Article

Satellite Remote Sensing False Forest Fire Hotspot Excavating Based on Time-Series Features

Haifeng Wang, Gui Zhang *, Zhigao Yang, Haizhou Xu, Feng Liu and Shaofeng Xie

Abstract: Satellite remote sensing has become an important means of forest fire monitoring because it has the advantages of wide coverage, few ground constraints and high dynamics. When utilizing satellites for forest fire hotspot monitoring, two types of ground hotspots, agricultural and other fire hotspots can be ruled out through ground object features. False forest fire hotspots within forested areas must be excluded for a more accurate distinction between forest fires and non-forest fires. This study utilizes spatio-temporal data along with time-series classification to excavate false forest fire hotspots exhibiting temporal characteristics within forested areas and construct a dataset of such false forest fire hotspots, thereby achieving a more realistic forest fire dataset. Taking Hunan Province as the research object, this study takes the satellite ground hotspots in the forests of Hunan Province as the suspected forest fire hotspot dataset and excludes the satellite ground hotspots in the forests such as fixed heat sources, periodic heat sources and recurring heat sources which are excavated. The validity of these methods and results was then analyzed. False forest fire hotspots, from satellite ground hotspots extracted from 2019 to 2023 Himawari-8/9 satellite images, closely resemble the official release of actual forest fires data and the accuracy rate in the actual forest fire monitoring is 95.12%. This validates that the method employed in this study can improve the accuracy of satellite-based forest fire monitoring.

Keywords: ground object features; forest fire monitoring; false forest fire hotspots; time-series classification scheme; Hunan province

1. Introduction

Forest wildfires present a significant threat to forestry resources, as well as to human lives and property [1]. As a result, timely access to hotspot information regarding forest wildfires is essential for enhancing early fire detection. The use of ground object features can distinguish satellite ground hotspots into three types of ground hotspots, such as forest fires, agricultural fires and other fires. In addition to real forest fire hotspots, there may also exist three types of false forest fire hotspots, such as fixed heat sources, periodic heat sources and recurring heat sources, which have obvious time-series characteristics. The exclusion of these heat sources from forest fire hotspots to enhance the accurate distinction between forest fires and non-forest fires is crucial for improving the accuracy of satellite monitoring of forest fires and is the goal of this study.

The widespread use of remote sensing satellites and geographic information systems (GIS) has led to the accumulation of significant amounts of spatial and temporal data, including meteorological, digital image, and geoscientific data [2]. Spatio-temporal data encompasses both temporal and spatial dimensions, exhibiting autocorrelation in both time and space. Spatio-temporal data mining, a new research field, is built on the basis of geoscience information sharing, and various technologies are used to mine hidden and rich knowledge from large-scale databases [3,4]. Spatio-temporal data analysis enables not...
only the mining of spatial and temporal characteristics of forest wildfires and identification of factors contributing to their occurrence, allowing for timely targeted interventions to mitigate wildfire risks [5,6], but also facilitates the excavation of false forest fire hotspots in satellite monitoring data, thereby enhancing the accuracy of forest fire monitoring [7]. Single images are utilized to distinguish forest fires from non-forest fires when forest fire monitoring is performed using methods such as sub-pixel mapping [8] and forest fire smoke detection [9]. In contrast, the spatio-temporal data used in this study contain multi data, which can more deeply excavate false forest fire hotspots hidden in the forest to more accurately distinguish forest wildfires from non-forest wildfires, and only need to exclude the false forest fire hotspots excavated in the next forest fire monitoring to obtain a more accurate forest fire. The method in this study improves the time-consuming problem compared to the methods used by previous studies [8,9]. However, the utilization of historical data on satellite ground hotspots through data mining techniques has not been fully leveraged. This study utilized data mining techniques along with satellite-derived ground hotspot data spanning from 2019 to 2023 in Hunan Province to identify and exclude false forest fire hotspots. These types of hotspots were removed to obtain a more accurate forest fire hotspots dataset. Subsequent forest fire monitoring can directly exclude false forest fire hotspots through the forest fire dataset and improve the accuracy of forest fire satellite monitoring.

In this study, satellite ground hotspots were extracted from Himawari-8/9 satellite images from 2019 to 2023 in Hunan Province. Based on ground object classification, the remaining satellite ground hotspots located in forests were utilized as a dataset representing suspected forest fire hotspots. The sliding window algorithm, dynamic time warping, statistical method and other time series classification methods were used to excavate the satellite ground hotspots such as fixed heat sources, periodic heat sources and recurring heat sources from the suspected forest fire hotspot dataset. These types of hotspots were then removed to obtain a more accurate forest fire hotspots dataset. The selected satellite ground hotspot data for validation were derived from Himawari-9 satellite images captured at 30 min intervals in Hunan Province between 8:00 and 17:30 on 12 January 2024, comprising 20 time points. These data were then compared with the forest fire hotspot data validated by the Hunan Provincial Department of Emergency Management to validate the methodology and conclusions of this study. The results indicate that the methodology employed in this study enhances the accuracy of satellite-based forest fire monitoring.

2. Materials and Methods

2.1. Study Area

Hunan Province has the highest incidence of forest wildfires nationwide, attributed to its extensive forest coverage, abundant combustible materials within forests, rugged terrain and diverse sources of ignition. Forest wildfires in this province exhibit typical spatial and temporal patterns. Hunan Province was chosen as the study area for extracting satellite-derived ground forest hotspots from 2019 to 2023, and the results and methodologies were subsequently validated.

Hunan Province is situated in the south-central part of China, spanning latitudes 24°38′ to 30°08′N and longitudes 108°47′ to 114°15′E. The central region of Hunan features undulating hills and mountains, whereas the northern lake basin plain is open and fertile, forming an asymmetrical horseshoe-shaped terrain extending northeastward, as illustrated in Figure 1. Hunan Province is situated in the subtropical evergreen broad-leaved forest region, characterized by forest and wetland ecosystems as the primary ecosystem types. Between the end of 2012 and 2022, the forest coverage increased from 57.34% to 59.98%. A total of 25,977 forest wildfires occurred between 2000 and 2018 in Hunan Province, with a total affected forest area of 106,220.78 hectares. Therefore, accurately distinguishing between forest and non-forest fires can enhance the accuracy of satellite-based forest fire monitoring, which is of practical importance in maintaining ecological security.
ince, with a total affected forest area of 106,220.78 hectares. Therefore, accurately distinguishing between forest and non-forest fires can enhance the accuracy of satellite-based forest fire monitoring, which is of practical importance in maintaining ecological security.

Figure 1. Overview map of the study area.

2.2. Data and Processing

2.2.1. Satellite Data

The data utilized for land classification in this study were obtained from the Sentinel-2 multispectral imaging satellite [10]. Due to its high spatial resolution, this satellite is extensively employed in ground monitoring, providing images of vegetation, soil and water cover and mapping land cover changes, which fully meet the needs for excavating features before forest fires occurred. This study obtained Sentinel-2 satellite remote sensing images from the Copernicus data space ecosystem (https://browser.dataspace.copernicus.eu/ (accessed on 8 February 2023)) as the basic dataset for land classification. The data is L1C atmospheric apparent reflectance products after ortho-rectification and subimage-level geometric fine correction. After using bands 2, 3 and 4, and utilizing the official Sen2Cor plugin for atmospheric correction, SNAP was applied to resample to a spatial resolution of 10 m and image cropping processing. Four classification methods were employed for land classification. The method demonstrating the highest accuracy was chosen to prepare for subsequent excavation experiments. The data from the ground object classification experiment are shown in Table 1.

Table 1. Experimental data and pre-processing steps.

<table>
<thead>
<tr>
<th>Area</th>
<th>Longitude and Latitude Range</th>
<th>Sentinel-2 (UTC)</th>
<th>Himawari-8/9 (CST)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xintian county</td>
<td>Latitudes 25°40' to 26°06'N and longitudes 112°02' to 112°23'E</td>
<td>14 October 2022 03:06 (Atmospheric correction, resample, crop)</td>
<td>/</td>
</tr>
<tr>
<td>Xinshao county</td>
<td>Latitudes 27°15' to 27°38'N and longitudes 111°8' to 111°50'E</td>
<td>14 October 2022 03:06 (Atmospheric correction, resample, crop)</td>
<td>/</td>
</tr>
<tr>
<td>Hunan province</td>
<td>/</td>
<td>1 December 2022–31 December 2022 (131 images)</td>
<td>1 January 2019–31 December 2023 8:00–17:30 (Interval 30 min) (Band extraction, crop)</td>
</tr>
</tbody>
</table>
The raw data utilized in this study to excavate false forest fire hotspots were sourced from the Himawari-8/9 meteorological satellite [11]. This satellite is capable of conducting full disc observations every 10 min [12], rendering it advantageous for forest fire monitoring; specific parameter descriptions are shown in Table 2. Channels 7 and 14 of the Himawari-8/9’s AHI sensor can be used for fire detection, using reflectivity information from channels 1, 3, 4 and 5 and brightness temperature information from channels 7, 14 and 16 to identify clouds, filter out flares and eliminate false alarms. The satellite is widely utilized for forest fire monitoring due to its high temporal resolution and data quality. In this study, the full-size observation data of Himawari-8/9 L1 NC (Network Common data Format) were downloaded from the Japan Meteorological Agency, and the satellite ground hotspot extraction and mining experiments were conducted after processing such as band extraction and image cropping, which were used to extract the satellite ground hotspot data as shown in Table 2.

Table 2. Himawari-8/9 parameters by band.

<table>
<thead>
<tr>
<th>Band</th>
<th>Centre Wavelength (µm)</th>
<th>Resolution (km)</th>
<th>Main Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.47</td>
<td>1.0</td>
<td>Oceanic water colors, Atmospheric environment</td>
</tr>
<tr>
<td>2</td>
<td>0.51</td>
<td>1.0</td>
<td>Oceanic water colors, Atmospheric environment</td>
</tr>
<tr>
<td>3</td>
<td>0.64</td>
<td>0.5</td>
<td>Land, Cloud</td>
</tr>
<tr>
<td>4</td>
<td>0.86</td>
<td>1.0</td>
<td>Oceanic water colors, Cloud</td>
</tr>
<tr>
<td>5</td>
<td>1.61</td>
<td>2.0</td>
<td>Land, Snow cover</td>
</tr>
<tr>
<td>6</td>
<td>2.26</td>
<td>2.0</td>
<td>Cloud</td>
</tr>
<tr>
<td>7</td>
<td>3.89</td>
<td>2.0</td>
<td>Surface temperature, Cloud top temperature</td>
</tr>
<tr>
<td>8</td>
<td>6.24</td>
<td>2.0</td>
<td>Cirrus, Atmospheric water vapor</td>
</tr>
<tr>
<td>9</td>
<td>6.94</td>
<td>2.0</td>
<td>Oceanic water colors, Phytoplankton</td>
</tr>
<tr>
<td>10</td>
<td>7.35</td>
<td>2.0</td>
<td>Oceanic water colors, Phytoplankton</td>
</tr>
<tr>
<td>11</td>
<td>8.59</td>
<td>2.0</td>
<td>Oceanic water colors, Phytoplankton</td>
</tr>
<tr>
<td>12</td>
<td>9.64</td>
<td>2.0</td>
<td>Atmospheric water vapor</td>
</tr>
<tr>
<td>13</td>
<td>10.41</td>
<td>2.0</td>
<td>Surface temperature, Cloud top temperature</td>
</tr>
<tr>
<td>14</td>
<td>11.24</td>
<td>2.0</td>
<td>Surface temperature, Cloud top temperature</td>
</tr>
<tr>
<td>15</td>
<td>12.38</td>
<td>2.0</td>
<td>Surface temperature, Cloud top temperature</td>
</tr>
<tr>
<td>16</td>
<td>13.28</td>
<td>2.0</td>
<td>Cloud top temperature</td>
</tr>
</tbody>
</table>

Table 3. Forest wildfire in Hunan Province, 2019–2022.

<table>
<thead>
<tr>
<th>Year</th>
<th>General Wildfires</th>
<th>Large Wildfires</th>
<th>Major/Particularly Important Wildfires</th>
<th>Area of Victimized Forests (Hectares)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>171</td>
<td>101</td>
<td>-</td>
<td>865</td>
</tr>
<tr>
<td>2020</td>
<td>33</td>
<td>26</td>
<td>-</td>
<td>284</td>
</tr>
<tr>
<td>2021</td>
<td>33</td>
<td>25</td>
<td>-</td>
<td>277</td>
</tr>
<tr>
<td>2022</td>
<td>86</td>
<td>86</td>
<td>1</td>
<td>1773</td>
</tr>
</tbody>
</table>

2.2.2. Satellite Data Processing

This study aims to enhance the accuracy of forest fire satellite monitoring by excluding false forest fire hotspots from satellite ground-based hotspots. The extracted satellite ground hotspots will serve as the experimental foundation for excavating false forest fire
hotspots. After pre-processing the Himawari-8/9 satellite images, satellite ground hotspots were extracted utilizing the spatial contextual method as outlined below:

Cloud detection: Clouds reduce brightness and increase pixel reflectivity. To enhance the monitoring accuracy of satellite ground hotspots, clouds in the image are identified and excluded using the following formula:

\[(\rho_3 + \rho_4 > 0.9) \text{ or } (BT_{15} < 265K) \text{ or } ((\rho_3 + \rho_4 > 0.7) \text{ and } (BT_{15} < 285K))\] (1)

Here, \(\rho_3\) refers to the reflectance of the 3rd band, \(\rho_4\) is the reflectance of the 4th band and \(T_{15}\) is the brightness temperature of the 15th band.

Identification of potential fire point pixels: To enhance the detection of potential forest fires to reduce the false alarm rate, the following formula can be employed to identify potential fire pixel elements in an image:

\[T_{07} > 280K \text{ and } (T_{07} - T_{14}) > 5K\] (2)

Here, \(T_{07}\) is the brightness temperature value of the central image element in band 7 of the Himawari-8/9 satellite, where \(T_{14}\) refers to the brightness temperature value of the central image element in band 14 of the Himawari-8/9 satellite.

Calculating the brightness temperature of background pixels and identifying fire point pixels: The initial window is centered on the potential fire point pixels, with a size of 3 × 3, to compute the mean and standard deviation of the high temperatures of background pixels within the window. If the proportion of background pixels is less than 20% of the total pixels in the window, the window size will be increased from 3 × 3 to 5 × 5, . . ., 11 × 11 for identification. If this criterion is still not met, the pixel will be labeled as a non-fire point. A potential fire point pixel is classified as a fire point pixel if it satisfies Formula (3).

\[\left\{\begin{array}{l}
T_{07} - \overline{T_{07}} > 4 \times \delta \times \overline{T_{07}} \\
\Delta T - \overline{\Delta T} > 3.5 \times \delta \times \overline{\Delta T}
\end{array}\right.\] (3)

Here, \(\overline{T_{07}}\) is the mean high temperature of the background image, \(\delta\) denotes the standard deviation of the brightness temperature within the window, \(\Delta T\) is the difference between the brightness temperature values of bands 7 and 14 in Himawari-8/9, and 4 and 3.5 are the thresholds for background coefficients, which are empirical thresholds, and according to the contextual method, the probability of fire point occurrence is 80% when 4 and 3.5 are taken during the daytime (4 and 3 are taken at night), and 70% if 3.5 and 3 are taken during the daytime (3 and 2.5 are taken at night).

2.3. Analysis of False Forest Fire Hotspots

Satellite monitoring of forest fire hotspots excludes ground hotspots caused by agricultural and other fires based on ground characteristics, thereby identifying forest fire hotspots. In addition to real forest fire hotspots, there may also exist three types of false forest fire hotspots, such as fixed heat sources, periodic heat sources and recurring heat sources, which have obvious time-series characteristics. Excluding these heat sources is crucial for enhancing forest fire monitoring accuracy. The subsequent section provides a pertinent characterization of false forest fire hotspots:

(1) Fixed heat source: Fixed heat sources in forests include dispersed small factories, forest barbecues and exposed rocks. These sources emit heat continuously and are geographically dispersed, making their positions difficult to pinpoint accurately. However, the continuous heat release from these hotspots aligns with satellite heat source extraction characteristics, leading to the identification of forest fire hotspots during satellite monitoring. Satellite ground hotspots extracted from forests over seven consecutive months are regarded as fixed heat sources.

(2) Periodic heat source: Periodic heat sources in forests refer to satellite ground hotspots with seasonal characteristics, which appear in the same season annually and are
relatively concentrated in location, producing heat source temperatures consistent with those extracted from satellite data. They are identified as forest fire hotspots during satellite monitoring. Satellite ground-based hotspots extracted consistently in the same season every year are classified as periodic heat sources.

(3) Recurring heat source: Recurring heat sources in forests are irregularly occurring hot spots that often manifest in the same location. These include activities such as rubbish burning, kiln burning and brick burning. These heat sources typically exhibit irregular and frequent behavior across various time periods. However, their high temperatures can lead to misclassification as real forest fires during forest fire detection. In fact, they are false forest fire hotspots, which somewhat compromise the accuracy of forest fire monitoring. Satellite ground hotspots that have accumulated more than 11 different days in the excavated forest are classified as recurring heat sources.

After processing Himawari-8/9 satellite images and Sentinel-2 satellite remote sensing images as well as analyzing false forest fire hotspots, the false forest fire hotspot mining experiments were started, and the experimental process and methods used in this study are shown in Figure 2.

**Figure 2.** Flow chart of the methodology.

2.4. Ground Object Classification Methods

This study excavates false forest fire hotspots with temporal characteristics in forests and extracts satellite ground hotspots in forests by classifying features in Hunan Province. Four methods, namely support vector machine, neural network, decision tree, and random forest, are employed for ground object classification in Xinshao and Xintian counties. Kappa coefficients, overall accuracy, producer accuracy and user accuracy are utilized as evaluation metrics for accuracy evaluation of these four classification methods. Then, the most accurate method will be selected to classify the ground object in Hunan Province.

For Sentinel-2 images of Xinshao and Xintian counties in ENVI 5.3 software, add six classes of land classes of training samples, and classification accuracy is greater than 1.8 in the Classification tool in the selection of support vector machine, neural network, decision tree and random forest in the four classification methods of land classification of the two areas. Each of the four classification methods possesses unique advantages. The support vector machine is a widely utilized algorithm in machine learning.
2.4. Ground Object Classification Methods

This study excavates false forest fire hotspots with temporal characteristics in forests and extracts satellite ground hotspots in forests by classifying features in Hunan Province. Four methods, namely support vector machine, neural network, decision tree, and random forest, are employed for ground object classification in Xinshao and Xintian counties. Kappa coefficients, overall accuracy, producer accuracy and user accuracy are utilized as evaluation metrics for accuracy evaluation of these four classification methods. Then, the most accurate method will be selected to classify the ground object in Hunan Province. For Sentinel-2 images of Xinshao and Xintian counties in ENVI 5.3 software, add six classes of land classes of training samples, six classes of samples classification accuracy are greater than 1.8 in the Classification tool in the selection of support vector machine, neural network, decision tree and random forest in the four classification methods of land classification of the two areas. Each of the four classification methods possesses unique advantages. The support vector machine is a widely utilized algorithm in machine learning and pattern recognition [13]. Based on statistical learning theory, the support vector machine utilizes optimal hyperplanes to classify in the sample or feature space, maximizing the spacing between different classes to enhance generalization. Feature classification with a support vector machine allows rapid classification with fewer training samples. The neural network approach is a mathematical model that mimics the synaptic connections of the brain to process information. The neural network approach to feature classification exhibits strong adaptability and high resistance to interference. Decision tree classification is a machine learning algorithm used for classifying and regressing tasks, primarily relying on remotely sensed image data and other spatial data through expert experience and simple mathematical statistics [14]. This classification is easy to understand, simple to model and its most significant advantage is its ability to handle uncorrelated feature data. Random forest is an ensemble learning model comprising decision trees as the basic classifiers, derived from the Bagging ensemble learning theory with the concept of feature subspace. Initially proposed by Breiman [15] in 2001, this method addresses the issue of decision trees being prone to overfitting, thereby improving their generalization performance and significantly enhancing classification.

2.5. Mining Methods

Satellite ground hotspots extracted from Himawari-8/9 satellite imagery were overlaid with the classified feature results. The remaining satellite ground hotspots located in forests were utilized as a dataset representing suspected forest fire hotspots. The suspected forest fire hotspots were exported from ArcGIS to Excel format. Leveraging the temporal dimension of spatio-temporal data, the suspected forest fire hotspots in Hunan from 2019 to 2023 are placed on a timeline to extract satellite ground hotspots with time-series characteristics. Time series data classification methods include similarity metrics, classification and visualization techniques [16]. Time series features are extracted to deeply excavate the potential information in satellite ground hotspots in Hunan Province from 2019 to 2023, thereby providing scientific support for decision-making.

There are two main approaches to common time series classification methods, time series classification methods based on computed distance and time series classification methods based on feature extraction [17,18]. Distance calculation involves determining the similarity between two time series based on their trends. The primary methods include Euclidean distance, edit distance and dynamic time warping. Feature extraction methods encompass those based on basic statistical methods and transform-based approaches. Time series feature representation methods encompass piecewise linear representation [19], piecewise aggregation approximation [20,21], symbolic aggregate approximation [22], etc., each exhibiting distinct differences and connections. Piecewise linear representation methods include sliding window, top-down, bottom-up, etc.

In this study, the sliding window algorithm, dynamic time warping and statistical method are selected to excavate the three types of heat sources based on their temporal
characteristics. A sliding window algorithm is employed for mining fixed heat sources, dynamic time warping is applied for mining periodic heat sources and a statistical method is utilized for mining recurring heat sources. To avoid duplicated ground hotspots among the three types of heat sources, they are mined sequentially as fixed, periodic and recurring heat sources when extracting satellite ground hotspots.

2.5.1. Sliding Window Algorithm

The sliding window algorithm [23] operates on a specific size of string characters or arrays, reducing the double nested loop problem to a single loop, thus effectively decreasing the time complexity. This algorithm is commonly employed to ascertain the characteristics of a continuous interval meeting specific conditions. The sliding window model is a data processing technique that keeps a dataset of fixed length constantly. In this model, the window slides with a constant length, continuously incorporating new data while discarding old data. With each slide of the window, there occurs data overlap, facilitating efficient segmentation of the data for processing. The operation of the sliding window is illustrated in Figure 3. When performing the sliding window algorithm, the input is an Excel format. It is an Excel sheet with information on the longitude, latitude, time and date of each satellite ground hotspot. The output Excel sheet is a satellite ground hotspot that matches the time-series characteristics. In this study, the sliding window algorithm is applied to excavate satellite ground hotspots with continuous features from the suspected forest fire hotspot dataset after overall dimensionality reduction. In Hunan Province, there are more cloudless periods from October to April compared to other months. The fixed heat source is time series features which possess continuity. Due to weather constraints, there are too many restrictions on extracting continuity features on a daily basis. Therefore, it is more feasible to choose satellite ground hotspots that have appeared continuously for more than 7 months as fixed heat sources. Firstly, the suspected forest fire hotspot data are grouped according to the hotspot, with a fixed window length of 12 months and a step size of 1 month. The original time series segmentation is performed based on each group of hotspot data. Following effective segmentation of each hotspot group using the sliding window algorithm, hotspots persisting for more than 7 months, exhibiting temporal continuity, are extracted from the sub-sequences.

Figure 3. Sliding Window Working Principle Diagram. We used suspected forest fire hotspot data in Excel format to mine fixed heat sources according to window settings and step size.

2.5.2. Dynamic Time Warping

Dynamic time warping matches time series by stretching or shrinking them to compute similarity between time-series data, thus enabling measurement of similarity for time series of varying lengths [24]. Fixed heat sources excavated from suspected forest fire hotspot datasets were removed. Dynamic time warping was then employed to excavate periodic heat sources exhibiting seasonal patterns, occurring annually during the same seasons from 2019 to 2023. Initially, a standard sequence is established, with the date of the middle day of the spring, summer, autumn and winter seasons of each year serving as the starting
point. The distance matrix between two sets of data, the time series and the standard series of the suspected forest fire hotspot data is computed using the following formula:

\[ d(i, j) = (s_i - t_j)^2, \quad 1 \leq i \leq n, 1 \leq j \leq m \] (4)

Secondly, the distance between the time series of suspected forest fire hotspot data and the standard series is computed using dynamic time warping, the cumulative distance matrix is calculated and the values in the lower right corner of the dynamic time warping matrix are extracted to determine the distance from the standard series at each time point. Next, the distances obtained are normalized by dividing them by the maximum value in the distance matrix. This normalization process aims to identify the optimal warping path that best aligns the shapes of the two time series, allowing for potential stretching or bending along the time axis, ultimately yielding the optimal warping path. The formula for computing the cumulative distance matrix is given by the following:

\[ D(i, j) = d(i, j) + \min\{D(i - 1, j), D(i, j - 1), D(i - 1, j - 1)\} \] (5)

The curved path, denoted as \( W = (w_1, w_2, \ldots, w_k) \), refers to the path with the minimum cumulative distance between two time series. The value of \( K \) ranges from \( \max(m, n) \leq K \leq m + n - 1 \), \( w_i = (i, j) \). The optimal warping path between two time series is represented by \( DTW(S, T) = D(n, m) \). Figure 4 illustrates the schematic diagram of the curved path. Bending paths possess the following three crucial properties [25]:

(1) Borderline: Any bending path \( w \) must satisfy the condition where the starting point is \( w_1 = (1, 1) \) and the ending point is \( w_k = (n, m) \), and the sequential order of the sequence cannot be altered during the construction of the optimal bending path using dynamic data bending.

(2) Monotonicity: If \( w_{k-1} = (i', j') \) and \( w_k = (i, j) \) are two neighboring points on the curved path, then \( 0 \leq (i - i') \) and \( 0 \leq (j - j') \) must be satisfied to ensure that the alignment process of the time series always maintains a unidirectional and incremental character, preventing cross-correspondence.

(3) Contiguity: If \( w_{k-1} = (i', j') \) and \( w_k = (i, j) \) are any two neighboring points on the bending path, then \( (i - i') \leq 1 \) and \( (j - j') \leq 1 \) must be satisfied, ensuring that the bending path can only move along the diagonal, horizontal or vertical directions, thus avoiding out-of-value matching.

![Figure 4. Schematic diagram of the curved path.](image-url)
Periodic heat sources exhibit seasonal characteristics, such as satellite ground hotspots monitored consistently in the same season each year. This study utilizes suspected forest fire hotspot data spanning five years from 2019 to 2023. According to the features of periodic heat sources, a hotspot data point occurring five times in any season (spring, summer, autumn or winter) in the standard sequence fulfills the conditions of a periodic heat source.

2.5.3. Statistical Method

Feature extraction using basic statistical methods is among the most straightforward approaches for time series classification [26,27]. A statistic is a generalized measure utilized to characterize the time series, and it is employed to estimate the parameters of the entire dataset, thereby extracting its characteristics. The temporal characteristics of recurring heat sources include irregular timing and frequent recurrence at specific locations. Recurring heat sources are identified by tallying the occurrences over multiple days. Before mining for recurring heat sources, fixed and periodic heat sources from the suspected forest fire hotspot dataset are excluded to ensure the uniqueness of satellite ground hotspots among the three types of heat sources. Forest wildfires are identified when they are observed over multiple consecutive days, even after the open flames are extinguished, and residual smoke and high heat emanating from the ground are still detected by satellites as high-temperature points for continued monitoring. The longest duration of a forest wildfire in Hunan Province between 2019 and 2023 was 10 days. To prevent satellite ground hotspots from such incidents from being included in the results of recurring hotspots, satellite ground hotspots monitored cumulatively for more than 11 days on various days in the suspected forest fire hotspot dataset for 2019–2023 were categorized as recurring hotspots. Firstly, the remaining suspected forest fire hotspot dataset is grouped based on hotspot and the preset conditions. Multiple sub-sequences are then derived. For each sub-sequence, satellite ground hotspots that have appeared for more than 11 days on different dates are extracted, thereby completing the feature extraction of the original sequence.

3. Results

Satellite ground hotspots extracted from Himawari-8/9 satellite images are superimposed with the feature classification outcomes. The remaining satellite ground hotspots located in forests are utilized as a dataset representing suspected forest fire hotspots. The sliding window algorithm, dynamic time warping, statistical method and other time-series classification methods are used to excavate the satellite ground hotspots such as fixed heat sources, periodic heat sources and recurring heat sources from the suspected forest fire hotspot dataset. These types of hotspots are then removed to obtain a more accurate forest fire hotspots dataset. The above methods and results are validated using forest fire hotspot data verified by the Hunan Provincial Department of Emergency Management.

3.1. Ground Object Classification Results

Ground objects in Hunan Province are classified into six categories according to the national standard Classification of Current Land Use Status (GB/T21010-2017) [28] and the Land Management Law of the People’s Republic of China [29], which are organized and revised by the Ministry of Land and Resources of China. These categories include forest, grassland, cropland, impervious, barren and water. Satellite ground hotspots occur in cropland during agricultural fires and in grassland, water, impervious and barren during other types of fires. Figure 5 displays the results map of four classification methods: Support Vector Machine, Neural Network, Decision Tree and Random Forest, used for classifying features in Xinshao County. Table 4 presents the evaluation of these methods in terms of kappa coefficients, overall accuracy, producer’s accuracy and user’s accuracy for the classification results in Xinshao County.
Figure 5. Map of land classification results for Xinshao County.

Table 4. Evaluation of forest classification accuracy results in Xinshao County.

<table>
<thead>
<tr>
<th>Classification Method</th>
<th>Overall Accuracy (%)</th>
<th>Kappa</th>
<th>Producer’s Accuracy (%)</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network</td>
<td>85.73</td>
<td>0.85</td>
<td>77.78</td>
<td>84.09</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>79.70</td>
<td>0.79</td>
<td>70.67</td>
<td>76.58</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>82.61</td>
<td>0.82</td>
<td>75.08</td>
<td>79.98</td>
</tr>
<tr>
<td>Random Forest</td>
<td>91.48</td>
<td>0.90</td>
<td>88.90</td>
<td>89.42</td>
</tr>
</tbody>
</table>

Figure 6 illustrates the outcomes of four classification methods—Support Vector Machine, Neural Network, Decision Tree and Random Forest—used to classify features in Xintian County. Table 5 presents the evaluation of these methods in terms of kappa coefficients, overall accuracy, producer accuracy and user accuracy for the classification results in Xintian County.

Table 5. Evaluation of forest classification accuracy results in Xintian County.

<table>
<thead>
<tr>
<th>Classification Method</th>
<th>Overall Accuracy (%)</th>
<th>Kappa</th>
<th>Producer’s Accuracy (%)</th>
<th>User’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural Network</td>
<td>86.55</td>
<td>0.86</td>
<td>79.04</td>
<td>84.23</td>
</tr>
<tr>
<td>Support Vector Machine</td>
<td>82.59</td>
<td>0.82</td>
<td>76.32</td>
<td>80.53</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>81.39</td>
<td>0.81</td>
<td>78.64</td>
<td>80.46</td>
</tr>
<tr>
<td>Random Forest</td>
<td>92.31</td>
<td>0.90</td>
<td>89.56</td>
<td>90.96</td>
</tr>
</tbody>
</table>
The evaluation metrics of the accuracy results in Tables 4 and 5 show that the Random Forest method has the best accuracy in classifying features in both regions. Xinshao County has an overall accuracy of 0.90, a kappa coefficient of 91.48 per cent, a producer accuracy of 88.90 per cent and a user accuracy of 89.42 per cent, and Xintian County has an overall accuracy of 0.90, a kappa coefficient of 92.31 per cent, a producer accuracy of 89.56 per cent and a user accuracy of 90.96 per cent in the results for the classification of the ground object in Xintian and Xinshao counties. Following accuracy comparison, the Random Forest, demonstrating the highest accuracy, is selected for feature classification in Hunan Province, as illustrated in Figure 7. Satellite ground hotspots extracted from Himawari-8/9 satellite images are overlaid on the feature classification results. Agricultural fires and other types of fires are excluded, and only satellite ground hotspots in the forests of Hunan Province are extracted, amounting to a total of 1247 satellite ground hotspots. Satellite ground hotspots in Hunan forests between 2019 and 2023 served as a dataset of suspected forest fire hotspots for the subsequent excavation of false forest fire hotspots.
Random Forest, demonstrating the highest accuracy, is selected for feature classification in Hunan Province, as illustrated in Figure 7. Satellite ground hotspots extracted from Himawari-8/9 satellite images are overlaid on the feature classification results. Agricultural fires and other types of fires are excluded, and only satellite ground hotspots in the forests of Hunan Province are extracted, amounting to a total of 1247 satellite ground hotspots. Satellite ground hotspots in Hunan forests between 2019 and 2023 served as a dataset of suspected forest fire hotspots for the subsequent excavation of false forest fire hotspots.

Figure 7. Map of the results of land classification of landforms in Hunan Province using random forests. The classification results of the six land classes in Hunan Province are shown in (a). The forest types are extracted from (a) to obtain (b). Figure (b) depicts the high forest coverage in Hunan Province predominantly concentrated in the western region, particularly in Huaihua City, Xiangxi Tujia and Miao Autonomous Prefecture and Zhangjiajie City. The northeastern region of the country is primarily characterized by extensive land use for watersheds and water management infrastructure, leading to a limited forested area.

3.2. Mining Results

3.2.1. Fixed Heat Source

Utilizing the dataset of suspected forest fire hotspots in Hunan from 2019 to 2023, the sliding window algorithm is employed to excavate fixed heat sources. The excavated outcomes are visualized using the ArcGIS platform, illustrated in Figure 8, with red dots representing the fixed heat sources. Through statistical analysis, a total of 38 satellite ground hotspots exhibiting continuity characteristics have been identified in the forested areas of the province. Figure 8 illustrates that the majority of fixed heat sources are concentrated in the cities of Chenzhou, Huaihua and Loudi, all of which have forest cover exceeding 50% of the city’s area.

3.2.2. Periodic Heat Source

Utilizing the dataset of suspected forest fire hotspots in Hunan from 2019 to 2023, the dynamic time warping is employed to excavate periodic heat sources, and the results are visualized in the ArcGIS platform. Figure 9 displays the excavation outcomes, with green dots indicating the periodic heat sources. Statistical analysis reveals 98 false forest fire hotspots recurring in the same seasons over a five-year period across the province. The seasonal statistics of the 98 periodic heat sources are graphically depicted in Figure 10.
3.2.2. Periodic Heat Source

Utilizing the dataset of suspected forest fire hotspots in Hunan from 2019 to 2023, the dynamic time warping is employed to excavate periodic heat sources, and the results are visualized in the ArcGIS platform. Figure 9 displays the excavation outcomes, with green dots indicating the periodic heat sources. Statistical analysis reveals 98 false forest fire hotspots recurring in the same seasons over a five-year period across the province. The seasonal statistics of the 98 periodic heat sources are graphically depicted in Figure 10.

Figure 9. Periodic Heat Source Excavation Results Map.

Figure 10 illustrates that the spring season has the highest number of periodic heat sources among the four seasons, with 50 satellite ground hotspots observed, comprising 51.02% of the total. However, winter has the fewest number, with six satellite ground hotspots detected, representing 6.1% of the total periodic heat source.

3.2.3. Recurring Heat Source

Utilizing the dataset of suspected forest fire hotspots in Hunan from 2019 to 2023, the statistical method is employed to excavate recurring heat sources. The visualization of these recurring heat sources is presented on ArcGIS, depicted in Figure 11, with blue dots.
Figure 10. Seasonal Statistical Map of Periodic Heat Source Hotspots.

3.2.3. Recurring Heat Source

Utilizing the dataset of suspected forest fire hotspots in Hunan from 2019 to 2023, the statistical method is employed to excavate recurring heat sources. The visualization of these recurring heat sources is presented on ArcGIS, depicted in Figure 11, with blue dots representing the recurring heat source hotspots. Utilizing the statistical method, 328 recurring heat sources were excavated, characterized by irregular temporal occurrences and frequent repetition in the same location.

Figure 11. Recurring Heat Source Excavation Results Map.
3.2.4. Analysis of Result

The results show that of the 1247 suspected forest fire hotspots in Hunan from 2019 to 2023, 38 fixed heat sources, 98 periodic heat sources and 328 recurring heat sources are excavated, and the remaining 783 are forest fires, with an average of 157 forest fires per year. The National Bureau of Statistics of the People’s Republic of China reports an average of 141 forest fires annually in Hunan Province from 2019 to 2022. This statistic excludes fires that go undetected, remain unreported, scheduled burnout, burn briefly, or cause no damage. The forest fires monitored in this study are generally similar to the actual situation.

3.3. Forest Fire Monitoring Application Case

This study established a dataset of false forest fire hotspots, including fixed, periodic and recurring heat sources, in Hunan forests from 2019 to 2023. As additional satellite ground hotspot data accumulates, this paper’s method enables the continuous mining and incorporation of false forest fire hotspots into the dataset, facilitating dynamic updates. In forest fire monitoring, suspected hotspots are extracted from each Himawari-9 image. A computer automatically compares suspected forest fire hotspots with the false forest fire hotspot dataset, and directly excludes false forest fire hotspots from the suspected data.

Himawari-9 satellite images of Hunan were chosen for 20 time points at 30 min intervals, ranging from 8:00 to 17:30 on 12 January 2024. The flowchart illustrating satellite forest fire monitoring at each time point is depicted in Figure 12.

![Forest fire monitoring flowchart](image)

**Figure 12.** Forest fire monitoring flowchart. When conducting forest fire monitoring, we are able to arbitrarily select a certain moment and first exclude satellite ground hotspots outside the forest based on the feature characteristics as the suspected forest fire hotspot dataset. The forest fire can be obtained by excluding the false forest fire hotspot dataset based on the false forest fire hotspot dataset constructed in this study. Emergency management forest fire verification data and comparison of land classes in high resolution imagery were also utilized in this study to verify the accuracy of forest fire monitoring.

Following the forest fire monitoring flowchart, satellite ground hotspots in Hunan are extracted from Himawari-9 images at 20 time points on 12 January 2024. Subsequently, a spatial distribution map of these hotspots is generated and overlaid with feature classification maps on the ArcGIS platform to identify suspected forest fire hotspots, as illustrated in Figure 13. On 12 January 2024, from 08:00 to 17:30, a total of 63 suspected forest fire hotspots were monitored in Hunan. A computer automatically compared these 63 hotspots with the false forest fire hotspot dataset from 2019 to 2023. The results show that, out of 63 suspected forest fire hotspots, 21 are periodic heat sources, 1 is a recurring heat source in the false forest fire hotspot dataset and the remaining 41 are forest fires, as shown in Table 6 and Figure 13.
The forest fire monitoring methodology and results were validated by superimposing the forest fire hotspot data, verified by the Hunan Provincial Department of Emergency Management on 12 January 2024, onto high-resolution imagery on the ArcGIS platform. The validation results indicate that out of the 41 forest fires monitored in this study, there were 2 false forest fire hotspots resulting from a feature classification error. There are 39 genuine forest fires confirmed, which match in number and similar location as those verified by the Hunan Provincial Department of Emergency Management. Consequently, the accuracy of forest fire monitoring is 95.12%.

**Figure 13.** Map of forest fire monitoring results. Figure (a) shows the distribution of suspected forest fire hotspots on 12 January 2024 in Hunan Province. Figure (b) shows 41 forest fires and 22 false forest fire hotspots in Hunan Province on 12 January 2024 monitored in this study. Figure (c) shows a close-up screenshot of some of the correctly classified forest fire hotspots monitored in this study in (b) with small light blue boxes located in Huaihua, Zhuzhou and Chenzhou. Figure (d) shows the small red boxes in (b), located in Xiangtan and Changde, and is a close-up screenshot of 2 of the 41 forest fires monitored in this study that are false forest fire hotspots.
Table 6. False forest fire hotspot results statistics.

<table>
<thead>
<tr>
<th>Juncture</th>
<th>Total Hotspots</th>
<th>Suspected Forest Fire Hotspots</th>
<th>False Forest Fire Hotspot Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fixed Heat Source</td>
</tr>
<tr>
<td>8:00</td>
<td>5</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>10:04</td>
<td>4</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>10:32</td>
<td>7</td>
<td>4</td>
<td>-</td>
</tr>
<tr>
<td>11:00</td>
<td>27</td>
<td>18</td>
<td>-</td>
</tr>
<tr>
<td>11:34</td>
<td>10</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>12:02</td>
<td>12</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>13:04</td>
<td>15</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>13:32</td>
<td>20</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>14:00</td>
<td>6</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>14:34</td>
<td>30</td>
<td>9</td>
<td>-</td>
</tr>
<tr>
<td>15:02</td>
<td>8</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>15:30</td>
<td>12</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td>16:04</td>
<td>7</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

4. Discussion

Excluding false forest fire hotspots is crucial for enhancing the accuracy of forest fire monitoring. When using the method of this study for forest fire monitoring, there were still two false forest fire hotspots caused by the error of ground object classification. For more accurate results, further optimization of the random forest model is needed, which is also a problem for further research in the future. Previous research has investigated the issue of false forest fire hotspots in fire detection. Early warning capabilities are attained through smoke detection systems [30], while UAVs are utilized for remote forest fire monitoring [31]. However, smoke detection systems and drone monitoring can be limited in eliminating false forest fire hotspots, as smoke targets may be too small or smoke features may be indistinct in complex natural environments. Single satellite imagery, on the other hand, is time-consuming to exclude false forest fire hotspots. In this study, we use multi satellite images combined with data mining techniques to deeply excavate false forest fire hotspots with temporal characteristics in forests and build a dataset, which is a more time-saving and precise solution.

Spatio-temporal data mining uncover valuable insights in spatio-temporal data. This data can be analyzed and modeled by either considering temporal and spatial dimensions separately, or integrating temporal and spatial features. Research in spatio-temporal data mining advances significantly across diverse fields. Spatio-temporal data enables the tracking of contemporary wildlife and facilitates the examination of changes in the marine environment [32,33]. Additionally, integrating satellite images with data mining techniques enables the automated extraction of spatial and temporal information from satellite image time series data [34]. This approach is also employed for studying land change [35,36]. However, there is still limited research on utilizing data mining techniques in conjunction with historical satellite ground hotspot data to excavate false forest fire hotspots. When monitoring forest fires with existing methods, single remote sensing images are used to distinguish forest fires from non-forest fires based on existing methods. In the next monitoring, forest fire monitoring methods such as cloud detection, potential fire point pixel recognition, background pixel brightness temperature statistics and fire point pixel confirmation need to be used for forest fire identification. Moreover, single remote sensing images cannot be used to monitor forest fires using temporal features. However, by combining the data mining technique with the satellite ground hotspot data, it becomes possible to comprehensively excavate the false forest fire hotspots hidden within forests. These excavated false forest fire hotspots are established as a false forest fire hotspot dataset. Such a dataset can not only distinguish non-forest fires from forest fires but also clarify the type of false forest fire hotspots. This study utilizes data mining techniques along with satellite ground hotspots extracted from Himawari-8/9 satellite images from 2019 to 2023.
in Hunan Province, based on feature classification. The aim is to excavate three types of false forest fire hotspots: fixed heat sources, periodic heat sources and recurring heat sources within forests, and to build up a dataset for false forest fire hotspots. In forest fire monitoring, satellite ground hotspots from the dataset of false forest fire hotspots can be directly excluded, enhancing the precision of detecting genuine forest fires. This approach improves the accuracy of forest fire monitoring and introduces a novel perspective for distinguishing false forest fire hotspots. Cloud and water reflections during satellite-based forest fire monitoring may lead to erroneous hotspots. Future research in this domain needs to be intensified to enhance the precision of forest fire monitoring.

5. Conclusions

Utilizing ground object features to exclude false forest fire hotspots facilitates the exclusion of satellite ground hotspots occurring outside forested areas. In addition to real forest fire hotspots, there are three types of false forest fire hotspots, including fixed heat sources, periodic heat sources and recurring heat sources. This study uses Hunan Province as the research object, and utilizes the optimal classification method based on ground object features to categorize and extract forest types. Optimal classification methods are chosen to pinpoint individual types of features and to more accurately extract satellite ground hotspots in the forest. Satellite ground hotspot data extracted from Himawari-8/9 satellite imagery for the years 2019–2023 are used in conjunction with data mining techniques to mine three types of false forest fire hotspots in forests with continuous characteristics, seasonal characteristics and irregular time characteristics. The sliding window algorithm, dynamic time warping, statistical method and other time-series classification methods are used to excavate the satellite ground hotspots such as fixed heat sources, periodic heat sources and recurring heat sources from the suspected forest fire hotspot dataset. The method and results of this study are implemented in forest fire monitoring with high accuracy, which proves that the mining method of this study can improve the accuracy of forest fire satellite monitoring. The method and results of this study provide a new direction and means for the exclusion of false forest fire hotspots.

Author Contributions: Conceptualization: G.Z.; supervision: G.Z. and Z.Y.; methodology, collected the study data, formal analysis and writing—original draft: H.W.; writing—review and editing: H.X., F.L. and S.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a project supported by the Scientific Research Fund of Hunan Provincial Education Department under Grant 23B0244 and 22A0194, the Science and Technology Innovation Platform and Talent Plan Project of Hunan Province under Grant 2017TP1022, the Natural Science Foundation of Hunan Province under Grant 2024JJ7645 and Field Observation and Research Station of Dongting Lake Natural Resource Ecosystem, Ministry of Natural Resources.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We thank the Department of Emergency Management of Hunan Province for their provision of forest fire hotspot data.

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

References


**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.