Iran.
[4] M.Sc, Faculty of Desert Studies, Semnan University, Semnan, Iran.
[5] M.Sc, Faculty of Desert Studies, Semnan University, Semnan, Iran.
[6] Associate professor, Faculty of Natural Resources, Semnan University, Semnan, Iran.

**** Correspondence***
Address: Associate Professor, Faculty of Desert Studies, Semnan University, Semnan, Iran.
Tell:+98928580616
Fax:02331535556
Email: azolfaghari@semnan.ac.ir

Article History
Received: September 18, 2024
Accepted: January 5, 2025
Published: February 28, 2025

ABSTRACT

Aims: This study assesses the impacts of natural and human factors on fire occurrences, identifies key contributors to fire susceptibility maps, and employs machine learning algorithms (MLAs) to enhance the spatiotemporal patterns of fire susceptibility maps.
Materials & Methods: Data were collected from 110 fire locations and 110 non-fire points from 2001 to 2022 at an annual scale. Various auxiliary variables were analyzed to model fire susceptibility, including climate data, terrain features, the Normalized Difference Vegetation Index (NDVI), and distance to roads. The study employed multiple MLAs, including Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Decision Trees (GBDT), to generate the fire susceptibility maps.
Findings: About 70% of fires occurred within 2 km of roads, indicating significant human influence. Grasslands had the highest fire rates, with over 25% of fires from 2001-2022 due to flammable fuels. The RF and mean models identified 0.4% and 1.31% of the area as very high susceptibility (38,800 km^2 and 12,600 km^2), while the GBDT and SVM models identified 2.42% and 1.86% (234,700 km^2 and 180,000 km^2). Though small in percentage, the very high susceptibility class covers large areas.
Conclusion: This research highlights the importance of integrating environmental and human factors to predict fire events in arid regions and develop comprehensive fire susceptibility maps, critical for protecting vulnerable ecosystems. These outcomes provide valuable tools for fire management and mitigation strategies within vulnerable ecosystems. Moreover, developing targeted fire management strategies focused on high-risk areas, such as juniper and broadleaf forests, must be a priority.

Keywords: Auxiliary Variables; Climatic Indicators; Juniper Forest; Distance to Road, Fire Susceptibility.

CITATION LINKS

[1] Marlon J., Bartlein P., ... [2] You C., Yao T., Xu C. Environmental signific... [3] Guo F., Su Z., Wang G., .. [4] Sevinc V., Kucuk O., Goltas M.A. Bayesian ne... [5] Su Z., Zheng L., Luo S., Tigabu M., Guo F. M... [6] Bowman D.M.J.S., Moreira-Muñoz A., Kolden C.... [7] Forkel M., Dorigo W., Lasslop G., Teubner I.... [8] Eskandari S., Pourghasemi H.R., Tiefenbacher... [9] Mallinis G., Mitsopoulos I., Chrysafi I. Eva... [10] Santos A.C. dos., Montenegro S. da. R., Ferr... [11] Arabameri A., Pal. S., Costache R.D., Saha A... [12] Ahmed I.A., Talukdar S., Shahfahad Parvez A.... [13] Akıncı H.A., Akıncı H. Machine learning base... [14] Alkan Akinci H., Akinci H., Zeybek M. Compar... [15] Novo A., Dutal H., Eskandari S. Fire suscept... [16] Tonini M., D'Andrea M., Biondi G., Degli Esp... [17] Liu Z., Peng C., Timothy W., Candau J.N., De... [18] Jahanbani M., Vahidnia M.H., ... [19] Janizadeh S., Chandra Pal S., Saha A., Chowd... [20] Mirzaei S., Vafakhah M.,... [21] Yang D., Zhang ... [22] Pourghasemi H.R., Gayen... [23] Barreto J., Armenteras... [24] Tan C., Feng Z. Mapping... [25] Chevitarese D.S., Szwarcman D., Silva R.G., ... [26] Hu Z.X., Wang Y., Ge M.... [27] Qu S., Guan Z., Verschuur... [28] Smith R., Mukerji T., Lupo T. Correlating ge... [29] Tse K.C., Chiu H.C., Tsang M.Y., Li Y., Lam ... [30] Zhang G., Wang Z., ... [31] Zhou K.B., Zhang Z.X., Liu J., Hu Z.X., Duan... [32] Barmpoutis P., Papaioannou P., Dimitropoulos... [33] Eskandari S., ... [34] Sadeghi A., Ahmadi... [35] Amiri M., Pourghasemi H.R. Chapter 25 - Pred... [36] Mayr M.J., Vanselow K.A., Samimi C. Fire reg... [37] Mandal P., Maiti A., Paul S., Bhattacharya S... [38] Raeesi M., Zolfaghari A., Rahimi M., Kaboli ... [39] Amini E., Zolfaghari A., Kaboli H., Rahimi M... [40] Wu Z., Li M., ... [41] Kopecký M., Macek M., Wild J. Topographic We... [42] Qin Z., Zhu Y., Li W., ... [43] Xu Ch., Li Y., Hu... [44] Kursa M.B., Rudnicki ... [45] Szul T., Tabor S., .. [46] Pourghasemi... [47] Breiman L. Random ... [48] Jain P., Coogan S.C.P., ... [49] Hawryło P., ... [50] Mustafa A., Rienow... [51] Lu J., Lu D., Zhang X., Bi Y., Cheng K., Zhe... [52] Touzani S., ... [53] Friedman J.H. :... [54] Tian Z., Xiao J., Feng H., Wei Y. Credit Ris... [55] Mosavi A., Golshan... [56] Tien Bui D., ... [57] Chang Y., Zhu Z., Bu R., Chen H., Feng Y., L... [58] Guo F., Zhang L., Jin, ... [59] McLauchlan K.K., Higuera P.E., Miesel J., Ro... [60] Van Etten E., Burrows N. On the Ecology of A... [61] Li W., Xu Q., Yi J., ... [62] Eskandari S., Oladi J., Jalilvand H., Saradj... [63] Pourghasemi H.R. ... [64] Pourtaghi Z.S., ... [65] Romero-Calcerrada R., Novillo C.J., Millingt... [66] Faraji F., Alijanpour ... [67] Akhzari D., Mohammadi E., Saedi K. Studying ... [68] Barreto J.S., Armenteras... [69] Iban M.C., Sekertekin A. Machine learning ba...

Introduction

Fire occurrence has a vital effect on ecosystem succession, carbon cycle, atmospheric chemical composition, and configuration [1,2]. With the advancement of fire susceptibility prediction studies, numerous analyses are conducted comprehensively using various factors, including meteorology, topography, vegetation cover and indices, and human activities [4,5] that determine fire potential in each region [6,7]. In recent years, fire incidents have increased significantly in Iran's forests and rangelands as human access to these natural areas expands [8]. Although fire monitoring, risk forecasting, and mapping have benefited significantly from using the earth observation data and models, every single monitoring, characterizing, and zoning of fires is challenging [9,10].

Hazard susceptibility mapping has recently seen widespread use of machine learning algorithms (MLAs) techniques worldwide [12]. MLA has been widely used in research on forest fires and flood susceptibility prediction [13], forest fire susceptibility map [11,14,15], wildfire susceptibility mapping [16], forest ecology [17], mapping the variability of flood hazard [19], and multi-hazard mapping [22]. Several researchers have employed MLA to model fire occurrences, such as Random Forest (RF) [4,16,24], Support Vector Machine (SVM) [24], and Gradient Boosting Decision Tree (GBDT) [24] receiving desirable outputs in solving classification issues [25-30]. Barmpoutis et al. (2020) presented a comprehensive review of fire detection systems, including ground, airborne, and spaceborne systems. They also illustrated the deep and classical machine learning models adopted to detect fire in each system. Eskandari and Chuvieco (2015) researched the fire danger systems to produce a fire danger probability map using a geographical database in Iran. They calculated the correlation coefficient between fire transmission probability and burned regions with a high significance of 69%. Research indicates increasing wildfire incidents in Lorestan Province, Iran, particularly in semiarid oak forests. Denser forests are found to be more vulnerable to fires, with the NDVI identified as a key indicator of wildfire susceptibility. Human activities near roads and urban areas significantly affect fire patterns. The study highlights over 1600 km^2 of highly susceptible regions, emphasizing the need for targeted conservation and community involvement in forest protection [34]. Another study highlighted forest-fire susceptibility (FFS) maps for Fars Province, Iran, using GIS-based machine-learning algorithms, with a boosted regression tree (BRT) achieving the highest accuracy (AUC = 88.90%). Key factors influencing FFS included land-use, annual mean rainfall, and slope angle. These findings can enhance forest resource management in the region.[46]. Mayr et al. (2018) researched Namibia and confirmed that mean annual rainfall is exceptionally significant for fire activity. However, human impacts are supplementary control of fuel accessibility. Eskandari et al. (2020) have demonstrated that fire danger is firmly connected to the distance from roads and climate conditions. They applied data mining techniques to model fire danger and mapping by using both climate and geomorphological factors to model fire danger in Koohdasht, Lorestan, Iran. Mandal et al. (2022) mapped the multi-hazards risks using Analytic Hierarchical Analysis (AHP), Random Forest (RF), and Artificial Neutral Network (ANN). They concluded that RF was the most accurate model in mapping multi-hazard risk. Tan & Feng (2023) developed a forest fire risk map using three machine learning methods, including RF, SVM, and GBDT, to predict the probability of a forest

fire. Results showed that RF performed the highest degree of accuracy (91.68%) and precision (92.78%). Moreover, they concluded that the main factors contributing to forest fires are meteorology and vegetation. Fire significantly threatens natural ecosystems in arid and semiarid regions like Semnan Province in central Iran. These areas are vulnerable due to climatic conditions and sparse vegetation, limiting fire fuel. Xerophytic plants can still suffer extensive damage, particularly in grasslands, rangelands, broadleaf forests, juniper forests, and Haloxylon habitats that protect soil from erosion. The high density of juniper trees raises concerns about fire frequency, even though fires are rare due to low rainfall and productivity. Our research hypothesizes that human factors, particularly distance to roads, significantly impact fire occurrence by increasing access and the risk of accidental ignitions. Understanding these spatial distributions is crucial for effective fire management. Our innovative study comprehensively assesses fire susceptibility in Semnan's fragile ecosystems. We generate detailed fire susceptibility maps using advanced MLAs that identify high-risk areas. This research enhances predictive accuracy and supports targeted resource allocation, addressing the urgent need to protect ecosystems increasingly threatened by fire incidents. We evaluated fire occurrences in semiarid areas, considering different land-uses and land-cover, particularly in slow-growing juniper forests. This research follows three specific goals: 1) evaluating the impacts of both natural environmental and geomorphological data (meteorology, topography, and vegetation cover) along with anthropogenic variables (distance to roads) on fire occurrences, 2) identifying the main factors in fire susceptibility map, and 3) highlighting the fire susceptibility maps using machine learning techniques, e.g., RF, SVM, GBDT, and mean of models.

Materials & Methods

Study Area

Semnan Province, encompassing an area of approximately 96,816 km^2, constitutes about 5.8% of Iran's total land area. It is geographically situated between the coordinates of 34°15' to 37°20' North latitude and 51°50' to 57°03' East longitude [38]. The province is positioned in the southern region of the Alborz Mountain range and includes a significant expanse of desert plain, which comprises more than half of its total area. The topography of Semnan Province exhibits a notable gradient in elevation that decreases from north to south, ranging from 640 meters to over 3,500 meters above sea level, with a mean elevation of approximately 1,067 meters (Figure 1).

The region can be categorized into three geomorphological zones: mountainous, sub-montane, and lowland desert plains. Climatically, Semnan has both arid and semiarid conditions characterized by low annual rainfall and a brief cold season. The mean annual precipitation is approximately 136 millimeters. Temperature variations within the study area include a minimum of 12.8°C, a maximum of 23.7°C, and a mean of 18.3°C [39]. Despite the predominantly arid climate, Semnan Province supports a diverse array of land-uses and land-cover types.

The land-use and land-cover (LULC) map of the study area highlights various land-cover types in Semnan Province (Figure 1). The map categorizes the landscape into several classes, including broadleaf forests, juniper forests, shrublands, grasslands, croplands, herbaceous wetlands, Haloxylon, permanent waterbodies, urban areas, bare soil, salty lands, gardens, and sparse vegetation. The dominant LULC cover type is juniper forests, prevalent in the north and significant for

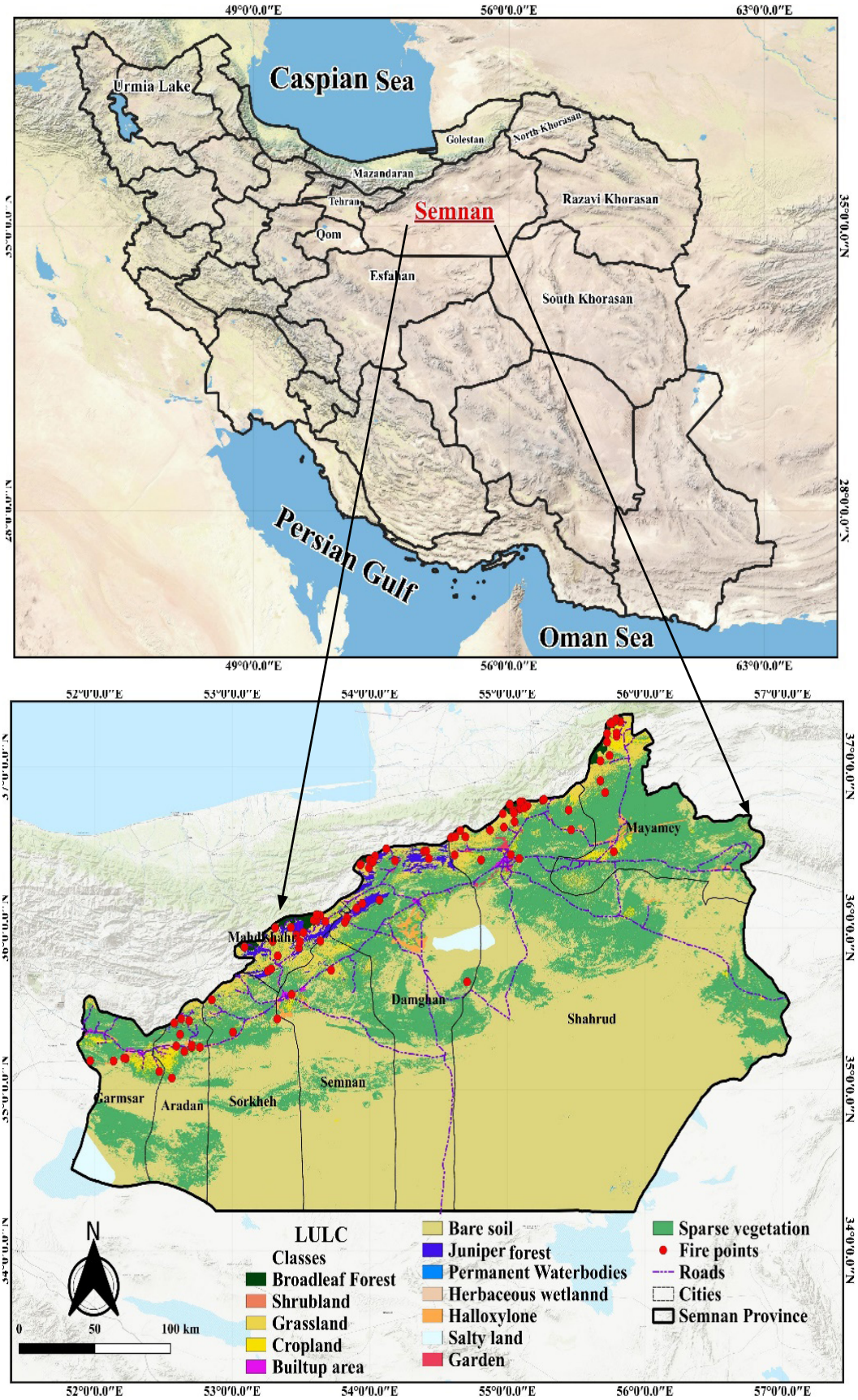

Figure 1) The location of the study area in Iran; the land-cover and land-use (LULC) map with fire occurrence points that occurred between 2001 and 2022 in the Semnan Province.

their ecological role and fire vulnerability. Other essential land-covers include shrublands and grasslands, enhancing the region's biodiversity and fire dynamics. Most agricultural land in Semnan Province, excluding the Hosein Abad Kalpush region in the north, is under irrigated cultivation, with a total area of approximately 200,000 hectares dedicated to this purpose (Figure 1).

Workflow

This study was conducted in four steps: i) determining the fire inventory map, ii) preparing the effective variables for prediction of the fire susceptibility map, iii) selecting effective variables, iv) generation of a fire susceptibility map using three different machine learning algorithms (Figure 2). The entire processing was conducted using R version 4.2. and QGIS version 3.36.

Auxiliary Variables for Mapping Fire Susceptibility

Several 110 observed fire locations were recorded between 2001 and 2022 in different LULC classes in Semnan Province. Moreover, 110 non-fire points were randomly identified in the study area. Consequently, 220 fire and non-fire points were used to model the fire susceptibility map. To do so, several auxiliary variables were considered, e.g., annual rainfall and temperature as climatic variables, digital

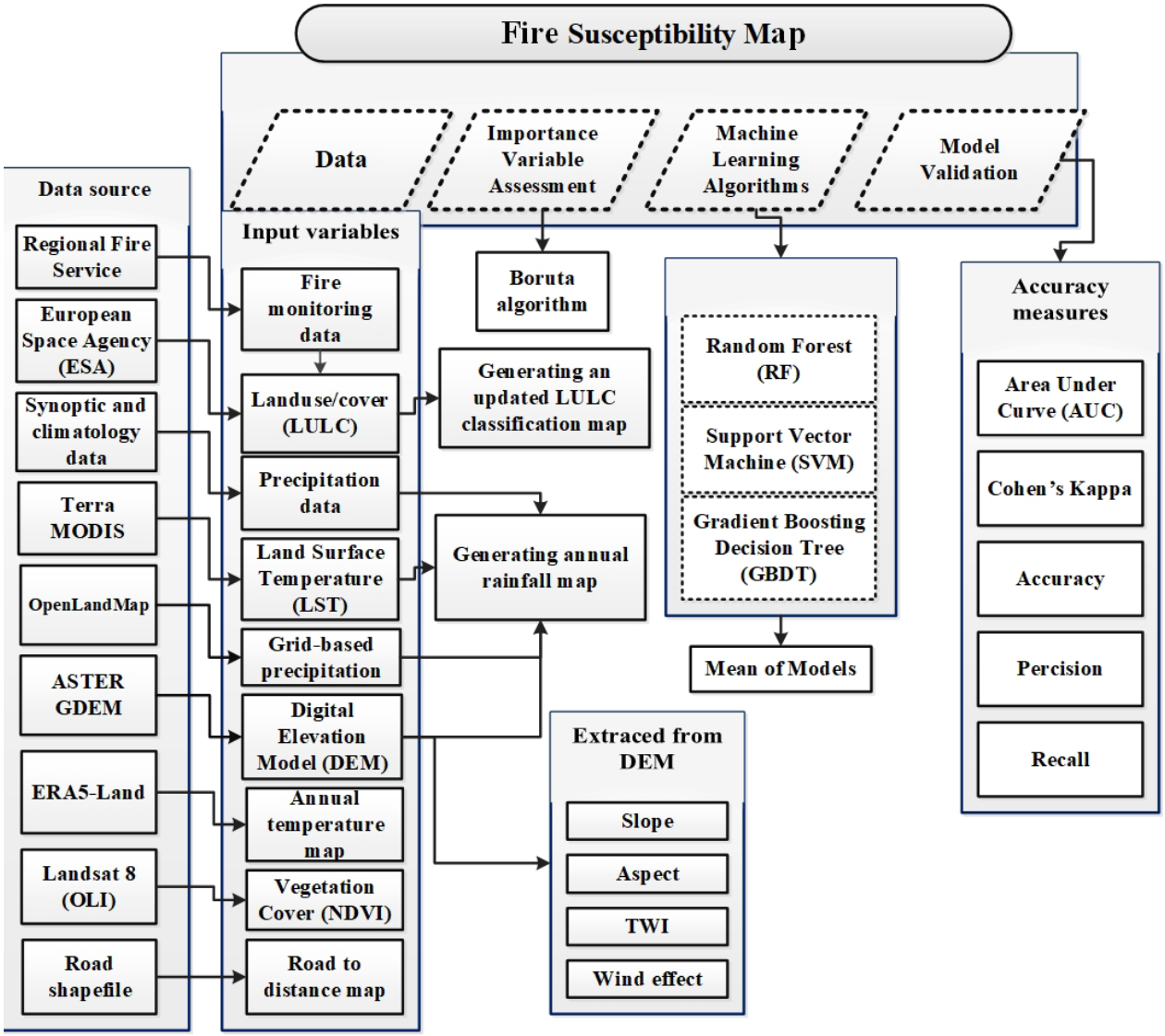

Figure 2) Flowchart of research processes and analysis.

elevation model and its terrain attributes (slope, aspect, and topographic wetness index), NDVI as vegetation cover, distance to roads, and LULC as anthropogenic variables (Table 1).

Land-Use/Cover (LULC) Map: We used the land-cover produced by the European Space Agency (ESA). This product provides a global land-cover map at 10 m resolution based on Sentinel-1 and Sentinel-2 data. The selected land-cover product was not sorted into several classes properly, and it considered Haloxylon and juniper forests as the same class in the rangelands class. Also, some cropland areas were not detected in the ESA landcover product. Thus, we updated the land-cover map and separated Haloxylon vegetation cover, juniper forest, and croplands into different classes according to the study area's environmental conditions (Figure 1).

Rainfall Data: In this study, the spatial distribution of mean annual rainfall varies significantly. An accurate rainfall map is crucial in modeling a fire susceptibility map. We utilized the predicted mean rainfall map at an annual scale created by Amini et al. (2022)many experimental models have been proposed, of which the Revised Universal Soil Loss Equation (RUSLE based on ground rainfall data from 9 synoptic stations and 44 rain gauge stations (2001-2022). Additionally, several auxiliary variables were employed to enhance the

Table 1) The primary selection of factors and data sources on the fire susceptibility map in the study area.

Row	Auxiliary Variables	Data Source	Impact
1	Mean annual rainfall (mm .year^{-1})	OLMP, LST (Terra MODIS), DEM	Decrease the spread of fire due to increasing soil moisture and vegetation growth
2	Mean annual temperature (°C)	ERA5-Land	Air surface temperature in participating fire ignition and spread
3	Distance from roads (km)	GIS-based data	Provide accessibility to human activities in fire frequency, spread, and occurrences in various areas
4	NDVI	Landsat 8 (OLI)	Vegetation cover and density that measures the availability of fire fuel
5	LULC	European Space Agency (ESA)	Determine the different land-use and land-cover involved in the exposure level of fire spread
6	Elevation/ Altitude(m)	ASTER DEM	Make a microclimate regarding vegetation distribution and types in controlling fire flammability and occurrences
7	Slope (%)	Terrain attributes	Slope controls vegetation cover distribution with a high impact on fire spreads, particularly at steep slopes.
8	Aspect	Terrain attributes	The hillside faces away from the direct sunshine and retains more moisture, supporting vegetation cover greenness.
9	TWI	Terrain attributes	TWI significantly influences fire occurrence and spread, as dry areas are prone to fire and spread rapidly. It indicates surface water content, representing the degree of humidity, especially in arid and semiarid regions.
10	Wind-effect	Terrain attributes	Topography/Terrain attributes offer fire protection by influencing terrain on climate state, vegetation, and as fuel breaks to fire spread.

accuracy of the annual rainfall map, including OpenLandMap precipitation (OLMP), digital elevation model (DEM), and land surface temperature (LST).

Annual Temperature Map: The annual air temperature is an effective factor in mapping fire susceptibility. However, the number of synoptic and climatic stations cannot map the study area's temperature. Therefore, the ERA5-Land reanalysis data were used to map air temperature in this research. The ERA5-Land data have biases and should be corrected before being used as measured data. In this study, ERA5-Land data were corrected using station data for each year using the quantile mapping (QM) method. This statistical approach adjusts the distribution of a variable to better align with the observed distribution. The QM method utilizes the Empirical Cumulative Distribution Function (ECDF) of both the measured and ERA5-Land data to reduce systematic biases in the ERA5-Land data. Equations (1-3) were applied to address these systematic biases in the spatial data by considering the probability of temperature occurrence ($Prob_{gp}$) for individual grid pixels derived from ERA5-Land data collected from 2000 to 2020.

$$Prob_{gp} = ECDF_{ERA5-Land}(P_{ERA5-Land,gp}) \quad \text{Eq. (1)}$$

$$CF_{gp} = ECDF^{-1}_{obs,station}(Prob_{gp}) - ECDF^{-1}_{ERA5-Land}(Prob_{gp}) \quad \text{Eq. (2)}$$

$$P_{gp,BC} = P_{ERA5-Land,gp} + CF_{gp} \quad \text{Eq. (3)}$$

$ECDF_{ERA5-Land}$ is the empirical cumulative distribution function (ECDF) of ERA5-Land reanalysis data at station points, computed annually to account for variations over time. The transfer function relating the ECDF of observed precipitation to ERA5-Land reanalysis data can be acquired from Eq.(2). Moreover, the correction factor (CF_{gp}) indicates the discrepancies between observed and ERA5-Land temperature data in the given pixel determined by the $Prob_{gp}$. Subsequently, the CF_{gp} is applied to adjust the ERA5-Land temperature data ($P_{ERA5-Land,gp}$) at each pixel, yielding the bias-corrected ERA5-Land reanalysis precipitation data ($P_{gp,BC}$), as detailed in Eq. (3). For next steps, the corrected ERA5-Land reanalysis data was considered as measured data (2001-2022) in the study area. The mean of ERA5-Lands data was used as a temperature map in machine learning modeling.

Topographic Data: Fire occurrences and spread are directly affected by topographic variations, which dominate spatial distribution and vegetation type composition. Moreover, DEM and its terrain attributes, e.g., slope and aspect, have been widely reported [24,40]. We downloaded ASTER DEM in 30 meters spatial resolution, and then slope angle, aspect, topographic wetness index (TWI), and wind effect were extracted as secondary features of DEM (Figure 3). TWI is mathematically expressed with the following equation [40].

$$TWI = \ln\left[\frac{A}{\tan\beta}\right] \quad \text{Eq. (4)}$$

Where A is the specific catchment area of a portion of land, and β is the slope gradients in degrees (angle).

Normalized Difference Vegetation Index (NDVI): The Normalized Difference Vegetation Index (NDVI) is used to measure vegetation cover, which is a good satellite-based indicator of vegetation on a landscape scale [42,43]. The mean NDVI values vary across different years; therefore, using the mean NDVI from multiple years to predict the fire susceptibility map is advisable. In the current study, the collection of Landsat satellite OLI data was processed to a time series layer stack of NDVI images from 2015 to 2022. Most of the annual plants in the study area have reached maximum growth

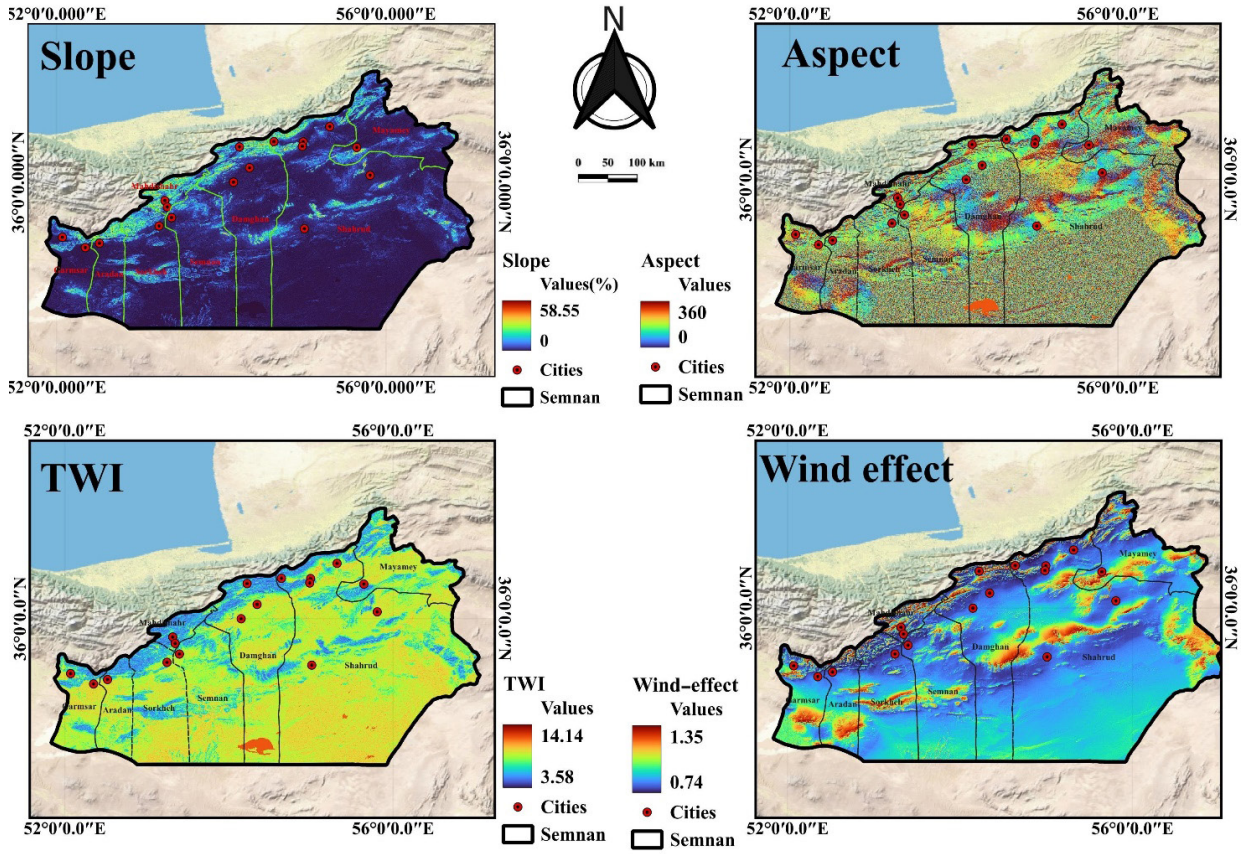

Figure 3) The practical driver factors in evaluating fire susceptibility extracting by Digital Elevation Model (DEM).

from March to June. Therefore, we used the median NDVI value of the first four months as an auxiliary variable for predicting the fire susceptibility map.

Distance to Roads: The influence of human activity on wildfire occurrence can be investigated using various spatial variables, including distance to roads, urban areas, villages, agricultural lands, and tourism regions. This study's initial analysis of fire incidents revealed that less than 5% occurred in agricultural land near villages and urban areas. Consequently, distance to populated areas was excluded as a predictor variable in the fire susceptibility model. However, distance to roads was incorporated as a key factor using a multiple buffer ring approach. Twelve buffer distance classes around roads were established: less than 50, 100, 200, 300, 400, 500, 700, 1000, 2000, 3000, 4000, and more than 4000 meters. These buffers were then dissolved to create contiguous zones for analysis.

Factor Analysis: Boruta, a machine-learning algorithm, is a feature-selection technique that can be used for regression and classification. It was developed as a wrapper-based approach based on a random forest classification. The Boruta algorithm functions by comparing the importance of each feature against that of randomly generated shadow features. A shadow attribute is created for each attribute, and its value is estimated by shuffling original attribute values across objects [44]. Subsequently, the shadow features are evaluated to identify the maximum Z-score (MZS). Any feature that scores better than MZS is assigned a hit. Features with importance scores significantly lower than

MZS (Z-scores< MZS) are considered irrelevant and rejected. Conversely, features with importance scores significantly higher than MZS (Z-scores> MZS) are deemed relevant and confirmed [45].

The Z-score is defined as:

$$Z-\text{Score}=\frac{\text{MDA}}{\text{SD}} \quad \text{Eq. (5)}$$

where MDA refers to Mean Decrease Accuracy, and SD is the Standard Deviation of accuracy losses.

Furthermore, a multicollinearity test was run between the variables (Table 1) using two important indices of variance inflation factor (VIF) and tolerance (T). VIF identifies the correlation between predictors or independent variables and the strength of that correlation. Based on the VIF value, a decision will be made to either exclude a variable from the modeling procedure or include it. If VIF >10, the multicollinearity test is considered high [46].

Data Processing

Machine Learning Algorithms (MLAs): Understanding the spatial distribution of fire susceptibility is vital for effective fire control and management. This study utilized three MLAs: Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Decision Trees (GBDT). Furthermore, we used the outputs of these models to create a new predictive model for mapping fire susceptibility in Semnan Province. Additional details are provided below.

Random Forest (RF): The RF algorithm is one of the non-parametric machine-learning techniques to appraise the relative significance of every variable in a predictive model. A considerable power of using the RF algorithm is its ability to execute both classification and regression processes [47]. It is an ensemble model with many individually trained Decision Trees (DTs). This algorithm's high level of performance is performed by reducing the correlations between trees while lessening model variance. Therefore, an abundance of diverse trees delivers more precision than each separated tree [48].

Support Vector Machine (SVM): The SVM algorithm is a nonlinear, binary classification process that aims to determine the thresholds that divide a training sample into predefined classes. The optimum separation minimizes misclassifications that usually occur during training [49]. The main benefit of SVM is the ability to convert models and solve nonlinear classification problems caused by a lack of prior knowledge of the modeling conditions [50].

Gradient Boosting Decision Tree (GBDT): Gradient boosting decision tree (GBDT) is a supervised classification used for prediction analysis [51,52], which can solve classification and regression problems [53]. It calculates the residuals between the current output and the true value of each weak learner. Then, it accumulates the residuals of each weak learner output to reduce the residuals in the training process and achieve the classification goal [54].

Fire susceptibility map: In this research, every single auxiliary variable consists of both natural environmental data (e.g. mean annual rainfall, bias-corrected mean annual temperature data, NDVI, LULC, elevation, slope's degree, aspect, TWI, and wind effect obtained from DEM) and anthropogenic covariates (distance to roads) were used to generate the final fire susceptibility maps in Semnan Province by using three machine learning methods: RF, SVM, and GBDT. The severity level of fire depends on the degree of influencing factors that originated from fire occurrences, and it is classified as low, moderate, high, and very high.

Mean of Models: The advantage of using the mean of models is that it improves risk assessment predictions and enhances the outputs of an individual model [55,56]. In the current study, the averaging of three different MLAs, e.g., RF, SVM, and GBDT, was applied to minimize the limitations of every single model and enhance the total accuracy of the fire susceptibility map.

Model Performance, Validation, and Accuracy Assessment: The performance of ML models are considered by their ability to detect burned areas accurately based on input maps and predict fire locations with the least errors. Therefore, the modeling results must be validated through several statistical analyses. In each modeling process, the sample points were accidentally divided into 70% training samples to build models, and the rest of 30% were utilized for testing models. Finally, modeling results must be validated and assessed through several statistical methods. Five metrics were used to validate the fire susceptibility models, e.g., Accuracy, Cohen's Kappa, Precision, the area under the curve (AUC), and Recall. Calculation of the five mentioned measures may be performed according to Eqs. (6) to (9).

$$\text{Accuracy} = \left(T_p + T_n\right) / \left(F_p + F_n\right) \quad \text{Eq. (6)}$$

$$\text{Cohen's Kappa} = \left(P_o - P_e\right) / \left(1 - P_e\right) \quad \text{Eq. (7)}$$

$$\text{Precision} = \frac{T_p}{T_p - F_p} \quad \text{Eq. (8)}$$

$$\text{Recall} = \frac{T_p}{T_p - F_n} \quad \text{Eq. (9)}$$

where T_p and T_n are true positive and true negative, respectively. F_p, F_n, P_o, and P_e are false positives, false negatives, relative agreements observed, and hypothetical probability of chance agreements, respectively.

Findings

Frequency of Fire Events in the Study Area: The frequency of fire occurrences was analyzed each month between 2001 and 2022 (Figure 4). An examination of temporal fire distribution demonstrates that most events occurred in 2013, 2016,

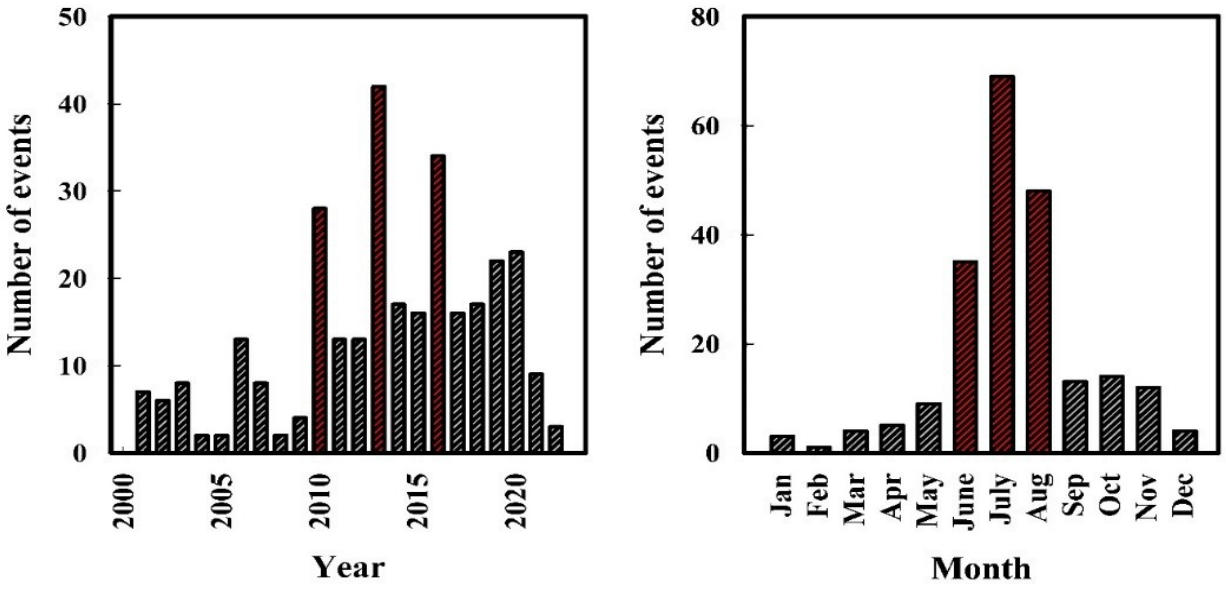

Figure 4) The temporal distribution of fire occurrences between the years 2001 and 2022 (left) and in each month (right).

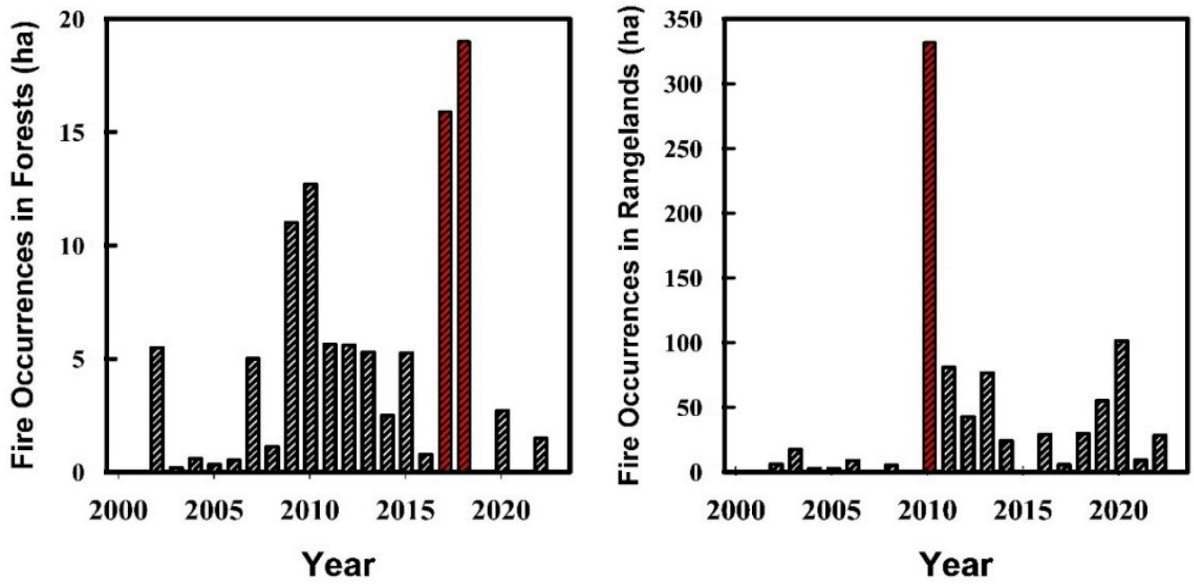

Figure 5) The affected area of fire occurrences in both classes of forests and rangelands (ha) from 2001 to 2022.

and 2010, respectively. Moreover, June, July, and August showed the highest rate of fire frequency (Figure 4). Because of summer's warm weather, annual vegetation cover does not grow anymore. Under this condition, the biomass is dry-off and becomes ignitable, and as a result, it supplies feasible fuel for fire occurrence.

Fire events, categorized by land-use and land-cover (LULC) classification, peaked in broadleaf forests in 2017 and 2018, with occurrences exceeding 15 hectares. In contrast, rangelands experienced significant fire activity only in 2010, when events surpassed 300 hectares in Semnan Province (Figure 5). Although grasslands, broadleaf forests, and sparse vegetation exhibited the highest rates of fire occurrences, herbaceous wetlands and shrublands recorded the lowest rates, each exceeding 5% (Figure 6). Between 2001 and 2022, over 25% of total fire events primarily occurred in grasslands, attributed to the presence of flammable surface fuels. Broadleaf forests, sparse vegetation, and juniper forests ranked second, third, and fourth in fire occurrences, comprising approximately 20%, 17%, and 13% of total events, respectively (Figure 6).

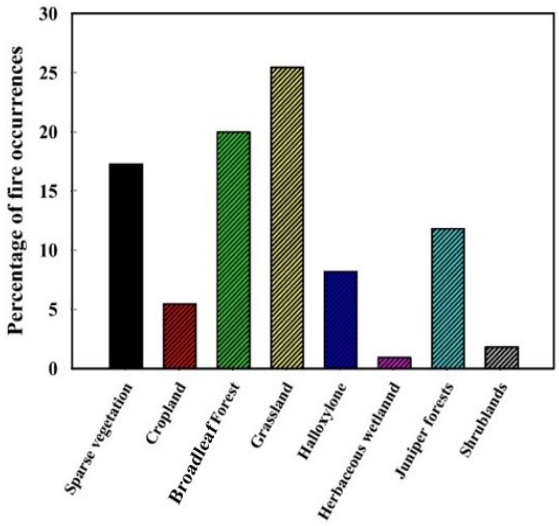

Figure 6) The spatial distribution of fire events in different Land-use/cover (LULC) classes from 2001 to 2022.

Evaluating Auxiliary Variables in Fire Frequency Occurrences: Annual rainfall in the study area ranges from approximately 70 to 460 mm .year^{-1}, categorized into eight distinct classes (Table 2). Results indicate that over half of the study area (55.15%) receives annual rainfall between 70 and 100 mm, primarily in the southern region

(Figure 7). The highest annual rainfall amounts are found in the northern areas, with more significant elevations. As elevation decreases towards the south, rainfall and NDVI values decline, particularly in the arid regions (Figure 7). In the northern part of the study area, annual rainfall ranges from 200 to 460 mm, characterizing a semiarid climate that accounts for approximately 6.54% of the total study area (Table 2, Figure 7). With its mountainous terrain, the northern edge supports broadleaf and juniper forests, where NDVI values exceed 0.6, and the mean annual temperature is recorded at a relatively low 8.66°C.

Table 2) Distribution of rainfall ranges and corresponding areapercentages.

Class Number	Rainfall Ranges (mm .year^{-1})	Area Percentage
1	70-100	55.15
2	100-150	29.26
3	150-200	9.06
4	200-250	2.21
5	250-300	2.23
6	300-350	0.75
7	350-400	1.02
8	400-460	0.33

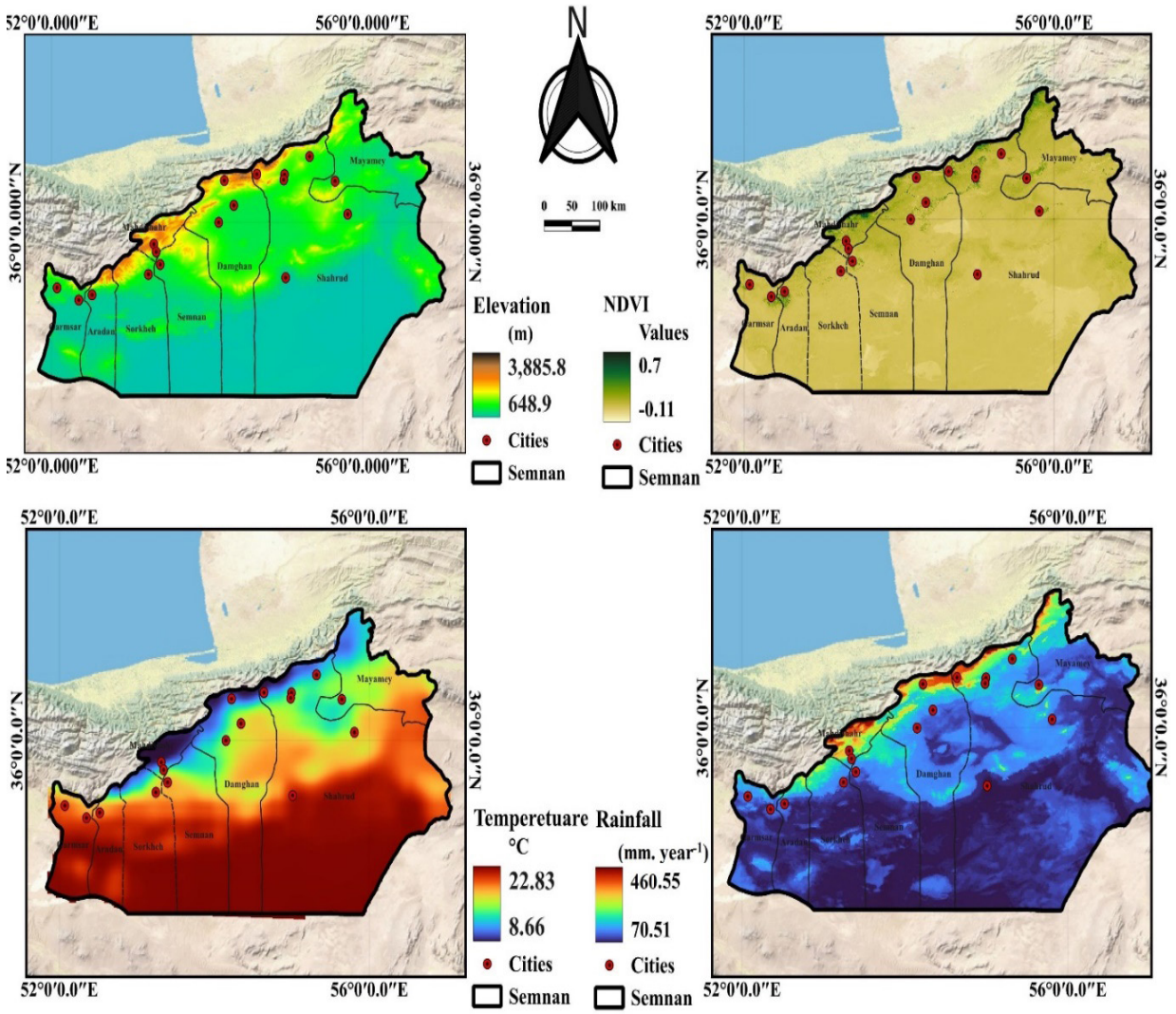

Figure 7) Auxiliary variables for fire susceptibility mapping in the study area.

Table 3 and Figure 8 present a frequency analysis of eight significant factors influencing fire events. The results indicate that half of the fire occurrences are in areas receiving more than 280 mm of rainfall annually. Additionally, 20% of the fires occurred in regions with rainfall between 330 and 380 mm .year^{-1}, primarily in the northern part of the study area, which is characterized by broadleaf forests, dense grasslands, and juniper forests. In this region, the annual temperature declines as elevation increases to between 2000 and 2500 meters, and vegetation cover becomes denser. Notably, the frequency of fire events has decreased with rising annual temperatures. In areas with rainfall between 70 and 100 mm .year^{-1}, Haloxylon

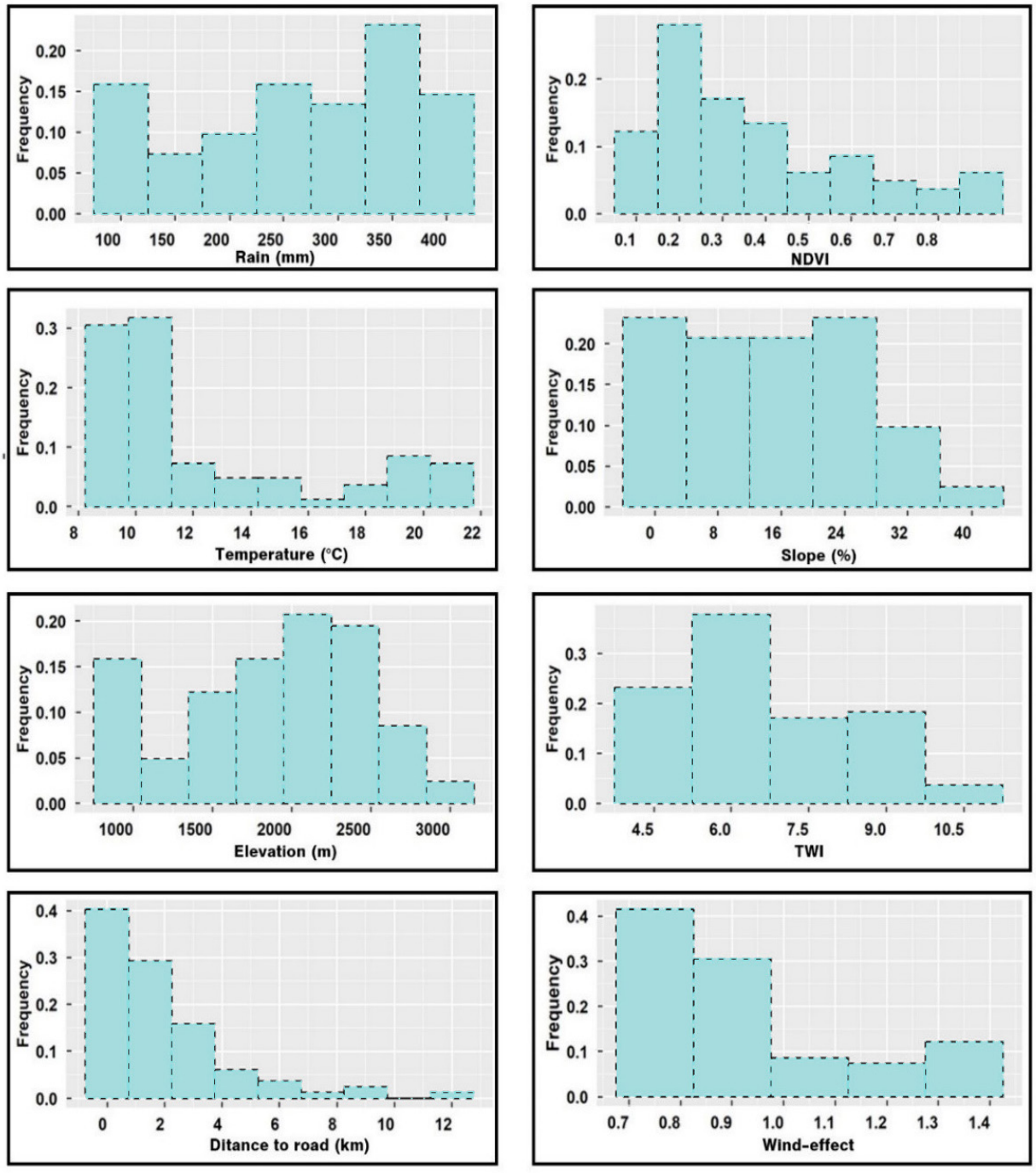

Figure 8) The frequency bar graphs of fire events in different input variables in the study

vegetation is present, accounting for over 15% of fire events (Figure 8). Furthermore, more than 60% of fire occurrences were recorded at low temperatures, ranging from 8°C to 11°C (Figure 8). The analysis of fire frequency revealed that over 50% of fires occurred within 1 km of roads (Table 3, Figure 9), and approximately 70% of fire events occurred within 2 km of roadways. Additionally, 75% of fires were recorded in areas with an NDVI greater than 0.19, predominantly covering grasslands and rangelands. In contrast, areas with NDVI values below 0.19, characterized by Haloxylon cover, accounted for 25% of fire occurrences.

Table 3) Frequency parameters of effective variables in fire occurrences area.

Variables	First Quartile	Median	Third Quartile
Rainfall (mm)	182.63	281.15	351.33
Temperature (°C)	9.01	10.65	14.2
Elevation (m)	1454.3	1973.8	2361.2
Distance to roads (km)	0.3	1.06	2.49
NDVI	0.19	0.29	0.51
Slope (%)	4.74	13.53	23.15
TWI	5.37	5.99	7.63
Wind-effect	0.80	0.84	0.99

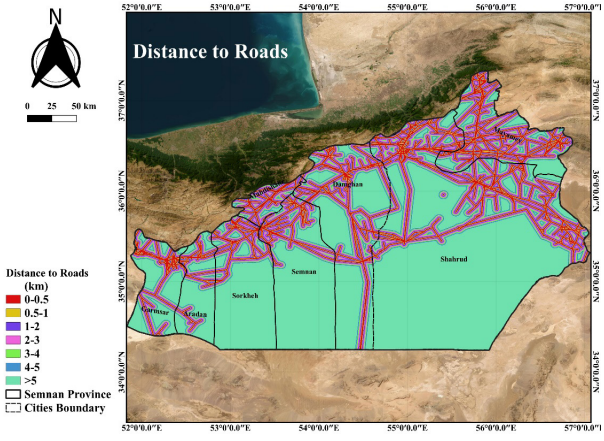

Figure 9) Distance to roads (km), the practical factor to assess fire susceptibility map.

The Importance of Auxiliary Variables: Figure 10 presents the correlation matrix among the auxiliary variables used in modeling fire susceptibility. The highest positive correlation was observed between annual rainfall and elevation (r = 0.81), indicating that as elevation increases, annual rain also tends to rise within the study area. Additionally, the correlation between annual rainfall and the NDVI was found to be 0.64, suggesting that increases in annual rainfall are associated with higher NDVI values. Conversely, the correlation between annual temperature and rainfall and elevation was calculated to be -0.89, demonstrating a strong inverse relationship with annual temperature.

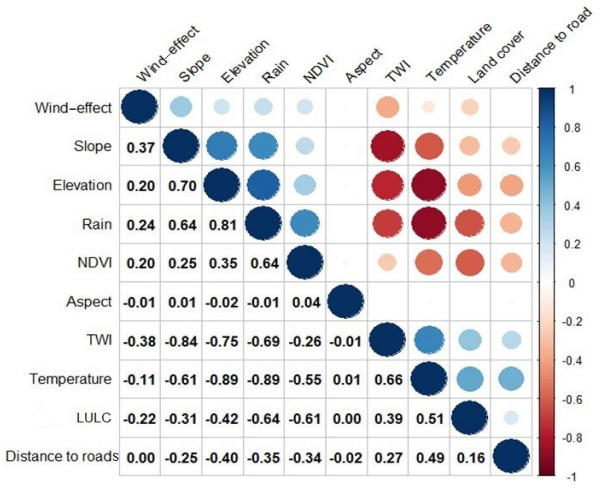

Figure 10) The matrix of correlation between variables.

Correlation Tests for Candidate Variables: In this study, the importance of auxiliary variables was detected using the Boruta machine-learning algorithm (Table 4, Figure 11). Annual temperature (mean importance: 17.59) was considered the decisive factor in fire events. The two other variables, NDVI and annual rainfall, were ranked second and third with a mean importance value of 16.60 and 13.69, respectively (Table 4, Figure 11). Moreover, the degree of elevation and distance to the road were counted as necessary, at 12.38 and 10.21, respectively. Generally, the variables, including temperature, NDVI, rain, elevation, and distance to the road, were ranked correspondingly as practical factors in fire susceptibility assessment based on the

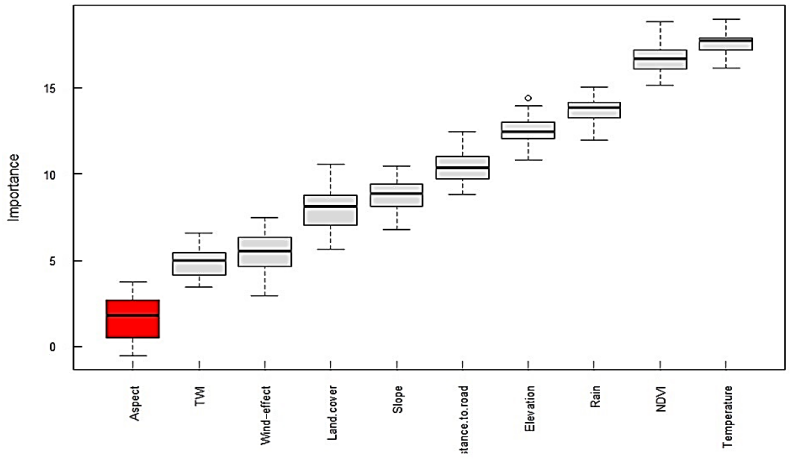

Figure 11) Boxplot for the importance of effective variables in fire susceptibility assessment using the Boruta algorithm.

Table 4) The Importance degree (Mean, Median, Minimum, and Maximum) for the effective variables in fire susceptibility assessment using the Boruta algorithm.

Variables	Mean	Median	Min	Max	Decision
Elevation	12.38	12.40	10.47	13.83	Confirmed
Slope	8.50	8.24	6.73	10.71	Confirmed
TWI	4.72	4.61	3.02	6.39	Confirmed
Wind effect	5.28	5.23	1.85	7.03	Confirmed
Distance to road	10.21	10.30	7.71	12.29	Confirmed
Land-cover	8.12	8.10	6.37	10.76	Confirmed
Rain	13.69	13.68	12.35	15.06	Confirmed
Temperature	17.59	17.58	16.43	19.22	Confirmed
NDVI	16.60	16.53	14.82	17.86	Confirmed
Aspect	1.71	1.67	-1.03	3.40	Rejected

Table 5) Multicollinearity tests of the effective variables.

Variable	Linear Regression Parameters				Collinearity Statistics	
	Estimate	S.E.	t value	sig	Tolerance (1/VIF)	VIF
Intercept	0.45	0.51	0.88	0.38	-	-
Elevation	0.00	0.00	-0.24	0.81	0.14	7.15
Slope	0.01	0.00	1.35	0.18	0.28	3.63
TWI	0.00	0.03	-0.04	0.97	0.22	4.59
Wind effect	-0.30	0.18	-1.70	0.09	0.73	1.36
Land-cover	0.01	0.00	4.10	0.00**	0.50	2.01
Rain	0.00	0.00	1.99	0.05*	0.12	8.00
Temperature	-0.03	0.01	-2.02	0.04*	0.10	7.25
NDVI	0.40	0.18	2.25	0.03*	0.38	2.62

** Significant in 99% confidence level * Significant in 95% confidence level

Boruta algorithm. At the same time, the aspect factor was rejected in the decision with a mean importance of 1.71 (Table 4, Figure 11).

The multicollinearity test was run to identify the practical factors in fire susceptibility assessment (Table 5). Generally, VIF can be explained using 10 as the critical value. There is no multicollinearity if VIF < 10, and for 10 ≤ VIF < 100, multicollinearity is high. When VIF ≥ 100, strong multicollinearity exists (57,58) and in prioritizing forest fuel treatments. In this paper, we chose easily obtained spatial variables pertaining to topography, vegetation types, meteorological conditions, climate, and human activity to predict forest fire ignition in Heilongjiang Province, China, using logistic regression. Results showed fire ignition prediction through logistic regression had good accuracy. Climatic variables (e.g., average annual mean temperature and precipitation. Results showed the VIF value for each variable counted below 10, meaning no collinearity was found between input variables. Consequently, all variables entered directly the importance test stage of the model (Table 5).

Validation Results: Machine learning models were developed to analyze the spatial relationship between contributing factors and fire occurrences. Model accuracy and goodness-of-fit were assessed using several evaluation metrics. All algorithms, including RF, SVM, GBDT, and the mean of models, performed well, achieving accuracy scores of 86%, 84%, 84%, and 84%, respectively. Precision, recall, and area under the ROC curve (AUC) values were also consistently high (Table 6, Figure 12). Notably, all models achieved AUC values above 0.89, indicating strong predictive ability. Furthermore, Kappa statistics exceeding 0.80 demonstrated strong agreement between observed and predicted fire occurrences.

The ROC curve (Figure 13) visually represents model performance. It plots the true positive rate (TPR) against the false positive rate (FPR). Curves positioned closer to the upper left corner indicate superior performance, reflecting higher TPR and lower FPR. In this study, all models exhibited good performance, with the RF and mean of models showing slightly better performance, as evidenced by their ROC curves being closer to the upper left corner.

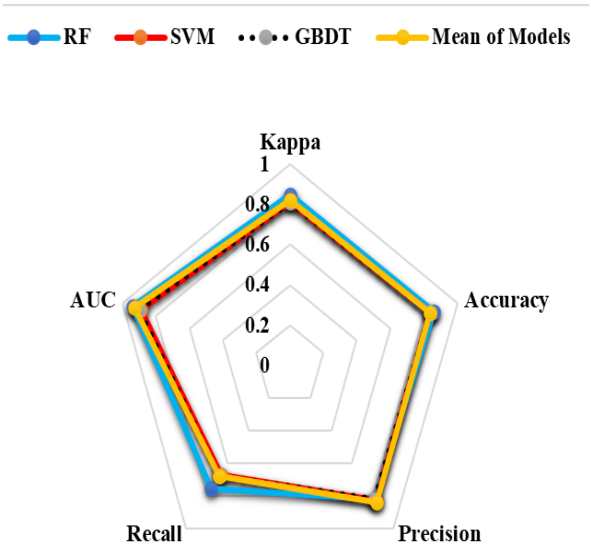

Figure 12) A comparison between the accuracy levels of three machine learning algorithms.

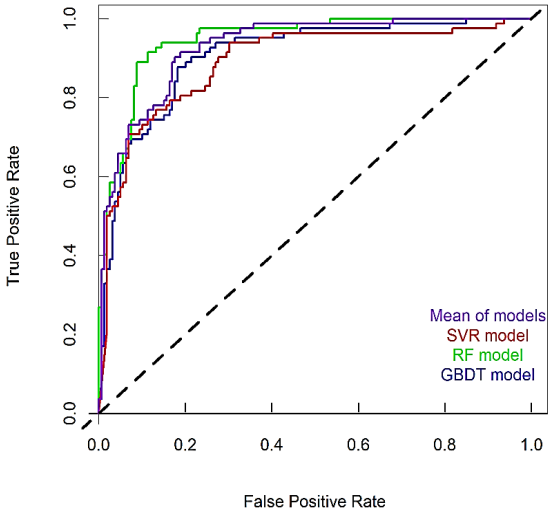

Figure 13) Validation of fir susceptibility maps applying ROC curves for GBDT, RF, SVR, and mean of models.

Fire Susceptibility Maps: Detecting the location of fire occurrences is fundamental for predicting the sensitivity to environmental change and ecological consequences. Fire susceptibility maps were created using three machine learning models, e.g., GBDT, SVM, and RF. The maps were classified into four classes: low, moderate, high, and very high classes, which demonstrate the area of concern (Figure 14). RF and mean models (average of three ML models) predicted 0.4 and 1.31% of the study area as a very highly susceptible class, which include areas about 38800 and 12600 km^2, respectively (Table 7). The other two models, GBDT and SVM, predicted 2.42 and 1.86% of the study area as very high-class, covering areas between 234700 and 180000 km^2. Although the lowest percentage of the total area is located in the very high class (Table 7, Figure 14), it covers relatively broad areas of the study area.

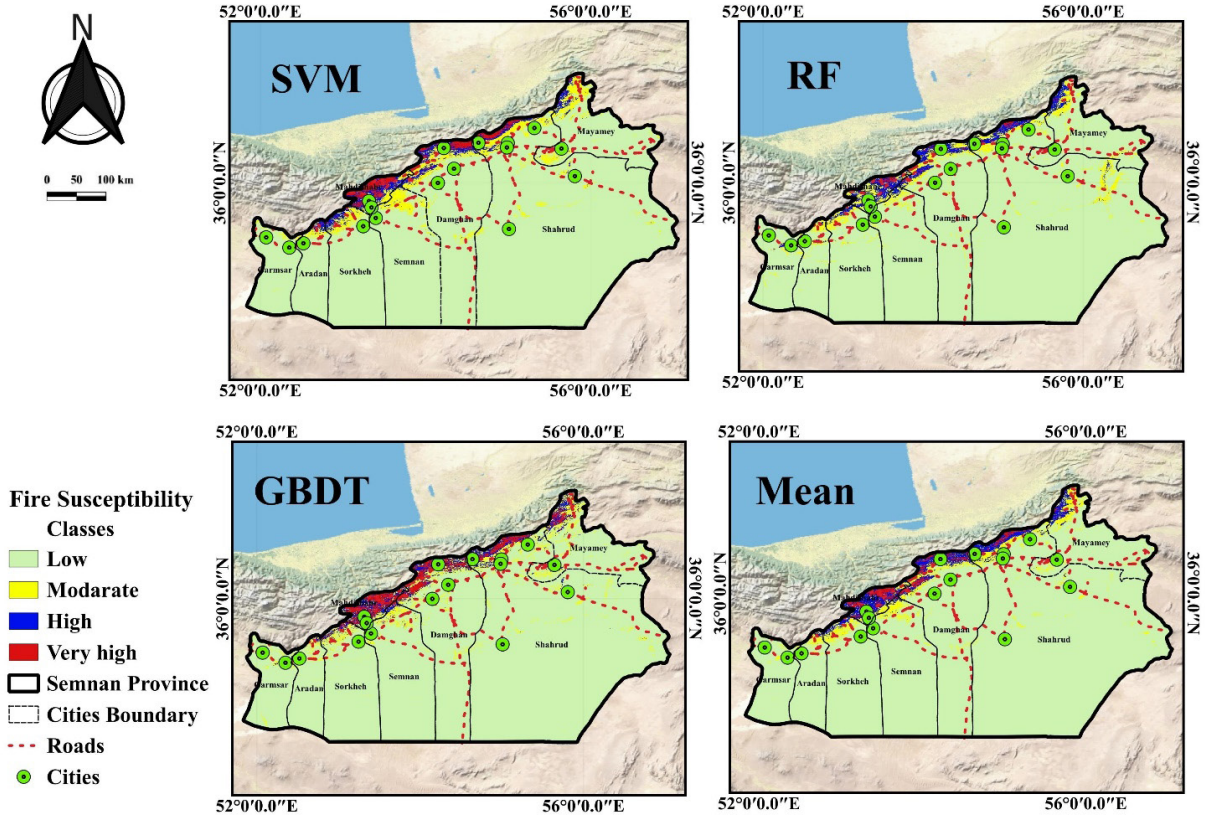

Figure 14) Prediction maps of fire susceptibility classes generated with RF, SVM, GBDT, and Mean of models in Semnan Province.

Table 6) Calculated measures for different machine learning models.

	Kappa	Accuracy	Precision	Recall	AUC
RF	0.85	0.86	0.82	0.76	0.94
SVM	0.81	0.84	0.83	0.67	0.89
GBDT	0.81	0.84	0.82	0.68	0.90
Mean of models	0.82	0.84	0.84	0.68	0.93

Table 7) Fire susceptibility classes in Semnan Province.

	ML Algorithms			
Danger Class	**GBDT**	**SVM**	**RF**	**Mean of Models**
Low	89.76	89.39	90.22	90.77
Moderate	5.13	6.18	6.17	4.69
High	2.68	2.58	3.20	3.33
Very high	2.42	1.86	0.40	1.31

Figure 14 illustrates the fire susceptibility maps generated using the individual machine learning models and the average of these models. Results indicated that three models predicted fire susceptibility mainly in the north and central parts of Semnan Province, especially in broadleaf forests, grasslands, and juniper forests. However, SVM predicted both moderate (in the north and central part) and very highly susceptible classes to a greater extent (in the north). The fire susceptible classified map indicated that most moderate, high, and very highly susceptible areas are located in the northern portion of the study area. Semnan's forest resources are primarily distributed in the north, and the southern parts do not have either forests or rangelands and fewer Haloxylon bushes.

Discussion

Our assessment had two components: first, fire frequency and occurrences were evaluated using fire occurrences data; second, the practical factors on fire susceptibility mapping were technically assessed in the study area. Results revealed that fire events surged in the summer warmest months, e.g., June, July, and August, recorded the highest fire frequency occurrences in the study area. In 2018 and 2017, the highest level of fire occurrences was recorded in broadleaf forests, with approximately 15 ha. Moreover, identifying 13% of juniper forests and 20% of broadleaf forests as sites of fire occurrences from 2001 to 2022 carries significant ecological implications at an annual scale. These vulnerabilities can lead to biodiversity loss and habitat degradation. Wildfires can destroy critical habitats, disrupt ecosystem dynamics, and reduce genetic diversity. Additionally, ongoing monitoring and adaptive management are essential to assess ecological conditions and adjust strategies based on observed outcomes.

The most significant fire events in 2010 took place in rangelands, covering an area of more than 300 ha, which may have been influenced by human activity in these regions. Temperature, rainfall, and distance to roads have been identified as key factors in fire susceptibility mapping. Consequently, climate change and human activities are expected to increase fire occurrences in the study area. Additionally, the rising annual temperatures reduce agricultural land in the southern part of Semnan Province. Meanwhile, intensified human activities, such as utilizing northern grasslands for agricultural practices, will likely exacerbate fire events in the northern region. While temperature and rainfall are essential predictors in our fire susceptibility model, they may not fully capture the impact of localized weather extremes like droughts and heat waves. However, these dynamic variables offer valuable insights for developing fire warning systems. For example, summer heatwaves and droughts often trigger fires in fire-prone areas. Therefore, integrating a fire susceptibility map with real-time monitoring of extreme weather events can enhance the development of dynamic fire warning systems.

About 75% of fires occurred with an NDVI value of 0.51, and more than 40% of fires were recorded with NDVI values between 0.2

and 0.4, where sparse vegetation, grasslands, rangelands, and Haloxylon cover are prevalent. Moving from the northern to the southern parts, there is a decrease in annual rainfall to 70 mm .year^{-1}, which is covered by Haloxylon, accounting for more than 15% of fire occurrences. These areas coincide with the traditional small ruminant grazing routes, which follow tracks and paths used by shepherds in a seasonal cycle, typically moving to lowlands in winter and highlands in summer. Shepherds often use Haloxylon bushes for ignition; as a human activity, they are the primary cause of fire events in these regions. In this study, vegetation is the primary fuel material influencing fire occurrences in Northern Semnan Province. Different vegetation types vary in combustibility, significantly impacting fire intensity and spread. The interplay between vegetation density, moisture content, and climatic conditions shapes fire dynamics in the region. Understanding these relationships is essential for assessing fire susceptibility, as denser and drier vegetation can exacerbate fire risk while well-managed areas may mitigate it. Approximately 50% of the observed fires occurred in areas with annual rainfall exceeding 280 mm. A study revealed that both meteorological and vegetation factors play crucial roles in influencing fire occurrences [24]. Additionally, other research has asserted that climate significantly influences short-term fire events [59]. Temporary above-average rainfall events are a typical driver of fire occurrences in arid regions [60], requiring fuel levels through the growth of annual plants, particularly grasses. More than 50% of fires occurred in the northern parts at elevations between 2000 and 2500 meters, covering approximately 1.5% of the total area, where the annual rainfall ranges between 350 and 450 mm. Also, the number of fire events in semiarid regions is similar to the results reported by [8]. Over half of the observed fire events occurred within 1 km of roads, and approximately 70% of events transpired within 0 to 2 km of roads more accessible to human activity. Other studies have supported this finding, indicating that the closer the road is to high-risk fire areas, the more likely human activity is associated with fire occurrences [61]. Research has demonstrated that areas with low elevation and short distances from roads exhibit a higher potential for fire danger [62–64]. Moreover, some research indicated a significant correlation between forest fire occurrences and meteorological and topographic factors, and vegetation type [61,65]. They found that the distance from roads has a negative impact on fire events, implying that the probability of forest fires decreases with an increase in distance from roads. Our findings align with research that noted elevation significantly influenced fire occurrence more than slope and aspect [24]. This study clearly showed the contributing factors in the fire susceptibility map, including the growth of vegetation cover as fire fuel ignition (0.2-0.4), relatively higher annual rainfall amounts (250-400 mm. year^{-1}), elevations ranging from 1000 to 2500 m, lower temperatures (approximately 8-11^{o}C), and an optimum distance to roads for human access (approximately 3 km). Results underscore the significant influence of climatic factors and human activity on the occurrence of fires. By incorporating these variables, our analysis provides a more nuanced perspective on how anthropogenic activities interact with ecological factors to influence fire risk. Fire occurrences stem from a variety of causes and have the potential to cause substantial damage. For this reason, incorporating a broader range of predictors and MLAs is vital for enhancing the accuracy of fire susceptibility assessments and mitigating the impact of wildfires. To protect

the forests and rangelands as one of the significant natural resources [66,67] in arid and semiarid regions, it is critical to implement measures to control such fires using machine learning techniques. The validation results indicated that the RF model demonstrated a 2% higher prediction accuracy than the SVM, GBDT, and mean models, all of which had an accuracy of 0.84. Other studies showed the superior performance of the RF model in modeling wildfire susceptibility in different regions [14,68,69]. The findings of this research underscore the necessity of considering both natural and human factors in predictive modeling, as this holistic approach can lead to more effective fire management strategies.

Conclusion

Since arid and semiarid regions are highly vulnerable ecosystems, any disruption can significantly impact their stability. In recent years, integrating spatiotemporal ground data with MLAs methods has become essential for fire monitoring, developing effective fire susceptibility maps, and implementing fire control strategies. This study highlights the importance of integrating natural ecological factors, human-related influences, and historical fire occurrence data in understanding fire susceptibility. It also focuses on identifying the key factors for predicting fire occurrences in Semnan Province while generating fire susceptibility maps using three different MLAs: Random Forest (RF), Support Vector Machine (SVM), and Gradient Boosting Decision Trees (GBDT), and the mean of these models. The classification of fire susceptibility indicated that a significant portion of the northern regions of the study area are at high and very high risk for fire occurrences. Notably, most Juniper forests, broadleaf forests, and grasslands in these areas are particularly prone to fires that negatively affect vegetation diversity and cause ecological disturbances. The hypothesis that distance to roads serves as an essential human factor influencing fire occurrence was supported in this study, highlighting its significance in fire susceptibility. Applying machine learning algorithms effectively generated fire susceptibility maps, enhancing our understanding of fire dynamics in the area. Resilience to fire events must be prioritized in arid and semiarid regions, necessitating enhanced security measures to prevent fire spread and implement resilience strategies to mitigate the adverse impacts. Further studies are needed to understand the mechanisms involved in the increasing frequency of fires, the recovery of vegetation cover, and the ecological balance post-fire. This research is distinctive for employing a combination of MLAs and the mean of models to enhance our understanding of fire patterns. The author acknowledges the limited research on predicting fire susceptibility using mean models incorporating uncertainty analyses, highlighting the need for focused scientific and emergency management strategies in northern Semnan. Future research should consider additional influencing factors, such as soil moisture data and socioeconomic variables, to enhance regional fire susceptibility assessments. Advancing our understanding of these factors will contribute to developing effective fire planning and management policies. To effectively manage fire risks in arid and semiarid regions like Semnan, it is crucial to establish enhanced monitoring systems that integrate remote sensing and machine learning for real-time tracking of ecological changes. Prioritizing fire management in high-risk areas, particularly juniper and broadleaf forests, while engaging local communities is essential. However, limitations in data quality and the study's specific focus on Semnan may affect broader applicability. Additionally, uncertainties

related to climate variability and human activities underscore the need for adaptive fire susceptibility assessments. Addressing these factors will strengthen resilience and inform effective fire management strategies, while continuously refining fire susceptibility maps using machine learning algorithms will improve predictive accuracy. This study modeled fire susceptibility using environmental variables such as climate data, terrain features, NDVI, and road proximity. However, anthropogenic factors like land-use changes and socioeconomic conditions can also significantly influence fire occurrence. Therefore, further research should analyze fire patterns in Semnan Province, incorporating these human factors. Semnan Province, situated in an arid and semiarid region, has a lower density of urban areas and villages than the Hyrcanian ecosystems located north of Semnan. A comparative analysis of fire data from semiarid and arid regions (e.g., Semnan Province) with data from the more humid Hyrcanian ecosystems could provide valuable insights into how wildfire patterns vary across diverse climatic regions.

Funding

The Department of Crisis Management of Semnan Governorate, [50.4.51254], Semnan Province, supported this work.

Conflict of Interests

The authors have no relevant financial or non-financial interests to disclose.

Authors Contributions

Ali Asghar Zolfaghari and Maryam Raeesi: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing, and Original draft preparation, Visualization, Investigation, Validation, Writing-Reviewing and Editing the final draft, **Azadeh Soltani:** Data collection. **Zahra Sheikh, Soghra Poodineh, Azadeh Soltani, and Mojtaba Amiri:** Reviewing original draft.

Acknowledgments

We would also like to thank the Department of Crisis Management of Semnan (CMS) Governorate and the individuals who provided financial support for this research. This work was supported by the Department of CMS Governorate, Semnan Province [50.4.51254]. Their contributions were critical to the success of this research, and we are deeply grateful for their dedication.

References

1. Marlon J., Bartlein P., Daniau A.L., Harrison S., Maezumi S., Power M. Global biomass burning: A synthesis and review of Holocene paleofire records and their controls. Quat. Sci. Rev. 2013;65(1):5–25.
2. You C., Yao T., Xu C. Environmental significance of levoglucosan records in a central Tibetan ice core. Sci. Bull. 2019; 64(2):122–127.
3. Guo F., Su Z., Wang G., Sun L., Tigabu M., Yang X. Understanding fire drivers and relative impacts in different Chinese forest ecosystems. Sci. Total Environ. 2017; 605(1):411–425.
4. Sevinc V., Kucuk O., Goltas M.A. Bayesian network model for prediction and analysis of possible forest fire causes. For. Ecol. Manag. 2020; 457:117723.
5. Su Z., Zheng L., Luo S., Tigabu M., Guo F. Modeling wildfire drivers in Chinese tropical forest ecosystems using global logistic regression and geographically weighted logistic regression. Nat. Hazard. 2021; 108(1):1317–1345.
6. Bowman D.M.J.S., Moreira-Muñoz A., Kolden C.A., Chávez R.O., Muñoz A.A., Salinas F. Human–environmental drivers and impacts of the globally extreme 2017 Chilean fires. Ambio. 2019; 48(4):350–362.
7. Forkel M., Dorigo W., Lasslop G., Teubner I., Chuvieco E., Thonicke K. Identifying required model structures to predict global fire activity from satellite and climate data. Geosci. Model Dev. Discuss. 2016; 1(1):1–35.
8. Eskandari S., Pourghasemi H.R., Tiefenbacher J.P. Relations of land-cover, topography, and climate to fire occurrence in natural regions of Iran: Applying new data mining techniques for modeling and mapping fire danger. For. Ecol. Manag. 2020; 473:118338.
9. Mallinis G., Mitsopoulos I., Chrysafi I. Evaluating and comparing Sentinel 2A and Landsat-8 Operational Land Imager (OLI) spectral indices for estimating fire severity in a Mediterranean pine

ecosystem of Greece. GISci. Remote Sens. 2018; 55(1):1–18.
10. Santos A.C. dos., Montenegro S. da. R., Ferreira M.C., Barradas A.C.S., Schmidt I.B. Managing fires in a changing world: Fuel and weather determine fire behavior and safety in the neotropical savannas. J. Environ. Manag. 2021; 289:112508.
11. Arabameri A., Pal. S., Costache R.D., Saha A., Rezaie F., Danesh A.l. Prediction of gully erosion susceptibility mapping using novel ensemble machine learning algorithms. Geomatic. Natur. Hazard. Risk. 2021; 12(1):469–498.
12. Ahmed I.A., Talukdar S., Shahfahad Parvez A., Rihan Mohd., Baig M.R.I. Flood susceptibility modeling in the urban watershed of Guwahati using improved metaheuristic-based ensemble machine learning algorithms. Geocarto Int. 2022; 37(26):12238–12266.
13. Akıncı H.A., Akıncı H. Machine learning based forest fire susceptibility assessment of Manavgat district (Antalya), Turkey. Earth Sci. Inform. 2023;16(1):397–414.
14. Alkan Akinci H., Akinci H., Zeybek M. Comparison of diverse machine learning algorithms for forest fire susceptibility mapping in Antalya, Türkiye. Adv. Space Res. 2024; 74(2):647–667.
15. Novo A., Dutal H., Eskandari S. Fire susceptibility modeling and mapping in Mediterranean forests of Turkey: a comprehensive study based on fuel, climatic, topographic, and anthropogenic factors. Euro-Mediterr J. Environ. Integr. 2024; 9(2):655–679.
16. Tonini M., D'Andrea M., Biondi G., Degli Esposti S., Trucchia A., Fiorucci P. A Machine Learning-Based Approach for Wildfire Susceptibility Mapping. The Case Study of the Liguria Region in Italy. Geosciences. 2020; 10(3):105.
17. Liu Z., Peng C., Timothy W., Candau J.N., Desrochers A., Kneeshaw D. Application of machine-learning methods in forest ecology: Recent progress and future challenges. Environ. Rev. 2018; 26(4): 339-350.
18. Jahanbani M., Vahidnia M.H., Aghamohammadi H., Azizi Z. Flood susceptibility mapping through geoinformatics and ensemble learning methods, with an emphasis on the AdaBoost-Decision Tree algorithm, in Mazandaran, Iran. Earth Sci. Inform. 2024; 17(2):1433–1457.
19. Janizadeh S., Chandra Pal S., Saha A., Chowdhuri I., Ahmadi K., Mirzaei S. Mapping the spatial and temporal variability of flood hazard affected by climate and land-use changes in the future. J. Environ. Manag. 2021; 298:113551.
20. Mirzaei S., Vafakhah M., Pradhan B., Alavi S.J. Flood susceptibility assessment using extreme gradient boosting (EGB), Iran. Earth Sci. Inform. 2021; 14(1):51–67.
21. Yang D., Zhang T., Arabameri A., Santosh M., Saha U.D., Islam A. Flash-flood susceptibility mapping: a novel credal decision tree-based ensemble approach. Earth Sci. Inform. 2023; 16(4):3143–3161.
22. Pourghasemi H.R., Gayen A., Edalat M., Zarafshar M., Tiefenbacher J.P. Is multi-hazard mapping effective in assessing natural hazards and integrated watershed management? Geosci. Front. 2020; 11(4):1203–1217.
23. Barreto J., Armenteras D. Open Data and Machine Learning to Model the Occurrence of Fire in the Ecoregion of "Llanos Colombo-Venezolanos." Remote Sens. 2020; 12(23):3921.
24. Tan C., Feng Z. Mapping Forest Fire Risk Zones Using Machine Learning Algorithms in Hunan Province, China. Sustainability. 2023; 15(7):6292.
25. Chevitarese D.S., Szwarcman D., Silva R.G., Brazil E.V. Deep learning applied to seismic facies classification: A methodology for training. In: Saint Petersburg 2018. Europ. Associat.Geoscientist. Engineer. 2018;2018(1):1–5.
26. Hu Z.X., Wang Y., Ge M.F., Liu J. Data-driven fault diagnosis method based on compressed sensing and improved multiscale network. IEEE Trans Ind Electron. 2019; 67(4):3216–3225.
27. Qu S., Guan Z., Verschuur D., Chen Y. Automatic high-resolution microseismic event detection via supervised machine learning. Geophysic. J. Int. 2020; 222(3):1881–1895.
28. Smith R., Mukerji T., Lupo T. Correlating geologic and seismic data with unconventional resource production curves using machine learning. Geophys. 2019; 84(2):O39–O47.
29. Tse K.C., Chiu H.C., Tsang M.Y., Li Y., Lam E.Y. Unsupervised learning on scientific ocean drilling datasets from the South China Sea. Front. Earth Sci. 2019; 13(1):180–190.
30. Zhang G., Wang Z., Chen Y. Deep learning for seismic lithology prediction. Geophys. J. Int. 2018; 215(2):1368–1387.
31. Zhou K.B., Zhang Z.X., Liu J., Hu Z.X., Duan X.K., Xu Q. Anode effect prediction based on a singular value thresholding and extreme gradient boosting approach. Meas. Sci. Technol. 2019; 30(1):015104.
32. Barmpoutis P., Papaioannou P., Dimitropoulos K., Grammalidis N. A Review on Early Forest Fire Detection Systems Using Optical Remote Sensing. Sen. 2020; 20(22):6442.
33. Eskandari S., Chuvieco E. Fire danger assessment in Iran based on geospatial information. Int. J. Appl. Earth Observ. Geoinform. 2015;1(42):57–64.
34. Sadeghi A., Ahmadi Nadoushan M., Ahmadi Sani N. Segment-level modeling of wildfire susceptibility in Iranian semiarid oak forests: Unveiling

the pivotal impact of human activities. Trees For. People. 2024; 15:100496.
35. Amiri M., Pourghasemi H.R. Chapter 25 - Predicting areas affected by forest fire based on a machine learning algorithm. In: Pourghasemi HR, editor. Computers in Earth and Environmental Sciences. Elsevier. 2022;1(1): 351–362.
36. Mayr M.J., Vanselow K.A., Samimi C. Fire regimes at the arid fringe: A 16-year remote sensing perspective (2000–2016) on the controls of fire activity in Namibia from spatial predictive models. Ecol. Indicat. 2018; 91(1):324–337.
37. Mandal P., Maiti A., Paul S., Bhattacharya S., Paul S. Mapping the multi-hazards risk index for coastal block of Sundarban, India using AHP and machine learning algorithms. Trop. Cyclone Res. Rev. 2022; 11(4):225–243.
38. Raeesi M., Zolfaghari A., Rahimi M., Kaboli S.H. Estimation of Vegetation Changes Concerning Annual Rainfall and Temperature in Semnan Province. E.E.R. 2023; 13(4):56–82.
39. Amini E., Zolfaghari A., Kaboli H., Rahimi M. Estimation of Rainfall Erosivity Map in Areas with Limited Number of Rainfall Stations (Case study: Semnan Province). Iran. J. Soil Water Res. 2022; 53(9):2027–2044.
40. Wu Z., Li M., Wang B., Quan Y., Liu J. Using artificial intelligence to estimate the probability of forest fires in Heilongjiang, northeast China. Remote Sens. 2021; 13(9):1813.
41. Kopecký M., Macek M., Wild J. Topographic Wetness Index calculation guidelines based on measured soil moisture and plant species composition. Sci. Total Environ. 2021; 757:143785.
42. Qin Z., Zhu Y., Li W., Xu B. Mapping vegetation cover of grassland ecosystem for desertification monitoring in Hulun Buir of Inner Mongolia, China. Proc. SPIE. 2008; 7104.
43. Xu Ch., Li Y., Hu J., Yang X., Sheng S., Liu M. Evaluating the difference between the normalized difference vegetation index and net primary productivity as the indicators of vegetation vigor assessment at landscape scale. Environ. Monit. Assess. 2012; 184(3):1275–1286.
44. Kursa M.B., Rudnicki W.R. Feature Selection with the Boruta Package. J. Stat. Softw. 2010; 36(11):1–13.
45. Szul T., Tabor S., Pancerz K. Application of the BORUTA Algorithm to Input Data Selection for a Model Based on Rough Set Theory (RST) to Prediction Energy Consumption for Building Heating. Energies. 2021; 14(10):2779.
46. Pourghasemi H.R., Gayen A., Lasaponara R., Tiefenbacher J.P. Application of learning vector quantization and different machine learning techniques to assessing forest fire influence factors and spatial modelling. Environ. Res. 2020; 184:109321.
47. Breiman L. Random Forest. Mach Learn. 2001; 45(1):5–32.
48. Jain P., Coogan S.C.P., Subramanian S.G., Crowley M., Taylor S., Flannigan M.D. A review of machine learning applications in wildfire science and management. Environ. Reviews. 2020; 28(4):478–505.
49. Hawryło P., Bednarz B., Wężyk P., Szostak M. Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. Eur. J. Remote Sens. 2018; 51(1):194–204.
50. Mustafa A., Rienow A., Saadi I., Cools M., Teller J. Comparing support vector machines with logistic regression for calibrating cellular automata land-use change models. Eur. J. Remote Sens. 2018; 51(1): 391–401.
51. Lu J., Lu D., Zhang X., Bi Y., Cheng K., Zheng M. Estimation of elimination half-lives of organic chemicals in humans using gradient boosting machine. BBA. General Subjects. 2016; 1860 (11, Part B): 2664–2671.
52. Touzani S., Granderson J., Fernandes S. Gradient boosting machine for modeling the energy consumption of commercial buildings. Energy Build. 2018; 158(1): 1533–1543.
53. Friedman J.H. Greedy function approximation: a gradient boosting machine. Ann. Statist. 2001; 1189–1232.
54. Tian Z., Xiao J., Feng H., Wei Y. Credit Risk Assessment based on Gradient Boosting Decision Tree. Procedia Comput. Sci. 2020; 174(1): 150–160.
55. Mosavi A., Golshan M., Janizadeh S., Choubin B., Melesse A., Dineva A. Ensemble models of GLM, FDA, MARS, and RF for flood and erosion susceptibility mapping: a priority assessment of sub-basins. Geocarto Int. 2020; 37(9):2541-2560.
56. Tien Bui D., Hoang N.D., Pham T.D., Ngo P.T.T., Hoa P.V., Minh N.Q. A new intelligence approach based on GIS-based Multivariate Adaptive Regression Splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area. J. Hydrol. 2019; 575(1):314–326.
57. Chang Y., Zhu Z., Bu R., Chen H., Feng Y., Li Y. Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China. Landscape Ecol. 2013; 28(10):1989–2004.
58. Guo F., Zhang L., Jin S., Tigabu M., Su Z., Wang W. Modeling Anthropogenic Fire Occurrence in the Boreal Forest of China Using Logistic Regression and Random Forests. Forests. 2016; 7(11):250.
59. McLauchlan K.K., Higuera P.E., Miesel J., Rogers B.M., Schweitzer J., Shuman J.K. Fire as a fundamental ecological process: Research advances and frontiers. J. Ecol. 2020; 108(5):2047–2069.
60. Van Etten E., Burrows N. On the Ecology of Aus-

tralia's Arid Zone: 'Fire Regimes and Ecology of Arid Australia.' In: On the Ecology of Australia's Arid Zone. Springer, Cham. 2018; 1(1):243–282.
61. Li W., Xu Q., Yi J., Liu J. Predictive model of spatial scale of forest fire driving factors: a case study of Yunnan Province, China. Sci. Rep. 2022; 12(1):19029.
62. Eskandari S., Oladi J., Jalilvand H., Saradjian M.R. Prediction of Future Forest Fires using the MCDM Method. Pol. J. Environ. Stud. 2015; 30 (24):2309–2314.
63. Pourghasemi H.R. GIS-based forest fire susceptibility mapping in Iran: a comparison between evidential belief function and binary logistic regression models. Scandinavian J. For. Res. 2016; 31(1):80–98.
64. Pourtaghi Z.S., Pourghasemi H.R., Aretano R., Semeraro T. Investigation of general indicators influencing on forest fire and its susceptibility modeling using different data mining techniques. Ecol. Indic. 2016; 64(1):72–84.
65. Romero-Calcerrada R., Novillo C.J., Millington J.D.A., Gomez-Jimenez I. GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). Landsc. Ecol. 2008; 23(3):341–354.
66. Faraji F., Alijanpour A., Sheidai Karkaj E., Motamedi J. The Consequences of Banqueting and Fire on Plant Functional Groups (Case Study: Atbatan Rangelands, Bostanabad County). ECOPERSIA 2020; 8(4):191–198.
67. Akhzari D., Mohammadi E., Saedi K. Studying the effect of fire on some vegetation and soil properties in a semiarid shrubland (Case study: Kachaleh Rangelands, Kamyaran Region). ECOPERSIA 2022; 10(1):27–35.
68. Barreto J.S., Armenteras D. Open Data and Machine Learning to Model the Occurrence of Fire in the Ecoregion of "Llanos Colombo–Venezolanos." Remote Sens. 2020; 12(23): 3921.
69. Iban M.C., Sekertekin A. Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin Provinces, Turkey. Ecol Inform. 2022; 69:101647.