

Research Article

Wildfire Susceptibility Mapping Using Five Boosting Machine Learning Algorithms: The Case Study of the Mediterranean Region of Turkey

Sohaib K. M. Abujayyab ^{1,2}, Moustafa Moufid Kassem ³, Ashfak Ahmad Khan,²
Raniyah Wazirali,⁴ Mücahit Coşkun,² Enes Taşoğlu ², Ahmet Öztürk,²
and Ferhat Toprak ²

¹Fire Safety Engineering, International College for Engineering and Management, Muscat 112, Oman

²Department of Geography, Faculty of Letters, Demir Celik Campus, Karabuk University, Karabuk 78050, Turkey

³School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, Penang 14300, Malaysia

⁴College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia

Correspondence should be addressed to Sohaib K. M. Abujayyab; sohaib@icem.edu.om

Received 22 August 2022; Revised 23 November 2022; Accepted 5 December 2022; Published 28 December 2022

Academic Editor: Zaher Mundher Yaseen

Copyright © 2022 Sohaib K. M. Abujayyab et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Forest fires caused by different environmental and human factors are responsible for the extensive destruction of natural and economic resources. Modern machine learning techniques have become popular in developing very accurate and precise susceptibility maps of various natural disasters to help reduce the occurrence of such calamities. The present study has applied and tested multiple algorithms to map the areas susceptible to wildfire in the Mediterranean Region of Turkey. Besides, the performance of XGBoost, CatBoost, Gradient Boost, AdaBoost, and LightGBM methods for wildfire susceptibility mapping is also examined. The results have revealed the higher testing accuracy of CatBoost (95.47%) algorithm, followed by LightGBM (94.70%), XGBoost (88.8%), AdaBoost (86.0%), and GBM (84.48%) algorithms. Resultant wildfire susceptibility maps provide proper inventories for forest engineers, planners, and local governments for future policies regarding disaster management in Turkey.

1. Introduction

Forest fires are critical natural disasters with severe ecological, economic, and social consequences worldwide [1, 2]. In recent years, a significant increase in the frequency of wildfire incidents and the extent of affected areas has been observed, indicating the sternness of the problem. According to the European Forest Fire Information System (EFFIS), an area of 831.46 km² in Turkey was affected by wildfires in 2019, nearly double that of 2018, while the figure reached 998.57 km² in 2020 [3]. These statistics show that Turkey's affected areas due to forest fires are increasing with time. Research reveals that anthropogenic factors [4, 5] and climate change [6–8] play a critical role in the frequency of wildfire occurrence and increase in affected areas. Wildfires

are not only responsible for the mass destruction of forests but they also have several adverse effects on the natural environment, such as increased erosion risk [9–12], poor water quality [9], changes in land use [13], and elimination of wildlife [14]. Nevertheless, many Mediterranean plant species have evolved some form of adaptation mechanism for survival in a fire [15]. For instance, in the Mediterranean Region of Turkey, plants like *Pinus brutia* have become resistant to fire regimes [16].

The findings of previous literature disclose that multiple environmental and human factors may be responsible for triggering forest fires in the Mediterranean Region [17, 18]. Different related studies have used data from various environmental parameters like elevation, aspect, slope, vegetation, temperature, humidity, and wind, along with human

parameters like distance to roads, distance to settlement, and population, to figure out the reasons for wildfires [19–22]. A review of related literature shows no consistently used model to evaluate fire risk analysis. Weights and variables in modelling indices may differ for different regions as wildfires have specific characteristics worldwide [23, 24]. According to the relevant literature, many other models have been used in forest fire risk analysis [25, 26]. In various studies, a Geographic Information System (GIS) program is frequently used to process large datasets and produce useful fire susceptibility maps [27–29]. In addition, the use of satellite images in a GIS-based program for forest fire risk analysis is empirically tested to produce more robust results [30, 31].

Statistical methods such as bivariate and multivariate analysis [32, 33], multiple linear regression [34, 35], and logistic regression [36–39] have been widely used for forest fire modelling. In recent years, machine learning algorithms in forest fire risk analysis have also gained popularity [40–44]. Rodrigues and De La Riva [45] developed random forest (RF), boosting regression trees (BRT), and support vector machine (SVM) algorithms in their study in the region covering almost the entire Spanish peninsula. It is concluded that the RF achieves the highest performance. Nelson et al. [46] compared CART, BRT, and RF in a study conducted in British Columbia, Canada, and found that the best performing model was BRT, followed by CART and RF. Analysis of related studies reveals that selecting an appropriate model for forest fire risk mapping is challenging as each algorithm's results vary from region to region [47].

Machine learning techniques have been applied and tested extensively in many empirical studies for developing susceptibility maps and accurate predictions of various natural calamities [48]. For example, Ma et al. [49] used an extreme gradient boosting (XGBoost) method for flash flood risk assessment in Yunnan Province of southwest China. The XGBoost method successfully identified the relationships between selected factors and flash flood events and outperformed the comparative LSSVM_RBF models. A study on landslide susceptibility mapping in Ulus district of Bartın Province in the Western Black Sea Region of Turkey [50] compared four new gradient boosting algorithms named gradient boosting machine (GBM), categorical boosting (CatBoost), extreme gradient boosting (XGBoost), and light gradient boosting machine (LightGBM). The accuracy results revealed the highest predictive capacity of the CatBoost model.

On the contrary, the RF method was found to have the lowest predictive ability compared to ensemble methods, Wu et al. [51] created a map of landslide susceptibility using an alternative decision tree (ADTree) in Longxian County (Shaanxi Province, China). Additionally, they used new GIS-based ensemble techniques, including ADTree with bootstrapping (Bagging), adaptive boosting (AdaBoost), and ADTree. The outcome showed that the ADTree-AdaBoost model had the best results. For the Three Gorges Reservoir area in China, Chen et al. [52] created a landslide susceptibility map using three advanced machine learning methods of gradient boosting decision tree (GBDT), random forest

(RF), and information value (InV) models. Among these compared models, the GBDT method showed the highest accuracy. Can et al. [53] used the extreme gradient boosting (XGBoost) method for landslide susceptibility mapping of the Atatürk Dam upper basin in Turkey. The performance of the XGBoost algorithm was found to be high in various metrics. In the Wanzhou section of the Three Gorges Reservoir area (China), a landslide susceptibility map was developed using the weighted gradient boosting decision tree (weighted GBDT) model by [54]. The logistic regression (LR) model and gradient boosting decision tree (GBDT) model were also used for comparison in the study. The results showed that the weighted GBDT model had the highest accuracy, followed by the GBDT and LR models. However, the weighted GBDT and GBDT models produced very similar results.

Dang et al. [55] used AdaBoost, XGBoost, RF, and multilayer perceptron (MLP) machine learning algorithms to identify commercial buildings at high fire risk in the Humberside area, UK. The results revealed that AdaBoost's performance was better than other algorithms. In the Yunnan Province of China, Zhou et al. [56] applied the CatBoost algorithm, for the risk estimation of forest fires. The analysis was made using five forest fire risk factors, and it was seen that the model and the actual fire points overlapped. Rosadi and Andriyani [57] compared the AdaBoost algorithm with decision tree and SVM method to predict the occurrence of forest fires. The study explained that fuzzy c-means clustering and AdaBoost methods provided good results in predicting forest fires. Michael et al. [58] used two satellite-derived measurements (NDVIW and NDVIT) in three machine algorithms (LR, RF, and XGBoost) to improve fire risk mapping. The research conducted in a region of Greece determined that XGBoost model produced the best results.

The review of recent studies reveals the application and performance of various algorithms to develop susceptibility maps of different natural disasters. However, most of these machine-learning methods are used in landslide susceptibility mapping. A literature gap is observed as very few studies have used and applied machine learning techniques to produce wildfire susceptibility maps. In addition, this research work is in the interest of and on the side of the United Nations' sustainable development 2030 agenda. The present study aims to apply and test various algorithms to map the areas susceptible to wildfire in the Mediterranean Region of Turkey. The formulation of susceptibility maps is essential to locate the areas prone to wildfire and take the necessary measures to avoid any mishaps in the future. The main contribution of this article is to evaluate the performance of XGBoost, CatBoost, Gradient Boost, AdaBoost, and LightGBM methods for wildfire susceptibility mapping. To our knowledge of previous literature, none of the earlier studies investigated these models' performance for wildfire susceptibility mapping. Therefore, finding the best performance model for wildfire susceptibility maps will help improve policymaking in the future to reduce the risks of forest fires.

2. Material and Methods

2.1. Selected Study Area. The study area includes five provinces of the Mediterranean Region of Turkey that were declared “disaster areas” by the Disaster and Emergency Management Presidency (AFAD) on July 31, 2021, during the recent wildfire incident in Turkey. These provinces include Muğla, Antalya, Mersin, Adana, and Osmaniye, stretching along the Taurus Mountains and the Mediterranean coastal belt south of Turkey. The study area extends from 36°1'14.884" N to 38°24'8.919" N latitude and 27°12'18.892" E to 36°42'19.828" E longitude (Figure 1).

The selected area remains under the influence of the Azores' High pressure in summer, while, in winter, typical climatic characteristics of the confluence of northern polar air mass and southern tropical air mass dominate the region [59, 60]. The area generally has a typical Mediterranean climate with hot and dry summers and mild, rainy winters. The altitude reaches approximately 3500 meters from the seacoast. The annual temperature varies from 12°C to 20°C, the annual precipitation average is 400–1200, and the relative humidity averages about 53–69 percent (<http://www.mgm.gov.tr>). In addition, prevailing dry and continuous north winds are considered highly responsible for forest fires, especially during fire periods. The area has the highest number of endemic plants in Turkey due to the limestone-covered lands and climatic conditions suitable for karstification [61–64].

The selected area extends over 66,014.26 km², of which approximately 54% is covered by forests (Table 1). The area is divided into different ecological regions based on the increase in altitude from the coastline. Some leading plant associations and forests in the study area are *Quercus coccifera*, *Olea europea*, *Arbutus andrachne*, *Laurus nobilis*, *Ceratonia siliqua*, *Pinus brutia*, *Pinus nigra*, *Abies cilicica subsp. cilicica*, *Abies cilicica subsp. isaurica*, *Cedrus libani*, *Juniperus excelsa*, *Juniperus foetidissima*, *Quercus libani*, *Quercus infectoria*, and *Quercus cerris*. *Pinus brutia* is a dominant species in the region with high fire sensitivity [65, 66].

2.2. Historical Forest Fires. The preparation of a historical forest fire inventory based on different sources (satellite images, fieldwork, historical archives, etc.) is the first step in modelling the susceptibility of forest fires [41, 67–70]. In this

research, 3256 samples of past forest fire events were determined. The historical forest fire dataset was generated using data from NASA's Fire Information for Resource Management System (FIRMS) (<https://earthdata.nasa.gov/firms>) and NASA's Earth Observing System Data and Information System (EOSDIS). The dataset of forest fire events was produced using near real-time (NRT) moderate resolution imaging spectroradiometer (MODIS), thermal anomalies/fire locations with 1 km spatial resolution from the Terra and Aqua platforms. Furthermore, the projection of this data is WGS84 and is known as MODIS Collection 61. The historical scope of the dataset covers the period from April 2021 to August 2021, as the area was declared “Disaster Areas Affecting General Life” by AFAD. According to the dataset, the majority of forest fire events occurred in Muğla in August and Antalya in July. Nevertheless, almost no forest fire events occurred in Mersin in April and Osmaniye in August as shown in Figure 2 and illustrated in Table 2.

2.3. Forest Fire Conditioning Factors. The possibility of a wildfire occurrence is based on the environmental conditions of any forest area. Environmental conditions can be topographical, climatic, vegetation-related, and human-related. These classes are known as wildfire conditioning factors and are essential to generate a final susceptibility map [71]. Thirteen wildfire conditioning factors including elevation, slope degree, slope aspect, Topographical Wetness Index (TWI), annual mean temperature, annual mean relative humidity, annual mean wind speed, land use, distance from water bodies, distance from residential areas, distance from roads, Normalized Difference Vegetation Index (NDVI), and land surface temperature (LST), were selected and produced in the GIS framework as geospatial database (Figures 3–5).

The LST is a critical indicator for a wide range of research topics. It represents the interaction of the atmosphere with the surface as well as the energy flow between them. The LST has been calculated using thermal infrared data from polar orbit satellites and geostationary satellites. Various algorithms have been developed to overcome external influences and retrieve LST data with high accuracy. In this research, the LST data from MODIS (moderate resolution imaging spectroradiometer) was calculated based on the generalized split-window (GSW) algorithm of [72]:

$$T_s = a_0 + \left(a_1 + a_2 \frac{1 - \varepsilon}{\varepsilon} + a_3 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{(T_i + T_j)}{2} + \left(a_4 + a_5 \frac{1 - \varepsilon}{\varepsilon} + a_6 \frac{\Delta \varepsilon}{\varepsilon^2} \right) \frac{(T_i + T_j)}{2}. \quad (1)$$

In the formula, T_s represents LST. $\varepsilon = (\varepsilon_i + \varepsilon_j)/2$, $\Delta \varepsilon = \varepsilon_i - \varepsilon_j$ (ε_i and ε_j represent land surface emissivities in channels i and j). T_i and T_j are top of atmosphere (TOA) brightness temperatures measured in channels i and j . $i = 31$ and $j = 32$ for MODIS data. $a_0 - a_6$ are coefficients obtained from simulated data.

All the spatial variables are represented based on WGS 1984 Mercator coordinate system. Raw datasets were obtained from various data sources as shown in Table 3.

Initially, the digital elevation model (DEM) was acquired from ASTER (advanced spaceborne thermal emission and reflection radiometer) as the GDEM (global digital elevation



FIGURE 1: Location map of the study area.

TABLE 1: Forest areas and rates by province (<http://www.ogm.gov.tr>).

Provinces	Total area (km ²)	Forest covered area (km ²)	Percentage
Adana	14,174.17	5,936.60	42%
Antalya	20,617.64	11,460.62	56%
Mersin	15,630.68	8,355.30	53%
Muğla	12,278.59	8,293.10	68%
Osmaniye	3,313.18	1,586.80	48%
Total	66,014.26	35,632.42	54%

model) with a 30 m spatial resolution. Topographical variables such as elevation, slope, aspect, and TWI (topographic wetness index) were generated from DEM. The following equation was used to generate the TWI:

$$TWI = \ln \frac{A_s}{\beta}. \quad (2)$$

In the equation, A_s indicates the specific catchment area (m/m²) and " β " shows the angle unit of the slope degree.

Climatic elements observations for an extended period were obtained from MGM (Directorate General of Meteorology) for each meteorological station. Selected climatic variables of annual mean temperature, annual mean relative humidity, and annual mean wind speed were joined to station locations in the GIS environment. Afterwards, the database of the variables was produced using IDW (inverse distance weighted) interpolation method. The vector-based data were acquired from the Turkish Ministry of Agriculture and Forestry to generate land use variables, distance from water bodies, and residential areas. Data were processed in GIS with tools such as rasterization for land use and Euclidean distance for the proximity of water reservoirs and settlements. The road network data was downloaded from OSM (open street map) to produce distances from roads

using the Euclidean distance tool in the GIS. All the variables were represented at the same spatial resolution (30 m). Contrariwise, the spatial variables obtained from MODIS (moderate resolution imaging spectroradiometer) were in different spatial resolutions, e.g., the resolutions of NDVI and LST were 250 m and 1 km, respectively.

2.4. Boosting Algorithms

(1) *Gradient Boosting Machine (GBM)*. Gradient boosting machines (GBMs) are one of the unique machine learning algorithms that have shown significant success in many types of research. It was produced with a formula based on the gradient descent of the boosting methods to establish a statistical association within the studies. These boosting methods and related algorithms were named gradient boosting machines [73, 74]. In GBMs, the learning system is concerned with sequentially fitting new models to ensure that the resulting response variable is a more accurate estimate. The main goal of the algorithm is to construct new base learners in such a way that they are maximally correlated with the negative gradient of the ensemble-related loss function. In this aspect, GBMs have a significant record of success both in practical applications and in various machine learning and data mining challenges [75].

(2) *Extreme Gradient Boosting (XGBoost)*. XGBoost is a gradient-boosting algorithm that combines weak learner predictions to get stronger learner predictions. This aspect has been frequently used by data scientists in research lately for better results [76]. In the XGBoost algorithm, CART acts as the base classifier. The input sample of the following decision tree and the training and prediction results of the previous decision tree are associated with each other and

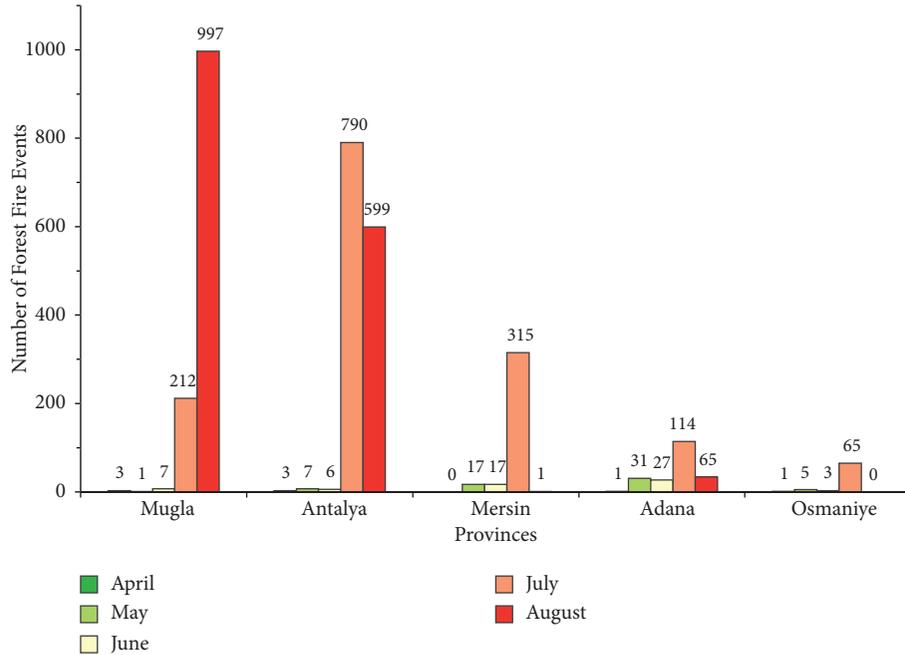


FIGURE 2: Monthly forest fire events according to the provinces.

TABLE 2: Historical events of fire forest in 2021.

Months/provinces	Muğla	Antalya	Mersin	Adana	Osmaniye
April	3	3	0	1	1
May	1	7	17	31	5
June	7	6	17	27	3
July	212	790	315	114	65
August	997	599	1	34	0

decided jointly. In addition to solving regression and classification problems, it is a flexible algorithm according to its intended use [49].

(3) *Light Gradient Boosting Machin (Light GBM)*. LightGBM is a kind of gradient boosting decision tree. This algorithm is mainly utilised in classification, sorting, and regression. LightGBM uses a histogram-based algorithm to increase computational speed and reduce complexity. It supports algorithms such as GBM, GBDT, GBRT, and MART, and its accuracy and efficiency are pretty high [77, 78]. In LightGBM, gradient-based one-side sampling (GOSS) is one of the methods used to calculate information gain, so that less-trained instances contribute more to information gain [79]. With these aspects, Light GBM provides quick practice and more extensive performance, low memory usage, good accuracy, support of GPU learning, and capacity to process large-scale data [80].

(4) *Categorical Boosting (CatBoost)*. Categorical boosting (CatBoost) was developed by [81]. CatBoost is a GBDT application in machine learning. This algorithm has two important features: ordered target statistics and ordered boosting. CatBoost is a good algorithm for solving complex data in a problem. However, it may not be very suitable for solving problems that are not too

complicated [82]. Another major feature of CatBoost in problem solving is that it captures high-degree dependencies and uses combinations of categorical features [83].

(5) *Adaptive Boosting (AdaBoost)*. Freund and Schapire developed the AdaBoost algorithm in 1995, whose weight can be regulated without the learner's requirement for prior knowledge [84]. Freund and Schapire developed the algorithm to solve the multiclass problem when there is a wide category in 1997 [85]. Since AdaBoost is an adaptive algorithm, it is among the most common boosting algorithms. In addition, AdaBoost is straightforward to use and practical in solving problems. It usually gives very effective results [86].

2.5. *Accuracy Evaluation*. To assess the performance of the LightGBM, GBM, XGBoost, AdaBoost, and CatBoost algorithms in forest fire susceptibility modelling, we used the receiver operating characteristic (ROC) curve as the first stage. ROC illustrates the true positives (TP), true negatives (TN), false positives (FP), and false-negative (FN) samples in machine learning models [41, 87]. After that, all the models were evaluated and compared with statistical methods such as overall accuracy (equation (3)), precision (equation (4)), recall (equation (5)), sensitivity (equation (6)), specificity (equation (7)), F1 measure (equation(8)), Kappa Index (equation (9)), and area under the curve (AUC):

$$\text{overall accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} - \text{FN} - \text{FP} + \text{TP}}, \quad (3)$$

$$\text{precision} = \frac{\text{TP}}{\text{FP} + \text{TP}}, \quad (4)$$

$$\text{recall} = \frac{\text{TP}}{\text{FN} + \text{TP}}, \quad (5)$$

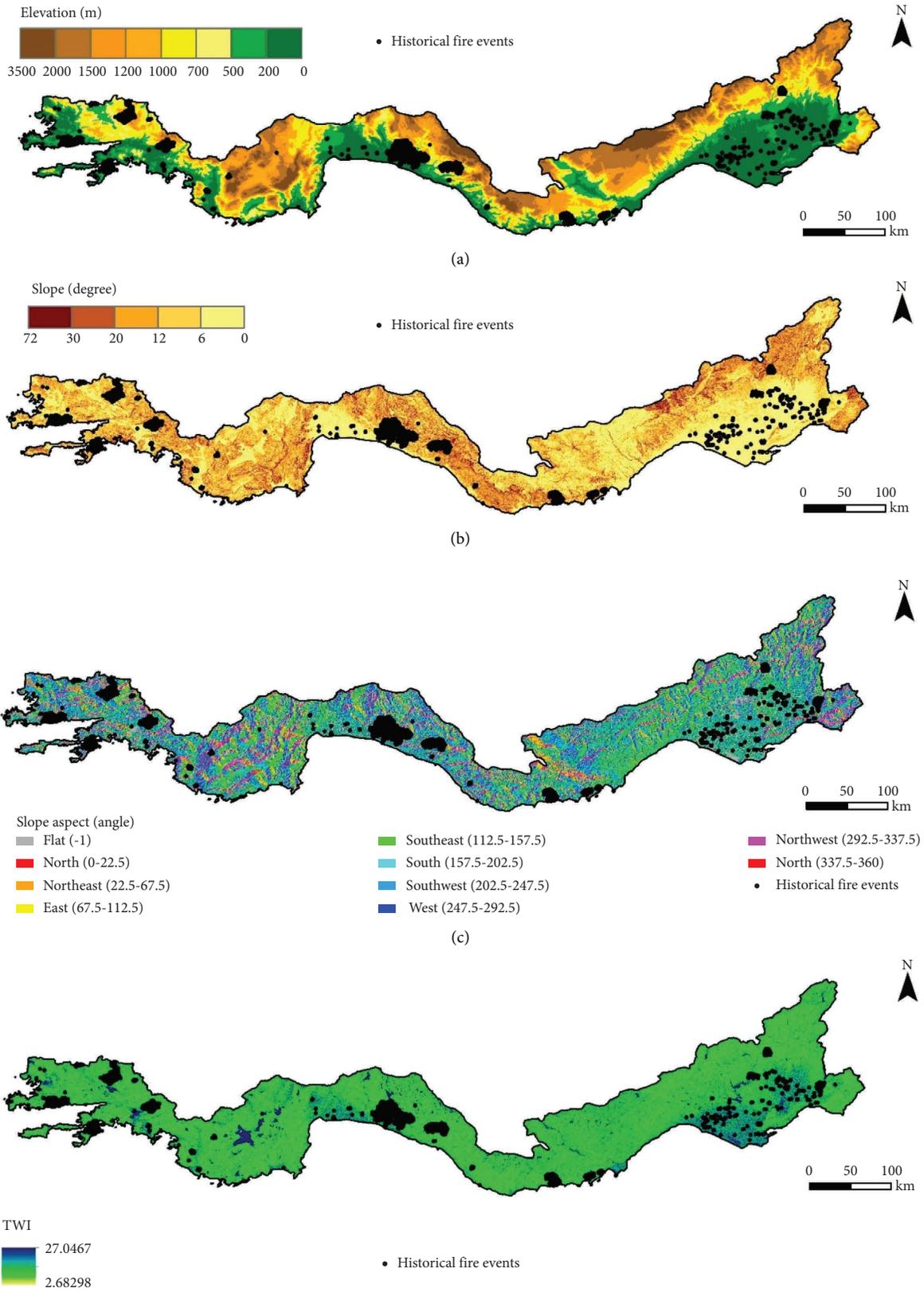


FIGURE 3: Continued.

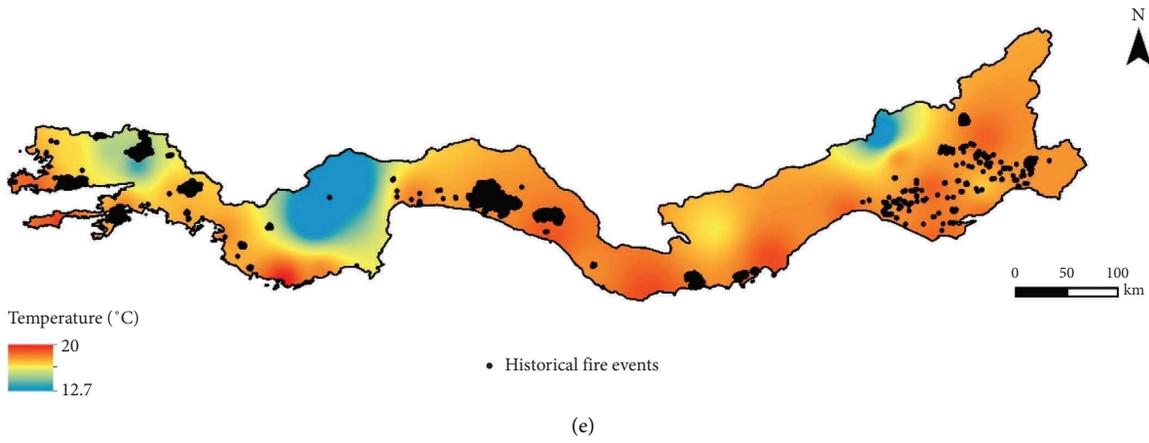


FIGURE 3: Forest fire conditioning factors: (a) elevation, (b) slope, (c) aspect, (d) TWI, and (e) temperature.

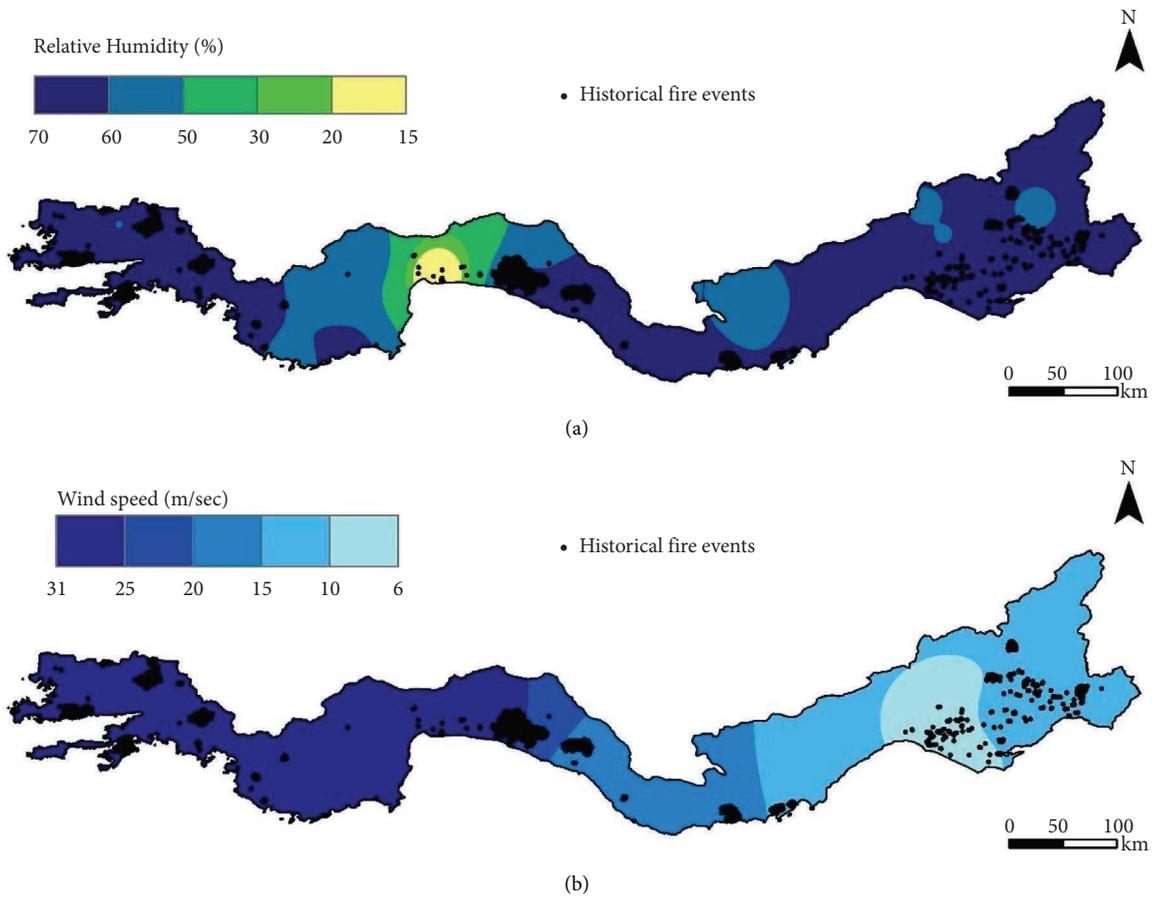


FIGURE 4: Continued.

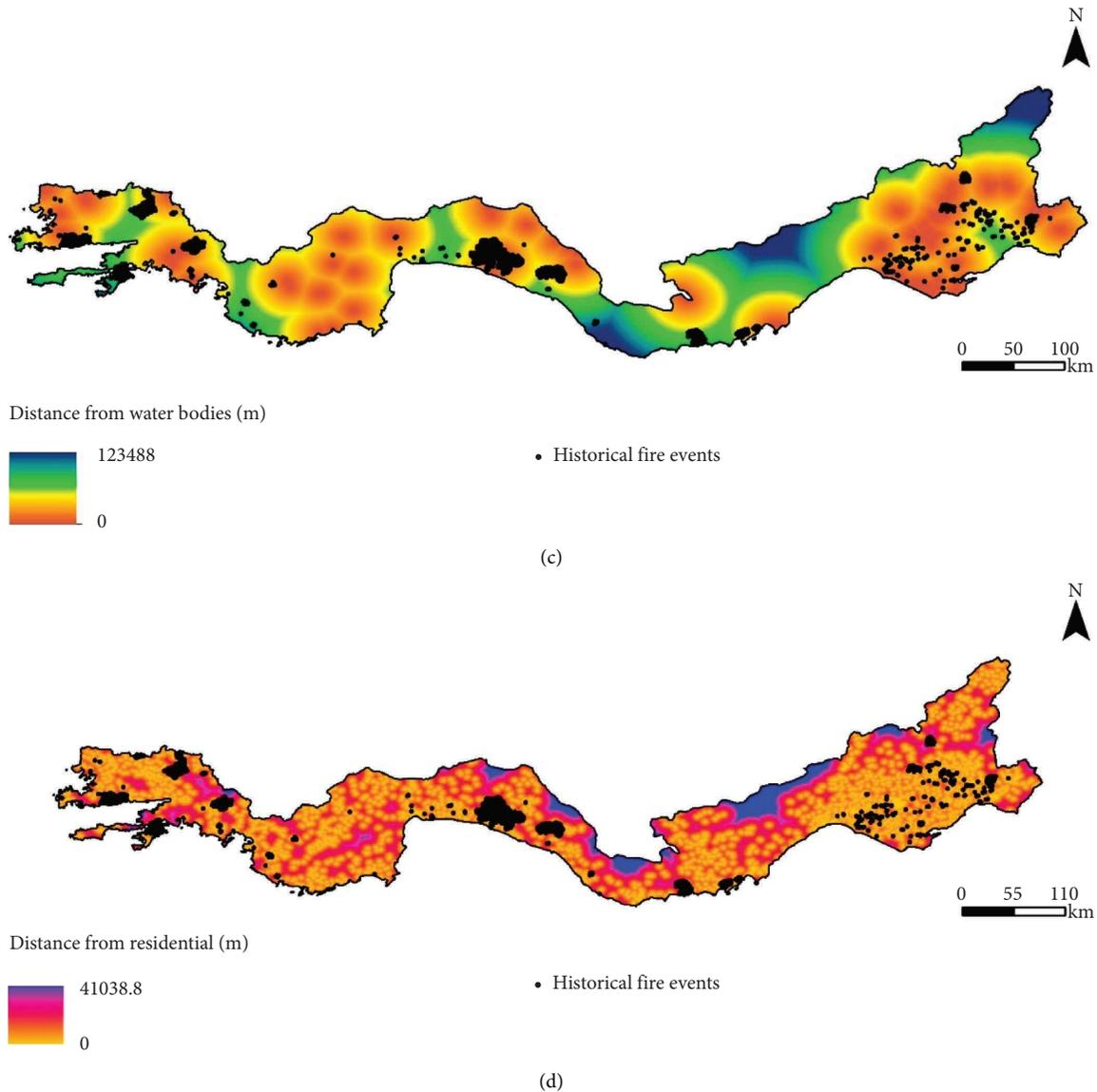


FIGURE 4: Forest fire conditioning factors: (a) relative humidity, (b) wind speed, (c) distance from water bodies, and (d) distance from residential.

$$\text{sensitivity} = \frac{TP}{TP + FN}, \quad (6)$$

$$\text{specificity} = \frac{TN}{TN + FP}, \quad (7)$$

$$F1 \text{ score} = \frac{2 \times (\text{precision} \times \text{recall})}{\text{precision} + \text{recall}}, \quad (8)$$

$$\text{Kappa} = \frac{\text{overall accuracy} - \text{random accuracy}}{1 - \text{random accuracy}}. \quad (9)$$

3. Results and Discussion

3.1. Importance of Wildfire Conditioning Factors. The importance degree of the wildfire conditioning factors used in the present study is given in Figure 6. It is observed that the

importance of the elements had the same ranking in all the models. Wind speed was observed as the most important factor in all the models, followed by humidity, temperature, LST, distance from water bodies, distance from residential, elevation, slope, land use, NDVI, distance from roads, TWI, and aspect, respectively. The AdaBoost model ignored the effects of some factors such as land use, NDVI, distance from roads, TWI, and aspect, while the GBM model disregarded the impact of the aspect.

3.2. Wildfire Susceptibility Models. In this study, we evaluated the predictive performance of five machine learning algorithms including LightGBM, GBM, XGBoost, AdaBoost, CatBoost in wildfire susceptibility mapping. The prediction performance of all models was compared with each other. Thirteen conditioning factors were prepared to analyze the wildfire susceptibility of the study area. Afterwards, all the factors were extracted to a total of

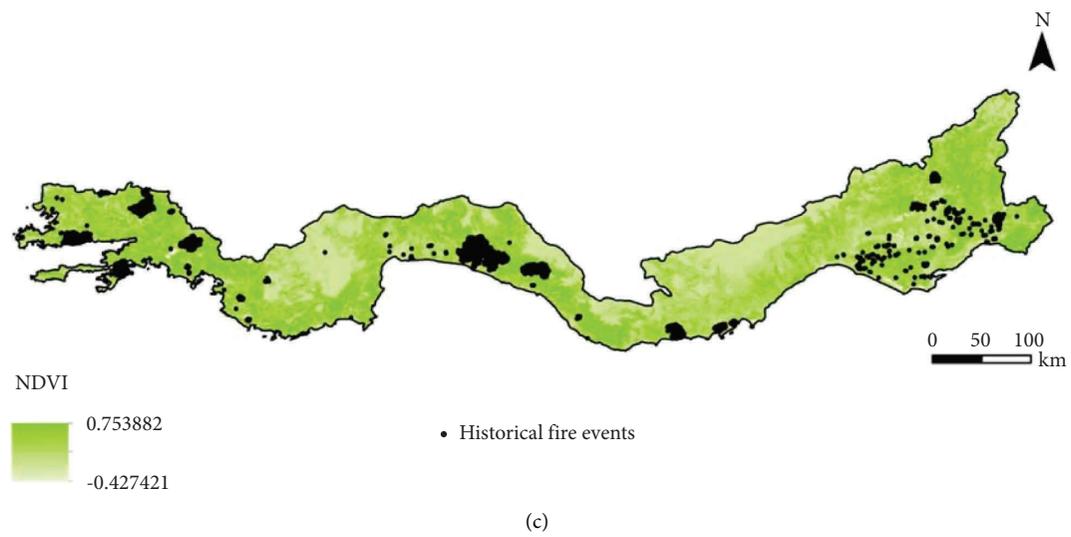
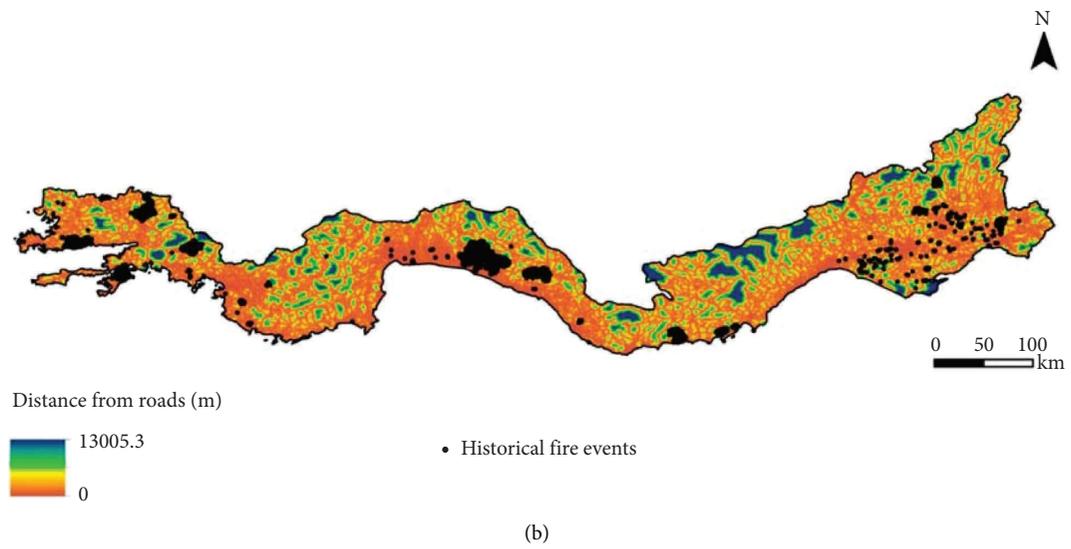
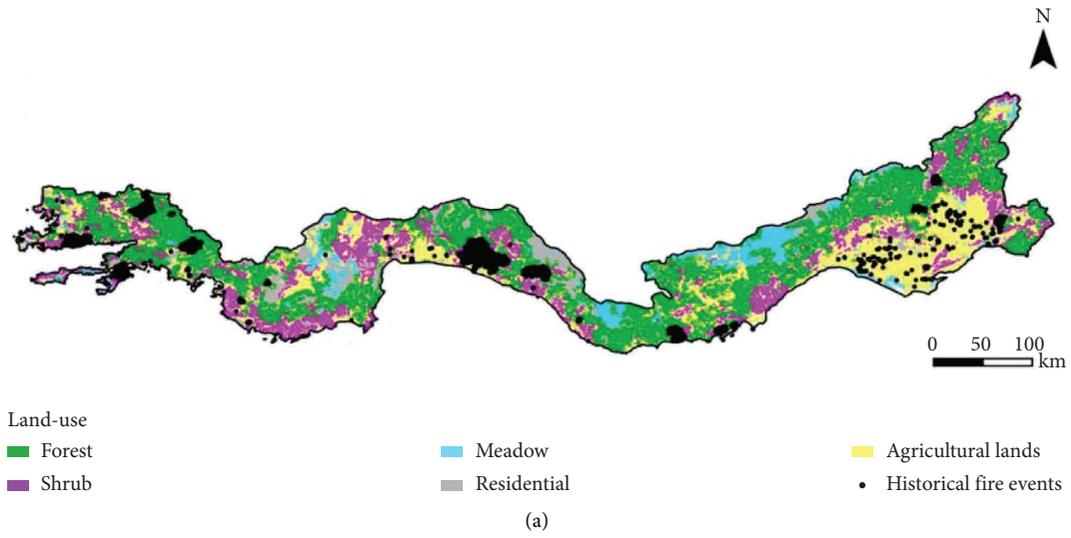


FIGURE 5: Continued.

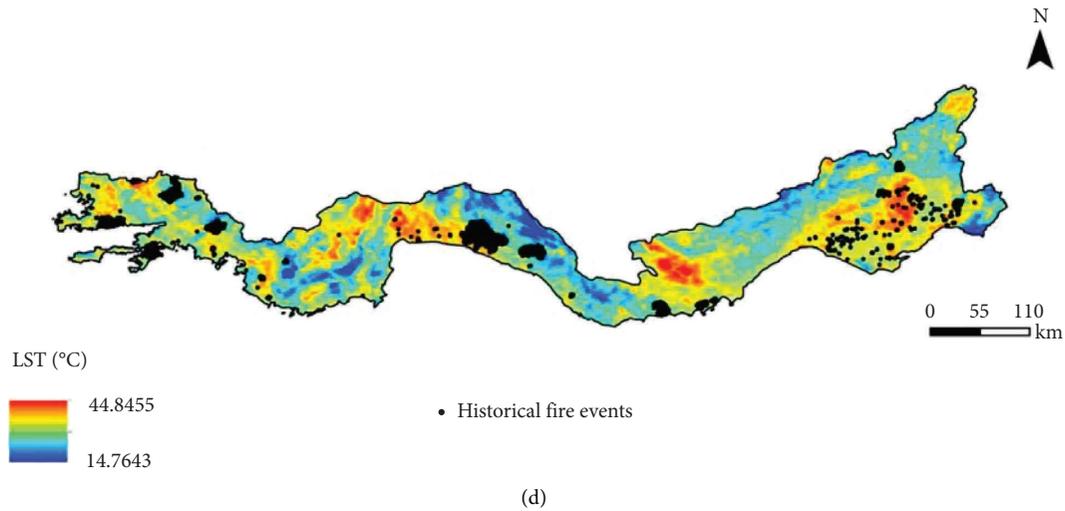


FIGURE 5: Forest fire conditioning factors: (a) land use, (b) distance from roads, (c) NDVI, and (d) LST.

TABLE 3: Selected fire ignition factors, data sources, data processing, and intervals.

Factors	Data sources	Processing	Intervals
Elevation	ASTER	Clipped with study area	0–3500 m
Slope degree	ASTER	Derived from DEM	0–72 degree
Slope aspect	ASTER	Derived from DEM	–1–360 angle
TWI	ASTER	Derived from DEM	2–27
Temperature	MGM	IDW	12–20°C
Humidity	MGM	IDW	15–69%
Wind speed	MGM	IDW	6–31 m/sec
Land use	Ministry of Agriculture and Forestry	Digitized	Categorical (18)
Distance from water bodies	Ministry of Agriculture and Forestry	Euclidean distance	0–123,488 m
Distance from residential	Ministry of Agriculture and Forestry	Euclidean distance	0–41038 m
Distance from roads	OSM	Euclidean distance	0–13005 m
NDVI	MODIS	Clipped with study area	–0.4–0.7
LST	MODIS	Average of multiple images	287–317 K

ASTER: advanced spaceborne thermal emission and reflection radiometer; MGM: Turkish Meteorology Service (Meteoroloji Genel Müdürlüğü); OSM: open street map, MODIS: moderate resolution imaging spectroradiometer; TWI: Topographical Wetness Index; NDVI: Normalized Difference Vegetation Index; LST: land surface temperature; DEM: digital elevation model; IDW: inverse distance weighted.

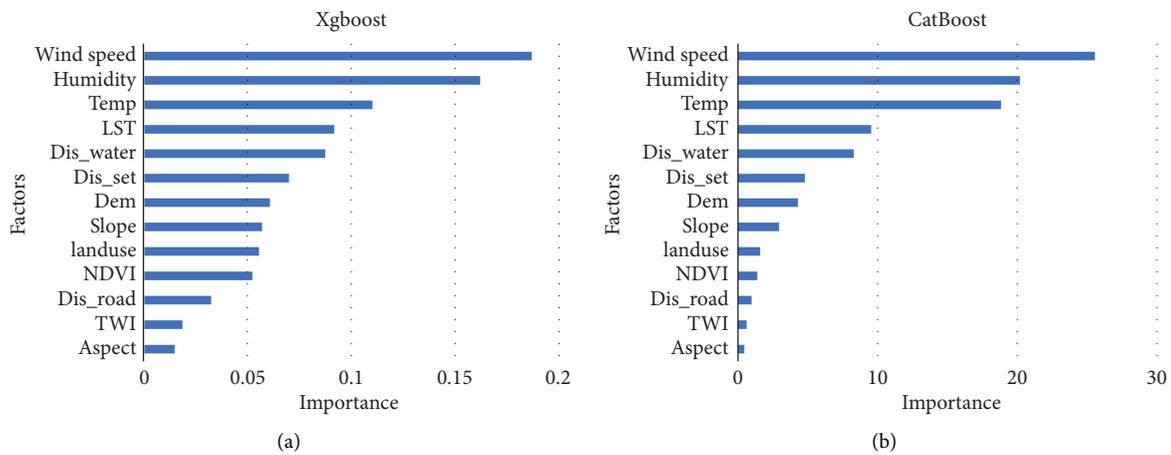


FIGURE 6: Continued.

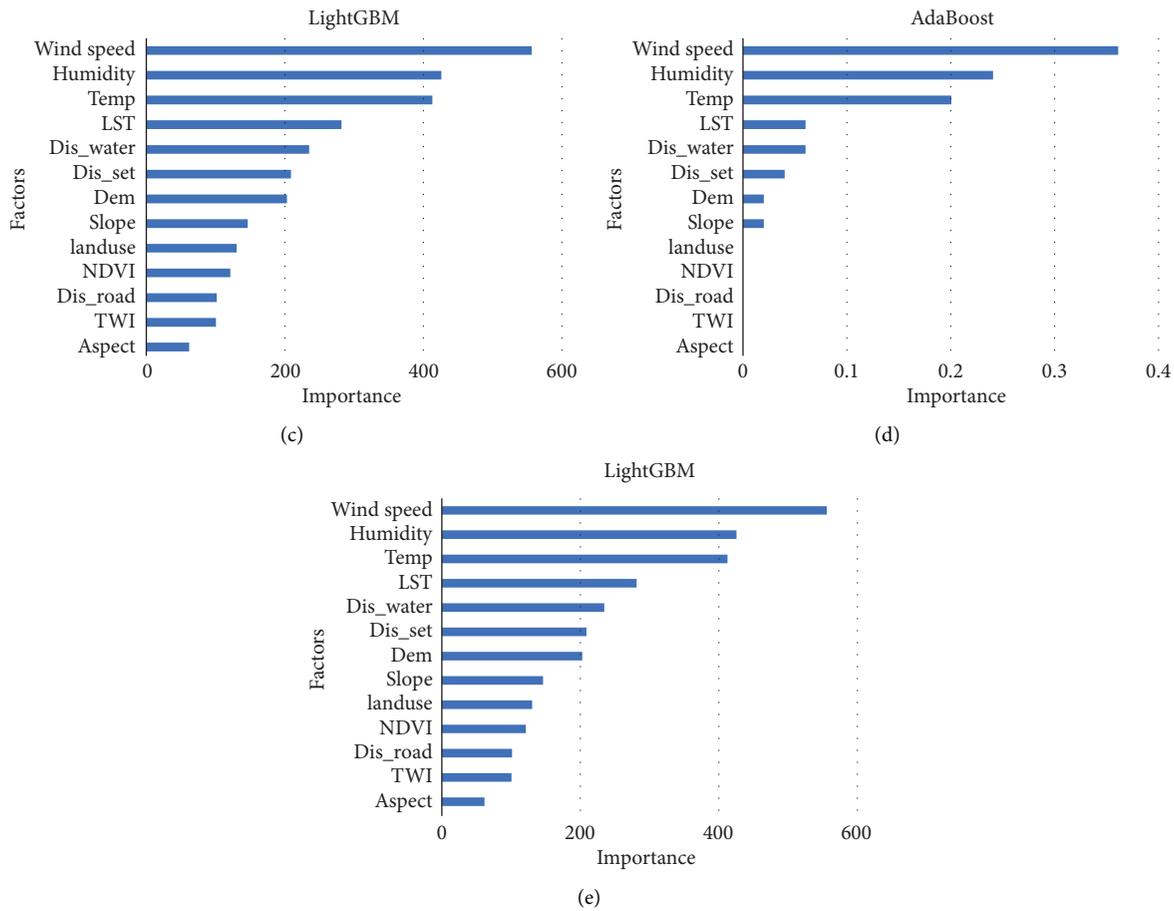


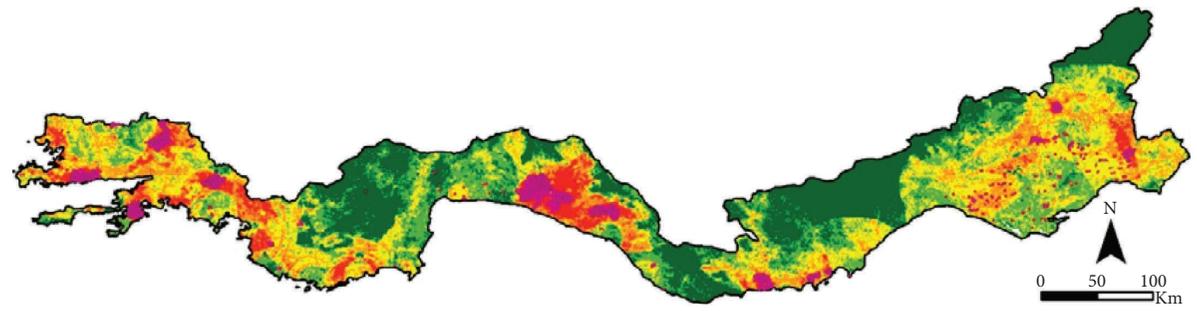
FIGURE 6: Importance degrees of the factors used in this study according to all the models: (a) XGBoost, (b) CatBoost, (c) GBM, (d) AdaBoost, and (e) LightGBM.

6292 target points, including historical forest fire points (3256) and nonfire points (3036). Also, target points were separated as binary into 0 for nonfire and 1 for historical fire samples. The models were trained with these sample points as the input dataset. The input dataset was divided into 70% for training and 30% for validation. After the analysis, maps of the results of all selected models were produced in ArcMap. The resultant wildfire susceptibility maps were classified into five classes: very low, low, medium, high, and very high, using the natural breaks classifier (Figure 7).

The spatial distribution of forest fire classes according to the models is presented in Table 4 and Figure 8. It has been observed that the high and very high susceptibility classes share the total area of 25%, 11%, 23%, 78%, and 10% with the methods of XGBoost, CatBoost, GBM, AdaBoost, and LightGBM, respectively.

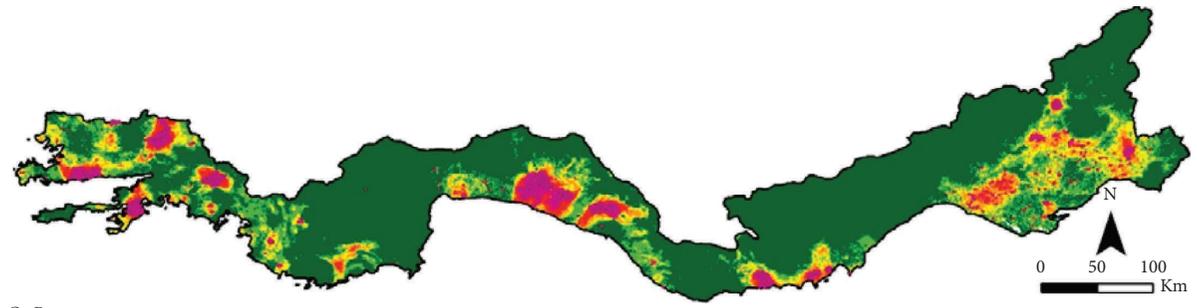
3.3. Evaluation and Comparison of the Wildfire Susceptibility Models. Statistical evaluation of the selected models is presented in Table 5. Accordingly, all the models have shown high and acceptable accuracy in training and testing scores. The training scores are found to be higher than the testing scores in all the models, validating that all the models have avoided the overfitting problem.

According to the statistical measure evaluations, the testing scores reveal the better performance of CatBoost algorithm than the other models followed by LightGBM, XGBoost, AdaBoost, and GBM algorithms. Overall accuracy scores demonstrate that the CatBoost model correctly classifies the samples with 95% accuracy. Also, the CatBoost model has a more relevant sampling rate of 0.951 in precision and 0.954 in recall, with higher accuracy compared to the other models. According to *F1* scores, the performance of the precision and recall measurements shows that the CatBoost (0.952) model reached higher accuracy, and it was followed up by the LightGBM (0.936), XGBoost (0.874), AdaBoost (0.864), and GBM (0.827). In addition, according to the specificity scores, the CatBoost model is better in classifying the TN samples more correctly, with a value of 0.954. The other models also perform well in classifying the TN samples (LightGBM 0.939, XGBoost 0.885, AdaBoost 0.873, and GBM 0.838). In classifying TP samples, the CatBoost and LightGBM models revealed equal performance with a value of 0.956. In contrast, the other models, such as XGBoost, GBM, and AdaBoost showed sensitivity scores of 0.892, 0.853, and 0.846, respectively. The Kappa Index shows a more balanced distribution of CatBoost model in classified samples with a value of



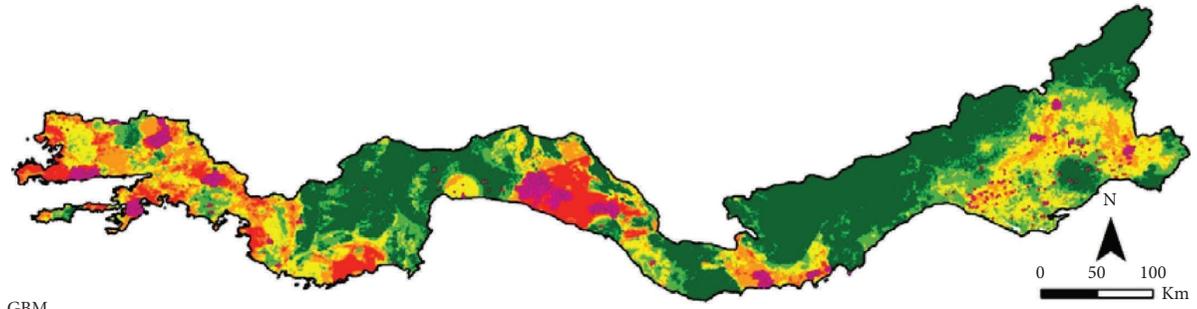
XGBoost
Forest fire probability
Very low (0.13 - 0.16)
Low (0.17 - 0.3)
Medium (0.31 - 0.44)
High (0.45 - 0.62)
Historical forest fire event
Very high (0.63 - 0.91)

(a)



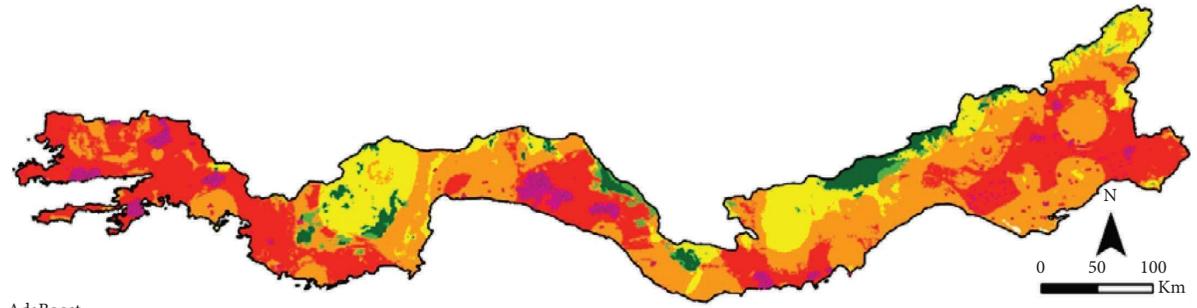
CatBoost
Forest fire probability
Very low (0 - 0.078)
Low (0.079 - 0.24)
Medium (0.25 - 0.47)
High (0.48 - 0.76)
Historical forest fire event
Very high (0.77 - 1)

(b)



GBM
Forest fire probability
Very low (0 - 0.14)
Low (0.15 - 0.31)
Medium (0.32 - 0.49)
High (0.5 - 0.71)
Historical forest fire event
Very high (0.72 - 0.95)

(c)



AdaBoost
Forest fire probability
Very low (0.26 - 0.3)
Low (0.31 - 0.33)
Medium (0.34 - 0.47)
High (0.48 - 0.49)
Historical forest fire event
Very high (0.5 - 0.53)

(d)

FIGURE 7: Continued.

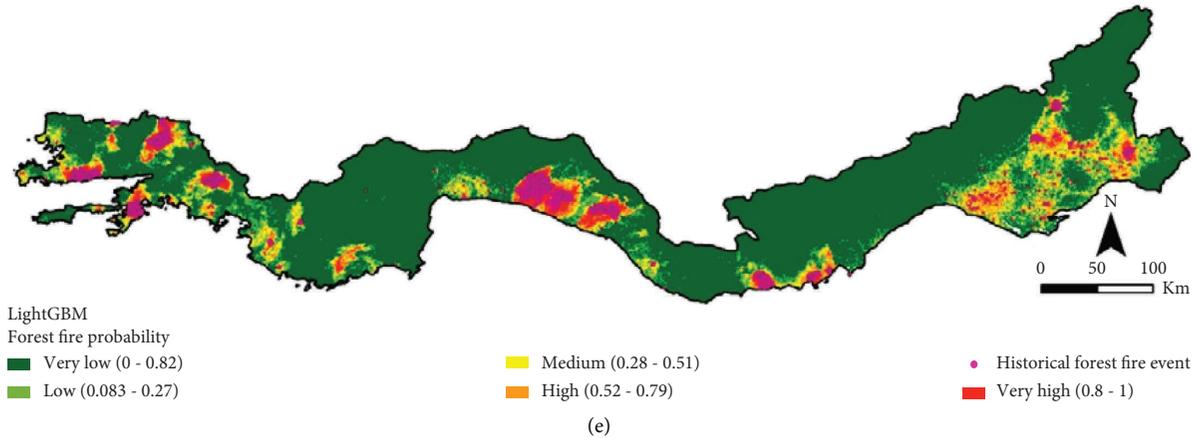


FIGURE 7: Result maps of all the models: (a) XGBoost, (b) CatBoost, (c) GBM, (d) AdaBoost, and (e) LightGBM.

TABLE 4: Area and percentage distribution of the susceptibility classes of all the models.

Models	Distribution	Susceptibility classes				
		Very low	Low	Medium	High	Very high
XGBoost	Area (km ²)	26349	27655	23716	17175	9258
	Percentage (%)	25	27	23	16	9
CatBoost	Area (km ²)	67201	17663	8170	5138	5981
	Percentage (%)	65	17	8	5	6
GBM	Area (km ²)	41579	20842	18171	13533	10029
	Percentage (%)	40	20	17	13	10
AdaBoost	Area (km ²)	4369	2633	16425	41285	39441
	Percentage (%)	4	3	16	40	38
LightGBM	Area (km ²)	73658	13388	6759	4321	6028
	Percentage (%)	71	13	6	4	6

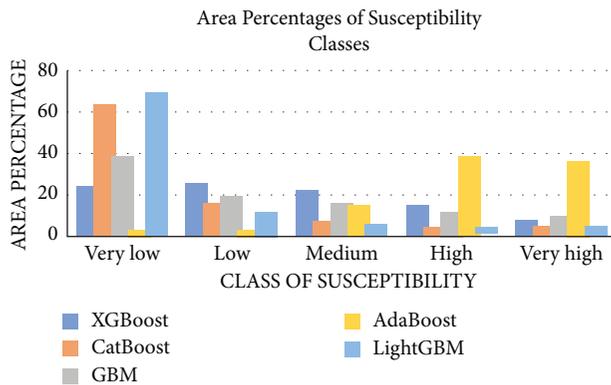


FIGURE 8: Area percentages of susceptibility classes.

0.909. Moreover, the AUC values indicate that the CatBoost model has a 0.955 testing score, and it was followed by LightGBM, XGBoost, AdaBoost, and GBM algorithms with 0.948, 0.888, 0.859, and 0.846 AUC values, respectively (Figure 9).

The algorithms used in the present study have been applied in many predictive mapping studies of various natural processes because of their solid and separative prediction performance as

an alternative to traditional statistical and machine learning methods. The results of the present study are parallel to many previous studies which compare the efficiency of different algorithms in susceptibility maps. Sahin [88] found that CatBoost model is superior in predicting landslide susceptibility areas in the Bolu region of Turkey. Similarly, Saber et al. [89] also appreciated CatBoost and LightGBM algorithms in flash flood susceptibility. Zhou et al. [56] proposed a fire prediction model using the CatBoost algorithm for Yunnan Province in China. The model that used inputs such as vegetation, meteorological, terrain, and human factors reached 0.83 AUC value. The results indicated that the CatBoost model effectively predicted the risk of forest fire occurrence.

The CatBoost algorithm is not tested extensively in previous studies on wildfire susceptibility despite its higher performance and accuracy in determining susceptibility maps. Hakim et al. [90] employed only AdaBoost and LogitBoost algorithms, while Arabameri et al. [91] tested VIKOR and Cforest models to create land subsidence susceptibility maps. Rosadi and Andriyani [57] applied the AdaBoost algorithm to predict the forest-fire occurrence and compared it with classical classification methods such as SVM (support vector machine) and decision tree methods. Michael et al. [58] investigated long-term vegetation condition effects on wildfire risk mapping by applying RF,

TABLE 5: The prediction performance results of the models according to the measurement methods.

Measurements	LightGBM		GBM		XGBoost		AdaBoost		CatBoost	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
Overall accuracy	99.886	94.703	85.675	84.488	90.443	88.801	86.430	86.021	99.126	95.473
Precision	0.999	0.933	0.822	0.816	0.871	0.863	0.840	0.855	0.989	0.951
Recall	0.999	0.939	0.844	0.838	0.884	0.885	0.854	0.873	0.990	0.954
Sensitivity	0.999	0.956	0.873	0.853	0.929	0.892	0.876	0.846	0.993	0.956
Specificity	0.999	0.939	0.844	0.838	0.884	0.885	0.854	0.873	0.990	0.954
AUC	0.999	0.948	0.858	0.846	0.907	0.888	0.865	0.859	0.991	0.955
F1	0.999	0.936	0.833	0.827	0.877	0.874	0.847	0.864	0.990	0.952
Kappa Index	0.998	0.894	0.712	0.689	0.808	0.774	0.728	0.719	0.982	0.909

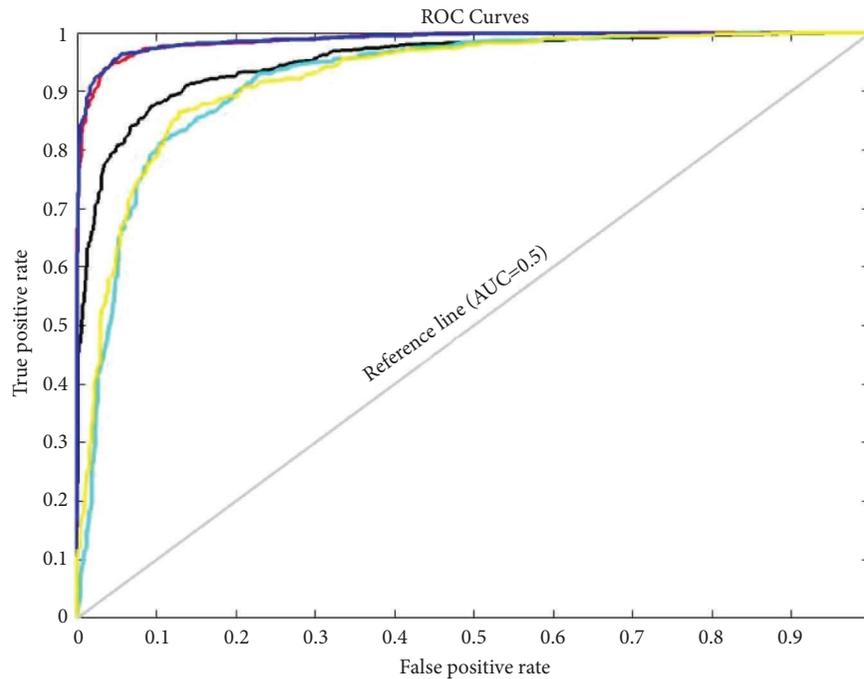


FIGURE 9: ROC curves and AUC values of all the models.

logistic regression, and XGBoost methods. Similarly, Dang et al. [55] used random forest (RF), multilayer perceptron (MLP), XGBoost, and AdaBoost to develop an annual fire prediction model for the Humberside area in the United Kingdom. The researchers reported that the AdaBoost algorithm outperformed compared to the other models. However, they did not use CatBoost model in their study, which has yielded more robust results in many previous studies, including the present research.

Hence, to the best of the authors' knowledge, no previous research has compared wildfire susceptibility mapping using the ML methods such as XGBoost, CatBoost, GBM, AdaBoost, and LightGBM. The present study's results help open further research by testing more innovative techniques in machine learning to enhance the accuracy of wildfire susceptibility maps.

4. Conclusion

Wildfires are one of the most dangerous natural hazards for forest areas and habitats. Due to global warming, the frequency of forest fires has increased in the last decades, especially in the Mediterranean climate zones. Predictive

susceptibility mapping for wildfires is an effective tool for planners and managers to prevent and protect against the undesirable effects of wildfires. The reliability of the susceptibility maps differs from its input parameters and the methodology used. In recent years, ML-based predictive mapping studies have been increasing rapidly and gaining the trust of researchers. In the present study, we compared state-of-the-art ML algorithms such as XGBoost, CatBoost, GBM, AdaBoost, and LightGBM to produce wildfire susceptibility mapping for the Mediterranean Region of Turkey. To our best knowledge, no study compared these algorithms before in wildfire susceptibility mapping literature. For the analysis, the thirteen input parameters were used: elevation, slope degree, slope aspect, TWI, temperature, humidity, wind speed, land use, distance from water bodies, distance from residential, distance from roads, and NDVI LST. According to the order of importance of the factors, while wind speed is the most crucial factor, the aspect is the least important in all the models. After producing the susceptibility maps, statistical accuracy assessment techniques such as overall accuracy, precision, recall, sensitivity, specificity,

AUC, F1 score, and Kappa Index were applied. The results showed that the CatBoost algorithm had higher accuracy than the other models, followed by LightGBM, XGBoost, AdaBoost, and GBM algorithms. However, all the models have revealed reasonably good AUC measurement performance: 0.955, 0.888, 0.859, and 0.846, respectively.

The present study has several limitations. First, the spatial interpolation technique produced wind speed, humidity, and temperature factors. The interpolation technique has disadvantages such as dependence on sample locations, generalization, and ignoring the geomorphological conditions. Therefore, some parameters may not fully reflect the actual climatic conditions of the selected area. Second, the present study does not depict human-induced factors of wildfire due to the absence of spatial dimension of such data. Third, the present study only provides an ML-based evaluation to give some idea about wildfire-sensitive areas in the study area. Therefore, future wildfires should be followed and examined, considering the maps produced in this study. One of the main limitations is the contradiction between the large-scale areas and the required details for the modelling. Therefore, explaining the factors during the model development becomes more complicated due to the missing details. In addition, the requirement of high computation capacity made the modelling challenging to deal with all the available resolutions.

The present study is considered novel research in producing hotspots and wildfire susceptibility maps of Turkey's Mediterranean Region. The maps are expected to provide valuable inventories for forest engineers, planners, and local governments for future policies regarding disaster management in Turkey. Besides, the present study also provides a comparative analysis of relatively new ML algorithms such as XGBoost, CatBoost, GBM, AdaBoost, and LightGBM for wildfire susceptibility mapping research. Suggestions for future research include comparing the methods used in this study with other statistical and ML-based methods and using different input parameters in the models. This research recommends future research to use the cloud computing platforms such as Google Earth engine, Google Colab, Amazon AWAS, and Kaggle for performing the modelling of wildfire susceptibility mapping.

Data Availability

All the relevant data are included in the article.

Disclosure

This work was performed as part of the duties of the authors as well as a collaboration between the International College for Engineering and Management (ICEM) and Karabuk University (KBU).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] K. Oğuz, E. Oğuz, and G. Çamalan, "İzmir-tirazli orman yangininin uydu ve model verileri ile analizi," *Ulusal Çevre Bilimleri Araştırma Dergisi, Sayı*, vol. 4, no. 1, pp. 1–12, 2021.
- [2] S. Oliveira, J. Rocha, and A. Sá, "Wildfire risk modeling," *Current Opinion in Environmental Science & Health*, vol. 23, 2021.
- [3] J. San-Miguel-Ayanz, T. Durrant, R. Boca et al., "Advance EFFIS report on forest fires in europe, Middle East and north africa 2019," in *Joint Research Center EC (Issue March) Publications Office of the European Union, Luxembourg*, 2020.
- [4] S. Ayanoglu, M. Dölarslan, and E. Gül, "Sadece bir yangın mı? Ekolojik ve sosyo-ekonomik açıdan orman yangınları," *Türk Bilimsel Derlemeler Dergisi*, vol. 10, no. 2, pp. 32–35, 2017.
- [5] N. Masoudvaziri, P. Ganguly, S. Mukherjee, and K. Sun, "Impact of geophysical and anthropogenic factors on wildfire size: a spatiotemporal data-driven risk assessment approach using statistical learning," *Stochastic Environmental Research and Risk Assessment*, vol. 4, 2021.
- [6] M. Goss, D. L. Swain, J. T. Abatzoglou et al., "Climate change is increasing the likelihood of extreme autumn wildfire conditions across California," *Environmental Research Letters*, vol. 15, no. 9, Article ID 094016, 2020.
- [7] J. E. Halofsky, D. L. Peterson, and B. J. Harvey, "Changing wildfire, changing forests: the effects of climate change on fire regimes and vegetation in the pacific northwest, USA," *Fire Ecology* [https](https://doi.org/10.1007/s11692-020-09401-6), vol. 16, 2020.
- [8] M. W. Jones, A. Smith, R. Betts, J. G. Canadell, I. C. Prentice, and C. Le Quéré, "Climate change increases the risk of wildfires rapid," *Science Brief Review*, vol. 116, 2020.
- [9] D. Blake, P. Nyman, H. Nice, F. M. L. D'souza, C. R. J. Kavazos, and P. Horwitz, "Assessment of post-wildfire erosion risk and effects on water quality in south-western Australia," *International Journal of Wildland Fire*, vol. 29, no. 3, pp. 240–257, 2020.
- [10] J. Llovet, M. Ruiz-Valera, R. Josa, and V. R. Vallejo, "Soil responses to fire in Mediterranean forest landscapes in relation to the previous stage of land abandonment," *International Journal of Wildland Fire*, vol. 18, no. 2, pp. 222–232, 2009.
- [11] R. A. Shakesby, "Post-wildfire soil erosion in the Mediterranean: review and future research directions," *Earth-Science Reviews*, vol. 105, no. 3–4, pp. 71–100, 2011.
- [12] D. A. Zema, J. P. Nunes, and M. E. Lucas-Borja, "Improvement of seasonal runoff and soil loss predictions by the MMF (Morgan-Morgan-Finney) model after wildfire and soil treatment in Mediterranean forest ecosystems," *Catena*, vol. 188, Article ID 104415, 2020.
- [13] T. C. J. Esteves, M. J. Kirkby, R. A. Shakesby et al., "Mitigating land degradation caused by wildfire: application of the PESERA model to fire-affected sites in central Portugal," *Geoderma*, vol. 191, pp. 40–50, 2012.
- [14] N. Öztürk and L. Özkan, "Orman yanginlari ve yaban hayati," *Doğanın Sesi*, vol. 4, pp. 67–74, 2019.
- [15] J. E. Keeley, W. J. Bond, R. A. Bradstock, J. G. Pausas, and P. W. Rundel, *Fire in Mediterranean Ecosystems: Ecology, Evolution and Management*, Cambridge University Press, Cambridge, UK, 2011.
- [16] A. Kavgaci, E. Örtel, I. Torres, and H. Safford, "Early postfire vegetation recovery of Pinus brutia forests: effects of fire severity, prefire stand age, and aspect," *Turkish Journal of Agriculture and Forestry*, vol. 40, no. 5, pp. 723–736, 2016.

- [17] A. Ganteaume, A. Camia, M. Jappiot, J. San-Miguel-Ayanz, M. Long-Fournel, and C. Lampin, "A review of the main driving factors of forest fire ignition over europe," *Environmental Management*, vol. 51, 2016.
- [18] M. Rodrigues, A. Jiménez, and J. de la Riva, "Analysis of recent spatial-temporal evolution of human driving factors of wildfires in Spain," *Natural Hazards*, vol. 84, no. 3, pp. 2049–2070, 2016.
- [19] C. Gai, W. Weng, and H. Yuan, "GIS-based forest fire risk assessment and mapping," in *Proceedings of the 2011 Fourth International Joint Conference on Computational Sciences and Optimization*, pp. 1240–1244, Kunming and Lijiang, China, April 2011.
- [20] D. Gülçin and B. Deniz, "Remote sensing and GIS-based forest fire risk zone mapping: the case of Manisa, Turkey," *Turkish Journal of Forestry | Türkiye Ormançılık Dergisi*, vol. 21, no. 1, pp. 15–24, 2020.
- [21] F. Guo, Z. Su, G. Wang, L. Sun, F. Lin, and A. Liu, "Wildfire ignition in the forests of southeast china: identifying drivers and spatial distribution to predict wildfire likelihood," *Applied Geography*, vol. 66, 2015.
- [22] R. K. Jaiswal, S. Mukherjee, K. D. Raju, and R. Saxena, "Forest fire risk zone mapping from satellite imagery and GIS," *International Journal of Applied Earth Observation and Geoinformation*, vol. 4, no. 1, pp. 1–10, 2002.
- [23] H. Adab, K. D. Kanniah, and K. Solaimani, "GIS-Based probability assessment of fire risk in grassland and forested landscapes of golestan province, Iran," in *Proceedings of the 2011 International Conference on Environmental and Computer Science*, vol. 19, pp. 170–175, Singapore, January 2011.
- [24] A. A. Gerdzheva, "A comparative analysis of different wildfire risk assessment models (a case study for smolyan district, Bulgaria)," *European Journal of Geography*, vol. 5, no. 3, pp. 22–36, 2014.
- [25] J. H. Scott, M. P. Thompson, and D. E. Calkin, "A wildfire risk assessment framework for land and resource management," USDA Forest Service - General Technical Report RMRS-GTR, 315 RMRS-GTR, U.S. Department of Agriculture, Forest Service, Washington, WA, USA, 2013.
- [26] D. M. G. Yakubu, I. Yakubu, and A. A. Duker, "Review of methods for modelling forest fire risk and hazard," *African Journal of Environmental Science and Technology*, vol. 9, no. 3, pp. 155–165, 2015.
- [27] H. Abedi Gheshlaghi, B. Feizizadeh, and T. Blaschke, "GIS-based forest fire risk mapping using the analytical network process and fuzzy logic," *Journal of Environmental Planning and Management*, vol. 63, no. 3, pp. 481–499, 2020.
- [28] A. E. Akay and A. Erdoğan, "GIS-based multi-criteria decision analysis for forest fire risk mapping," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, no. 4W4, pp. 25–30, 2017.
- [29] A. C. Teodoro and L. Duarte, "Forest fire risk maps: a GIS open source application – a case study in Norwest of Portugal," *International Journal of Geographical Information Science*, vol. 27, no. 4, pp. 699–720, 2013.
- [30] B. Leblon, J. San-Miguel-Ayanz, L. Bourgeau-Chavez, and M. Kong, "Remote sensing of wildfires," *Land Surface Remote Sensing*, vol. 2016, pp. 55–95, 2016.
- [31] A. Parajuli, A. P. Gautam, P. Sundar et al., "Forest fire risk mapping using gis and remote sensing in two major landscapes of nepal," *Geomatics, Natural Hazards and Risk*, vol. 11, 2020.
- [32] A. Jaafari, D. Mafi-Gholami, B. T. Pham, and D. T. Bui, "Remote sensing wildfire probability mapping: bivariate vs. multivariate statistics," *Remote Sensing*, vol. 11, 2019.
- [33] D. Liu and Y. Zhang, "Research of regional forest fire prediction method based on multivariate linear regression," *International Journal of Smart Home*, vol. 9, no. 1, pp. 13–22, 2015.
- [34] S. Oliveira, F. Oehler, J. San-Miguel-Ayanz, A. Camia, and J. M. C. Pereira, "Modeling spatial patterns of fire occurrence in mediterranean europe using multiple regression and random forest," *Forest Ecology and Management*, vol. 275, 2012.
- [35] H. H. d. G. Viganó, C. C. d. Souza, J. F. Reis Neto, M. F. Cristaldo, and L. d. Jesus, "Previsão e Modelagem das ocorrências de incêndios no pantanal," *Revista Brasileira de Meteorologia*, vol. 33, no. 2, pp. 306–316, 2018.
- [36] F. Mohammadi, M. P. Bavaghar, and N. Shabani, "Forest fire risk zone modeling using logistic regression and GIS: an Iranian case study," *Small-Scale Forestry*, vol. 13, no. 1, pp. 117–125, 2014.
- [37] M. Rodrigues, A. Jiménez-Ruano, D. Peña-Angulo, and J. De La Riva, "A comprehensive spatial-temporal analysis of driving factors of human-caused wildfires in spain using geographically weighted logistic regression," *Journal of Environmental Management*, vol. 225, 2018.
- [38] N. Sahani and T. Ghosh, "GIS-based spatial prediction of recreational trail susceptibility in protected area of Sikkim Himalaya using logistic regression, decision tree and random forest model," *Ecological Informatics*, vol. 64, no. June, Article ID 101352, 2021.
- [39] Z. X. Zhang, H. Y. Zhang, and D. W. Zhou, "Using GIS spatial analysis and logistic regression to predict the probabilities of human-caused grassland fires," *Journal of Arid Environments*, vol. 74, no. 3, pp. 386–393, 2010.
- [40] S. Bar, B. R. Parida, and A. C. Pandey, "Landsat-8 and Sentinel-2 based Forest fire burn area mapping using machine learning algorithms on GEE cloud platform over Uttarakhand, Western Himalaya," *Remote Sensing Applications: Society and Environment*, vol. 18, 2020.
- [41] M. Mohajane, R. Costache, F. Karimi et al., "Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area," *Ecological Indicators*, vol. 129, no. June, Article ID 107869, 2021.
- [42] Z. Mohammed, C. Hanae, and S. Larbi, "Comparative study on machine learning algorithms for early fire forest detection system using geodata," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 5, pp. 5507–5513, 2020.
- [43] B. Thai Pham, A. Jaafari, M. Avand et al., "Performance evaluation of machine learning methods for forest fire modeling and prediction maps provide a basis for developing more efficient fire-fighting strategies and reorganizing policies in favor of sustainable management of forest resources," *Symmetry*, vol. 12, 2020.
- [44] S. K. M. Abujayyab, M. A. S. Ahamad, A. S. Yahya, and A. M. H. Y. Saad, "A new framework for geospatial site selection using artificial neural networks as decision rules: a case study on landfill sites," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-2/W2, pp. 131–138, 2015.
- [45] M. Rodrigues and J. De La Riva, "An insight into machine-learning algorithms to model human-caused wildfire occurrence," *Environmental Modelling & Software*, vol. 57, 2014.

- [46] T. A. Nelson, M. L. Bourbonnais, and M. A. Wulder, *Mapping Forest Landscape Patterns*, Springer, New York, NY, USA, pp. 1–326, 2017.
- [47] P. Jain, S. C. P. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, and M. D. Flannigan, “A review of machine learning applications in wildfire science and management,” *Environmental Reviews*, vol. 28, no. 4, pp. 478–505, 2020.
- [48] Q. B. Pham, Y. Achour, S. A. Ali et al., “A comparison among fuzzy multi-criteria decision making, bivariate, multivariate and machine learning models in landslide susceptibility mapping,” *Geomatics, Natural Hazards and Risk*, vol. 12, no. 1, pp. 1741–1777, 2021.
- [49] M. Ma, G. Zhao, B. He et al., “XGBoost-based method for flash flood risk assessment,” *Journal of Hydrology*, vol. 598, Article ID 126382, 2021.
- [50] E. K. Sahin, “Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping,” *Geocarto International*, vol. 37, 2020.
- [51] Y. Wu, Y. Ke, Z. Chen, S. Liang, H. Zhao, and H. Hong, “Application of alternating decision tree with adaBoost and bagging ensembles for landslide susceptibility mapping,” *Catena*, vol. 187, Article ID 104396, 2020.
- [52] T. Chen, L. Zhu, R. Q. Niu, C. J. Trinder, L. Peng, and T. Lei, “Mapping landslide susceptibility at the Three Gorges Reservoir, China, using gradient boosting decision tree, random forest and information value models,” *Journal of Mountain Science*, vol. 17, no. 3, pp. 670–685, 2020.
- [53] R. Can, S. Kocaman, and C. Gokceoglu, “A comprehensive assessment of XGBoost algorithm for landslide susceptibility mapping in the upper basin of Ataturk dam, Turkey,” *Applied Sciences*, vol. 11, no. 11, p. 4993, 2021.
- [54] Y. Song, R. Niu, S. Xu et al., “Landslide susceptibility mapping based on weighted gradient boosting decision tree in Wanzhou section of the three Gorges reservoir area (China),” *ISPRS International Journal of Geo-Information*, vol. 8, no. 1, p. 4, 2018.
- [55] T. T. Dang, Y. Cheng, J. Mann, K. Hawick, and Q. Li, “Fire risk prediction using multi-source data: a case study in humberside area,” in *Proceedings of the 2019 25th IEEE International Conference on Automation and Computing ICAC*, Lancaster, UK, September 2019.
- [56] F. Zhou, H. Pan, Z. Gao et al., “Fire prediction based on catboost algorithm,” *Mathematical Problems in Engineering*, vol. 2021, Article ID 1929137, 9 pages, 2021.
- [57] D. Rosadi and W. Andriyani, “Prediction of forest fire using ensemble method,” *Journal of Physics: Conference Series*, vol. 1918, no. 4, Article ID 042043, 2021.
- [58] Y. Michael, D. Helman, O. Glickman, D. Gabay, S. Brenner, and I. M. Lensky, “Forecasting fire risk with machine learning and dynamic information derived from satellite vegetation index time-series,” *Science of the Total Environment*, vol. 764, Article ID 142844, 2021.
- [59] İ. Atalay, *Uygulamalı Klimatoloji. Meta Basım*, İzmir: Meta Basım ve Matbaacılık Hizmetleri, izmir, Turkey, 2010.
- [60] T. Varol and M. Ertuğrul, “Climate change and forest fire trend in the Aegean and Mediterranean regions of Turkey,” *Fresenius Environmental Bulletin*, vol. 24, no. 10B, pp. 3436–3444, 2015.
- [61] H. Muminjanov and A. Karagöz, “Biodiversity of turkey: contribution of genetic resources to sustainable agriculture and food systems,” *Food and Agriculture Organization of the United Nations*, 2018, <https://www.fao.org/3/ca1517en/CA1517EN.pdf>.
- [62] L. Nazik, M. Poyraz, and M. Karabiyiçoğlu, “Karstic landscapes and landforms in Turkey BT-landscapes and landforms of Turkey,” in *Landscapes and Landforms of Turkey. World Geomorphological Landscapes*, C. Kuzucuoğlu, A. Çiner, and N. Kazancı, Eds., Springer International Publishing, Cham Switzerland, pp. 181–196, 2019.
- [63] M. Z. Öztürk, “Fluvio-karstic evolution of the taşeli plateau (central Taurus, Turkey),” *Turkish Journal of Earth Sciences*, vol. 29, no. 5, pp. 733–746, 2020.
- [64] Ç. Şenkul and S. Kaya, “Türkiye endemi?k bi?tki?leri?ni?n coğrafi? dagilislari,” *Türk Coğrafya Dergisi*, vol. 69, pp. 109–120, 2017.
- [65] İ. Atalay, *Türkiye'nin Ekolojik Bölgeleri. Meta Basım*, İzmir: Meta Basım ve Matbaacılık Hizmetleri, izmir, Turkey, 2014.
- [66] M. Bugala, Z. Parobeková, J. Parobek, and A. E. Akay, “Influence of the ecological conditions on the productive potential of grey alder,” *Sumarski List*, vol. 141, no. 7–8, pp. 371–379, 2018.
- [67] P. Banerjee, “MODIS-FIRMS and ground-truthing-based wildfire likelihood mapping of Sikkim Himalaya using machine learning algorithms,” *Natural Hazards*, vol. 110, no. 2, pp. 899–935, 2021.
- [68] A. Jaafari, E. K. Zenner, and B. T. Pham, “Wildfire spatial pattern analysis in the Zagros Mountains, Iran: a comparative study of decision tree based classifiers,” *Ecological Informatics*, vol. 43, pp. 200–211, 2018.
- [69] H. V. Le, D. A. Hoang, C. T. Tran et al., “A new approach of deep neural computing for spatial prediction of wildfire danger at tropical climate areas,” *Ecological Informatics*, vol. 63, Article ID 101300, 2021.
- [70] M. A. Storey, M. Bedward, O. F. Price, R. A. Bradstock, and J. J. Sharples, “Derivation of a Bayesian fire spread model using large-scale wildfire observations,” *Environmental Modelling & Software*, vol. 144, Article ID 105127, 2021.
- [71] K. Gholamnia, T. Gudiyangada Nachappa, O. Ghorbanzadeh, and T. Blaschke, “Comparisons of diverse machine learning approaches for wildfire susceptibility mapping,” *Symmetry*, vol. 12, no. 4, pp. 604–620, 2020.
- [72] Z. Wan and J. Dozier, “A generalized split-window algorithm for retrieving land-surface temperature from space,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 4, pp. 892–905, 1996.
- [73] J. Friedman, “Greedy boosting approximation: a gradient boosting machine,” *Annals of Statistics*, vol. 148, pp. 148–162, 2001.
- [74] J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression,” *Annals of Statistics*, vol. 28, no. 2, pp. 337–374, 2000.
- [75] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial,” *Frontiers in Neuroinformatics*, vol. 7, no. DEC, p. 21, 2013.
- [76] T. Chen and C. Guestrin, “XGBoost: a scalable tree boosting system,” *Il Friuli Medico*, vol. 19, no. 6, pp. 785–794, 2016.
- [77] M. Al-kasassbeh, M. A. Abbadi, and A. M. Al-Bustanji, “LightGBM algorithm for malware detection,” *Advances in Intelligent Systems and Computing*, vol. 1230, pp. 391–403, 2020.
- [78] E.-A. Minastireanu and G. Mesnita, “Light GBM machine learning algorithm to online click fraud detection,” *Journal of Information Assurance & Cybersecurity*, vol. 2019, pp. 1–12, 2019.
- [79] G. Ke, Q. Meng, T. Finley et al., “LightGBM: a highly efficient gradient boosting decision tree,” in *Proceedings of the 31st International Conference on Neural Information Processing*

- Systems*, pp. 3147–3155, Long Beach CA USA, December 2017.
- [80] M. Corporation, *LightGBM*, 2021, https://lightgbm.readthedocs.io/_/downloads/en/v3.2.1/pdf/ (Access at 20.12.2022).
- [81] A. V. Dorogush, V. Ershov, and A. Gulin, “CatBoost: gradient boosting with categorical features support,” 2018, http://learningsys.org/nips17/assets/papers/paper_11.pdf.
- [82] J. T. Hancock and T. M. Khoshgoftaar, “CatBoost for big data: an interdisciplinary review,” *Journal of Big Data*, vol. 7, no. 1, p. 94, 2020.
- [83] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “Catboost: unbiased boosting with categorical features,” in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 6638–6648, Montréal, Canada, December 2018.
- [84] C. Tu, H. Liu, and B. Xu, “Adaboost typical algorithm and its application research,” *MATEC Web of Conferences*, vol. 139, 2017.
- [85] Y. Freund and R. E. Schapire, “A decision-theoretic generalization of on-line learning and an application to boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.
- [86] D. J. Fleet and M. Brubaker, *AdaBoost*, pp. 123–129, Department of Computer and Mathematical Sciences University of Toronto Scarborough, Toronto, Canada, 2015.
- [87] A. Jaafari, M. Panahi, B. T. Pham et al., “Meta optimization of an adaptive neuro-fuzzy inference system with grey wolf optimizer and biogeography-based optimization algorithms for spatial prediction of landslide susceptibility,” *Catena*, vol. 175, no. November 2018, pp. 430–445, 2019.
- [88] E. K. Sahin, “Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping,” *Geocarto International*, vol. 37, no. 9, pp. 2441–2465, 2020.
- [89] M. Saber, T. Boulmaiz, M. Guermoui et al., “Examining LightGBM and CatBoost models for wadi flash flood susceptibility prediction,” *Geocarto International*, vol. 0, no. 0, pp. 1–26, 2021.
- [90] W. L. Hakim, A. R. Achmad, and C. W. Lee, “Land subsidence susceptibility mapping in jakarta using functional and meta-ensemble machine learning algorithm based on time-series insar data,” *Remote Sensing*, vol. 12, no. 21, p. 3627, 2020.
- [91] A. Arabameri, S. Chandra Pal, F. Rezaie et al., “Comparison of multi-criteria and artificial intelligence models for land-subsidence susceptibility zonation,” *Journal of Environmental Management*, vol. 284, no. February, Article ID 112067, 2021.