

Article

Mapping Forest Fire Risk Zones Using Machine Learning Algorithms in Hunan Province, China

Chaoxue Tan ¹  and Zhongke Feng ^{1,2,*} ¹ Precision Forestry Key Laboratory of Beijing, Beijing Forestry University, Beijing 100083, China² Intelligent Forestry Key Laboratory of Haikou City, School of Forestry, Hainan University, Haikou 570228, China

* Correspondence: zhongkefeng@bjfu.edu.cn

Abstract: Forest fire is a primary disaster that destroys forest resources and the ecological environment, and has a serious negative impact on the safety of human life and property. Predicting the probability of forest fires and drawing forest fire risk maps can provide a reference basis for forest fire control management in Hunan Province. This study selected 19 forest fire impact factors based on satellite monitoring hotspot data, meteorological data, topographic data, vegetation data, and social and human data from 2010–2018. It used random forest, support vector machine, and gradient boosting decision tree models to predict the probability of forest fires in Hunan Province and selected the RF algorithm to create a forest fire risk map of Hunan Province to quantify the potential forest fire risk. The results show that the RF algorithm performs best compared to the SVM and GBDT algorithms with 91.68% accuracy, 91.96% precision, 92.78% recall, 92.37% F1, and 97.2% AUC. The most important drivers of forest fires in Hunan Province are meteorology and vegetation. There are obvious differences in the spatial distribution of seasonal forest fire risks in Hunan Province, and winter and spring are the seasons with high forest fire risks. The medium- and high-risk areas are mostly concentrated in the south of Hunan.

Keywords: forest fire occurrence probability model; forest fire risk rating; machine learning



Citation: Tan, C.; Feng, Z. Mapping Forest Fire Risk Zones Using Machine Learning Algorithms in Hunan Province, China. *Sustainability* **2023**, *15*, 6292. <https://doi.org/10.3390/su15076292>

Academic Editors: Aqil Tariq and Na Zhao

Received: 6 February 2023

Revised: 16 March 2023

Accepted: 31 March 2023

Published: 6 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As an important component of terrestrial ecosystems, forests are energy reservoirs, gene banks, water reservoirs, and carbon reservoirs on Earth, and play a vital role in maintaining the ecological balance of the planet and improving the ecological environment [1]. Forest fires, as a global phenomenon [2], pose a serious threat to the ecological environment as well as the safety of human life and property [3–6]. In recent years, the frequency of forest fires has increased due to global warming and frequent human activities [7]. Forest fire risk is defined as the likelihood of fire occurrence and its consequences [2]. Forest fire risk assessment is a scientific method to quantify the level of forest fire risk [8]. Forest fire risk level zoning map is an important part of forest fire risk assessment. It is based on the probability of forest fire occurrence at a specific threshold [9] and provides an effective map for resource allocation for forest fire risk management. Therefore, in order to protect forest resources and forest ecosystem functions, it is important to map the forest fire risk level zones for forest fire prevention and control [10].

Forest fires are influenced by a variety of factors [11,12]. With the advancement of forest fire risk prediction studies, at present, most analyses are carried out in a comprehensive manner with multiple factors, such as meteorology, topography, vegetation, and human activities [13–15]. Meteorology is considered to be a determinant of forest fires, which mainly affects forest fires in two ways: by influencing the frequency of forest fire weather and the water content of combustible materials [16]. Differences in topography can influence wind, water, and heat transfer between sites [17]. Furthermore, they have

an impact on the composition of vegetation types as well as the spatial distribution of combustibles, influencing the occurrence and spread of forest fires. Human activities are considered to be a major factor affecting forest fires, including mainly the distance of roads, the distance of settlements, and the location of recreational areas [18]. Vegetation is a source of fuel for forest fires and has a direct impact on their ability to catch fire [19].

Currently, forest fire risk prediction has evolved from statistical analysis models to more sophisticated models [20]. Simple statistical and empirical methods include analytic hierarchy [21], weight of evidence [22], multiple linear regression [23], Poisson regression [24], and logistic regression [25–27]. However, forest fires are a typically non-linear and complex process, and the above methods do not always achieve satisfactory results [28,29]. To address these problems, machine learning has been applied to modeling forest fire occurrence, such as random forests (RFs) [15,19,30], support vector machines (SVMs) [31,32], multilayer perceptron neural networks (MLP) [20,33], artificial neural networks (ANNs) [34–36], and adaptive neuro-fuzzy inference systems (ANFISs) [37,38]. In addition, hybrid and ensemble models have also received attention from researchers in order to obtain higher prediction accuracy. For example, Bui et al. [28] used a novel hybrid AI approach combining a neuro-fuzzy inference system (NF) and particle swarm optimization (PSO) for forest fire risk prediction in Lam Dong Province, Vietnam, and they found that the model performed well on both the training dataset (AUC = 93.2%) and the validation dataset (AUC = 91.6%), and outperformed SVM and RF. Moayedi et al. [39] used ant colony optimization (ACO) and biogeography-based optimization (BBO) algorithms to optimize an ANN for forest fire prediction accuracy, and the results of the study showed that BBO and ACO could improve ANN accuracy from 81.3% to 84.0% and 83.9%, respectively. Jaafari et al. [40] improved the forest fire prediction capability of an ANFIS using an integrated combination of ANFIS with applied genetic algorithm (GA) and firefly algorithms (FA). They demonstrated that these two algorithms improve the prediction efficiency of ANFIS from 77.6% to 88.1% and 90.8%, respectively. Similarly, Moayedi et al. [41] used the three wise meta-heuristics of genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) in combination with ANFIS for forest fire sensitivity mapping in the Golestan Province, Iran. Their study found that GA-ANFIS obtained optimal accuracy on training (AUC = 91.25%) and prediction datasets (AUC = 85.03%). In fact, there is currently no consensus on the choice of a model for forest fire risk prediction. No study has proven that a particular method is applicable to all areas in different environments [42,43]. Different models differ in predicting the probability of forest fire occurrence and mapping forest fire risk zones.

Hunan Province is one of the key forestry provinces in the south of China, rich in forest resources and with a high frequency and intensity of fires. In the past, Guo Haifeng et al. [44] used principal component analysis to establish a weighted forest fire weather index model and to determine forest fire weather classes. Wang Shuang et al. [45] used logistic model to study the forest fire risk occurrence pattern in Hunan Province. However, due to the nonlinear complexity of forest fire occurrence, simple empirical methods or linear regression models can no longer meet the needs of managers. Few studies have compared multiple machine learning methods for mapping seasonal forest fire risk levels in the region. This study utilized random forests, support vector machines, and gradient boosting trees to predict forest fires in Hunan Province. The primary goals of this study are to (1) assess the predictive ability of random forest, gradient boosting tree, and support vector machine in Hunan Province forest fire risk, (2) produce accurate and reliable forest fire risk zoning maps using optimal models, and (3) assess the importance of various influencing factors in forest fire risk prediction.

2. Materials and Methods

2.1. Study Area

Hunan Province is located in the middle reaches of the Yangtze River, between 108°47' and 114°15', 24°38' and 30°08' N, with a total land area of 21.18×10^4 km² (see Figure 1).

The region has a continental subtropical monsoonal humid climate with a warm climate, concentrated rainfall, and abundant light and heat resources, with an average annual temperature between 16 and 18.5 °C and an average annual precipitation between 1200 and 1700 mm [45]. Its topography consists of plains, hilly land, mountains, basins and rivers, and lakes, with mountains and hills dominating, together accounting for 66.62% of the total area. The province has a vast number of planted forests and primarily evergreen broad-leaved forests, mainly concentrated in southwestern, southern, northwestern, and eastern Hunan [44].

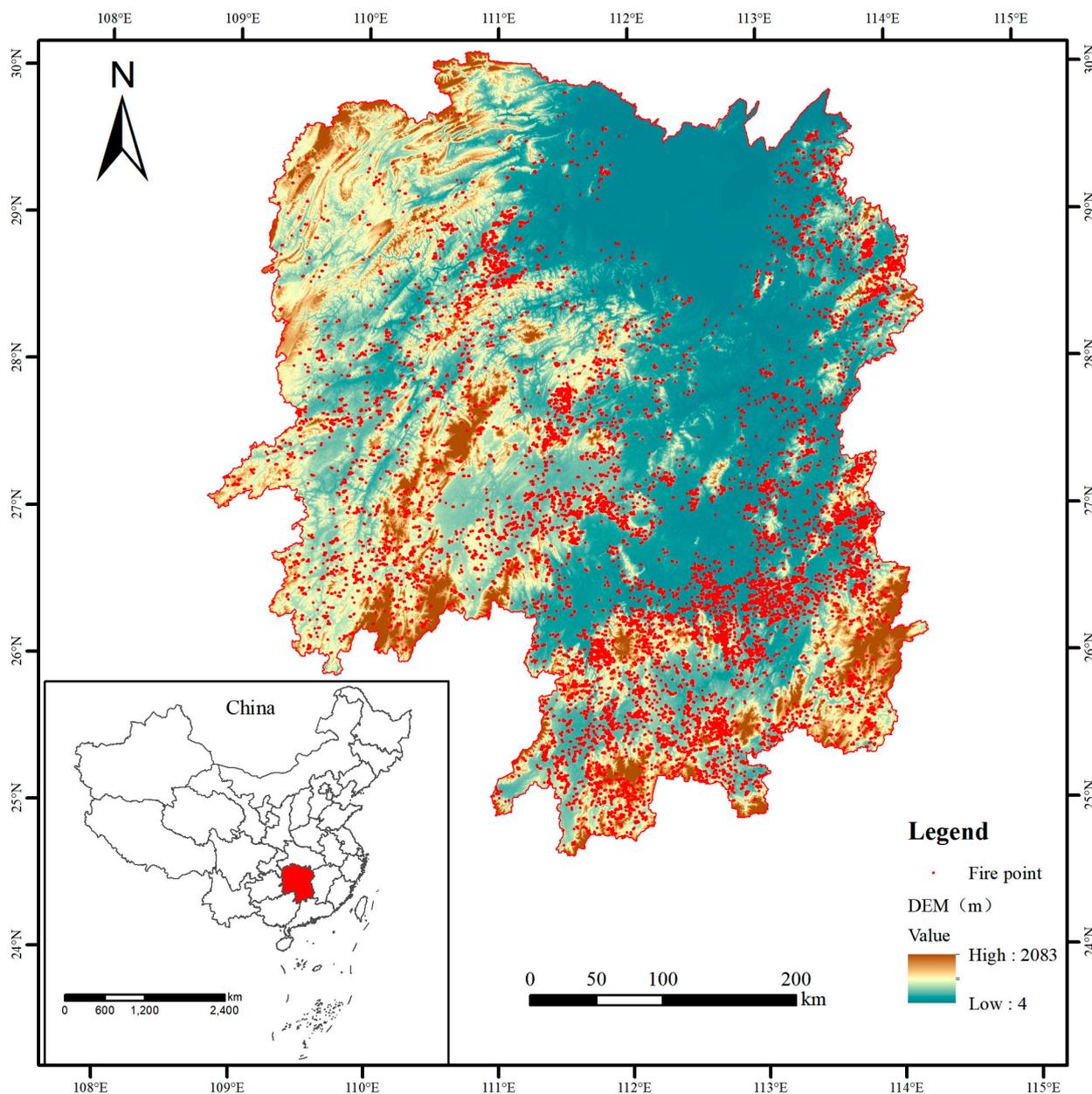


Figure 1. Study area and distribution of provincial forest hotspot data in Hunan Province, 2010–2018.

2.2. Forest Fire Data

The Department of Fire Prevention and Control Management, Ministry of National Emergency Management, China, provided the satellite monitoring hotspot data from 2010 to 2018. Fire point data were obtained from this data, and the abnormal samples in the original dataset were removed before screening the forest fire data with the land type of forest land. When modelling forest fire prediction, because the dependent variable

under study is a binary variable, a certain number of random points need to be created to participate in the modelling as non-fire points. We created random points (non-fire points) at a 1:1 ratio within the forested area of the 30 m surface cover data for Hunan Province in 2020, and performed a 500 m buffer zone analysis on fire points to avoid random points being located at or near the same location as fire points. The fire point is set to 1 and the random point is set to 0. The random point follows the principle of double randomness in time and space [19,30]. The woodland data were obtained from the 30 m global land cover dataset GlobeLand30 from the Global Geographic Information Public Product (<http://www.globallandcover.com/> accessed on 16 January 2022), and the number of fire points and random points were 12,815 and 10,539, respectively.

2.3. Forest Fire Impact Factor Data

The fire variables of the forest fire risk model mainly include meteorology, topography, vegetation, and human activities. We chose 22 forest fire variables as the initial variables influencing the occurrence of forest fires in Hunan Province for this study, and detailed descriptions of the categories are shown in Table 1. In this study, all were continuous variables except for aspect and special festival, which were categorical variables.

Table 1. The initial driving factors of forest fire.

Influencing Factors	Independent Variable	Symbol	References
Location	Longitude (°)	Lon	[31]
	Latitude (°)	Lat	
	Altitude (m)	Alt	[2,5,8,10]
	Slope (°)	Slo	
	Aspect	Asp	
Infrastructure	Closest distance of fire point to residential area(m)	Set	[15,28,46]
	Distance from the fire point to the highway (m)	Hig	
	Nearest distance of fire point to railway (m)	Ral	
Social humanity	Special festival	Spe	[30,31]
	Population	Pop	[18,47,48]
Vegetation	GDP	GDP	
	NDVI	NDVI	[7,32,41]
Meteorology	Average surface temperature (°C)	Ast	[11,19,49]
	Daily maximum surface temperature (°C)	Mast	
	Cumulative precipitation at 20–20 (mm)	Pre	
	Average station pressure (hPa)	Spr	
	Average relative humidity (%)	Arh	
	Minimum relative humidity (%)	Mrh	
	Average temperature (°C)	Ate	
	Daily maximum temperature (°C)	Mate	
Average wind speed (m/s)	Aws		
	Hours of sunshine (h)	Suh	

2.3.1. Meteorological Data

The meteorological data were obtained from the China Meteorological Data Network (<https://data.cma.cn/> accessed on 30 June 2021), which includes daily value dataset (V3.0) of Chinese surface climate data for 8 years from 2010 to 2018. After pre-processing the meteorological data, we finally selected 10 meteorological factors, including daily average surface temperature, daily maximum surface temperature, cumulative precipitation at 20–20 (the 24-h cumulative precipitation from 20:00 pm to 20:00 pm the following day), daily average air pressure, daily average relative humidity, daily minimum relative humidity,

hours of sunshine, daily average temperature, daily maximum temperature, and average wind speed, as the initial forest fire meteorological variables.

2.3.2. Topographic Data

The incidence and spread of forest fires are impacted by topographic variations. Differences in topography play an important role in the composition of vegetation types and the spatial distribution of combustible materials, which directly affect the occurrence and spread of forest fires, for which elevation, slope, and aspect have been widely reported [17]. To collect altitude, slope, and aspect information for Hunan Province, the DEM data with a spatial resolution of 30 m was acquired from the Geospatial Data Cloud website (<http://www.gscloud.cn/> accessed on 16 January 2022). We divided the aspect into nine categories as shown in Table 2.

Table 2. Aspect classification.

Aspect	Aspect Range (Degrees)	Classification	Description
Gentle slope	−1	0	
North	0~22.5/337.5~360	1	
Northeast	22.5~67.5	2	
East	67.5~112.5	3	
Southeast	112.5~157.5	4	
South	157.5~202.5	5	
Southwest	202.5~247.5	6	
West	247.5~292.5	7	
Northwest	292.5~337.5	8	

2.3.3. Vegetation Data

Changes in NDVI values can indicate changes in water and nutrient availability, plant diseases, and other stressors, which in turn are indicators of vegetation vulnerability to fire [50]. As a result, the vegetation data in this study were expressed by NDVI (Normalized Vegetation Index). The Resource Environment Science and Data Center (<http://www.resdc.cn/> accessed on 30 June 2021) provided the spatial distribution dataset of China Quarterly Vegetation Index (NDVI) with a spatial resolution of 1 km. The seasons were separated into four groups based on the vegetation status: spring (March–May), summer (June–August), autumn (September–November), and winter (December–February).

2.3.4. Social and Humanistic Data

The basic geographic data were obtained from the National Basic Geographic Database of 1:250,000 from the website of National Geographic Information Resource Catalog System (<http://www.webmap.cn/> accessed on 30 June 2021). Based on ArcGIS software, the shortest distance from sample points to infrastructures, such as railroads, roads, and settlements, was calculated. Socioeconomic data included population density, gross domestic product (GDP) per capita, and special festival. GDP and population density were uploaded from the National Earth System Science Data Center (<http://www.geodata.cn/> accessed on 30 June 2021) for the 2015 spatial distribution of population and GDP on a kilometer grid with a resolution of 1 km. Since there are certain traditional Chinese holidays where people burn paper money to pay respects to their deceased relatives that may lead to forest fires, Chinese New Year's Eve, the first day of Chinese New Year, the second day of Chinese New Year, the Lantern Festival, the Tomb Sweeping Festival, and the Zhongyuan Festival (i.e., July 15 of the lunar calendar) were set as special festival days and denoted as 1; non-special festival days were denoted as 0.

2.4. Data Processing

2.4.1. Normalization

The magnitudes and magnitude units of the forest fire inciting elements vary, which will have an impact on the analysis of the data. The normalization of the data is required so that each factor is in the same order of magnitude in order to avoid the influence of magnitudes among indicators and the issue of excessive differentiation of output data magnitudes. The normalization formula is as follows.

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where x_i and x_i^* are the values before and after normalization of the data, respectively; and x_{max} and x_{min} are the maximum and minimum values of the sample data, respectively.

2.4.2. Multiple Collinearity Test

The assessment of multicollinearity of the independent variables can provide their respective importance and positional positioning in the optimal model construction [51]. Therefore, the variance inflation factor (VIF) was applied in this study to step out the independent variables with significant covariance. Generally, when $VIF > 10$, it indicates that the independent variables should be excluded because of their significant covariance. Since the diagnosis of multicollinearity is only applicable to continuous variables, not to categorical variables, aspect and special festival did not perform multicollinearity diagnosis, and these two variables entered directly the importance test stage of the model. After the test, after excluding the three variables of daily average surface temperature, daily average temperature, and minimum relative humidity (VIF values of 76.849, 89.026, and 14.605, respectively), the VIF values of the remaining 17 continuous variables were less than 10, and there was no multicollinearity (see Table 3). Finally, 17 continuous variables and the 2 categorical variables of aspect and special festival, a total of 19 feature variables, entered the model fitting stage.

Table 3. The results of the multicollinearity diagnosis.

Independent Variable	VIF
Lon	1.274
Lat	1.329
Alt	1.818
Slo	1.319
Set	1.135
Hig	1.248
Ral	1.095
GDP	3.807
Pop	4.137
NDVI	1.861
Mast	8.634
Pre	1.179
Spr	1.521
Arh	2.283
Suh	2.623
Mate	7.146
Aws	1.216

2.5. Methods

2.5.1. Random Forest

RF is a popular machine learning algorithm proposed by Breiman in 2001 [52], which is an inheritance and improvement of the traditional decision tree, capable of analyzing and evaluating the relative importance of the input factors with high classification accuracy

and computational speed as well as robustness to outliers [47]. The performance of RF is influenced by two important parameters: the number of trees in the forest (ntree) and the number of random variables per split node (mtry). Therefore, these two parameters must be set appropriately beforehand. As ntree is less sensitive to classification accuracy, we set ntree to 500 trees [20] and used five-fold cross-validation to determine the optimal parameters for the model mtry, finally settling on mtry = 4.

2.5.2. Support Vector Machine

SVM is a machine learning method that is applicable to classification and regression. Its basic idea is to maximize the gap between different classes of samples by finding an optimal hyperplane in the feature space as the basis for classification [53]. The prerequisite for classification using SVM is that the training sample space is linearly divisible, but the actual data may be complex. To solve nonlinear problems in classification or regression, kernel functions are introduced into SVM classification methods. Kernel functions can map the original input space to a new feature space, making samples that are otherwise linearly indistinguishable potentially distinguishable in the kernel space. The kernel functions are mainly divided into kernel functions with linear kernel function, polynomial kernel function, radial basis kernel function (RBF), etc. In this study, the parameters of C and g and the optimal model were determined based on grid search, and then the optimal values of C and g were determined to be 100 and 0.01, respectively, and the RBF kernel function was selected to build the model.

2.5.3. Gradient Boosting Decision Tree

Gradient boosting decision tree (GBDT) calculates the residuals between the current output and the true value by each weak learner, and then accumulates the residuals of each weak learner output to reduce the residuals in the training process to achieve the classification goal [54]. The GBDT algorithm has the advantages of high prediction accuracy, robustness, and the ability to handle both continuous and discrete data [55]. The main purpose of the GBDT algorithm is to solve the optimization of the loss function, using the negative gradient of the loss function to fit the residuals of the previous round of weak learners, and the training process can be represented by the following equations [56].

$$f_M(x) = \sum_{m=1}^M T(x, \theta_m) \quad (2)$$

where M is the number of iterations, $T(x, \theta_m)$ is the weak classifier generated at each iteration, and θ_m is the loss function, which can be expressed as:

$$\theta_m = \operatorname{argmin} \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + T(x_i, \theta_m)) \quad (3)$$

where $F_{m-1}(x_i)$ is the current iteration.

This study determined a definite learning rate of 0.1 and a number of weak learners of 190 for the GBDT model through a five-fold cross-validation and grid search.

2.5.4. Model Performance Evaluation

In this study, the classification ability of different machine learning methods was evaluated using five metrics: accuracy, precision, recall, F1 (H-mean), and area under curve (AUC) [31,48]. F1 is used to assess precision and recall. Accuracy is the proportion of correctly classified samples in the total sample, while precision is the proportion of positive samples in the sample that are predicted to be true, and recall is the proportion of positive samples in the sample that are actually true [57]. The relationship between sensitivity and specificity is represented by the receiver operating characteristic curve (ROC), and the area of the lower part of the ROC curve is known as the AUC. This area is frequently used to assess the predictive power of classification models, and the closer its value is to 1, the more

accurate the mode prediction [49]. The accuracy, precision, recall, and F1 can be expressed by the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

where TP (true positive) was predicted by the model as the number of positive samples in the positive category, FP (false positive) was predicted by the model as the number of negative samples in the positive category, TN (true negative) was predicted by the model as the number of negative samples in the negative category, and FN (false negative) was predicted by the model as the number of positive samples in the negative category.

3. Results

In this study, the original sample data were randomly divided into 70% training samples (for models building) and 30% test samples (for models testing).

3.1. Comparison and Validation of the Three Models

This research conducted a grid search and five-fold cross-validation on each classifier to measure the predictive accuracy of the model. Using the training dataset, the final RF, SVM, and GBDT models were trained. Five evaluation metrics are then used to validate the performance of these three machine learning algorithms: accuracy, precision, recall, F1 value, and AUC. The fit results of all three models were good ($AUC > 0.85$), according to the test findings of the validation dataset (see Figures 2 and 3). The higher accuracy value indicates the stronger predictive ability of the model, and the order of accuracy values of the three models was $RF > GBDT > SVM$. Therefore, RF was found to be the best method for predicting forest fire risk in Hunan.

Among the three machine learning algorithms, the RF algorithm performed the best, outperforming the other two algorithms in every evaluation index with 91.68% accuracy, 91.96% precision, 92.78% recall, 92.37% F1, and 97.2% AUC. RF algorithms can tolerate outliers and noise, and have the ability to handle redundant attributes and good generalization [58,59]. This was followed by GBDT with 89.38% accuracy, 88.56% precision, 92.36% recall, 90.42% F1, and 95.83% AUC. SVM had the worst performance with an accuracy of 88.88%, precision of 87.07%, recall of 93.38%, F1 value of 90.11%, and AUC of 95.29%. Although SVM also shows good classification and generalization capabilities, it is very time consuming to calibrate [48,59].

3.2. Importance of Feature Factors

The RF algorithm is able to automatically identify the relative importance of feature variables by mean decrease accuracy [2]; therefore, the importance of the mean decrease accuracy of the 19 drivers was ranked. The results show (see Figure 4) that the importance of average relative humidity is significantly greater than that of the other variables, and its importance is ranked first, followed by the maximum daily temperature and hours of sunshine. Higher temperatures and long periods of sunshine tend to reduce the water content of vegetation and increase the likelihood of forest fires. Among the meteorological factors, the average air pressure has a relatively small effect on the occurrence of forest fires in Hunan Province. In this study, NDVI was the fourth most important factor influencing the occurrence of forest fires in Hunan, followed by the daily maximum surface

temperature and latitude, and the influence of longitude was relatively small compared to latitude. Longitude and latitude reflect, to some extent, differences in forest and tree species categories, as well as differences in the degree of flammability of forests of different tree species. Among the topographical factors, elevation has a greater influence on the occurrence of forest fires in Hunan compared to slope and aspect. The influence of human activities on the occurrence of forest fires in Hunan Province is ranked as follows: GDP, nearest distance of fire point to railway, population density, special festival, closest distance of fire point to residential area, and distance from the fire point to the highway.

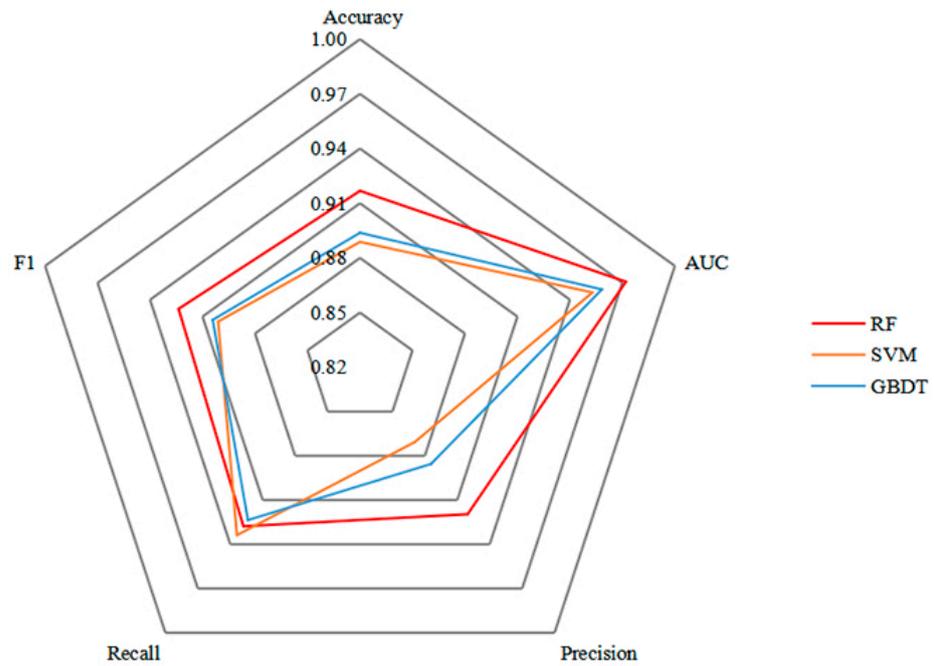


Figure 2. Comparison of the accuracy levels of the 3 machine learning methods.

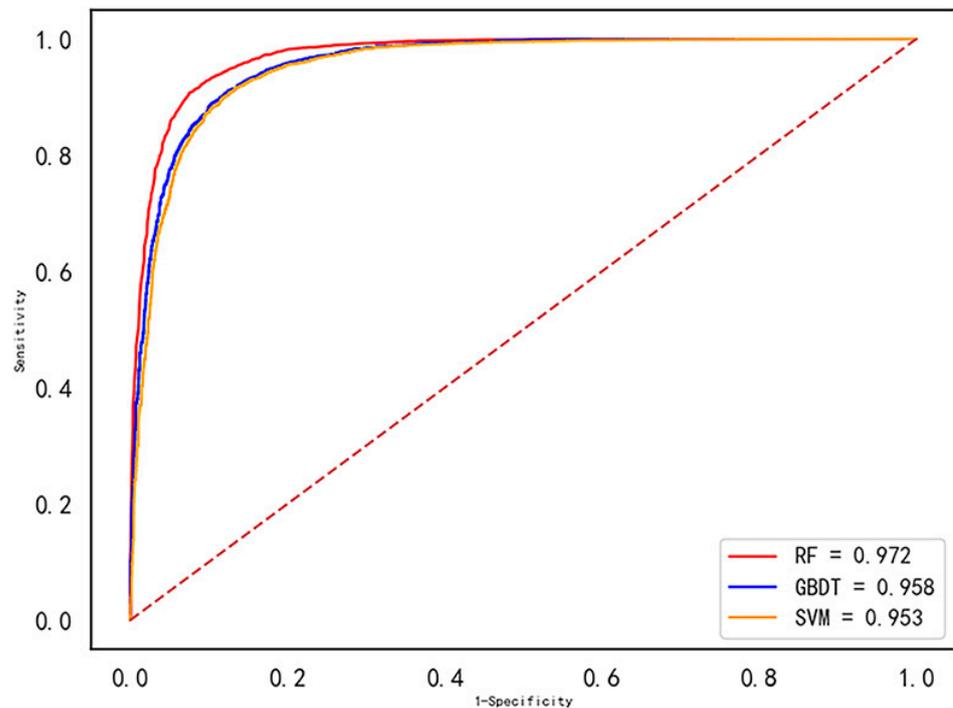


Figure 3. ROC curve.

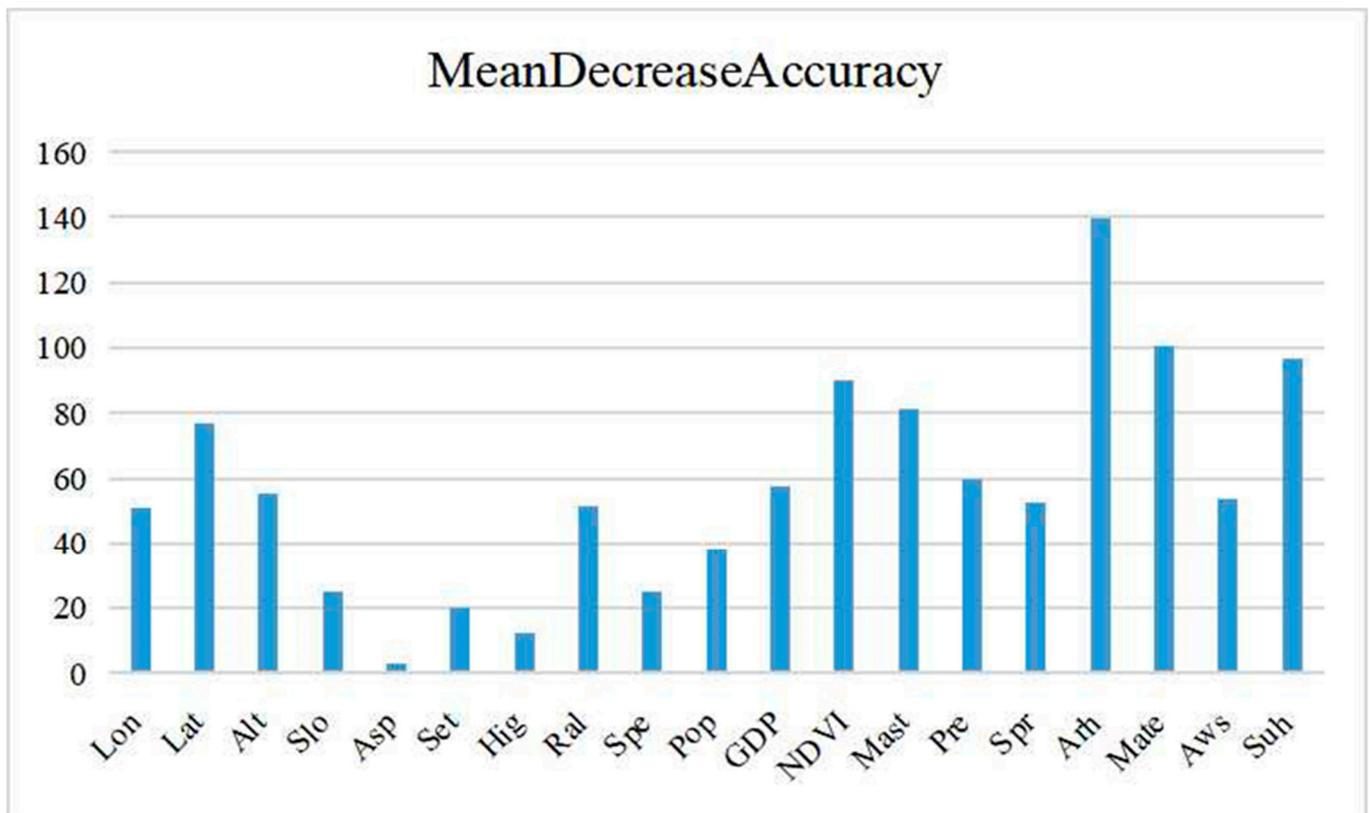


Figure 4. The importance of variables based on the random forest algorithm.

3.3. Seasonal Fire Zoning Map of Hunan Province

The spatial distribution of forest fire occurrence is crucial for forest fire prevention and control as well as fire management. In order to achieve the optimal allocation of firefighting resources, the best performing RF model's forest fire occurrence probability prediction results were selected for this study to map the fire risk in Hunan Province over four seasons. The kriging interpolation method in the ArcGIS 10.4 software was used to interpolate the fire prediction probabilities. In this study, the fire risk zones in Hunan Province were classified into five categories (I–V). I: forest fire probability range of 0.0–0.2 represents the very-low-risk zone, i.e., forest fires are basically unlikely to occur; II: forest fire probability range of 0.2–0.4 represents the low-risk zone, i.e., forest fires are unlikely to occur; III: forest fire probability range of 0.4–0.6 represents the medium-risk zone, i.e., forest fires are likely to occur; IV: forest fire probability range of 0.6–0.8 represents the high-risk zone, i.e., a forest fire is likely to occur; and V: forest fire probability range of 0.8–1.0 represents the very-high-risk zone, i.e., forest fire is very likely to occur.

Figure 5 illustrates the stark differences in the spatial range of the risks of seasonal forest fires in Hunan Province. Among them, winter and spring are the seasons with high forest fire risks. The areas of medium and high fire risks are relatively large, mainly concentrated in southern Hunan. There are relatively few forest fires in autumn and summer. Yongzhou City, Chenzhou City, Hengyang City, Zhuzhou City, the center portion of Loudi City, the southeast of Shaoyang City, and the east and south of Yueyang City are the main locations of the highly high-risk zones in winter. The central and southern parts of Yongzhou City, the eastern part of Shaoyang City, the central part of Loudi City, the southern and eastern parts of Hengyang City, and the northern and central parts of Huaihua City are the main locations of the highly high-risk zones in spring. In the fall, the southeast of Hengyang City, Zhuzhou City, and the center of Yongzhou City are the main distribution points for the extremely high-risk zones. In the fall, the southeast of Hengyang City, Zhuzhou City, and the center of Yongzhou City are the main distribution points for

the extremely high-risk zones. The south of Hengyang City and the east of Shaoyang City are mostly where the severely high-risk zones in the summer are located. The relevant management authorities in Hunan Province should step up their fire prevention efforts in spring and winter, especially in the major cities mentioned above.

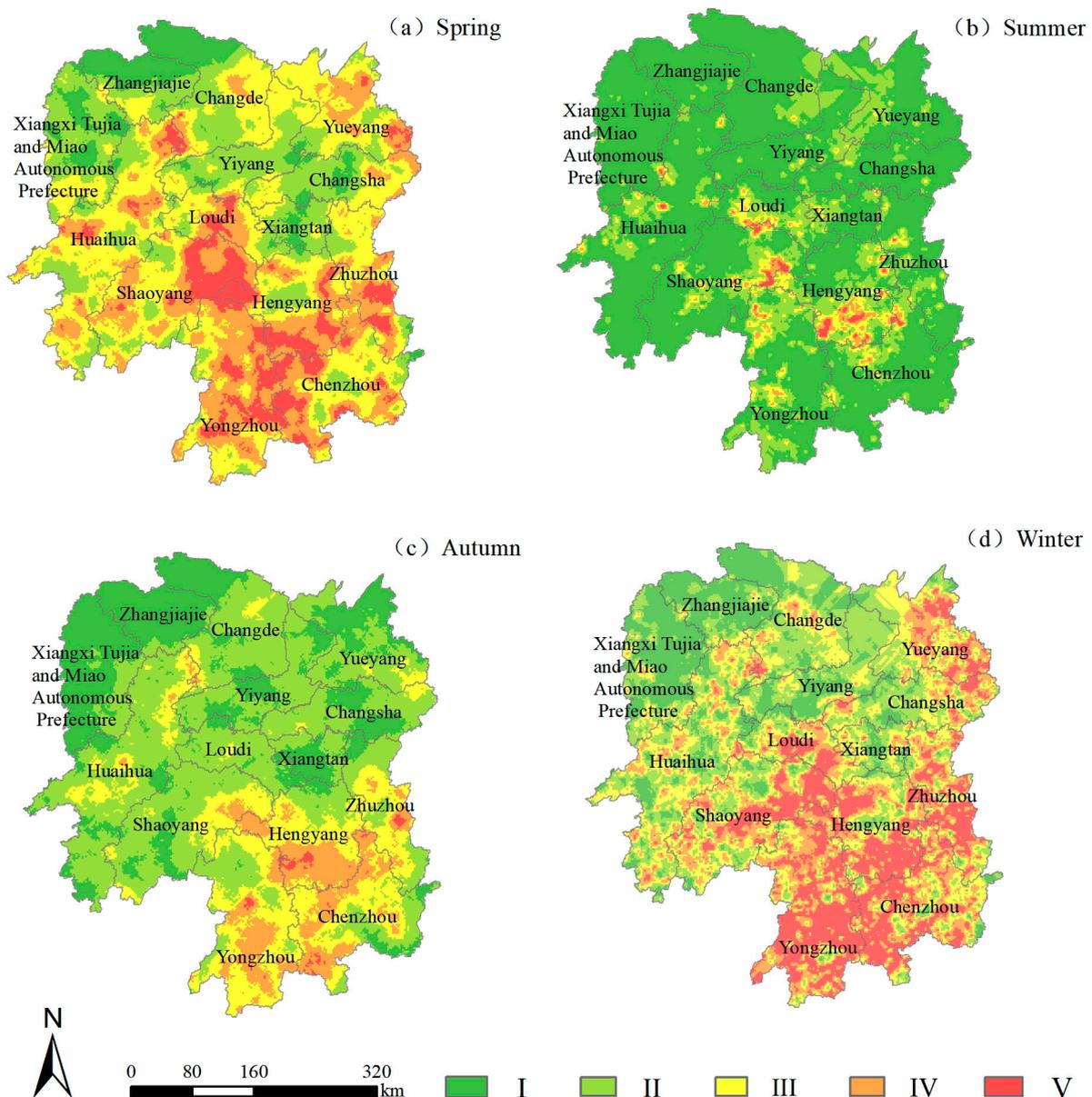


Figure 5. Seasonal forest fire risk zoning maps of Hunan Province ((a) spring, (b) summer, (c) fall, and (d) winter).

4. Discussion

Three machine learning methods (RF, SVM, and GBDT) and 19 forest-fire-driving factors were used in this study to predict the likelihood of forest fire occurrence in Hunan Province. The results demonstrate that all three models are suitable for predicting forest fire occurrence in Hunan Province (prediction accuracy is greater than 85%), but RF has a higher generalization ability than GBDT and SVM. The optimal model's accuracy is 91.68%, precision is 91.96%, recall is 92.78%, F1 is 92.37%, and the AUC is 97.2%, indicating that the stochastic forest model is more appropriate for this assignment. The results can provide a reference for future forest fire modeling in Hunan. RF is able to operate on

large datasets with a large number of feature variables, has a high tolerance to noise and missing data, and can efficiently assess complex interactions and nonlinearities among explanatory variables [2,30]. Due to its powerful functions and high usability, RF has become one of the most popular machine learning methods. SVM has the advantage of nonlinear mapping without excessive interference from noisy data and is not prone to overfitting; however, it requires considerable time to test different kernel functions and model parameters to find the best model, making this approach impractical for dealing with large sample datasets [31,50]. GBDT is an additive model consisting of multiple CART regression trees, which improves the accuracy of prediction by updating the residuals and continuously reducing them with the number of training rounds. This study concluded that SVM is less suitable for predicting forest fire incidence in Hunan compared to the other two methods, mainly because it is difficult to calibrate, too time-consuming to optimize, and does not achieve the accuracy of RF or GBDT. Because each model's prediction accuracy depends heavily on the input data and the adjusted parameters, different results may be achieved for various study areas and datasets [48,49]. In addition, we compared a number of other studies related to forest fire risk prediction in Hunan Province and found that the optimal model in this study, RF, achieved a high prediction accuracy, as detailed in Table 4.

Table 4. Comparison of the prediction accuracy of some relevant studies on forest fire risk in Hunan Province.

Investigator	Method Description	Impact Factor	Precision
Guo et al. [44]	Combined with the principal component analysis method, a weighted forest fire risk weather index model was established to determine the forest fire risk weather level according to the weather index.	Meteorology (5 factors)	AUC = 74.2%
Wang et al. [45]	The logistic model was used to predict the probability of forest fire risk to classify the forest fire risk level in Hunan Province.	Meteorology, vegetation, topography, social/humanity (7 factors)	AUC = 77.9%
Yang et al. [60]	Construction of the Maxent wildfire risk assessment model using GIS to analyze the contribution, importance, and response of environmental variables to wildfire in Hunan Province.	Meteorology, vegetation, topography, social/humanity (12 factors)	AUC = 80.2%
This study	This study used random forest, support vector machine, and gradient boosting tree for forest fire prediction in Hunan Province and selected the optimal model to map the seasonal forest fire risk level in the region.	Meteorology, vegetation, topography, social/humanity (19 factors)	AUC = 97.2%

The analysis of the significance of the characteristic factors revealed that slope, among the topographic factors, has the least influence on the incidence of forest fires in Hunan Province, followed by human activity infrastructures. However, elevation, among the topographic factors, has a greater influence on forest fire occurrence in Hunan. The higher the elevation, the higher the relative humidity and the less likely fires are to occur [61,62], while the lower the elevation, the higher the human accessibility and the more significant the human activities; thus, the more likely forest fires are to occur [48]. GDP and population density in human activities have a greater influence on the occurrence of forest fires in Hunan Province, most likely because the development of the forestry industry is closely linked to human activities, and an increase in population density and GDP promotes the occurrence of forest fires [63,64]. Meteorological factors and vegetation factors are the most important influencing factors of forest fires in Hunan Province. There is a great deal of weight assigned to meteorological elements, including mean relative humidity,

daily maximum temperature, and sunlight hours. The environment's water and heat conditions as well as the moisture content of forest fuels are all impacted by meteorological elements [65,66], which is one of the main causes of forest fires. Among the vegetation factors, NDVI has an important influence. Although longitude and latitude reflect the differences in forest and tree species categories to some extent, they have less influence relative to NDVI, mainly because NDVI can directly reflect the amount of fuel for forest fires to occur, and the amount of fuel directly determines the fire capacity of the forest. Although latitude and longitude and NDVI reflect the situation of combustibles to a certain extent, they have certain limitations. Soil moisture affects the physiological activity of vegetation and is related to vegetation water content and soil moisture. Therefore, in the next work, we hope to incorporate accurate combustible material data and soil moisture data for forest fire prediction modeling.

The Hunan Province's forest fire risk level map indicates that the majority of Hunan's medium- and high-risk areas for forest fires are located in the southern portion of the province, especially in Hengyang City, Shaoyang City, Yongzhou City, Chenzhou City, central Loudi City, and southern Zhuzhou, which are essential locations for monitoring and forecasting forest fires, which is largely consistent with previous studies [45]. The main reason for this situation is that, due to the influence of monsoons and the topography, Hunan is under the influence of winter wind in winter, and the geomorphological characteristics of being surrounded by mountains in the southeast and west and open to the north are conducive to the long drive of cold air, with the general trend of temperature distribution being high in the south and low in the north. In addition, Hunan's forest resources are mainly distributed in the south, and the northern region has a wide water area with small forest coverage and fewer fires. The eastern part of Yueyang City (Linxiang City and Pingjiang County) in this study is also an area with a high incidence of forest fires, probably because of the developed man-made infrastructures and large population in the area, and the large forested area with poor fire resistance of forest tree species. It is advised to raise investments to enhance comprehensive prevention and control of forest fires and to safeguard the security of forest resources in Hunan Province due to the risk of forest fires listed above in high-risk areas. Hunan Province experiences a large number of forest fires in the winter and spring, so the forestry department should concentrate its fire prevention efforts during these two seasons. Hunan is under the control of the winter monsoon in winter, and the winter is dry, which increases the risk of forest fires. In spring, the temperature rises, human activities such as agricultural production and Qingming festival sacrifices increase, and the frequency of forest fires is high. Therefore, the forestry department should strengthen the management of human activities and fire prevention publicity and education, and raise awareness of forest fire prevention among all people. In addition, when a forest fire occurs, structural changes occur in the forest ecosystem. How it recovers is a complex process that includes a series of events, actions or changes, and the role of humans [67]. Resilience can be seen as a key parameter in decision-making processes [68], such as event mitigation following forest fires. In order to minimize the impact of forest fires and reduce the recovery period, resilience in forest fire-prone areas needs to be assessed. In future research, we hope to explore forest fire resilience in order to aid the decision-making processes of local management bodies.

5. Conclusions

Accurately predicting the probability of forest fires and mapping scientific forest fire risk levels can help forestry management departments to make scientific and effective forest resource management decisions. In order to achieve these goals, this study used the 2010–2018 satellite monitoring hotspot data provided by the Department of Fire Prevention and Control Management, Ministry of National Emergency Management, China, taking into account meteorological, terrain, vegetation, and socio-human factors, and using three machine learning methods (RF, SVM, and GBDT) to evaluate and map forest fire risk zones. The model performance comparison showed that the RF model was more suitable for

forest fire occurrence prediction in Hunan Province, with the optimal model having 91.68% accuracy, 91.96% precision, 92.78% recall, 92.37% F1, and 97.2% AUC. In addition, the main characteristic factors of forest fires in Hunan Province were meteorological factors and vegetation factors by RF importance ranking. The drawn forest fire risk level zoning map showed that there are obvious differences in the spatial distribution of seasonal forest fire risk in Hunan Province, among which winter and spring are the seasons with high forest fire risk. The high-risk area of forest fires is mainly concentrated in the south of Hunan Province, and the prevention of forest fires should focus on these areas, and the authorities at all levels should develop scientific management strategies and make reasonable emergency resource allocation according to the local conditions. The results of this study can provide some reference basis for future forest fire management and prevention and control in Hunan Province.

Since the geographical distribution of forest fires and their influencing factors is highly heterogeneous in space, the relationship between them has significant spatial instability [69]. Therefore, in future work, we will consider adding a geographically weighted regression model for comparative studies, which incorporates spatial location information in the regression parameters and is capable of conducting the spatial analysis of the influencing factors and spatial prediction of forest fires. In addition, in future studies, we expect to add more accurate combustible data, soil moisture data, and different types of socio-economic factors to better support forest fire risk assessment in Hunan Province.

Author Contributions: Conceptualization, C.T. and Z.F.; Data curation, C.T.; Formal analysis, C.T.; Software, C.T.; Validation, C.T. and Z.F.; Visualization C.T.; Methodology, C.T. and Z.F.; Writing—original draft preparation, C.T.; Writing—review and editing, C.T. and Z.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Beijing Natural Science Foundation (8232038, 8234065), the Key R & D Projects in Hainan Province (ZDYF2021SHFZ256), and Natural Science Foundation of Hainan University, grant number KYQD (ZR) 21115.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The fire point data used in this study were obtained from the Department of Fire Prevention and Control Management, Ministry of National Emergency Management, China, and are available from the corresponding authors upon request. Meteorological data were obtained from the China Meteorological Data Network (<https://data.cma.cn/> accessed on 30 June 2021). DEM was taken from the Geospatial Data Cloud (<http://www.gscloud.cn/> accessed on 16 January 2022). The China Quarterly Vegetation Index (NDVI) spatial distribution dataset was obtained from the Resource Environment Science and Data Center (<http://www.resdc.cn/> accessed on 30 June 2021). Forest land data were obtained from the Global Geographic Information Public Product (<http://www.globallandcover.com/> accessed on 16 January 2022). GDP and population data were obtained from the National Center for Earth System Science and Data (<http://www.geodata.cn/> accessed on 30 June 2021). Road and residential datasets were obtained from the National Geographic Information Resource Catalog System (<https://www.webmap.cn> accessed on 30 June 2021).

Acknowledgments: We would like to thank all the faculty and students who contributed to this study, as well as the anonymous reviewers for their helpful comments to improve this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Feng, J.G.; Ding, L.B.; Wang, J.S.; Yao, P.P.; Yao, S.C.; Wang, Z.K. Case-based evaluation of forest ecosystem service function in China. *Chin. J. Appl. Ecol.* **2016**, *5*, 1375–1382.
2. Milanović, S.; Marković, N.; Pamučar, D.; Gigović, L.; Kostić, P.; Milanović, S.D. Forest Fire Probability Mapping in Eastern Serbia: Logistic Regression versus Random Forest Method. *Forests* **2021**, *12*, 5. [[CrossRef](#)]
3. Hering, A.S.; Bell, C.L.; Genton, M.G. Modeling spatio-temporal wildfire ignition point patterns. *Environ. Ecol. Stat.* **2009**, *16*, 225–250. [[CrossRef](#)]

4. Modugno, S.; Balzter, H.; Cole, B.; Borrelli, P. Mapping regional patterns of large forest fires in Wildland–Urban Interface areas in Europe. *J. Environ. Manag.* **2016**, *172*, 112–126. [[CrossRef](#)]
5. Deng, O.; Li, Y.Q.; Feng, Z.K.; Dong, Z.Y. Model and zoning of forest fire risk in Heilongjiang province based on spatial Logistic. *Trans. Chin. Soc. Agric. Eng.* **2012**, *28*, 200–205.
6. Argañaraz, J.P.; Radeloff, V.C.; Bar-Massada, A.; Gavier-Pizarro, G.I.; Scavuzzo, C.M.; Bellis, L.M. Assessing wildfire exposure in the Wildland-Urban Interface area of the mountains of central Argentina. *J. Environ. Manag.* **2017**, *196*, 499–510. [[CrossRef](#)]
7. Zhang, G.L.; Wang, M.; Liu, K. Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China. *Int. J. Disast. Risk Sc.* **2019**, *10*, 386–403. [[CrossRef](#)]
8. Naderpour, M.; Rizeei, H.M.; Ramezani, F. Forest Fire Risk Prediction: A Spatial Deep Neural Network-Based Framework. *Remote Sens.* **2021**, *13*, 2513. [[CrossRef](#)]
9. Naderpour, M.; Rizeei, H.M.; Khakzad, N.; Pradhan, I. Forest Fire Induced Natech Risk Assessment: A Survey of Geospatial Technologies. *Reliab. Eng. Syst. Safe* **2019**, *191*, 106558. [[CrossRef](#)]
10. Mohajane, M.; Costache, R.; Karimi, F.; Bao Pham, Q.; Essahlaoui, A.; Nguyen, H.; Laneve, G.; Oudija, F. Application of remote sensing and machine learning algorithms for forest fire mapping in a Mediterranean area. *Ecol. Indic.* **2021**, *129*, 1–17. [[CrossRef](#)]
11. Guo, F.T.; Wang, G.Y.; Su, Z.W.; Liang, H.L.; Wang, W.H.; Lin, F.F.; Liu, A.Q. What drives forest fire in Fujian, China? Evidence from logistic regression and Random Forests. *Int. J. Wildland Fire* **2016**, *25*, 505. [[CrossRef](#)]
12. Ganteaume, A.; Camia, A.; Jappiot, M.; San-Miguel-Ayanz, J.; Long-Fournel, M.; Lampin, C. A Review of the Main Driving Factors of Forest Fire Ignition Over Europe. *Environ. Manag.* **2013**, *51*, 651–662. [[CrossRef](#)]
13. Guo, F.T.; Su, Z.W.; Wang, G.Y.; Sun, L.; Tigabu, M.; Yang, X.J.; Hu, H.Q. Understanding fire drivers and relative impacts in different Chinese forest ecosystems. *Sci. Total Environ.* **2017**, *605–606*, 411–425. [[CrossRef](#)]
14. Su, Z.W.; Zheng, L.J.; Luo, S.S.; Tigabu, M.; Guo, F.T. Modeling wildfire drivers in Chinese tropical forest ecosystems using global logistic regression and geographically weighted logistic regression. *Nat. Hazards* **2021**, *108*, 1317–1345. [[CrossRef](#)]
15. 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. [[CrossRef](#)]
16. Li, S.; Wu, Z.W.; Liang, Y.; He, H.S. A Review of Fire Controlling Factors and Their Dynamics in Boreal Forests. *World For. Res.* **2017**, *30*, 41–45.
17. Wu, Z.C.; Li, M.Z.; Wang, B.; Quan, Y.; Liu, J.Y. Using Artificial Intelligence to Estimate the Probability of Forest Fires in Heilongjiang, Northeast China. *Remote Sens.* **2021**, *13*, 1813. [[CrossRef](#)]
18. Maingi, J.K.; Henry, M.C. Factors influencing wildfire occurrence and distribution in eastern Kentucky, USA. *Int. J. Wildland Fire* **2007**, *16*, 23. [[CrossRef](#)]
19. Ma, W.Y.; Feng, Z.K.; Cheng, Z.X.; Chen, S.L.; Wang, F.G. Identifying Forest Fire Driving Factors and Related Impacts in China Using Random Forest Algorithm. *Forests* **2020**, *11*, 507. [[CrossRef](#)]
20. Nguyen, N.T.; Dang, B.N.; Pham, X.; Nguyen, H.; Bui, H.T.; Hoang, N.; Bui, D.T. Spatial pattern assessment of tropical forest fire danger at Thuan Chau area (Vietnam) using GIS-based advanced machine learning algorithms: A comparative study. *Ecol. Inf.* **2018**, *46*, 74–85.
21. Akay, A.E.; Şahin, H. Forest Fire Risk Mapping by using GIS Techniques and AHP Method: A Case Study in Bodrum (Turkey). *Eur. J. For. Eng.* **2019**, *5*, 25–35. [[CrossRef](#)]
22. Amatulli, G.; Pérez-Cabello, F.; de la Riva, J. Mapping lightning/human-caused wildfires occurrence under ignition point location uncertainty. *Ecol. Model.* **2007**, *200*, 321–333. [[CrossRef](#)]
23. Oliveira, S.; Oehler, F.; San-Miguel-Ayanz, J.; Camia, A.; Pereira, J.M.C. Modeling spatial patterns of fire occurrence in Mediterranean Europe using Multiple Regression and Random Forest. *For. Ecol. Manag.* **2012**, *275*, 117–129. [[CrossRef](#)]
24. Wotton, B.M.; Martell, D.L.; Logan, K.A. Climate Change and People-Caused Forest Fire Occurrence in Ontario. *Clim. Chang.* **2003**, *60*, 275–295. [[CrossRef](#)]
25. Mohammadi, F.; Bavaghar, M.P.; Shabanian, N. Forest Fire Risk Zone Modeling Using Logistic Regression and GIS: An Iranian Case Study. *Small-Scale* **2014**, *13*, 117–125. [[CrossRef](#)]
26. Pan, J.H.; Wang, W.G.; Li, J.F. Building probabilistic models of fire occurrence and fire risk zoning using logistic regression in Shanxi Province, China. *Nat. Hazards* **2016**, *81*, 1879–1899. [[CrossRef](#)]
27. Chang, Y.; Zhu, Z.; Bu, R.; Chen, H.; Feng, Y.; Li, Y.; Hu, Y.; Wang, Z. Predicting fire occurrence patterns with logistic regression in Heilongjiang Province, China. *Landsc. Ecol.* **2013**, *28*, 1989–2004. [[CrossRef](#)]
28. Tien Bui, D.; Bui, Q.; Nguyen, Q.; Pradhan, B.; Nampak, H.; Trinh, P.T. A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area. *Agr. For. Meteorol.* **2017**, *233*, 32–44. [[CrossRef](#)]
29. Tien Bui, D.; Le, K.; Nguyen, V.; Le, H.; Revhaug, I. Tropical Forest Fire Susceptibility Mapping at the Cat Ba National Park Area, Hai Phong City, Vietnam, Using GIS-Based Kernel Logistic Regression. *Remote Sens.* **2016**, *8*, 347. [[CrossRef](#)]
30. Ma, W.Y.; Feng, Z.K.; Cheng, Z.X.; Wang, F.G. Study on driving factors and distribution pattern of forest fires in Shanxi province. *J. Cent. South Univ. For. Technol.* **2020**, *40*, 57–69.
31. Li, Y.D.; Feng, Z.K.; Chen, S.L.; Zhao, Z.Y.; Wang, F.G. Application of the Artificial Neural Network and Support Vector Machines in Forest Fire Prediction in the Guangxi Autonomous Region, China. *Discret. Dyn. Nat. Soc.* **2020**, *2020*, 1–14. [[CrossRef](#)]

32. Hong, H.Y.; Tsangaratos, P.; Ilija, I.; Liu, J.Z.; Zhu, A.; Xu, C. Applying genetic algorithms to set the optimal combination of forest fire related variables and model forest fire susceptibility based on data mining models. The case of Dayu County, China. *Sci. Total Environ.* **2018**, *630*, 1044–1056. [[CrossRef](#)] [[PubMed](#)]
33. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Aryal, J. Forest Fire Susceptibility and Risk Mapping Using Social/Infrastructural Vulnerability and Environmental Variables. *Fire* **2019**, *2*, 50. [[CrossRef](#)]
34. Satir, O.; Berberoglu, S.; Donmez, C. Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. *Geomat. Nat. Hazards Risk* **2016**, *7*, 1645–1658. [[CrossRef](#)]
35. Wang, D.D.; Rong, H.J.; Liu, W.; Meng, Q.L.; Tian, P.F. The Prediction of the Forest Fire Based on the Artificial Neural Network. *J. Northwest For. Univ.* **2010**, *25*, 143–146.
36. Bisquert, M.; Caselles, E.; Sánchez, J.M.; Caselles, V. Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data. *Int. J. Wildland Fire* **2012**, *21*, 1025. [[CrossRef](#)]
37. Razavi-Termeh, S.V.; Sadeghi-Niaraki, A.; Choi, S. Ubiquitous GIS-Based Forest Fire Susceptibility Mapping Using Artificial Intelligence Methods. *Remote Sens.* **2020**, *12*, 1689. [[CrossRef](#)]
38. Mabdeh, A.N.; Al-Fugara, A.K.; Khedher, K.M.; Mabdeh, M.; Al-Shabeeb, A.R.; Al-Adamat, R. Forest Fire Susceptibility Assessment and Mapping Using Support Vector Regression and Adaptive Neuro-Fuzzy Inference System-Based Evolutionary Algorithms. *Sustainability* **2022**, *14*, 9446. [[CrossRef](#)]
39. Moayedi, H.; Khasmakhi, M.A.S.A. Wildfire susceptibility mapping using two empowered machine learning algorithms. *Stoch. Environ. Res. Risk A* **2023**, *37*, 49–72. [[CrossRef](#)]
40. Jaafari, A.; Razavi Termeh, S.V.; Bui, D.T. Genetic and firefly metaheuristic algorithms for an optimized neuro-fuzzy prediction modeling of wildfire probability. *J. Environ. Manag.* **2019**, *243*, 358–369. [[CrossRef](#)]
41. Moayedi, H.; Mehrabi, M.; Bui, D.T.; Pradhan, B.; Foong, L.K. Fuzzy-metaheuristic ensembles for spatial assessment of forest fire susceptibility. *J. Environ. Manag.* **2020**, *260*, 109867. [[CrossRef](#)] [[PubMed](#)]
42. Pham, B.T.; Jaafari, A.; Avand, M.; Al-Ansari, N.; Dinh Du, T.; Yen, H.P.H.; Phong, T.V.; Nguyen, D.H.; Le, H.V.; Mafi-Gholami, D.; et al. Performance Evaluation of Machine Learning Methods for Forest Fire Modeling and Prediction. *Symmetry-Bp.* **2020**, *12*, 1022. [[CrossRef](#)]
43. Tuyen, T.T.; Jaafari, A.; Yen, H.P.H.; Nguyen-Thoi, T.; Phong, T.V.; Nguyen, H.D.; Van Le, H.; Phuong, T.T.M.; Nguyen, S.H.; Prakash, I.; et al. Mapping forest fire susceptibility using spatially explicit ensemble models based on the locally weighted learning algorithm. *Ecol. Inform.* **2021**, *63*, 101292. [[CrossRef](#)]
44. Guo, H.F.; Yu, W. Study weather grade prediction model of forest-fire risk in Hunan province. *J. Cent. South Univ. For. Technol.* **2016**, *36*, 44–47.
45. Wang, S.; Zhang, G.; Tan, S.Q.; Wang, P.; Wu, X. Assessment of forest fire risk in Hunan province based on spatial logistic model. *J. Cent. South Univ. For. Technol.* **2020**, *40*, 88–95.
46. Guo, F.; Innes, J.L.; Wang, G.; Ma, X.; Sun, L.; Hu, H.; Su, Z. Historic distribution and driving factors of human-caused fires in the Chinese boreal forest between 1972 and 2005. *J. Plant Ecol.* **2015**, *8*, 480–490. [[CrossRef](#)]
47. Su, Z.W.; Hu, H.Q.; Wang, G.Y.; Ma, Y.F.; Yang, X.J.; Guo, F.T. Using GIS and Random Forests to identify fire drivers in a forest city, Yichun, China. *Geomat. Nat. Hazards Risk* **2018**, *9*, 1207–1229. [[CrossRef](#)]
48. Shao, Y.K.; Feng, Z.K.; Sun, L.H.; Yang, X.H.; Li, Y.D.; Xu, B.; Chen, Y. Mapping China's Forest Fire Risks with Machine Learning. *Forests* **2022**, *13*, 856. [[CrossRef](#)]
49. Su, Z.W.; Liu, A.Q.; Guo, F.T.; Liang, H.L.; Wang, W.H.; Lin, F.F. Driving factors and spatial distribution pattern of forest fire in Fujian Province. *J. Nat. Disasters* **2016**, *25*, 110–119.
50. Bajocco, S.; Dragoz, E.; Gitas, I.; Smiraglia, D.; Salvati, L.; Ricotta, C. Mapping Forest Fuels through Vegetation Phenology: The Role of Coarse-Resolution Satellite Time-Series. *PLoS ONE* **2015**, *10*, e119811. [[CrossRef](#)]
51. Kalantar, B.; Ueda, N.; Idrees, M.O.; Janizadeh, S.; Ahmadi, K.; Shabani, F. Forest Fire Susceptibility Prediction Based on Machine Learning Models with Resampling Algorithms on Remote Sensing Data. *Remote Sens.* **2020**, *12*, 3682. [[CrossRef](#)]
52. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–23. [[CrossRef](#)]
53. Xu, Z.Q.; Su, X.Y.; Zhang, Y. Forest Fire Prediction Based on Support Vector Machine. *Chin. Agric. Sci. Bull.* **2012**, *28*, 126–131.
54. Tian, Z.Y.; Xiao, J.L.; Feng, H.N.; Wei, Y.T. Credit Risk Assessment based on Gradient Boosting Decision Tree. *Procedia Comput. Sci.* **2020**, *174*, 150–160. [[CrossRef](#)]
55. Ma, L.F.; Xiao, H.M.; Tao, J.W.; Su, Z.X. Intelligent lithology classification method based on GBDT algorithm. *Pet. Geol. Recovery Effic.* **2022**, *29*, 21–29.
56. Rong, G.; Alu, S.; Li, K.; Su, Y.; Zhang, J.; Zhang, Y.; Li, T. Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China. *Water-Sui.* **2020**, *12*, 3066. [[CrossRef](#)]
57. Hu, J.; Min, J. Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model. *Cogn. Neurodynamics* **2018**, *12*, 431–440. [[CrossRef](#)]
58. Abid, F. A Survey of Machine Learning Algorithms Based Forest Fires Prediction and Detection Systems. *Fire Technol.* **2021**, *57*, 559–590. [[CrossRef](#)]
59. Xie, Y.; Peng, M. Forest fire forecasting using ensemble learning approaches. *Neural Comput. Appl.* **2019**, *31*, 4541–4550. [[CrossRef](#)]

60. Yang, X.; Jin, X.; Zhou, Y. Wildfire Risk Assessment and Zoning by Integrating Maxent and GIS in Hunan Province, China. *Forests* **2021**, *12*, 1299. [[CrossRef](#)]
61. Han, J.; Shen, Z.; Ying, L.; Li, G.; Chen, A. Early post-fire regeneration of a fire-prone subtropical mixed Yunnan pine forest in Southwest China: Effects of pre-fire vegetation, fire severity and topographic factors. *Forest Ecol. Manag.* **2015**, *356*, 31–40. [[CrossRef](#)]
62. Abdollahi, M.; Dewan, A.; Hassan, Q. Applicability of Remote Sensing-Based Vegetation Water Content in Modeling Lightning-Caused Forest Fire Occurrences. *Isprs Int. J. Geo.-Inf.* **2019**, *8*, 143. [[CrossRef](#)]
63. Syphard, A.D.; Radeloff, V.C.; Keeley, J.E.; Hawbaker, T.J.; Clayton, M.K.; Stewart, S.I.; Hammer, R.B. Human influence on California fire regimes. *Ecol. Appl.* **2007**, *17*, 1388–1402. [[CrossRef](#)] [[PubMed](#)]
64. Pereira, M.G.; Malamud, B.D.; Trigo, R.M.; Alves, P.I. The history and characteristics of the 1980–2005 Portuguese rural fire database. *Nat. Hazard Earth Sys.* **2011**, *11*, 3343–3358. [[CrossRef](#)]
65. Rollins, M.G.; Morgan, P.; Swetnam, T. Landscape-scale controls over 20th century fire occurrence in two large Rocky Mountain (USA) wilderness areas. *Landsc. Ecol.* **2002**, *17*, 539–557. [[CrossRef](#)]
66. Wu, Z.C.; Li, M.Z.; Wang, B.; Tian, Y.P.; Quan, Y.; Liu, J.Y. Analysis of Factors Related to Forest Fires in Different Forest Ecosystems in China. *Forests* **2022**, *13*, 1021. [[CrossRef](#)]
67. Forcellini, D. A Resilience-Based Methodology to Assess Soil Structure Interaction on a Benchmark Bridge. *Infrastructures* **2020**, *5*, 90. [[CrossRef](#)]
68. Scott, B.M.; Stephanie, E.C. ResilUS: A Community Based Disaster Resilience Model. *Cart. Geogr. Inf. Sc.* **2011**, *38*, 36–51.
69. 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.-UK* **2022**, *12*, 19029. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.