Article

Integrating Multi-Source Remote Sensing Data for Forest Fire Risk Assessment

Xinzhu Liu [1], Change Zheng [1,*], Guangyu Wang [2], Fengjun Zhao [3,*], Ye Tian [1] and Hongchen Li [1]

1 School of Technology, Beijing Forestry University, Beijing 100083, China; lxz0327@bjfu.edu.cn (X.L.)
2 Heilongjiang Ecological Engineering Vocational College, Harbin 150025, China; nimi416@163.com
3 Key Laboratory of Forest Protection of National Forestry and Grassland Administration, Ecology and Nature Conservation Institute, Chinese Academy of Forestry, Beijing 100091, China
* Correspondence: zhengchange@bjfu.edu.cn (C.Z.); zhaofj@caf.ac.cn (F.Z.)

Abstract: Forest fires are a frequent and destructive phenomenon in Southwestern China, posing significant threats to ecological systems and human lives and property. In response to the growing need for effective forest fire prevention, this study introduces an innovative method for predicting and assessing forest fire risk. By integrating multi-source data, including optical and microwave remote sensing, meteorological, topographic, and human activity data, the approach enhances the sensitivity of risk models to vegetation water content and other critical factors. The vegetation water content is derived from both Vegetation Optical Depth and optical remote sensing data, allowing for a more accurate assessment of changes in vegetation moisture that influence fire risk. A time series prediction model, incorporating attention mechanisms, is used to assess the probability of fire occurrence. Additionally, the method includes fire spread simulations based on Cellular Automaton and Monte Carlo approaches to evaluate potential burn areas. This combined approach can provide a comprehensive fire risk assessment using the probability of both fire occurrence and potential fire spread. Experimental results show that the integration of microwave data and attention mechanisms improves prediction accuracy by 2.8%. This method offers valuable insights for forest fire management, aiding in targeted prevention strategies and resource allocation.

Keywords: forest fire risk prediction; potential spread; vegetation optical depth (VOD); remote sensing

1. Introduction

Forest fires are significant natural disasters that have emerged as a major global concern. These fires present substantial threats to ecological environments, economic development, and human life and property, posing challenges for many countries in recent decades. Forest fire risk assessment is instrumental in mitigating these risks and plays a crucial role in forest fire management [1,2]. By evaluating forest fire risk, managers can implement fire prevention measures such as fire risk zoning and fuel management to reduce fire losses [3–5].

Risk assessment can use probability to express the uncertainty of the occurrence or the intensity of hazardous events [6]. Predicting the probability of forest fires in a region requires identifying factors influencing fire occurrence, such as climate, topography, and fuel conditions [7–10], and determining the connection between these factors and the probability of fire occurrence [11]. These factors, commonly designated as driving factors, affect the occurrence and spread of forest fires. Vegetation, as the primary fuel source in forest fires, plays a pivotal role as a driving factor and has therefore been extensively studied. Jiang et al. [12] studied the application of vegetation water content (VWC) in forest fire risk assessment, and Luo [13] explored forest fire risk using the MODIS Leaf Area Index (LAI) and Live Fuel Moisture Content (LFMC) data as vegetation driving factors. All these indices aim to represent vegetation's role as a combustible material in fires, where its internal moisture and carbon content significantly impact fire behavior. However, in dense

forests, optical vegetation indices saturate with canopy quality [14]. Optical remote sensing primarily detects the canopy fuel moisture content, which limits its ability to sense surface fuel moisture content in dense forests. In fire risk assessment, the surface fuel moisture content is more critical. Microwave remote sensing technology has unique advantages in acquiring vegetation water content because the dielectric constant of vegetation is primarily influenced by moisture content information, which in turn determines the absorption and scattering of microwave signals [15,16]. Microwave remote sensing signals can penetrate the canopy to sense the moisture content of ground fuels. Wang et al. [17–19] utilize microwave remote sensing technology to retrieve fuel moisture content (FMC), employing different microwave remote sensing data and models to enhance the accuracy of FMC estimation. The results demonstrate that microwave remote sensing has advantages over traditional optical methods for monitoring FMC under all-weather conditions, confirming its effectiveness for FMC retrieval in grassland and forest environments. Mei [20] used RADARSAT-2 dual-polarization Synthetic Aperture Radar (SAR) data combined with the water cloud model and bare soil scattering model to retrieve the fuel moisture content of forest margin grassland. These studies demonstrated the model's feasibility for moisture retrieval. However, fewer studies have used data from microwave remote sensing to assess forest fire risk. The Vegetation Optical Depth (VOD), a microwave index derived from microwave remote sensing data, has been shown to be proportional to the water content of above-ground vegetation [21]. Leveraging this relationship, we attempt to improve the model's sensitivity to the vegetation water status by combining optical and microwave remote sensing data.

To enhance the accuracy of forest fire prediction, many scholars have proposed different methods. Eskandari and Miesel [22], for the high-risk areas of forest fires in Iran, proposed that knowledge-based hierarchical analysis (AHP) and fuzzy set methods can effectively identify the locations where forest fires are likely to occur. Nami et al. [23] employed the Evidential Belief Function (EBF) method to predict fire occurrence probability in northern Iran's Hyrcanian ecological zone, showing significant effectiveness with an area under the curve (AUC) of 84.14%. Farahmand et al. [24] developed the Fire Danger from Earth Observation (FDEO) algorithm to predict fire danger in the contiguous United States with a lead time of up to two months, achieving an overall accuracy of up to 75% during the fire season. In recent years, machine learning techniques have been widely used to analyze complex environmental data and predict fire occurrences [25,26]. Specifically, the application of deep learning models in forest fire occurrence probability research has become increasingly common. For instance, Abdollahi and Pradhan [27] utilized deep neural networks (DNNs) in conjunction with an explainable artificial intelligence (XAI) to develop predictive models to map wildfire susceptibility and identify key contributing components within them. Lin et al. [28] applied the long-term and short-term time series network (LSTNet) model, incorporating convolutional and recurrent layers, for forest fire prediction. This method captures long-term patterns missed by traditional models, recognizing that the occurrence of forest fires is influenced by both short-term and long-term data variations, thereby achieving high accuracy (ACC 0.941) and demonstrating effectiveness in spatial predictions of forest fire susceptibility using time series data.

In natural disaster research, risk is defined as the expected loss or benefit, including the probability and potential impact of natural disaster events [29]. Fire occurrence and spread require a certain amount of time, and, if quickly extinguished after the initial occurrence, severe consequences can be avoided. However, in large-scale or severe forest fires, rapid spread and expansion accompany fire occurrences [30]. Simulating fire spread can reflect the combined influence of factors on forest fires, thereby comprehensively reflecting regional fire risks [31,32]. To predict and simulate forest fire situations, various model techniques with different predictive variables have been developed. Cellular Automaton (CA) models and Monte Carlo (MC) simulations are widely used in fire spread modeling and forest fire research. CA models simulate fire spread by representing the landscape as a grid of cells, each of which can be in different states (e.g., unburned, burning, burned). The state of

each cell evolves based on predefined rules and the states of neighboring cells [33]. MC simulations add a probabilistic element to this process, allowing for the incorporation of randomness and uncertainty in fire behavior, making the models more realistic and robust [30]. Xuezheng et al. [32] used the Burn-P3 model to simulate burn probability, potential fire intensity, spread speed, and fire occurrence types at the landscape scale and calculate fire exposure, using the AHP to assess forest fire risk. Carmel Y et al. [34] used the FARSITE model for Monte Carlo simulation of fire spread, generating high-resolution forest fire risk maps in Mount Carmel, Israel.

To address the complex and dynamic nature of forest fires, many studies have developed various methods for assessing fire risks and predicting occurrences. Building on these advances, this paper seeks to further improve the accuracy and comprehensiveness of forest fire risk assessments by integrating multi-source data. This paper integrates multi-source remote sensing, meteorological, topographic, and social data through deep-learning-based time series predictions and potential fire spread simulation techniques to establish a forest fire risk assessment method. The innovation of this work lies in the following: (1) merging the microwave index VOD with optical remote sensing data to obtain vegetation water information, (2) predicting forest fire occurrence probability using time series methods and incorporating an attention mechanism to better focus on and weigh driving factors in the input data, learning key features from data during the week before the fire and capturing trends and changes in meteorological and vegetation factors before fire occurrences, (3) assessing forest fires through an integrated assessment that includes both the probability of fire occurrence and the potential fire spread probability to provide an overall measure of the fire risk.

2. Materials and Methods

2.1. Study Area and Fire Data

The study area (98° E–104° E, 24° N–30° N) is located at the border of Sichuan Province and Yunnan Province in the southwest region of China, covering Liangshan Prefecture, Panzhihua City in Sichuan Province, and Diqing Prefecture, Lijiang City, Dali City, Chuxiong City, and Kunming City in Yunnan Province (Figure 1a).
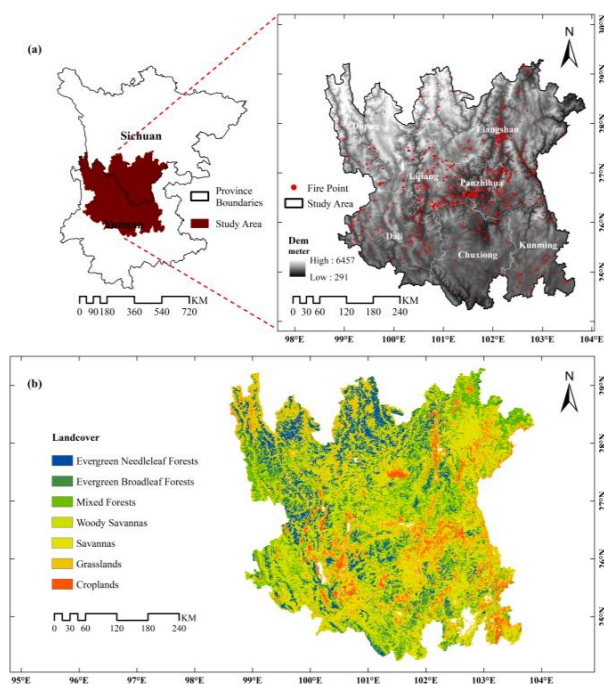

Figure 1. (**a**) Study area showing the historical (2015–2018) fire point extracted from the NASA website and the DEM (Digital Elevation Model) as the background image. (**b**) Land classification in the study area extracted from MCD12Q1.

The terrain in this region is characterized by complex topography, steep mountains, and elevations ranging from 350 to 6400 m. The forests are primarily composed of subtropical forests, and the main vegetation types are shown in Table 1 and Figure 1b. The climate belongs to a subtropical monsoon climate, characterized by distinct wet and dry seasons. The annual average temperature is 15–20 °C, and the annual temperature varies between 15 °C and 25 °C. The annual average precipitation is 800–1000 mm, with 75%–85% of the annual precipitation occurring from May to October. The precipitation in winter is scarce, and temperatures rise quickly, leading to severe drought and dry vegetation during winter and spring.

Table 1. The main vegetation type in the study area.

Fuel Class	IGBP	Coverage Area (km^2)
Forest	Evergreen Needleleaf Forests	35,641.75
	Evergreen Broadleaf Forests	5977.50
	Deciduous Needleleaf Forests	3.75
	Deciduous Broadleaf Forests	5339.00
	Mixed Forests	31,276.75
Grassland	Woody Savannas	55,528.75
	Savannas	54,512.50
	Grasslands	17,862.50
	Permanent Wetlands	228.00
	Croplands	12,663.50
	Cropland/Natural Vegetation Mosaics	2373.50
Shrubland	Closed Shrublands	70.50
	Open Shrublands	1.00
Other	Snow and Ice	48.50
	Barren or Sparsely Vegetated	530.00
	Water Bodies	719.75
	Urban and Built-Up Lands	1970.50

The coverage area was calculated based on the IGBP (International Geosphere-Biosphere Programme) classification scheme of the MODIS product MCD12Q1.

The study area is a high-fire-risk zone in China, where forest fires mainly occur in winter and spring [35]. Our analysis of the data reveals a similar seasonal pattern, where the forest fire occurrence frequency across different months exhibits a clear trend (Figure 2a), with peak occurrences particularly from December to May (Figure 2b). In contrast, the fire frequency from June to November is lower and more concentrated.
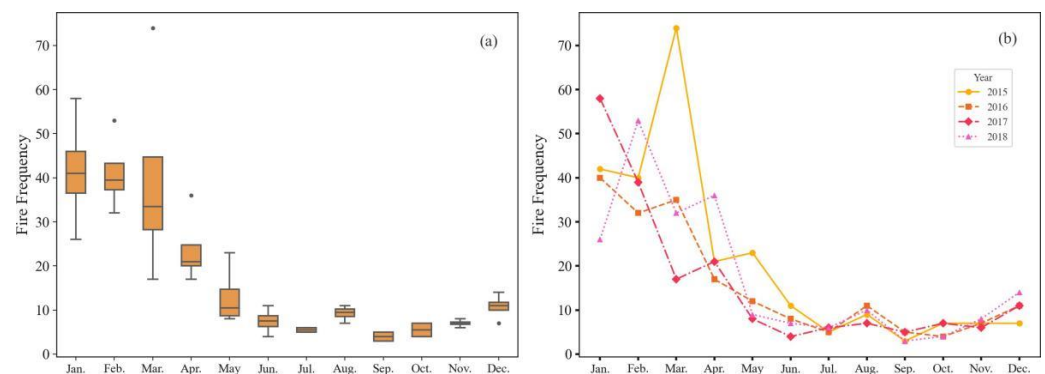

Figure 2. (**a**) Monthly and (**b**) yearly fire frequency from 2015 to 2018, calculated from NASA website data in the study area, with trends highlighting high fire frequency from January to May.

In the study, forest fire data were collected through the NASA Fire Information for Resource Management System (FIRMS) (https://firms.modaps.eosdis.nasa.gov/, accessed on 16 December 2023) using the Moderate Resolution Imaging Spectroradiometer (MODIS). Based on the main vegetation types shown in Table 1 and Figure 1b, forest fire incidents within forest areas were selected. A total of 783 forest fire incidents were gathered from 2015 to 2018, including notable fires such as the forest fire in Lijiang City in March 2015; the forest fire in Miyi County, Panzhihua City, in February 2016; and the forest fire in Muli County, Liangshan Prefecture, in November 2016.

2.2. Fire Point Unbalanced Data Preprocessing

A significant challenge in forest fire prediction is the data imbalance between fire point and non-fire point. Fire events are relatively rare, leading to far fewer fire point data compared to non-fire point data, which can affect the model's accuracy and predictive performance for minority classes [36,37]. In this study, non-fire point data were selected based on the semi-variogram function method [38]. A buffer zone was established for each fire point using a semi-variogram function based on a spherical model (Equation (1)). The distance at which the model first flattens is called the range, i.e., the buffer radius. Sample positions separated by distances within the range are spatially correlated, while sample positions separated by distances beyond the range are not spatially correlated.

$$\gamma(h) = \begin{cases} C_0 + C\left(1.5\,\frac{h}{a} - 0.5\,\frac{h^3}{a^3}\right) & 0 \le h < a \\ C + C_0 & h \ge a \end{cases} \tag{1}$$

where $\gamma(h)$ is the semi-variogram function value, C_0, C, a are model parameters, and h is the spatial distance.

Due to seasonal variations in vegetation and meteorological conditions, conditions at the same time each year are similar. Therefore, positions with high fire risk at a particular time are also likely to have high fire risk at similar times in adjacent periods and years. Using multi-temporal data, a fire buffer overlay image was established, selecting only data outside any buffer zones as the range for non-fire point data selection. Within this range, typical non-fire points were randomly generated at a ratio of 1:1.2 to fire points. Then, the influencing factor data corresponding to these typical non-fire points were extracted for the corresponding dates.

2.3. Driving Factors

The damage caused by forest fires is influenced by various factors, including types of vegetation, fuel characteristics, topography, weather and climatic conditions, and human activities [39]. In this study, we considered 11 important parameters for the years from 2015 to 2017 from four categories (meteorological, topographic, vegetation factors, and anthropogenic factors). Figure 3 illustrates these factors, spatially mapped across the study area.

Vegetation Factors: The rates of forest fire occurrences are heavily affected by land cover and types of vegetation [12]. We used the 500 m spatial resolution MODIS Normalized Difference Vegetation Index (NDVI) dataset to analyze the vegetation status in the study area. Additionally, to gain a more comprehensive understanding of vegetation factors, we supplemented with microwave data VOD. VOD data were sourced from the VODCA product provided by the University of Texas, Vienna [40], which integrates and reprocesses VOD data from multiple microwave sensors. Bilinear interpolation was used to resample VOD data to the 500 m spatial resolution consistent with the NDVI in this study.

Meteorological Factors: These factors include temperature, precipitation, humidity, and wind speed, and play key roles in the occurrence and spread of forest fires, impacting forest fire differently under varying conditions. In conditions of low precipitation, low relative humidity, and high temperatures, forest surface fuels are more likely to ignite [41]. Wind not only accelerates the evaporation and drying process of soil and surface fuels before

fires, but also injects fresh fuel for flames after forest fire occurrence, influencing the spread speed. When wind speeds increase, and dry and hot conditions occur simultaneously, forest fires can be ignited quickly and spread rapidly [42]. All meteorological data in this study were from the European Centre for Medium-Range Weather Forecast (ECMWF) global atmospheric reanalysis dataset. Relative humidity data were calculated using the dew point and air temperature, while wind speed data were obtained by combining meridional (U) and zonal (V) wind speeds to calculate horizontal 10 m wind speeds. All factors were adjusted to match the resolution of the MODIS vegetation indices.
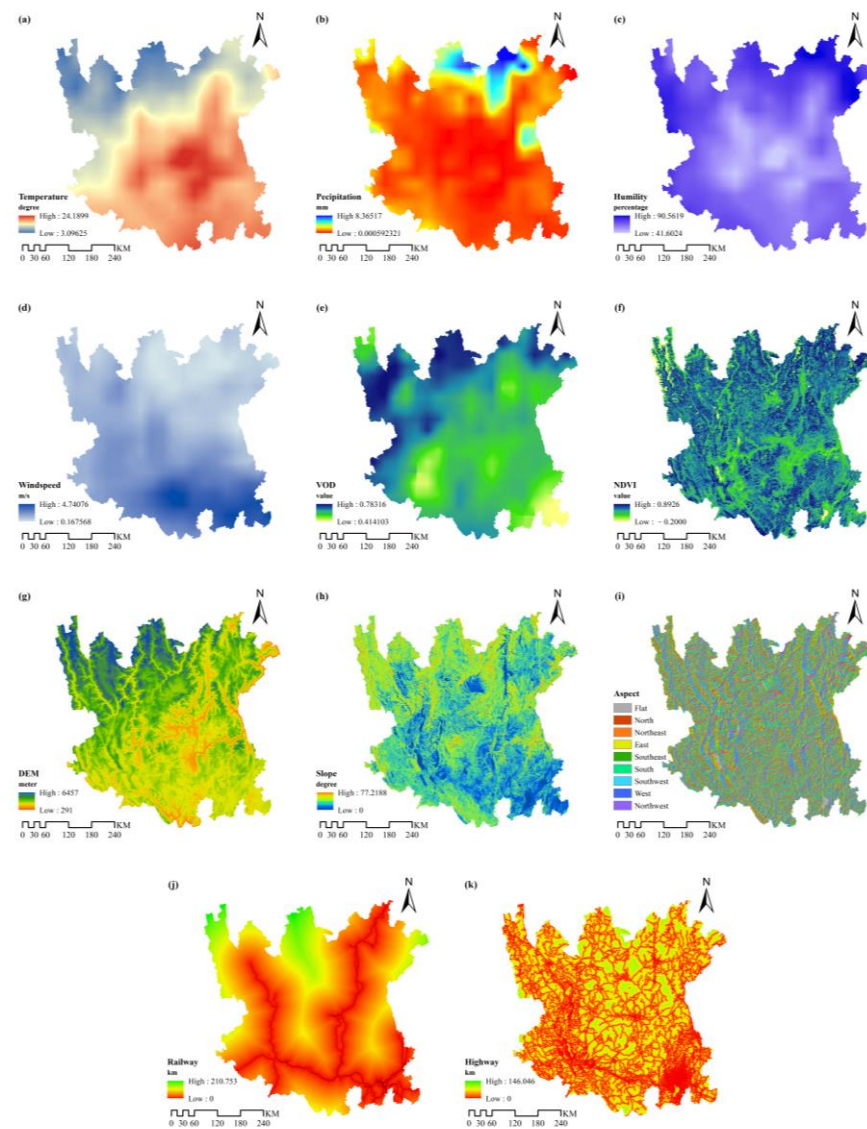

Figure 3. Driving factors in predicting forest fire occurrence: (**a**) temperature, (**b**) precipitational, (**c**) humidity, (**d**) wind speed, (**e**) VOD, (**f**) NDVI, (**g**) DEM, (**h**) slope, (**i**) aspect, (**j**) railway, and (**k**) highway.

Topographic Factors: Topography significantly influences local climate conditions such as wind and precipitation, which in turn affect forest fire behavior. For example, variations in elevation and slope can alter wind flow and humidity levels, making certain areas more susceptible to be ignited. Mountain regions are often more sensitive to temperature changes at higher elevations, which can intensify fire behavior by altering local temperature and precipitation patterns [43]. In this study, we considered slope, aspect, and elevation. The derived variables such as slope and aspect were calculated from the Digital Elevation Model (DEM). The DEM data used were sourced from the Shuttle Radar Topography Mis-

sion (SRTM) and accessed via the NASA Earth Data platform (https://earthdata.nasa.gov/, accessed on 17 December 2023.), with a resolution of approximately 90 m.

Anthropogenic Factors: Human activities significantly impact forest fires, with areas near transportation infrastructure experiencing frequent human activities, increasing fire risk [44]. In our study, we selected the distance to railways and highways as anthropogenic driving factors. Euclidean distances to railways (Figure 3j) and highways (Figure 3k) were calculated using buffering tools.

2.4. Training and Testing Datasets

The training samples in our experiment included 5078 labeled locations, consisting of 2308 fire event data points and 2770 non-fire event data points. Each data point included a 7-day record of relevant driving factors for the respective area. Time series information was extracted from these labeled locations, with data combinations every 7 days forming a forest fire dataset with time series influences. To train the forest fire occurrence probability model, we used two sets of labels 0 and 1, with 1 indicating fire events and 0 indicating no fire events. During training, data from 2015 to 2017 were split into 80% for model training and 20% for validating the model's performance, while the data from 2018 were used for prediction.

2.5. Forest Fire Occurrence Prediction Models

In this study, the Long Short-Term Memory (LSTM) algorithm, augmented with an attention mechanism, was presented to develop a predictive model for forest fire occurrence probability. LSTM is a type of Recurrent Neural Network (RNN) specifically designed to overcome the limitations of traditional RNNs in handling long-term dependencies. It achieves this goal through a special gating mechanism, which helps retain important information over longer periods while reducing the risk of vanishing gradients. The method's structure is built on a multi-layer neural network, combining the LSTM's time series processing capability and the attention mechanism's feature focusing ability, providing a powerful solution for binary classification tasks. This model captures dependencies in time series data through LSTM layers and enhances prediction performance by focusing on key features with the attention mechanism. As shown in Figure 4, the architecture was composed of an input layer L_i, an output layer L_o, and LSTM layers X_i (where $i \in \{1, 2,\ldots, 6\}$), with an Attention Module between the input and output layers. The input layer L_i received time series data with multiple features at each time step. The LSTM layers, each employing the tanh activation function (as defined in Equation (2)), processed the input and returned the complete output sequence. Following this step, a Dropout layer was applied to prevent overfitting. The Attention layer focused on important features by calculating attention weights for time steps, enhancing the model's learning ability. The output from the Attention layer was flattened and passed into the dense layer using the tanh activation function for subsequent processing. The Sigmoid function is advantageous for binary classification due to its smooth gradient and output range of (0, 1) [45]. Therefore, the output layer L_o used the Sigmoid activation function to convert the output to binary classification probability values.

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$
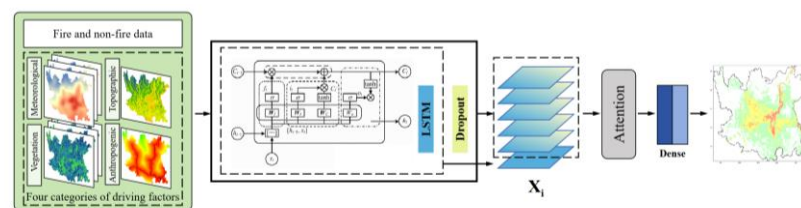

Figure 4. Deep learning model framework for predicting forest fire occurrence probability.

2.6. Fire Risk Combined Assessment

In this study, we considered forest fire risk as the combined result of forest fire occurrence probability and potential burn probability. Using fire occurrence probability and potential burn probability, forest fire risk was estimated using Equation (3), as follows:

$$\text{Fire Risk (\%)} = \text{OP} \times \text{BP (\%)} \tag{3}$$

where OP is the fire occurrence probability ($0 < \text{OP} < 1$), and BP is the fire potential burn probability ($0 < \text{BP} < 100\%$).

The potential burn probability (BP), calculated based on the outcomes of fire spread simulations according to Equation (4), represents the proportion of times a given location burns out of the total number of simulations. To determine the BP, we randomly selected points from the high and very-high forest fire occurrence probability areas predicted by the forest fire occurrence probability model as fire points. This random selection is an inherent feature of the Monte Carlo simulation, which relies on stochastic sampling within a defined range to account for variability and uncertainty in fire spread. These points were then utilized for Cellular Automaton-Monte Carlo simulations, excluding locations in non-combustible areas like roads, residential zones, and water bodies. The inputs for the Cellular Automaton model included temperature, humidity, wind speed, and NDVI data. Monte Carlo simulations were used to randomly select ignition points within the high and very-high forest fire occurrence probability areas, simulating potential fire spread behavior. In this study, the total number of simulations was 100, which ensures a good balance between statistical stability and computational efficiency, meaning that 100 fire spread simulation experiments were conducted.

$$\text{BP(\%)} = \frac{\text{Number of burns}}{\text{Total number of simulations}} \times 100\% \tag{4}$$

2.7. Evaluation Metrics

To validate the accuracy of the proposed forest fire occurrence prediction models, three common metrics were applied in our study: Accuracy, F1 score, and the AUC, as shown in Table 2.

Table 2. The evaluation metrics and the formulas.

Evaluation Metrics	Formula
Accuracy	$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$
F1 Score	$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$
AUC	$\text{AUC} = \int_0^1 \text{TPR(FPR)}d(FPR)$

Accuracy is an indicator of the model's overall prediction accuracy, defined as the ratio of correctly predicted cases, True Positive (TP) and True Negative (TN), to all cases, True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The F1 score is the harmonic mean of precision and recall, used to comprehensively evaluate the model's performance. It balances precision and recall and is commonly used when there is class imbalance. A high F1 score (typically above 0.8) indicates good performance in both precision and recall.

The AUC is used to evaluate the performance of classification models, where an AUC value close to 0 indicates random prediction, while an AUC value close to 1 indicates high prediction accuracy. The AUC is derived from the Receiver Operating Characteristic (ROC) curve, which illustrates the trade-off relationship between the false positive rate ($\text{FPR} = \frac{FP}{FP+TN}$) and the true positive rate ($\text{TPR} = \frac{TP}{TP+FN}$), quantifying the model's overall performance.

3. Results

In the study, we established and optimized a forest fire occurrence probability model to evaluate its predictive capabilities for fire occurrence. As mentioned above, to better assess the effectiveness of the forest fire occurrence probability model, Table 3 presents the quantitative results of different evaluation metrics for five different models. The Multi-Layer Perceptron (MLP) model serves as a baseline for comparison in this study, to highlight the effectiveness of using LSTM for a time series data analysis.

Table 3. Quantitative results of the proposed model.

Model	Accuracy (%)	AUC	F1 Score
MLP	84.46	0.9603	0.8613
LSTM	86.02	0.9456	0.8640
LSTM+Attention	87.27	0.9537	0.8779
LSTM+VOD	87.04	0.9579	0.8735
LSTM+VOD+Attention	**88.82**	**0.9608**	**0.8869**

As shown in Table 3, the forest fire occurrence probability model combining microwave remote sensing data and the attention mechanism performed the best, achieving a prediction accuracy of 88.82%, an AUC of 0.9608, and an F1 score of 0.8869. Compared to the model without VOD data, the prediction accuracy increased by 1.55%, and the inclusion of the attention mechanism enhanced the model's generalization ability, improving the prediction accuracy by 1.78%. Using time series for prediction, the model significantly outperformed models without such data, with an accuracy of 84.46%. Time series played a crucial role in capturing trends and changes in meteorological and vegetation factors shortly before fire occurrences, significantly enhancing the model's predictive performance. The ROC curve for the best-performing model, which achieved a prediction accuracy of 88.82%, is depicted in Figure 5. This ROC curve illustrates the model's excellent ability to distinguish between different classes, demonstrating its robustness and reliability in practical applications.
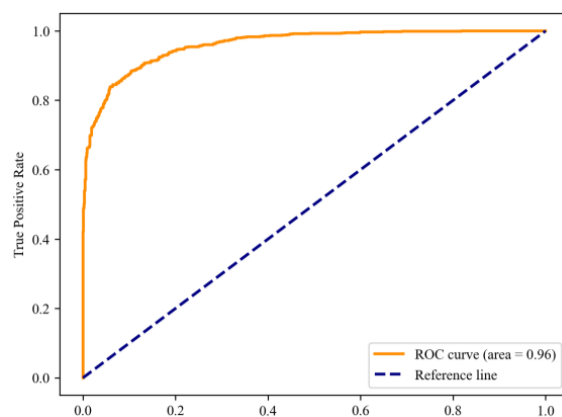

Figure 5. The resulting ROC curve of the proposed mode.

To validate the practical application of the model, we used the final optimized model to predict fire risk in the study area and generated a fire occurrence probability map. We employed five classes: very-high, high, moderate, low, and very-low, established using the natural break algorithm [46], a standard method for grouping similar values to classify fire occurrence probability. The results of the classification are shown in Figure 6.
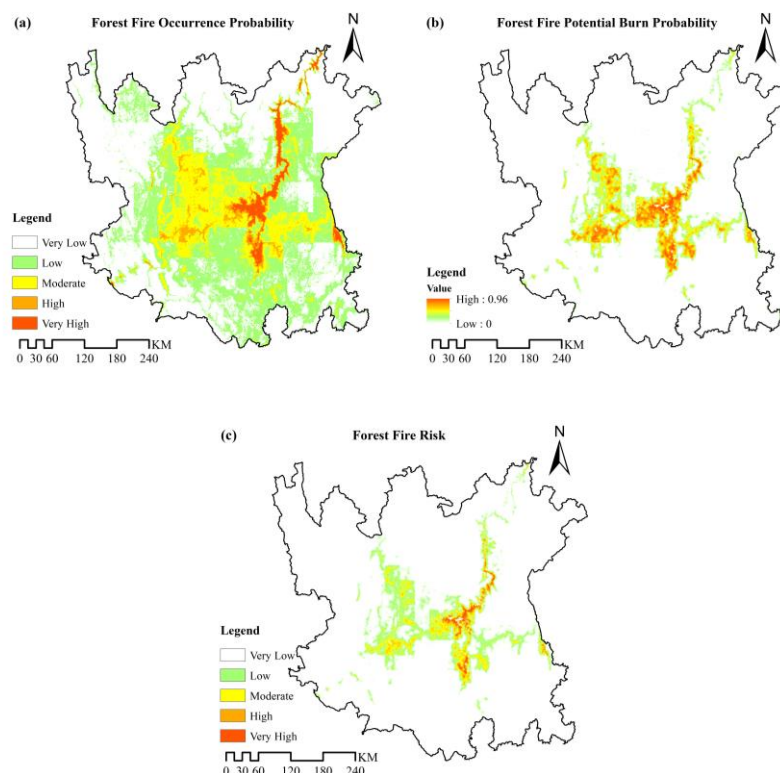
Figure 6. (**a**) The forest fire occurrence probability map using the proposed mode. (**b**) The forest fire potential burn probability using simulation. (**c**) The forest fire risk in the study area.

Analyzing the fire occurrence probability map reveals that most areas within the study region have a low and very-low probability of forest fire occurrence. The high and very-high probability areas are mainly concentrated in densely vegetated and dry regions, particularly near some railways and highways. Conversely, low probability areas are primarily distributed in sparsely vegetated or moist regions. The high temperatures in the region have increased the probability of forest fires. Due to the distribution of forests, the probability of forest fires is relatively high in the areas of Liangshan Prefecture and Panzhihua. Meanwhile, the probability of forest fires is slightly higher in the areas near the Sichuan–Yunnan border compared to locations further away.

By applying the fire risk combined assessment method, which integrates the potential burn probability (Figure 6b) with the fire occurrence probability, the resulting fire risk map (Figure 6c) shows a significant reduction in the proportion of high and very-high risk areas compared to the fire occurrence probability alone (Table 4). The very-high risk areas are more clearly defined, allowing for more targeted fire prevention measures.

Table 4. The area of each forest fire risk class in the study area.

Risk Class	Forest Fire Occurrence Probability		Forest Fire Risk	
	Area (km^2)	Area (%)	Area (km^2)	Area (%)
Very Low	106,010.25	49.84	197,224.25	92.83
Low	67,475.50	31.72	10,204.50	4.80
Moderate	28,808.50	13.54	3639.00	1.71
High	5872.50	2.76	1223.50	0.58
Very High	4542.75	2.13	169.00	0.08

4. Discussion

In the study, we proposed a new method for forest fire risk assessment by integrating multi-source remote sensing data (including microwave and optical remote sensing) with

deep learning models. The results indicate that the method has advantages in predicting forest fire occurrence probabilities and assessing fire risk. By introducing VOD as a vegetation influencing factor, the model's sensitivity to forest fire occurrence was improved. Compared to traditional methods relying solely on optical remote sensing data, VOD data better reflect vegetation water content, enhancing the model's ability to capture fire occurrence conditions. Additionally, the deep learning model combining LSTM networks with attention mechanisms effectively handled time series data, capturing trends and changes in meteorological and vegetation factors before fire occurrences. The evaluation metrics, including accuracy, F1 score, and the AUC, all showed high accuracy, further validating the model's effectiveness.

This study comprehensively utilized various data sources, including optical remote sensing data, microwave remote sensing data, meteorological data, topographic data, and anthropogenic factors. This diversity enriched the feature dimensions of the model inputs and enhanced the model's adaptability to the complex forest fire environment. Furthermore, the inclusion of distances to railways and highways as anthropogenic factors allowed the model to reflect the impact of human activities on forest fire risk, highlighting the significance of human-caused fires in forests.

We utilized a combination of cellular automaton and Monte Carlo simulation methods to perform fire spread simulations and calculate potential burn probability for points within the study area with predicted forest fire occurrence probabilities in the high and very-high regions. Figure 6c shows the results of fire risk assessments based on two different fire risk-related factors (fire occurrence probability in Figure 6a and potential burn probability in Figure 6b). The final results indicate that fire risk is significantly influenced by the combined effects of fire occurrence probability and potential burn probability. The very-high risk areas show high fire occurrence probabilities and high potential burn probabilities. Medium risk and low risk areas are characterized by high and moderate fire occurrence probabilities, but lower potential burn probabilities reduce the overall fire risk. Very-low risk areas are mainly distributed in sparsely vegetated or moist regions, with both low fire occurrence probabilities and potential burn probabilities. This comprehensive assessment method enables us to more accurately identify high fire risk areas, making forest fire prevention measures more targeted. By comprehensively assessing fire occurrence probabilities and potential burn probabilities, high fire risk areas within the study region were effectively identified. It provides scientific support for forest fire management departments to develop targeted fire prevention measures and resource allocation strategies. Additionally, the results of fire spread simulations help quick responses in the early stages of fire occurrences, reducing the damage to ecosystems and human society.

However, our study has some limitations. Firstly, the study area is limited to specific regions in Southwestern China, and the model's applicability and generalizability need further validation. Secondly, for potential burn probability studies, we used a combination of cellular automaton and Monte Carlo simulation methods, which can be influenced by initial conditions and rule settings [33]. This model lacks a detailed consideration of environmental factors and complex terrains, making it difficult to accurately reflect the interactions of multiple factors during an actual fire spread [47]. Additionally, the impact of lightning-caused fires has not been adequately accounted for in this study. Although the proportion of lightning-induced forest fires in Southwestern China, such as in Muli County, Sichuan (46.7%) [48], is lower compared to northern regions, such as the Greater Khingan Mountains (68.28%) [49], lightning-caused fires remain a significant factor that should not be overlooked. Future studies could incorporate more detailed lightning data and consider its spatial and temporal distribution to improve the accuracy of fire occurrence models in different regions. Furthermore, training and tuning the deep learning model for forest fire occurrence probability is complex and requires significant computational resources, posing potential challenges for practical applications. Future research can improve potential spread simulations by combining deep learning spread prediction methods to obtain more accurate potential burn probabilities.

5. Conclusions

Forest fire risk assessment is a key aspect of forest emergency management. A detailed fire risk assessment of forest areas can effectively reduce the losses caused by natural disasters, provide crucial support for firefighters' emergency response and post-disaster recovery, and is of great significance for protecting property safety and ensuring human life. We proposed a forest fire risk assessment method based on multi-source remote sensing data and deep learning models, establishing a comprehensive evaluation model of forest fire occurrence probability using vegetation, meteorological, topographic, and anthropogenic data and potential burn probability, obtained through fire spread simulations. The method improved the prediction accuracy by 2.8%. The results demonstrate that the method enhances the accuracy and predictability of fire risk assessments, providing new technical means for forest fire early warning and management.

References

1. Ajin, R.; Ciobotaru, A.-M.; Vinod, P.; Jacob, M.K. Forest and wildland fire risk assessment using geospatial techniques: A case study of Nemmara forest division, Kerala, India. *J. Wetl. Biodivers.* **2015**, *5*, 29–37.
2. Ager, A.A.; Preisler, H.K.; Arca, B.; Spano, D.; Salis, M. Wildfire risk estimation in the Mediterranean area. *Environmetrics* **2014**, *25*, 384–396. [CrossRef]
3. Jaiswal, R.K.; Mukherjee, S.; Raju, K.D.; Saxena, R. Forest fire risk zone mapping from satellite imagery and GIS. *Int. J. Appl. Earth Obs. Geoinf.* **2002**, *4*, 1–10. [CrossRef]
4. Mitsopoulos, I.; Trapatsas, P.; Xanthopoulos, G. SYPYDA: A software tool for fire management in Mediterranean pine forests of Greece. *Comput. Electron. Agric.* **2016**, *121*, 195–199. [CrossRef]
5. Noonan-Wright, E.K.; Opperman, T.S.; Finney, M.A.; Zimmerman, G.T.; Seli, R.C.; Elenz, L.M.; Calkin, D.E.; Fiedler, J.R. Developing the US wildland fire decision support system. *J. Combust.* **2011**, *2011*, 168473. [CrossRef]
6. Alcasena, F.J.; Salis, M.; Ager, A.A.; Arca, B.; Molina, D.; Spano, D. Assessing landscape scale wildfire exposure for highly valued resources in a Mediterranean area. *Environ. Manag.* **2015**, *55*, 1200–1216. [CrossRef]
7. Turco, M.; Llasat, M.C.; Hardenberg, J.; Provenzale, A. Impact of climate variability on summer fires in a Mediterranean environment (northeastern Iberian Peninsula). *Clim. Chang.* **2013**, *116*, 665–678. [CrossRef]
8. Zumbrunnen, T.; Pezzatti, G.B.; Menéndez, P.; Bugmann, H.; Bürgi, M.; Conedera, M. Weather and human impacts on forest fires: 100 years of fire history in two climatic regions of Switzerland. *For. Ecol. Manag.* **2011**, *261*, 2188–2199. [CrossRef]
9. Mollicone, D.; Eva, H.D.; Achard, F. Human role in Russian wild fires. *Nature* **2006**, *440*, 436–437. [CrossRef]
10. Vasilakos, C.; Kalabokidis, K.; Hatzopoulos, J.; Matsinos, I. Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Nat. Hazards* **2009**, *50*, 125–143. [CrossRef]
11. Naderpour, M.; Rizeei, H.M.; Ramezani, F. Forest fire risk prediction: A spatial deep neural network-based framework. *Remote Sens.* **2021**, *13*, 2513. [CrossRef]
12. Jiang, M.; Hu, Z.; Ding, Y.; Fang, D.; Li, Y.; Wei, L.; Guo, M.; Zhang, S. Estimation of vegetation water content based on MODIS: Application on forest fire risk assessment. In Proceedings of the 2012 20th International Conference on Geoinformatics, Hong Kong, 15–17 June 2012; IEEE: New York, NY, USA, 2012; pp. 1–4.
13. Luo, K. Research on Global Wildfire Risk Spatiotemporal Mining and Prediction Methods. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2021.

14. Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [CrossRef]

15. El-Rayes, M.A.; Ulaby, F.T. Microwave dielectric spectrum of vegetation-Part I: Experimental observations. *IEEE Trans. Geosci. Remote Sens.* **1987**, *25*, 541–549. [CrossRef]

16. Ulaby, F.T.; El-Rayes, M.A. Microwave dielectric spectrum of vegetation-Part II: Dual-dispersion model. *IEEE Trans. Geosci. Remote Sens.* **1987**, *25*, 550–557. [CrossRef]

17. Wang, L.; Quan, X.; He, B.; Yebra, M.; Xing, M.; Liu, X. Assessment of the dual polarimetric Sentinel-1A data for forest fuel moisture content estimation. *Remote Sens.* **2019**, *11*, 1568. [CrossRef]

18. Wang, L.; He, B.; Quan, X.; Xing, M.; Liu, X. Estimation of Fuel Moisture Content Based on Quad Polarimetric Decomposition Parameters of Radarsat-2 Data. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; IEEE: New York, NY, USA, 2019; pp. 9934–9937.

19. Wang, L.; He, B.; Quan, X.; Xing, M.; Zhang, H. First Assessment of Dual Polarization Sentinel-1A Data for Fuel Moisture Content Retrieval. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; IEEE: New York, NY, USA, 2019; pp. 9902–9905.

20. Mei, H. Remote Sensing Inversion of Forest Margin Grassland Living Combustibles Moisture Content Based on RADARSAT-2 Dual-Polarization SAR Data. Master's Thesis, Inner Mongolia Normal University, Hohhot, China, 2022.

21. Jackson, T.J.; Schmugge, T.J. Vegetation effects on the microwave emission of soils. *Remote Sens. Environ.* **1991**, *36*, 203–212. [CrossRef]

22. Eskandari, S.; Miesel, J.R. Comparison of the fuzzy AHP method, the spatial correlation method, and the dong model to predict the fire high-risk areas in hyrcanian forests of Iran. *Geomat. Nat. Hazards Risk* **2017**, *8*, 933–949. [CrossRef]

23. Nami, M.; Jaafari, A.; Fallah, M.; Nabiuni, S. Spatial prediction of wildfire probability in the hyrcanian ecoregion using evidential belief function model and GIS. *Int. J. Environ. Sci. Technol.* **2018**, *15*, 373–384. [CrossRef]

24. Farahmand, A.; Stavros, E.N.; Reager, J.T.; Behrangi, A. Introducing Spatially Distributed Fire Danger from Earth Observations (FDEO) Using Satellite-Based Data in the Contiguous United States. *Remote Sens.* **2020**, *12*, 1252. [CrossRef]

25. Kim, S.J.; Lim, C.-H.; Kim, G.S.; Lee, J.; Geiger, T.; Rahmati, O.; Son, Y.; Lee, W.-K. Multi-temporal analysis of forest fire probability using socio-economic and environmental variables. *Remote Sens.* **2019**, *11*, 86. [CrossRef]

26. Singh, K.R.; Neethu, K.P.; Madhurekaa, K.; Harita, A.; Mohan, P. Parallel SVM model for forest fire prediction. *Soft Comput. Lett.* **2021**, *3*, 100014. [CrossRef]

27. Abdollahi, A.; Pradhan, B. Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. *Sci. Total Environ.* **2023**, *879*, 163004. [CrossRef] [PubMed]

28. Lin, X.; Li, Z.; Chen, W.; Sun, X.; Gao, D. Forest Fire Prediction Based on Long- and Short-Term Time-Series Network. *Forests* **2023**, *14*, 778. [CrossRef]

29. Finney, M.A. The challenge of quantitative risk analysis for wildland fire. *For. Ecol. Manag.* **2005**, *211*, 97–108. [CrossRef]

30. Woo, H.; Chung, W.; Graham, J.M.; Lee, B. Forest fire risk assessment using point process modelling of fire occurrence and Monte Carlo fire simulation. *Int. J. Wildland Fire* **2017**, *26*, 789–805. [CrossRef]

31. Talukdar, S.; Das, T.; Naikoo, M.W.; Rihan, M.; Rahman, A. Forest fire susceptibility mapping by integrating remote sensing and machine learning algorithms. *Adv. Remote Sens. For. Monit.* **2022**, *16*, 179–195.

32. Xuezheng, Z.; Xiaorui, T.; Shuai, M.; Chang, L. Quantitative assessment for forest fire risk based on fire simulation: Taking the Subtropical Forest Experimental Center of Chinese Academy of Forestry as an example. *J. Beijing For. Univ.* **2022**, *44*, 83–90.

33. Karafyllidis, I.; Thanailakis, A. A model for predicting forest fire spreading using cellular automata. *Ecol. Model.* **1997**, *99*, 87–97. [CrossRef]

34. Carmel, Y.; Paz, S.; Jahashan, F.; Shoshany, M. Assessing fire risk using Monte Carlo simulations of fire spread. *For. Ecol. Manag.* **2009**, *257*, 370–377. [CrossRef]

35. Tian, X.R.; Shu, L.; Zhao, F.; Wang, M. Dynamic characteristics of forest fires in the main ecological geographic districts of China. *Sci. Silvae Sin.* **2015**, *51*, 71–77.

36. Japkowicz, N.; Stephen, S. The class imbalance problem: A systematic study. *Intell. Data Anal.* **2002**, *6*, 429–449. [CrossRef]

37. Sun, Y.; Wong, A.K.C.; Kamel, M.S. Classification of imbalanced data: A review. *Int. J. Pattern Recognit. Artif. Intell.* **2009**, *23*, 687–719. [CrossRef]

38. Wen, C. Methodology and Application for Wildfire Risk Assessement Based on Remote Sensing Techniques. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2020.

39. 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] [PubMed]

40. Moesinger, L.; Dorigo, W.; de Jeu, R.; van der Schalie, R.; Scanlon, T.; Teubner, I.; Forkel, M. The global long-term microwave Vegetation Optical Depth Climate Archive (VODCA). *Earth Syst. Sci. Data* **2020**, *12*, 177–196. [CrossRef]

41. Dennison, P.E.; Brewer, S.C.; Arnold, J.D.; Moritz, M.A. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* **2014**, *41*, 2928–2933. [CrossRef]

42. Jolly, W.M.; Cochrane, M.A.; Freeborn, P.H.; Holden, Z.A.; Brown, T.J.; Williamson, G.J.; Bowman, D.M. Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* **2015**, *6*, 7537. [CrossRef]

43. Pepin, N.C.; Arnone, E.; Gobiet, A.; Haslinger, K.; Kotlarski, S.; Notarnicola, C.; Palazzi, E.; Seibert, P.; Serafin, S.; Schöner, W.; et al. Climate changes and their elevational patterns in the mountains of the world. *Rev. Geophys.* **2022**, *60*, e2020RG000730. [CrossRef]

44. Syphard, A.D.; Radeloff, V.C.; Keuler, N.S.; Taylor, R.S.; Hawbaker, T.J.; Stewart, S.I.; Clayton, M.K. Predicting spatial patterns of fire on a southern California landscape. *Int. J. Wildland Fire* **2008**, *17*, 602–613. [CrossRef]

45. Dubey, S.R.; Singh, S.K.; Chaudhuri, B.B. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing* **2022**, *503*, 92–108. [CrossRef]

46. Febrianto, H.; Fariza, A.; Hasim, J.A.N. Urban flood risk mapping using analytic hierarchy process and natural break classification (Case study: Surabaya, East Java, Indonesia). In Proceedings of the 2016 International Conference on Knowledge Creation and Intelligent Computing (KCIC), Manado, Indonesia, 15–17 November 2016; IEEE: New York, NY, USA, 2016; pp. 148–154.

47. Finney, M.A.; McHugh, C.W.; Grenfell, I.C.; Riley, K.L.; Short, K.C. A simulation of probabilistic wildfire risk components for the continental United States. *Stoch. Environ. Res. Risk Assess.* **2011**, *25*, 973–1000. [CrossRef]

48. Deng, X.; Xia, X. Forest fire condition analysis and defense in Muli. *Sichuan For. Technol.* **2009**, *30*, 126–130.

49. Zhang, Z.; Tian, Y.; Wang, G.; Zheng, C.; Zhao, F. A Forest Fire Prediction Method for Lightning Stroke Based on Remote Sensing Data. *Forests* **2024**, *15*, 647. [CrossRef]