remote sensing

Article

A Universal Framework for Near-Real-Time Detection of Vegetation Anomalies from Landsat Data

Yixuan Xie [1,2], Zhiqiang Xiao [1,*], Juan Li [1], Jinling Song [1], Hua Yang [1] and Kexin Lv [1]

1 State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
2 China Siwei Surveying and Mapping Technology Co., Ltd., Beijing 100190, China
* Correspondence: zhqxiao@bnu.edu.cn

Abstract: Vegetation anomalies are frequently occurring and may greatly affect ecological functions. Many near-real-time (NRT) detection methods have been developed to detect these anomalies in a timely manner whenever a new satellite observation is available. However, the undisturbed vegetation conditions captured by these methods are only applicable to a particular pixel or vegetation type, resulting in a lack of universality. Also, most methods that use single characteristic parameter may ignore the multi-spectral expression of vegetation anomalies. In this study, we developed a universal framework to simultaneously detect various vegetation anomalies in NRT from Landsat observations. Firstly, Landsat surface reflectance data from the Benchmark Land Multisite Analysis and Intercomparison of Products (BELMANIP) sites were selected as a reference vegetation dataset to calculate the normalized difference vegetation index (NDVI) and the normalized burn ratio (NBR), which describe vegetation conditions from the perspectives of greenness and moisture, respectively. After the elimination of cloud-contaminated pixels, the high-quality NDVI and NBR data over the BELMANIP sites were further normalized in order to remove the differences in the growth of the varying vegetation. Based on the normalized NDVI and NBR, kernel density estimation (KDE) was used to create a universal measure of undisturbed vegetation, which described the uniform spectral frequency distribution of different undisturbed vegetation with a series of accumulated probabilities on a monthly basis. Whenever a new Landsat observation is collected, the vegetation anomalies are determined according to the universal measure in NRT. To demonstrate the potential of this framework, three study areas with different anomaly types (deforestation, fire event, and insect outbreak) in distinct ecozones (rainforest, coniferous forest, and deciduous broad-leaf forest) were used. The quantitative analyses showed generally high overall accuracies (>90% with the kappa > 0.82). The user accuracy for the fire event and the producer accuracy for the earlier insect infestation were relatively lower. The accuracies may be affected by the complexity of the land surface, the quality of the Landsat image, and the accumulated probability threshold.

Keywords: anomaly detection; near-real-time; kernel density estimation; Landsat

Academic Editor: Lênio Soares Galvão

Received: 4 December 2024
Revised: 28 January 2025
Accepted: 31 January 2025
Published: 3 February 2025

Citation: Xie, Y.; Xiao, Z.; Li, J.; Song, J.; Yang, H.; Lv, K. A Universal Framework for Near-Real-Time Detection of Vegetation Anomalies from Landsat Data. *Remote Sens.* **2025**, *17*, 520. https://doi.org/10.3390/rs17030520

1. Introduction

Increasingly active human activities have become one of the primary drivers of ecological environmental changes [1]. In addition, surface discrete events such as locust plagues [2], wildfires [3], and floods [4] are occurring with greater frequency in the context of global warming and frequent climate extremes [5]. As one of the major components of the land surface, forest and other plants play a vital role in the carbon and water cycles.

In this study, vegetation anomalies are defined as instances of poor growth or even cover loss due to human or natural factors such as deforestation [6,7], fire events [8,9], insect outbreaks [10,11], hurricanes [12], and so on. Large-scale and long-duration vegetation anomalies can have a detrimental impact on the stability of ecosystem services, as they impede the photosynthesis and transpiration processes. Consequently, it is of paramount importance to detect these vegetation anomalies in an accurate and timely manner [13].

Remote sensing has been the primary technology employed in vegetation monitoring for decades. Numerous algorithms have been devised to detect anomalies in vegetation from satellite observations. Early methods commonly detect anomalies by single image analysis or bi-temporal differences. For example, a threshold was set for an anomaly-sensitive index on a single image to detect the European spruce bark beetle attack [14]. In the case of identifying the burned pixels in the western boreal of Canada, a threshold was set for the discrepancy between pre-fire and post-fire images [15]. In addition to the use of threshold judgment, classification methods are also employed. For example, the change areas in the northern forest of New England were identified by comparing the bi-temporal classification results [16]. Based on the no-forest change and forest change training samples, the forest anomaly maps in the Cascade Range of Oregon were created by directly classifying the paired bi-temporal images [17,18]. Although these methods are capable of accurately detecting, they are highly reliant on the region's specific characteristics when determining the threshold value and collecting training samples.

Applying cloud-free time series stacks in vegetation anomaly detection has been extensively studied over the past decade. The Vegetation Change Tracker (VCT) method [17] has demonstrated high accuracy in detecting harvest, fire, and urban development in the United States. The method calculates an integrated forest z-score (IFZ) for the years being monitored based on the forest samples. Whether the forest has changed can be discovered by the number of consecutive high or low IFZ values. Algorithms such as Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr) [19] and Detecting Breakpoints and Estimating Segments in Trend (DBEST) [20] segment the trend series into a sequence of linear fitted straight lines. Breakpoints connecting neighboring segments are regarded as potential vegetation cover changes. LandTrendr is designed for Landsat and performed well on several hundred points across western Oregon and Washington in the United States. DBEST, in contrast to LandTrendr, is not constrained by any specific sensor or spatio-temporal scale. Other methods, such as Breaks for Additive Season and Trend (BFAST) [21], use an additive decomposition model to decompose the entire time series into trend, seasonal, and residual components and detect breakpoints in the first two components. Smith et al. [22] utilized the BFAST method in tropical dry forest deforestation in Mexico and Costa Rica. The Bayesian Estimator of Abrupt change, Seasonal change, and Trend (BEAST) method [23] is similar to BFAST. The difference is that the BEAST method integrates all potential decomposition models and offers an anomaly probability. These algorithms are effective when capturing vegetation dynamics over a long period, but they aim to reconstruct anomaly history, which means retrospectively detecting anomalies in an existing historical period [13]; thus, they are lacking instantaneity.

Recently, researchers have devoted greater attention to the near-real-time (NRT) detection of vegetation anomalies from remote sensing data. NRT detection methods emphasize recognizing vegetation anomalies from each new satellite observation. This is significant in timely grasping anomaly extents and releasing early warnings to prevent further loss [24]. Parametric approaches are dominant in the relevant literature. The fundamental principle of these approaches is commonly that using parametric functions such as cosine or sine, double logistic or harmonic cycles to fit a stable periodic time series for each pixel in order to capture the undisturbed vegetation condition [25–30]. When a new satellite image is collected, the anomaly is identified by the residual between the new observation and the prediction of the fitted time series. Representative approaches are as below: BFAST Monitor [22,31] is a modification of BFAST and has been used to monitor drought disturbance in NRT in Somalia. Continuous Change Detection and Classification (CCDC) enables NRT anomaly detection although it is also applicable for the above retrospective detection [32]. Modified from CCDC, NRT-CCDC [33] improves data input, anomaly judgment criterion and anomaly probability offering. The COntinuous Monitoring of Land Disturbance (COLD) [30] also improves upon the CCDC by optimizing data input and filtering out irrelevant changes. Moreover, Stochastic Continuous Change Detection (S-CCD) [13] improved upon the COLD algorithm by incorporating the state space theory to fit the temporal dynamics of satellite observations in a recursive manner, thereby enhancing the accuracy and computational efficiency of NRT anomaly detection.

The nonparametric kernel density estimation theory was firstly employed in the field of vegetation anomaly detection by Estay et al. [34]. It was integrated into the R package 'npphen', which has been examined in insect outbreak cases [35,36] and drought evaluation of a wildfire event [37]. This method is appropriate for both NRT and retrospective detection. For each pixel in the study area, the 'npphen' package allows for the estimation of the spectral frequency distribution within one undisturbed phenological annual cycle from its stable historical records. Then, anomalies can be determined according to the new observation's position relative to this distribution. This method is free from the errors that arise from time series fitting due to a few outliers and missing data. Currently, the Anomaly Vegetation Change Detection (AVOCADO) algorithm [26] provides an enlightening idea of NRT continuous monitoring based on the 'npphen' package. Differentiating itself from 'npphen', the historical records in AVOCADO are derived from the neighboring pixels that share the same vegetation type as the study area. The AVOCADO attempts to estimate a unified phenological distribution for the entire study area in advance; therefore, there is no need to set aside and process the historical baseline for each pixel during the detection process, which significantly accelerates the detection process. The performance of the AVOCADO algorithm was evaluated in three tropical forest ecosystems with anomalies such as selective logging and shifting agriculture.

Parametric approaches may not be optimal for semi-arid and arid ecosystems, where vegetation phenology cycles cannot be represented by regular annual waves [38]. This is due to the fact that mathematical functions with explicit seasonal components may not be appropriate for such ecosystems. In addition, they are computationally expensive and may be heavily dependent on the quality and speed of accurate fitting [13,39]. In contrast, the 'npphen' package is more flexible and can be applied to any ecosystem, as it captures the vegetation phenology directly from statistical estimation. However, it is still necessary to estimate the undisturbed vegetation condition separately at the pixel scale, in a manner analogous to the parametric approaches. Although the AVOCADO algorithm represents a significant advance over the 'npphen' package, utilizing a reference vegetation pixel dataset to create an undisturbed vegetation distribution for a specific study area, there are still some limitations. Firstly, the representative reference vegetation must be re-selected when applied to other areas with different vegetation types. Moreover, the detection process will be laborious when applied to larger areas with diverse vegetation types. In such instances, it is necessary to determine the specific vegetation type for each pixel and to create the undisturbed distribution for each vegetation type [26]. In essence, the current methods are inadequate for some special ecosystems. Furthermore, the undisturbed vegetation conditions captured by these methods vary from pixels and vegetation types, which limit their universality.

In terms of variable input, instead of directly utilizing multi-band spectral reflectance [30,32], a large part of current algorithms just employ single index related to vegetation condition, like the NDVI [27,40], the EVI [36,39,41], the RGI [25], the NDMI [26,42], etc. In fact, some studies reported that combining multiple bands or indices can provide more comprehensive information to realize accurate anomaly recognition in contrast to using just a single band or index, because vegetation anomalies usually have a multi-spectral expression [32,43,44].

This study proposes a universal framework for vegetation anomaly detection (UFVAD) to robustly detect anomalies in NRT from Landsat observations. UFVAD uses vegetation parameters calculated from observations over the BELMANIP sites. It develops a universal measure of undisturbed vegetation via kernel density estimation (KDE). This measure monthly describes the undisturbed vegetation distributions with a series of accumulated probabilities. Based on this measure, UFVAD can detect anomalies in NRT whenever a new observation is available. To evaluate the performance of UFVAD, it was applied to detect various vegetation anomalies (deforestation, fire event, and insect outbreak) in three forest areas with different ecozones.

2. Study Area and Landsat Data

2.1. Study Area

Three study areas were selected according to their ecozones and vegetation anomaly types (shown in Figure 1). The first area (denoted by SA) is located in Guaviare, Southeastern Colombia. The most important vegetation type in Guaviare is the tropical humid forest. Driven by armed groups' illegal occupation, grazing, road construction and other factors, Guaviare became one of the departments with the highest deforestation rate in Colombia [45]. From 2002 to 2022, Guaviare has lost 301 kha of humid primary forest, corresponding to a 6.1% reduction [46]. SA is near the western boundary of Guaviare (center coordinate: 1°56′N, 73°28′W), where small-scale rainforest deforestation has occurred since 2018.

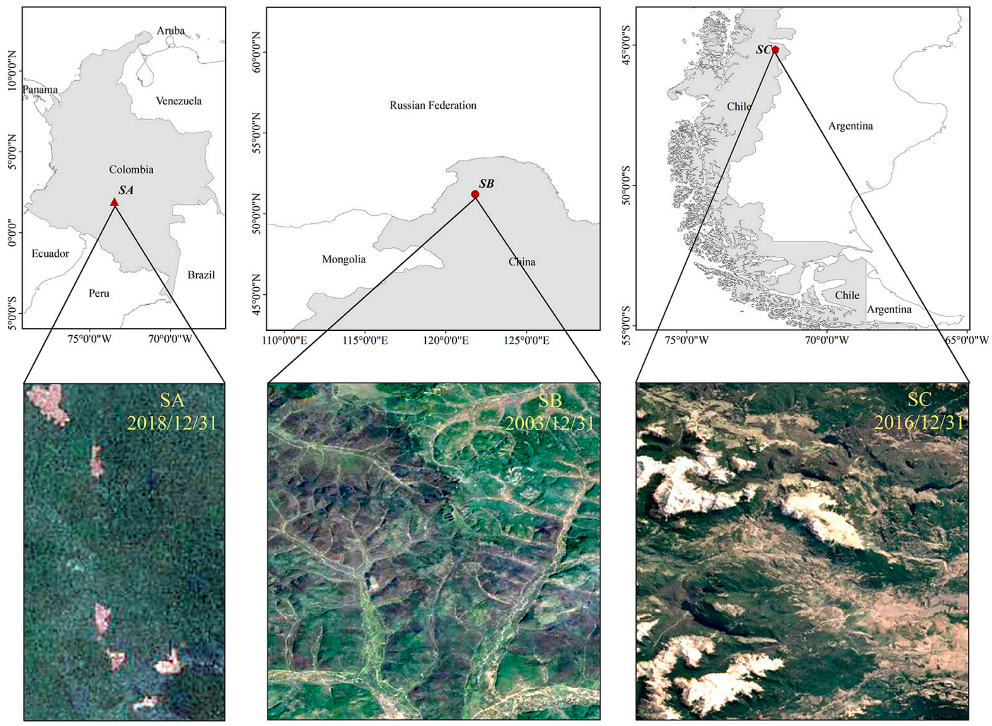

Figure 1. Overview of the three study areas. SA (red triangle), located in Guaviare, Colombia, is a tropical humid forest area with deforestation patches since 2018. SB (red circle), located in Genhe, Inner

Mongolia, China, is a frigid coniferous forest area with a catastrophic fire event in 2003. SC (red pentagon), located in Aysén Region, Chilean Patagonia, is a deciduous broad-leaf forest area with a large-scale forest defoliation event caused by an *Ormiscodes amphimone* outbreak during the 2015 to 2016 growing season. Optical images in the bottom row are from Google historical records in the year in which the vegetation anomaly occurred.

The second area (denoted by SB) is located in Genhe, Inner Mongolia, China, in the northern segment of the Greater Khingan Mountains. Vegetation types in Genhe are mainly coniferous forests such as Larix gmelinii and Pinus sylvestris var. mongolica, which account for more than 76% of the total forest area [47]. The climate is characterized by short summer and extremely cold and long winter with the snowfall period from late September to early May of the following year, which leads to a short vegetation growing season [48]. The center coordinate of SB is (51°12′N, 121°49′E). In 2003, a catastrophic fire event happened in this area, causing severe damage and loss to the ecosystem and economy.

The third area (denoted by SC) is located in the Mañihuales watershed, Aysén Region, Chilean Patagonia. In the western part of Aysén Region, it is mountainous and densely forested, while in the east, the terrain is flat, and the vegetation type is mainly grassland. Primary forests cover 47% of the Mañihuales watershed, and 67% of these forests are nothofagus pumilio deciduous broadleaf forest [39,49]. The center coordinate of SC is (45°9′S, 71°50′W). A large-scale forest defoliation event caused by an *Ormiscodes amphimone* outbreak occurred in this area during the 2015 to 2016 growing seasons [39]. This kind of insect is considered detrimental to tree growth and can cause crown dieback if defoliation is severe [50].

2.2. Landsat Surface Reflectance Product

Landsat archive can simultaneously provide global records with relatively high spatial resolution (30 m) and sufficient time series of more than 40 years compared to MODIS and Sentinel-2 [51], making it suitable for mining temporal information and detecting prevalent finer change patches. Thus, it is often regarded as the best free data source for vegetation monitoring [52,53]. The Landsat surface reflectance product has six common spectral bands including Blue, Green, Red, Near-Infrared (NIR) and two Shortwave Infrared bands (SWIR1, SWIR2). Meanwhile, it also provides the reflectance quality assessment band (QA band) which contains information on cloud and shadow. QA band can be applied to roughly screen and mask cloud, shadow and snow pixels. The current version of the Landsat surface reflectance product is Collection 2. The Collection 2 surface reflectance product was released in early 2021. Compared to Collection 1 surface reflectance product, the Collection 2 surface reflectance product has substantial enhancements in terms of geometric accuracy, radiometric calibration, and so on. In this study, the Collection 2 surface reflectance product generated from Landsat 4-5 TM, Landsat 7 ETM+, and Landsat 8 OLI was used.

3. Methodology

UFVAD is designed to be a universal framework to detect different vegetation anomalies robustly from Landsat observations. Figure 2 shows the flowchart of UFVAD, which consists of two major processes: creating the universal measure of undisturbed vegetation and detecting vegetation anomaly in NRT. Landsat observations over the BELMANIP sites were selected as a reference dataset to calculate vegetation characteristic parameters. After being composited into monthly intervals and eliminated the cloud-contaminated pixels, the vegetation characteristic parameters were further normalized to remove the differences in growth among various vegetation types. Then, the KDE theory was used to create the measure of undisturbed vegetation based on the normalized vegetation characteristic parameters. When a newly collected Landsat image is available, vegetation anomaly is identified in NRT according to this universal measure.

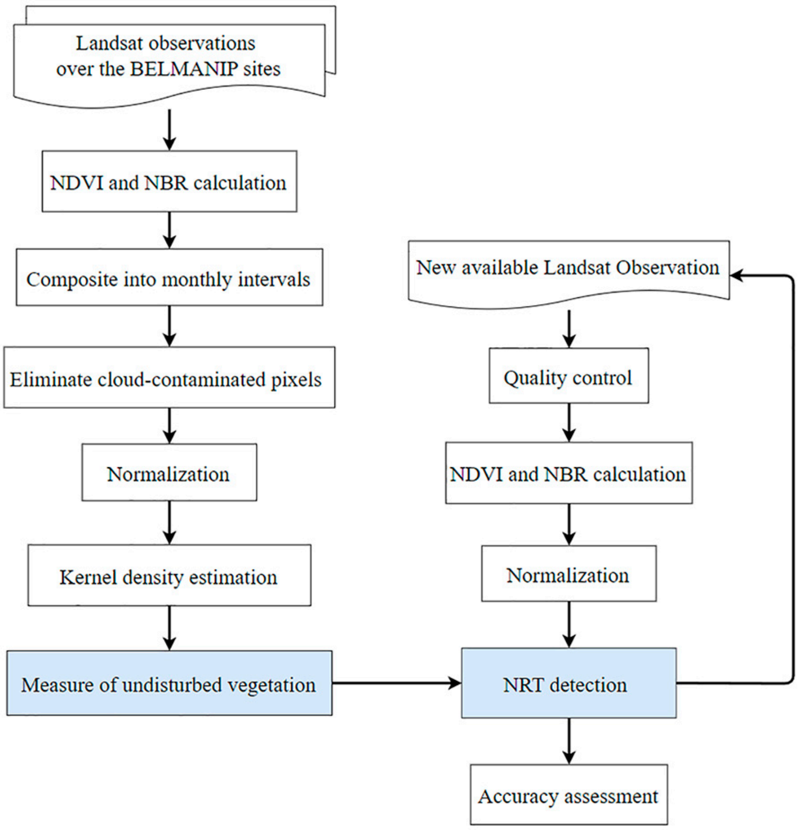

Figure 2. The workflow of the UFVAD framework.

3.1. Creating the Universal Measure of Undisturbed Vegetation

3.1.1. Selection of Reference Data and Vegetation Characteristic Parameters

Vegetation's phenology is diverse in regions with different climates and terrain conditions. In order to create the universal measure of undisturbed vegetation, the reference data should be globally representative of surface types and conditions. The Benchmark Land Multisite Analysis and intercomparison of Products (BELMANIP) network contains 402 sites. These sites provide a good sampling of biome types and conditions throughout the world [54] and are widely used for global LAI product intercomparison [55]. In this study, Landsat surface reflectance data from 1985 to 2022 over the BELMANIP sites were selected as reference data.

There is a wide variety of characteristic parameters derived from satellite observations to characterize vegetation conditions. In this study, NDVI and the normalized burn ratio (NBR) are selected to describe vegetation conditions from greenness and moisture perspectives. NDVI is a greenness index. It is highly correlated with parameters such as green biomass and absorbed radiation by photosynthetically active plant canopies [56] and is widely used in vegetation monitoring. NDVI is calculated using Formula (1):

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{Red}}} \tag{1}$$

where ρ_{NIR} and ρ_{Red} are the surface reflectance of NIR and Red bands.

NBR is a moisture index. It is primarily sensitive to moisture content of leaves and soils, as well as char and ash [57]. In addition to being commonly used in burn severity estimation, NBR can also be applied in other anomaly cases [22]. NBR is calculated using Formula (2):

$$\mathrm{NBR} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{SWIR2}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{SWIR2}}} \tag{2}$$

where ρ_{NIR} and ρ_{SWIR2} are the surface reflectance of NIR and SWIR2 bands.

In this study, the NDVI and NBR values were calculated from the reference data to create the universal measure of undisturbed vegetation.

3.1.2. Elimination of Cloud-Contaminated Pixels

Generally, cloud contamination is considered noise for vegetation anomaly detection. To minimize cloud and atmospheric contamination in the NDVI and NBR data over the BELMANIP sites, the NDVI and NBR data were composited into monthly intervals. In this study, the maximum value composite (MVC) method was used to composite the NDVI and NBR data. According to QA band of the Landsat surface reflectance data, an initial screening process was applied to determine whether the NDVI and NBR values were valid. The NDVI and NBR values were assumed to be invalid if the QA values indicated cloud or cloud shadow. In this study, only valid NDVI and NBR values were used for compositing. The NDVI and NBR data in a month were formed into monthly composite images. If the number of valid NDVI and NBR values over a monthly compositing period was two or more, the MVC method was used, which selects the NDVI and NBR values with the highest NDVI values in a month. If only one valid NDVI and NBR value was available over the composite period, these NDVI and NBR values were automatically selected to represent the period. If all NDVI and NBR values during a monthly period were invalid, the NDVI and NBR values with the highest NDVI were selected.

However, some composite NDVI and NBR values may still be affected by clouds which usually appear as negatively biased noises, for instance, abrupt drops in the time series of the NDVI [58]. Therefore, further processing is needed to remove the NDVI and NBR values contaminated by clouds.

Generally, the NDVI and NBR time series exhibit different seasonal changes, which make it difficult to identify the NDVI and NBR values contaminated by clouds accurately. In this study, linear functions were applied to fit the composite NDVI values from the same month of all years to detect the cloud-contaminated NDVI and NBR values. Let $(t_i, ndvi_i),\ i = 1, \cdots, m$ be a series of NDVI values for a pixel, where t_i is time, and $ndvi_i$ is the NDVI values. Let w_i be a weight corresponding to the NDVI value $ndvi_i$. The weight can be seen as a reliability measure of the NDVI value when estimate the linear model. A linear function, $f(t) = at + b$, was fitted to all n NDVI values using the least squares method, i.e., minimizing Formula (3):

$$J(a,b) = \sum_{i=1}^{m} w_i (ndvi_i - f(t_i))^2 \tag{3}$$

Then, the coefficients, a and b, of the linear function can be calculated from Formulas (4) and (5):

$$a = \frac{\sum w_i ndvi_i \sum w_i t_i{}^2 - \sum w_i ndvi_i t_i \sum w_i t_i}{\sum w_i \sum w_i t_i{}^2 - (\sum w_i t_i)^2} \tag{4}$$

$$b = \frac{\sum w_i ndvi_i \sum w_i t_i - \sum w_i ndvi_i t_i \sum w_i}{(\sum w_i t_i)^2 - \sum w_i t_i{}^2 \sum w_i} \tag{5}$$

Figure 3 is an example of detecting cloud-contaminated NDVI and NBR values by fitting time series NDVI values with a linear function. It consists of two iterations. The first iteration is to find cloud-contaminated candidates. The initial weights are set equally for all NDVI values (all equal to one) to fit the time series NDVI values. The NDVI values below the linear function of the first fit are considered as the cloud-contaminated candidates, and their weights are updated to 0.1. The second iteration is to determine the cloud-contaminated NDVI values from these candidates. According to the updated weights, the linear function is applied to fit the NDVI values again. This two-iteration procedure leads to a straight line adapted to the upper envelope of the time series NDVI values. The fitted linear function is used to calculate NDVI values of the upper envelope, denoted by $ndvi\_env_i, i = 1, \cdots, m$. We assume that the cloud-contaminated NDVI values should satisfy the following condition:

$$ndvi\_env_i - ndvi_i > ndvi_{envi} * \delta \quad (6)$$

where δ is a threshold and is set to 0.25 in this study after several experiments, which can better determine the cloud-contaminated NDVI values. Then, referring to the cloud-contaminated NDVI values, the cloud-contaminated NBR values were determined. In our study, the cloud-contaminated NDVI and NBR values were removed, and only the NDVI and NBR values with high quality were reserved to further remove seasonal and inter-annual variations in growth among different vegetation types.

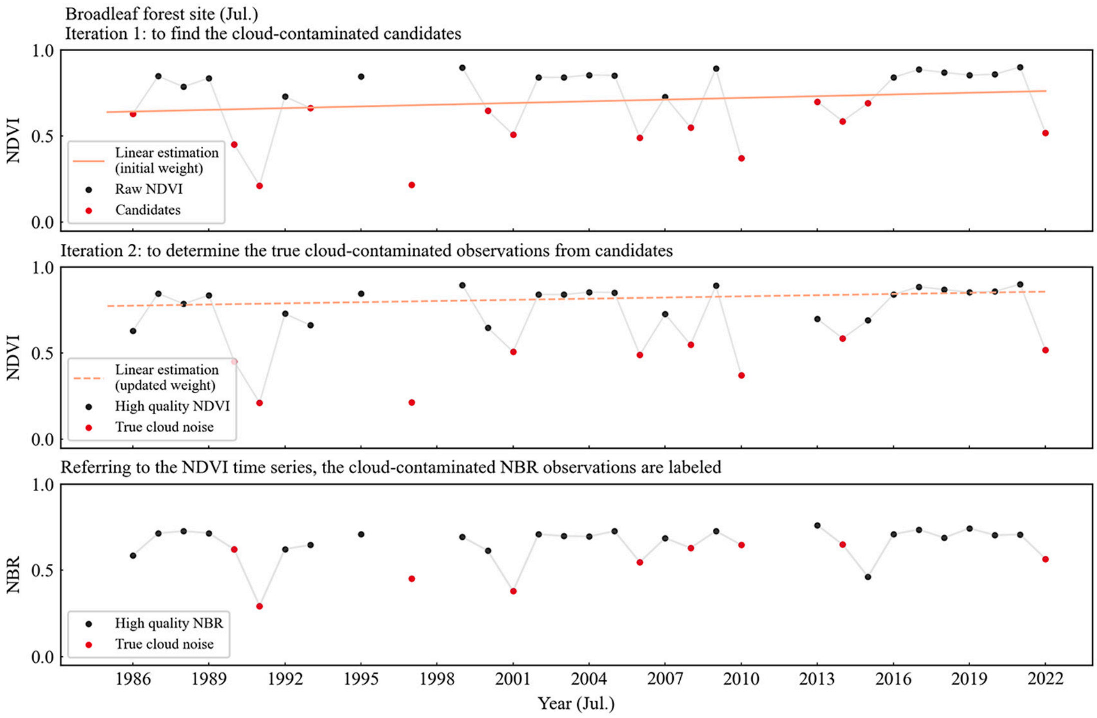

Figure 3. An example of detecting cloud-contaminated NDVI and NBR values in July from 1986 to 2022 at a broadleaf forest site. The highest possible cloud-contaminated NDVI and NBR values appear as abrupt drops in the time series, which can be determined after performing weighted linear fitting in two iterations.

3.1.3. Normalization of NDVI and NBR Values

In order to reduce the impact of seasonal changes on the detection of vegetation anomalies and to ensure the universality and robustness of the anomaly detection algorithm, the high-quality NDVI and NBR values are normalized to remove seasonal and inter-annual variations in growth among different vegetation types.

The normalization of NDVI values were used as an example. For each site, the time series of high-quality NDVI values from the same month of all years is fitted by a linear function, and the NDVI values calculated by the linear function at time t_i is denoted by $ndvi\_est_i$. Then, the high-quality NDVI values $ndvi_i$ at time t_i is normalized by subtracting $ndvi\_est_i$ from $ndvi_i$, resulting in what we refer to as the normalized NDVI values, $\Delta ndvi_i$.

$$\Delta ndvi_i = ndvi_i - ndvi\_est_i \tag{7}$$

Similarly, the high-quality NBR value nbr_i at time t_i was normalized to obtain the normalized NBR value, Δnbr_i, by utilizing the same method.

The normalized NDVI and NBR values range from −1 to +1 regardless of the vegetation type. Figure 4 shows the time series of the NDVI and NBR values with high quality and their normalized NDVI and NBR values at the Walker Branch, South Hill, and Turco sites. The biome types of these sites are broadleaf forest, needleleaf forest and shrub, respectively. The time series of the NDVI and NBR values at the Walker Branch and South Hill sites show obvious seasonal variations. Compared to the Walker Branch and South Hill sites, the phenological cycles at the Turco site are relatively less obvious. Although the NDVI and NBR values at the three sites show different amplitudes and phenological cycles, the normalized NDVI and NBR values are all within the same range.

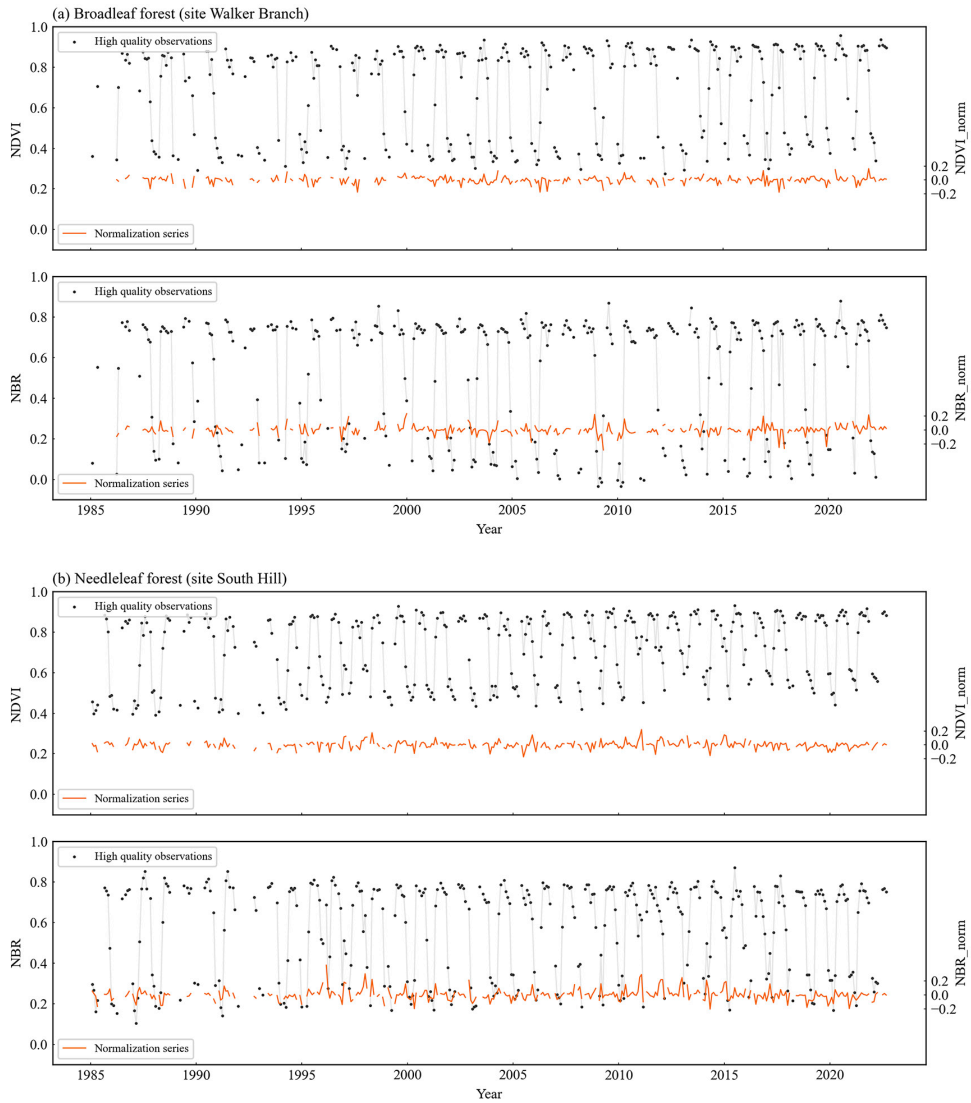

Figure 4. *Cont.*

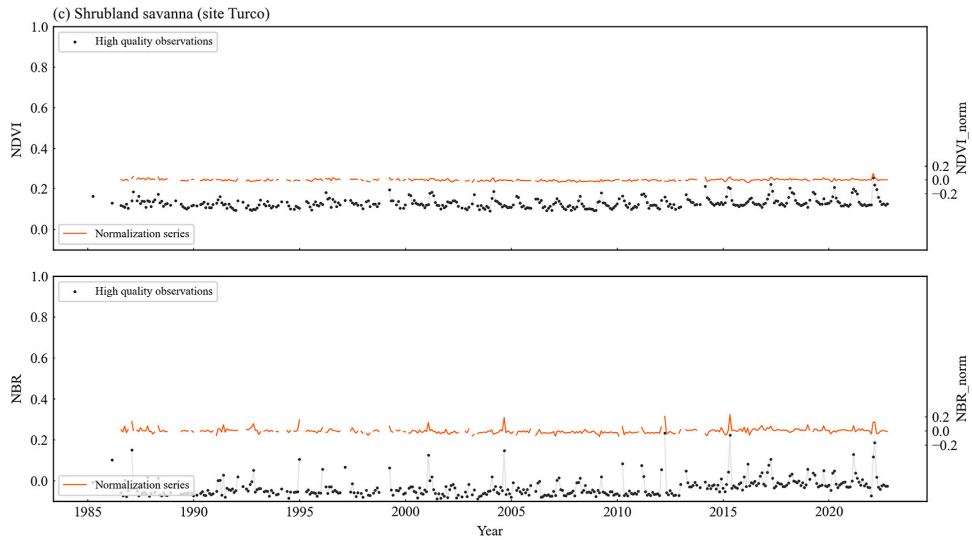

Figure 4. Time series of the NDVI and NBR values with high quality and their normalized NDVI and NBR values at the (**a**) Walker Branch, (**b**) South Hill, and (**c**) Turco sites. The biome types of these sites are broadleaf forest, needleleaf forest, and shrub, respectively.

The above methods were used to normalize the high-quality NDVI and NBR values of each month over the BELMANIP sites, and all the normalized NDVI and NBR values were used in Section 3.1.4 to create the universal measure of undisturbed vegetation.

3.1.4. Probability Calculation Based on Kernel Density Estimation

The measure of undisturbed vegetation in this study is essentially the distribution of the normalized NDVI and NBR values of each month over the BELMANIP sites, which can describe the undisturbed vegetation condition. The distribution of the normalized NDVI and NBR values is completely characterized by its probability density function. Therefore, an estimation of the probability density function yields estimates for different distribution characteristics. In this study, the KDE is used to estimate the probability density function.

Let $x \in R^N$ be a vector containing N parameters. For n samples $X_1, X_2, \ldots, X_n$ in R^N, the estimation of the probability density function $f(x)$ can be defined as follows:

$$\hat{f}(x) = \frac{1}{nh^N} \sum_{i=1}^{n} K(\frac{||x - X_i||}{h}) \tag{8}$$

where h is the bandwidth determined by Scott's rule [59], K represents the kernel function, and $||x - X_i||$ represents the distance between x and X_i in the R^N space.

In this study, vector x includes the parameters $\Delta ndvi$ and Δnbr, so the size of N is equal to 2. The KDE was performed on a 500 × 500 grid with Gaussian kernel function to estimate the probability density function of the normalized NDVI and NBR values of each month over the BELMANIP sites.

According to [60], the continuous probability density function $f(x)$ satisfies the following equations at a given probability $1 - \alpha$:

$$\mathcal{L}(f(x); c_\alpha) = \left\{ x \in R^N : f(x) \geq c_\alpha \right\} \tag{9}$$

$$\int_{\mathcal{L}(f(x);c_\alpha)} f(x)dx = P[x \in \mathcal{L}(f(x); c_\alpha)] = P[f(x) \geq c_\alpha] \geq 1 - \alpha \quad \alpha \in (0, 1) \tag{10}$$

where c_α can be regarded as a density level which indicates the region $\mathcal{L}(f; c_\alpha)$ intending to contain at least $1 - \alpha$ accumulated probability. If we know $f(x)$, c_α is precisely the lower

α-quantile of $f(x)$. Therefore, the region $\mathcal{L}(f;c_\alpha)$ corresponding to $1-\alpha$ accumulated probability can also be obtained by $f(x)$. Since this is not the case, $f(x)$ was replaced with the estimation of $\hat{f}(x)$, and the estimated $\hat{c}_\alpha$ can be approximated efficiently as the lower α quantile of sample $\hat{f}(X_1), \hat{f}(X_2), \ldots, \hat{f}(X_n)$ [60]. For example, if the accumulated probability is 95%, then α is 5% and $\hat{c}_\alpha$ is the 5th percentile of $\hat{f}(X_1), \hat{f}(X_2), \ldots, \hat{f}(X_n)$, so the region $\mathcal{L}(\hat{f};\hat{c}_\alpha)$ corresponding to 95% probability can also be obtained. Different $\hat{c}_\alpha$ values yield a range of accumulated probabilities; thus, finally, the measure of undisturbed vegetation can be expressed by accumulative probabilities to show the distribution of undisturbed vegetation.

Regions within different probabilities describe different degrees of undisturbed vegetation state. In this study, we calculated several estimated density levels $\hat{c}_\alpha$ under the accumulated probabilities of 95%, 90%, 75%, and 50% for each month, which are shown as irregular circles in Figure 5. For each month, the distribution of undisturbed vegetation in the space of normalized NDVI and NBR values appears as an irregular inclined ellipse. From May to September, the distributions of normalized NDVI and NBR values are more compact in the short-axis direction, while, for the remaining months the distributions, they are relatively scattered, with larger areas and more irregular boundaries in regions corresponding to 95%, 90%, and 75% probabilities. It is necessary to make a monthly reference to the measure of undisturbed vegetation when detecting anomalies in near real time.

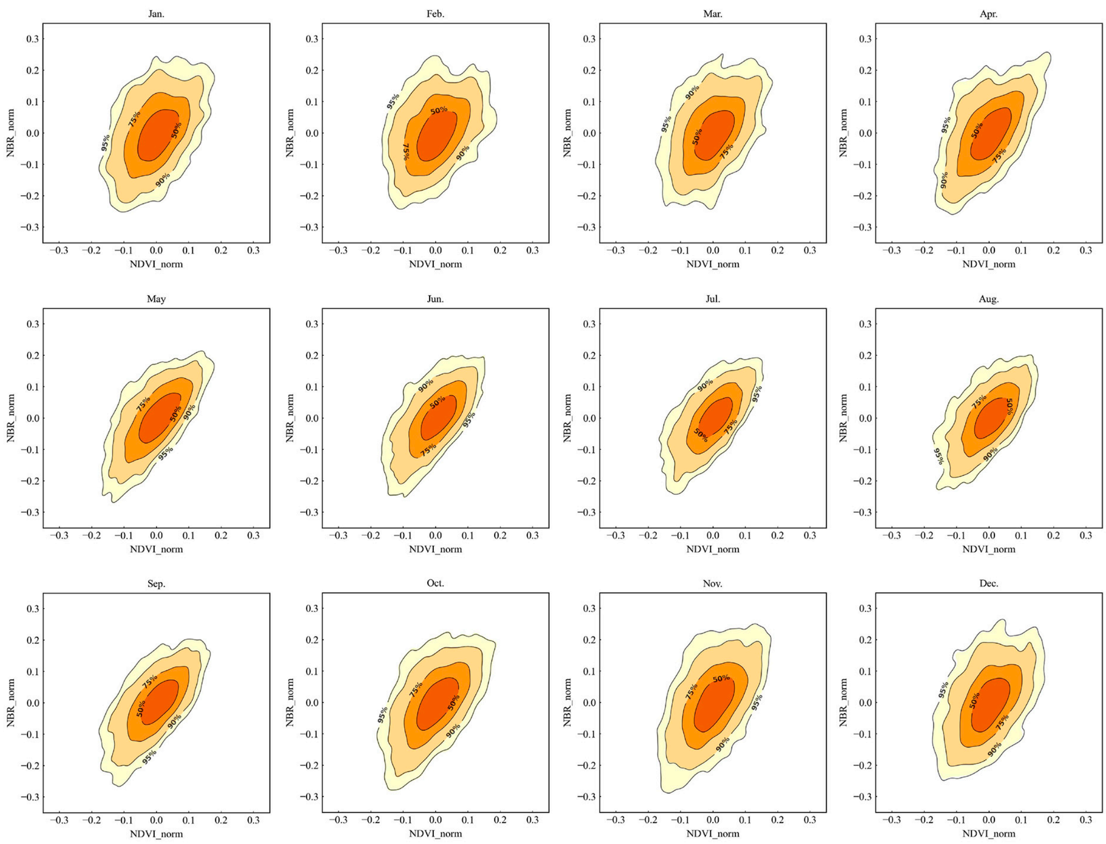

Figure 5. The measure of undisturbed vegetation for each month based on the normalized NDVI and NBR values. Irregular circles indicate the $\hat{c}_\alpha$ under the accumulated probability of 95%, 90%, 75%, and 50%, respectively.

3.2. NRT Vegetation Anomaly Detection

NRT anomaly detection is executed at the pixel scale. When a new Landsat image is available, the surface reflectance data contaminated by the clouds were firstly removed according to the QA band of the Landsat image. Then, the NDVI and NBR values are calculated from the cloud-free surface reflectance data. Since the universal measure of undisturbed vegetation is created by the normalized NDVI and NBR values, the NDVI and NBR values from the new Landsat image should also be normalized to remove the differences in vegetation growth. The method described in Section 3.1.3 is used to obtain a new vector X containing the normalized NDVI and NBR values.

For each pixel, we input this vector X into the corresponding KDE model for the same month, and each pixel shares the same anomaly criterion. Whether the vegetation is anomalous is determined by Formula (11):

$$\hat{f}(X) < \hat{c}_{\alpha} \tag{11}$$

In this study, $\hat{c}_{\alpha}$ was set as the estimated density level corresponding to 95% probability threshold. We assumed that regions with density values under $\hat{c}_{\alpha}$ belongs to the lowest representation of undisturbed vegetation. Therefore, the new normalized observation in this region is flagged as anomalous. Near-real-time detection is moving forward by continuously collecting the new available images and repeating the above processing steps.

Figure 6 shows an example of NRT detection when UFVAD was applied on an image on 26 May 2003 in SB. Burned pixels (black points within red circles) are identified as anomalous under the $\hat{c}_{\alpha}$ density level corresponding to the 95% threshold.

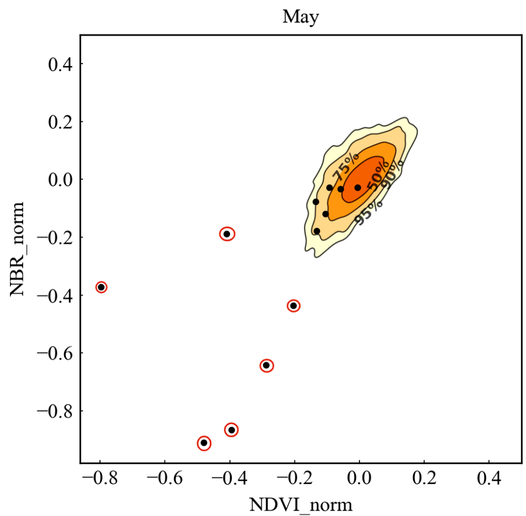

Figure 6. An example of vegetation anomaly detection using UFVAD. The normalized NDVI and NBR values (named NDVI_norm, NBR_norm) come from several pixels of the Landsat image on 26 May 2003 in SB. The red circles show the normalized NDVI and NBR values outside the accumulated probability of 95%. The vegetation at these pixels is flagged as anomalies if we set $\hat{c}_{\alpha}$ as density value corresponding to 95% probability threshold.

3.3. Accuracy Assessment

3.3.1. Sampling Design

The anomaly detection accuracy of UFVAD was assessed over the three study areas. The assessment method is based on samples selected through stratified random sampling. A 5-pixel buffer around each anomaly area was applied to better describe the omission errors occurring in spatial proximity [61]. Thus, three strata were created: anomaly, non-anomaly within the buffer (NAWB), and non-anomaly outside the buffer (NAOB) (Table 1).

Table 1. The mapped area and the number of samples of each stratum.

Study Site	Image	Area [Pixels]			Sample Units			
		NAOB	NAWB	Anomaly	NAOB	NAWB	Anomaly	Total
SA	2018/01/29	23,182	287	460	387	149	238	774
	2018/02/22	24,451	654	770	454	209	246	909
	2018/03/26	24,142	788	945	455	207	248	910
	2018/12/31	24,888	622	365	453	285	167	905
SB	2003/05/26	410,829	262,393	748,247	266	170	436	872
	2003/06/11	445,815	308,065	776,289	298	206	505	1009
	2003/08/14	698,671	193,683	542,982	337	93	430	860
SC	2016/01/08	1,367,892	79,658	22,861	431	311	119	861
	2016/02/02	1,152,224	205,214	157,471	466	268	197	931
	2016/03/05	1,152,793	182,123	176,868	462	245	216	923

According to [62,63], the optimal sample size s was estimated as follows:

$$s \approx \left(\frac{\sum W_i \sqrt{U_i(1-U_i)}}{S(\hat{O})} \right)^2 \tag{12}$$

where W_i is the mapped proportion of area of stratum i, $S(\hat{O})$ is the standard error of the estimated overall accuracy that we would like to achieve, and U_i is conjectural user's accuracy of stratum i. In this study, $S(\hat{O})$ was set to 0.01, the conjectural user accuracies will be 0.9 for NAOB, 0.88 for NAWB, and 0.88 for anomaly.

Sample units were distributed equally between the largest stratum and the remaining two strata. Then, for the remaining two strata, the distribution of samples was proportional to the area of each stratum. This may avoid under sample in small areas and, at the same time, considered the difference among strata [63]. In Table 1, the number of sample units of each strata can be found.

3.3.2. Sample Interpretation and Accuracy Calculation

The manual interpretation of the anomaly and non-anomaly sample units was usually performed using data sources with higher spatial quality to obtain the ground truth. For deforestation cases in SA, we interpreted each validation sample using Sentinel-2A images near the detection dates. However, for wildfires in SB and insect outbreak in SC, we depended on the Landsat imagery itself due to the lack of available Sentinel data with high spatial resolution.

Using the sample-count-based confusion matrix, we calculated accuracy evaluation indexes: the user's accuracy (UA), the producer's accuracy (PA), the overall accuracy (OA), and the Kappa coefficient for each detection map [64].

4. Results

UFVAD was used to detect vegetation anomalies in NRT manner over the three study areas to evaluate the performance of UFVAD.

4.1. SA

In SA, the anomaly maps detected on four dates in 2018 (Figure 7B) all show a high degree of consistency with the deforestation cases (visually purple patches on false-color maps) shown in the corresponding Landsat images (Figure 7A). The deforestation here is almost clear-cutting. The forest cleared patches are characterized as relatively regular and small with a similar spectrally visual interpretation to bare land, making them distinctly different from the surrounding stable forests. There is a continuous expansion of deforestation among the four detection dates. The new deforested areas were generally adjacent to the old ones. From 29 January to 22 February, three clear-cutting patches first appeared in the south, and meanwhile a smaller patch in the upper middle of the study area was precisely captured by UFVAD, which expanded on 26 March (Figure 7(B2,B3). By 31 December (Figure 7(A4,B4)), some of the patches identified as anomalous had re-expressed vegetation characteristics (light green on the false-color map). Secondary succession occurred here, and shallow herbaceous began to grow. They are no longer labeled as vegetation anomalies of concern by UFVAD.

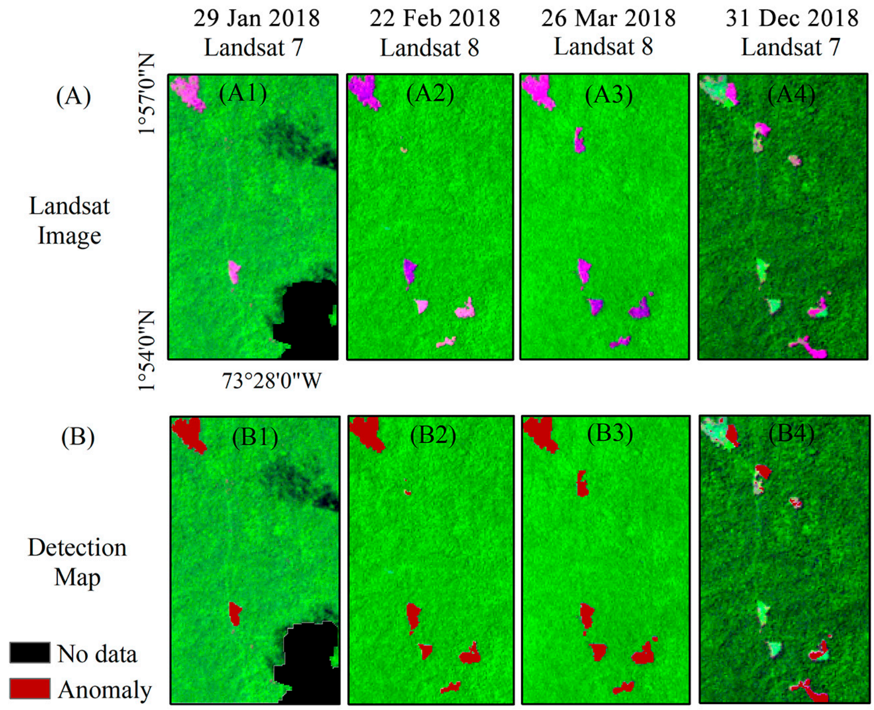

Figure 7. The anomaly detection maps for deforestation events in SA using UFVAD from high quality Landsat images. The first row ((**A**) A1–A4) is false-color composite (R: SWIR1; G: NIR; B: Red). The second row ((**B**) B1–B4) is the corresponding vegetation anomaly detected by UFVAD using the false-color composite as base maps. The legend 'No data' represents masked atmospheric noise pixels or masked non-vegetation pixels.

The accuracy of the deforestation anomaly detected by UFVAD in SA is given in Table 2. The OA of clear-cutting identification for the four detection dates was greater than 95%, with Kappa coefficients greater than 0.94. The UA and PA were close to 100% for the NAOB stratum and in the 89% to 96% range for the NAWB and anomaly strata. According to the confusion matrix, the identification errors were mainly concentrated between the NAWB and the anomaly strata.

Table 2. The confusion matrix and accuracy of the deforestation anomaly detected by UFVAD in SA.

Strata	NAOB	NAWB	Anomaly	OA (kappa)	UA	PA
SA 29 Jan 2018				97.3 (0.96)		
NAOB	384	0	0		99.2	100.0
NAWB	0	141	10		94.6	93.4
Anomaly	3	8	228		95.8	95.4
SA 22 Feb 2018				96.9 (0.95)		
NAOB	454	0	0		100.0	100.0
NAWB	0	192	11		91.9	94.6
Anomaly	0	17	235		95.5	93.3

Table 2. *Cont.*

Strata	NAOB	NAWB	Anomaly	OA (kappa)	UA	PA
SA 26 Mar 2018						
NAOB	455	0	0	95.9 (0.93)	100.0	100.0
NAWB	0	192	22		92.8	91.1
Anomaly	0	15	226		89.7	93.8
SA 31 Dec 2018				96.0 (0.94)		
NAOB	453	0	0		100.0	100.0
NAWB	0	266	17		93.3	94.0
Anomaly	0	19	150		89.8	88.8

4.2. SB

Figure 8 shows the anomaly maps for a fire event detected by UFVAD from Landsat images in SB. The Landsat image of 26 May 2003 is the first high-quality data acquired after the wildfire. In contrast to the clear-cutting in SA, the severely burned area appears dark red on the false-color image (Figure 8(A1)). The forest tree slash is evident in the NIR and SWIR bands, which can be captured by the NBR index. The fire event is distributed in the oblique lower half of SB, occupying more than 50% of the whole area. The anomaly map detected by UFVAD is generally consistent with the burned area (Figure 8(B1)). Surrounding the forest tree slash, some canopies were not completely disturbed (shown as dark gray-green on Figure 8(A1)), where moisture stress may be more evident relative to the greenness loss. This phenomenon was also detected as an anomaly by UFVAD. According to Figure 8(A2,A3,B2,B3), there was no further expansion or new occurrence of the fire event from 11 June to 14 August, but the canopies that were not completely burnt regrew to an undisturbed condition during the growing season, so they were no longer labeled as anomalous. However, there are also some detection noises in areas that are actually non-anomalous.

The confusion matrix and accuracy of the fire anomaly detection in SB are shown in Table 3. The OA and Kappa coefficients for the three detection dates were 89.1% (0.82), 96.9% (0.95), and 94.7% (0.9), respectively. For the detection anomaly maps on 26 May and 11 June, the UA for the anomaly stratum (87.4% on 26 May and 87.7% on 11 June) and the PA for the NAOB stratum (84.5% on 26 May and 84.6% on 11 June) are relatively lower. This indicates an overestimation of the anomaly and a lower estimation of NAOB. There are some omission errors on 26 May and 14 August. From the confusion matrix, samples identified as NAWB were actually anomalous, resulting in lower UA for NAWB. In general, the detection anomaly maps by UFVAD in SB are reliable.

Table 3. The confusion matrix and accuracy of the fire anomaly detected by UFVAD in the SB.

Strata	NAOB	NAWB	Anomaly	OA (kappa)	UA	PA
SB 26 May 2003				89.1(0.82)		
NAOB	251	0	46		94.4	84.5
NAWB	0	145	9		85.3	94.2
Anomaly	15	25	381		87.4	90.5
SB 11 Jun 2003				96.9 (0.95)		
NAOB	291	0	53		97.7	84.6
NAWB	0	191	9		93.2	95.5
Anomaly	7	14	443		87.7	95.5
SB 14 Aug 2003						
NAOB	328	0	14	94.7 (0.9)	97.3	95.9
NAWB	0	69	4		74.2	94.5
Anomaly	9	24	412		95.8	92.6

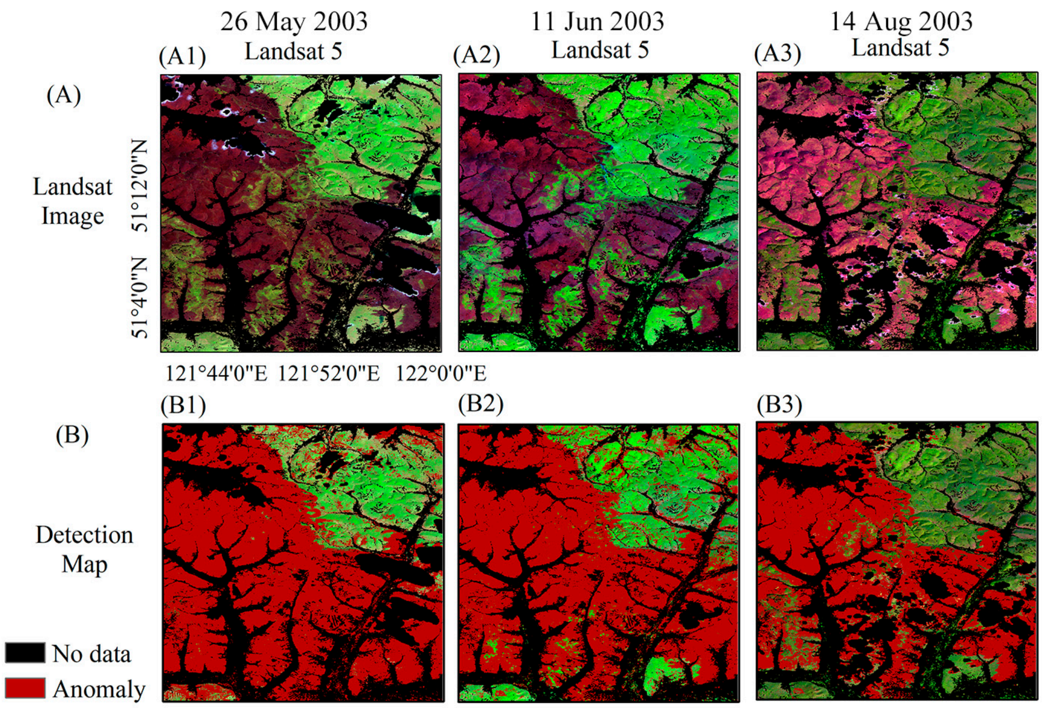

Figure 8. The anomaly detection maps by UFVAD from Landsat images for a fire event in SB. The first row ((**A**) A1–A3) is false-color composite (R: SWIR1; G: NIR; B: Red). The second row ((**B**) B1–B3) is the corresponding anomaly maps detected by UFVAD using the false-color composite as base maps. Non-vegetation pixels are masked according to the classification of a high-quality image during the historical period. The legend 'No data' represents masked atmospheric noise pixels or masked non-vegetation pixels.

4.3. SC

Figure 9 shows the anomaly maps detected by UFVAD from Landsat images for *O. amphimone* insect outbreak in SC in 2016. Three detection maps in Figure 9(B1–B3) show the process from early insect erosion to outbreak and diffusion. UFVAD can detect the vegetation anomaly caused by the insect accurately, as visually compared with the corresponding Landsat images Figure 9(A1–A3). On 8 January, vegetation anomaly caused by the insect was only recognized in two major local areas. But one month later, an outbreak led to severe defoliation and vegetation anomaly were finally detected in most of the southwest, northwest, and northeast areas (shown in Figure 9(B1,B2)). In the detection anomaly map on 5 March near the end of summer, disturbances were further spread on the old basis in the northeast region (Figure 9(B3)). In addition to insect outbreak anomaly, there are also other small-scale grassland degradation cases in the central and southeast of SC.

Table 4 demonstrates that the OA is greater than 90% for all three detection dates, with Kappa coefficients of 0.87, 0.93, and 0.93, which indicates the overall good performance in insect erosion detection by UFVAD. On 8 January, the UA and PA of the anomaly stratum are relatively lower, especially for the PA, which is only 66.9%. Identification errors are due to the underestimation of anomalies and the overestimation of non-anomalies. According to the confusion matrix, a large proportion of the anomaly samples are identified as NAOB and NAWB. In general, the PA for the two non-anomaly strata yields good results (greater than 95%), but the UA for NAWB is relatively lower than the PA. This indicates omission errors near the anomaly boundary.

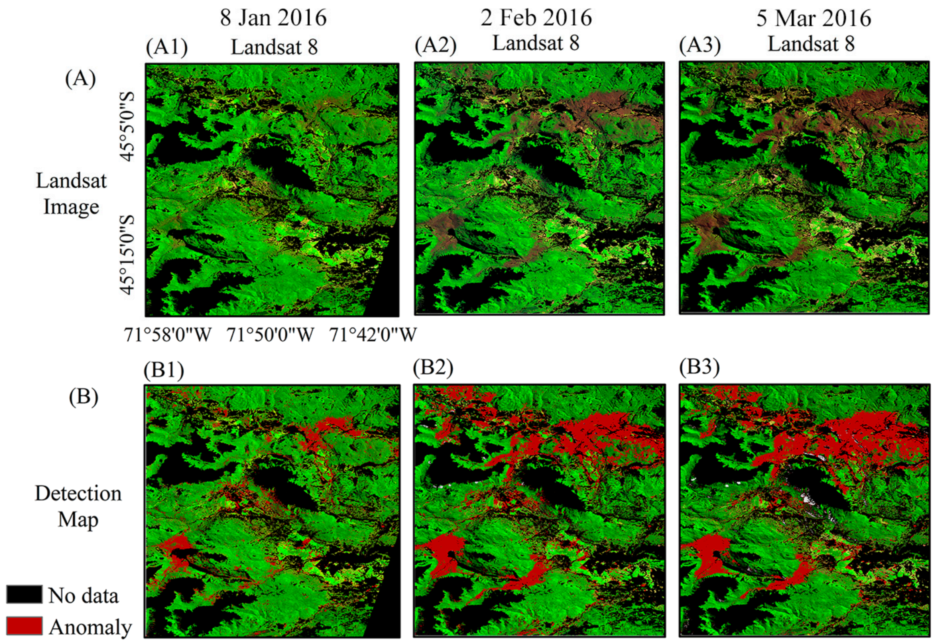

Figure 9. The anomaly maps detected by UFVAD from Landsat images for *O. amphimone* insect outbreak in SC. The first row ((**A**) A1–A3) is false-color composite (R: SWIR1; G: NIR; B: Red). The second row ((**B**) B1–B3) is the corresponding anomaly maps detected by UFVAD using the false-color composite as base maps. Non-vegetation pixels are masked according to the classification of a high-quality image during the historical period. The legend 'No data' represents masked atmospheric noise pixels or masked non-vegetation pixels.

Table 4. The confusion matrix and accuracy of the insect anomaly detected by UFVAD in the SC.

Strata	NAOB	NAWB	Anomaly	OA (kappa)	UA	PA
SC 8 Jan 2016				91.9 (0.87)		
NAOB	411	0	19		95.4	95.6
NAWB	0	283	3		91.0	99.0
Anomaly	20	28	97		81.5	66.9
SC 2 Feb 2016				95.6 (0.93)		
NAOB	456	0	5		97.9	98.9
NAWB	0	243	1		90.7	99.6
Anomaly	10	25	191		97.0	84.5
SC 5 Mar 2016						
NAOB	448	0	2	95.9 (0.93)	97.0	99.6
NAWB	0	223	0		91.0	100.0
Anomaly	14	22	214		99.1	85.6

5. Discussion

5.1. The Robustness of UFVAD in Detecting Different Vegetation Anomalies

The anomaly maps detected by UFVAD over the three study areas showed good performance with some variation among them, which verified the potential of UFVAD to detect various vegetation anomalies with robustness, at least in forest cases.

In SA, the accuracy is relatively the highest compared to SB and SC. Pixels outside the buffer are almost completely identified correctly (UA and PA for the NAOB strata are almost 100%). In addition to the distinctive difference between clear-cutting and stable forest, its homogeneous land cover and dense forest distribution also help improve its accuracy because of the less mixed pixels. A few errors between NAWB and anomaly strata mainly come from the boundary pixels within the buffer. Referring to the Sentinel-2A images with higher spatial resolution, the edge of clear-cutting patch on Landsat images is clearer on the Sentinel-2A images, which leads to the ambiguity in interpretation [65].

In SB, the Landsat images on detection dates contain a small amount of clouds and shadows. Atmospheric noises that are not sufficiently masked on the newly collected image will still be involved in the NRT detection, being identified as anomalies even though they are spurious. This is because they also show a negative bias in the vegetation index, which results in their much lower negative normalized values than before [58]. Consequently, they are likely to fall outside the 95% probability range. On the other hand, the land cover in SB is more complex, with roads and the transition areas where scattered trees, shrubs, grass lands and bare lands mixed [66,67]. Figure 10(A1,A2,B1,B2) show the major areas of misdetection on 11 June 2003 in SB and the high-resolution images of two local regions. In the vegetation-sparse transition areas, the inadequately masked non-vegetation pixels, and the mixed pixels with bare land change irregularly and frequently in time series under moisture variation [13,68]. As a result, the anomaly detection uncertainty is increased due to those pixels' low-reliability normalized values. Due to the factors above, the anomaly was overestimated (lower PA), and most of them are spurious.

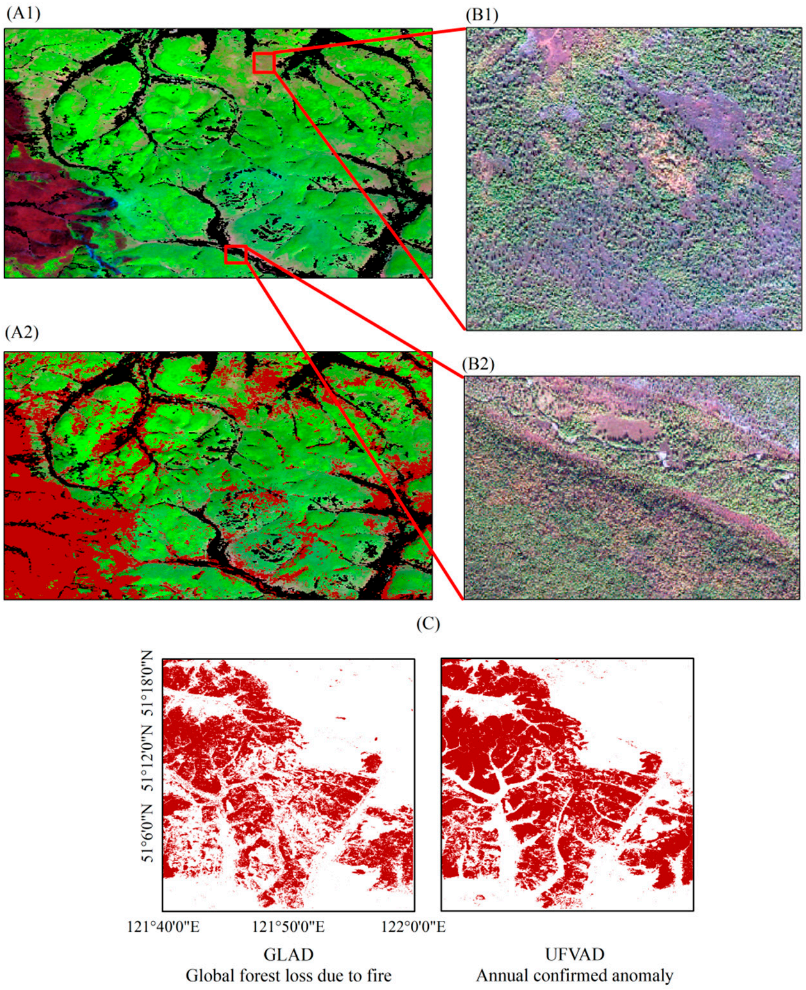

Figure 10. The major areas of detection noises on 11 June 2003 in SB (**A1**,**A2**) and (**B1**,**B2**), and the comparison of confirmed anomaly maps in 2003 with GLAD forest loss due to fire dataset (**C**). Detection noise commonly occurred beside the masked bare land where sparse vegetation distributed (shown in high resolution images (**B1**,**B2**), resulting in more mixed pixels varying easily due to the effect of soil and background).

However, there are few methods to exclude these pseudo-anomalies just on the current image. Typically, the detection results of three to six consecutive observations are used, as pixels that appear anomalous in multiple continuous images are more likely to be genuine anomalies [26,40,65,69,70]. Since fire detection dates are exactly within the growing season and the rest of 2003 is during the snowing season, we made an annual composite following the rule that pixels with all three detection dates labeled as anomaly are considered as truly anomalous. Meanwhile, the composite result was compared to an annual dataset, global forest loss due to fire [71], to further confirm the robustness of the detection results in SB (Figure 10C). The dataset originated from the Global 30 m forest cover loss map [72], and the forest loss areas due to fire driver with median and high certainty in SB were extracted. The composite anomaly map indeed excludes most spurious anomalies and yields a good consistency of 85.07% with the dataset on 1500 random points. However, this approach may result in a confirmation delay, particularly in humid tropical regions where there may not be enough available data to confirm anomalies in a short time [13,29].

Of importance to the anomaly detection results is the $\hat{c}_{\alpha}$ value corresponding to a probability threshold, which was set to 95% in our study. In SC, although some forests were infested on 8 January, the erosion degree may not be enough to cause significant defoliation [73,74]. Their normalized values tend to be consistent with those before. As a result, these subtle erosion changes were easily identified as non-anomalies by UFVAD under the 95% threshold, which led to the lowest PA (66.7%) for the anomaly stratum. The relatively high omission errors for detecting subtle anomalies like early infection have been reported in most threshold-dependent algorithms, and differentiated thresholds for distinct anomalies are suggested [75]. Indeed, the probability threshold may lead to the over- or underestimation of an anomaly. A higher threshold (extremely 100%) takes into account the 'complete' anomalies, i.e., values that have never existed before [39], and may thus result in an underestimation. However, a smaller threshold may help to better detect those small disturbance events or earlier events that may be hidden in the natural fluctuations of vegetation. Considering the deforestation, fire, and insect outbreak cases, 95% is a good representation of the undisturbed vegetation condition and upper anomaly frontier in this study. Users are allowed to make appropriate adjustments to the 95% threshold according to their interests.

These factors above produce some errors in the anomaly detection results and affect the accuracy of UFVAD, but, in general, UFVAD is robust to detect different vegetation anomalies. In this study, we did not analyze the temporal accuracy but only focused on the universality and robustness of UFVAD and whether it has the ability to achieve detection as long as new images are available. This is because the temporal accuracy depends on the quality of the new image and the temporal resolution of the satellite sensor. In principle, anomaly detection could be achieved every eight days after 2013 using Landsat imagery with the harmonization of the ETM+ and OLI sensors [76]. However, this requires that the newly collected image is of high quality with a low cloud fraction, so that enough pixels can be included in UFVAD without being masked; otherwise, the time to detect the anomaly for each pixel will be prolonged. If we have high-quality data with a high observing frequency, the temporal accuracy of UFVAD will naturally be higher. To improve the observation frequency, multi-source data may be an option. For instance, harmonized Landsat and Sentinel-2 data can reduce the revisit time to three days and increase the possibility of obtaining clear observations [29]. Combining optical and SAR images can help to overcome the limitations of cloudy and rainy weather and recover the true surface situation [70,77].

5.2. Strength and Defects to UFVAD

The main advantage of UFVAD is its ability to create the universal measure suitable for different pixels, vegetation types or regions. Unlike existing algorithms, UFVAD no longer captures undisturbed condition and makes anomaly criteria, respectively, for each under-detected pixel. This speeds up the detecting process. Compared to other approaches utilizing the KDE theory, we have improved the selection of reference vegetation and the dimensionality of input parameters. Therefore, it is not necessary to consider the vegetation type and seek its representative samples for KDE when detecting different areas of interest.

The universality of the measure of undisturbed vegetation is manifested in the following aspects. First, the data source for creating the measure is reliable and with high quality. BELMANIP network contains typical biome types that are distributed globally, making it highly representative for different vegetation types. When attempting to obtain high-quality Landsat series, we assigned a lower weight to potential noise and a higher weight to potential high-quality data in linear fitting. This approach is more accurate than ordinary linear regression in filtering any remaining cloud noise and reducing errors in subsequent normalization step. Despite the possibility of some remaining normalization noise, the flexibility of the KDE ensures the accuracy of the universal measure of undisturbed vegetation. Based on the universal measure, we can obtain the unified anomaly criterion. Secondly, the universal measure is developed based on two-dimensional indices: NDVI and NBR. They can adequately characterize vegetation conditions in terms of greenness and moisture, which enables them to capture the multi-spectral expression of vegetation anomalies. A fire-affected pixel in SB is used as an example (Figure 11); the normalized NDVI in 2003 is around -0.1, while the normalized NBR is around -0.25. If only the single NDVI is considered, this pixel will not be identified as anomaly unless the probability threshold is lower. However, when combining the NBR and NDVI, the moisture anomaly can be captured even though the greenness anomaly is not evident. All this makes UFVAD a choice for NRT detection.

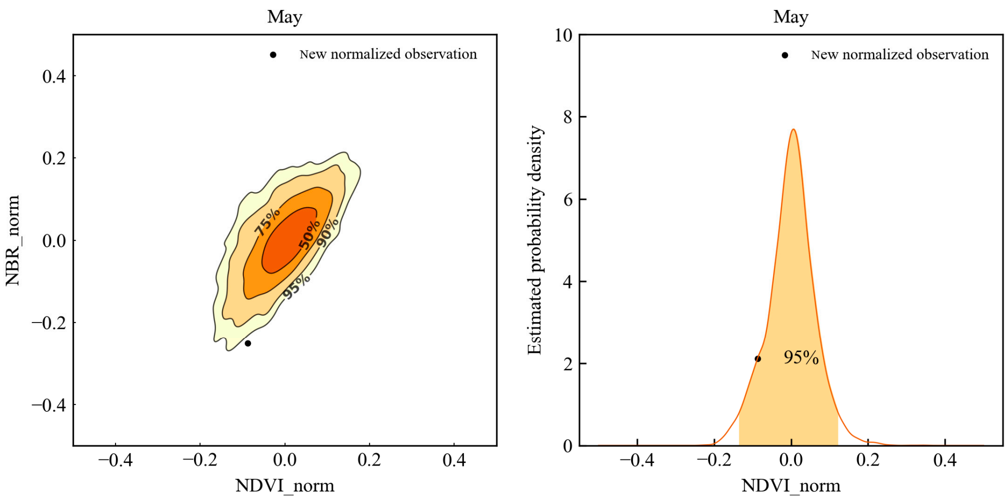

Figure 11. An example of the anomaly detection of a pixel in SB on 26 May 2003. The normalized NBR value clearly deviates from the undisturbed condition, while the normalized NDVI value is insignificant. The density value of this pixel is smaller than that with 95% probability threshold in the normalized NDVI and NBR space, while still larger than that with 95% threshold in single normalized NDVI KDE space.

The limitation is that UFVAD still requires a part of the time series as a historical baseline to calculate the normalization values. On the one hand, a large memory space is needed to host massive historical image stacks when detecting in larger areas. On the other hand, the quality and length of the time series may affect the normalization values. The normalization accuracy may decline with more noise and shorter time series [26]. In addition, UFVAD has not yet resolved the pseudo-anomalies in the current image, even though they may account for a relatively small proportion.

5.3. Improvements and Future Plans

In the future, more research is needed to assess the performance of UFVAD in other vegetation anomaly contexts such as hurricanes and grassland plagues or to extend UFVAD to water environment anomaly detection [78,79] and impervious cover anomaly detection [80]. Additionally, a comprehensive study on which index or band combinations work best is also significant [30]. The use of NDVI and NBR in UFVAD has shown good performance. One note is that other index or band combinations representing different vegetation conditions can also be used in UFVAD, for example, other greenness indices such as EVI, DVI, and SR; other moisture indices, such as NDMI and NBR2; and other indices that include both visible and shortwave infrared bands, such as DSWI. In the future, a comparison of the performances of different index or band combinations could significantly contribute to the literature.

There are three possible orientations to improve UFVAD: (1) adding an anomaly confirming module to remove pseudo-anomalies from the current detected image; (2) exploring multi-source data input, as harmonized optical and SAR data are worth trying to increase observation frequency and overcome cloudy condition; and (3) adding an anomaly type recognition module. In practical applications, while the anomaly location can be detected, the specific anomaly type needs to be further explored.

6. Conclusions

In this study, a universal framework (UFVAD) is developed to detect various vegetation anomalies from Landsat data in NRT. The main innovations of UFVAD are that (a) it creatively introduces the network of BELMANIP and normalized the seasonal and inter-annual differences in growth among vegetation of the same species or not using high-quality observations; (b) it emphasizes multi-dimensional input and utilizes the combination of NDVI and NBR to describe greenness and moisture conditions; (c) based on the two-dimensional normalization results, universal measures are created using KDE, which is the critical step and main purpose of UFVAD. Meanwhile, each pixel under detected shares the same anomaly criterion.

We demonstrate its good performance in areas that are complementary in ecozones and anomaly regimes, which has proved the potential to detect various vegetation anomalies in any region with robustness. The characteristics of the areas of interest, the quality of Landsat images such as residual atmospheric noises may affect its performance to some extent. The above-mentioned future work and improvements are worth trying to further examine and promote UFVAD in the future.

Author Contributions: Conceptualization, Z.X. and Y.X.; methodology, Z.X., Y.X., J.S. and H.Y.; software, Y.X. and J.L.; formal analysis, Y.X., Z.X., J.S., H.Y. and J.L.; data curation, Y.X., Z.X. and J.L.; writing—original draft preparation, Y.X.; writing—review and editing, Z.X., Y.X. and K.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under grant 41771359.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The author Yixuan Xie was employed by the China Siwei Surveying and Mapping Technology Co. Ltd. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Chughtai, A.H.; Abbasi, H.; Karas, I.R. A review on change detection method and accuracy assessment for land use land cover. *Remote Sens. Appl. Soc. Environ.* **2021**, *22*, 100482. [CrossRef]
2. Iplak, B. Locust and Grasshopper Outbreaks in the Near East: Review under Global Warming Context. *Agronomy* **2021**, *11*, 111. [CrossRef]
3. Westerling, A.L.; Turner, M.G.; Smithwick, E.A.H.; Romme, W.H.; Ryan, M.G. Continued warming could transform Greater Yellowstone fire regimes by mid-21st century. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 13165–13170. [CrossRef] [PubMed]
4. Mirza, M.M.Q. Global warming and changes in the probability of occurrence of floods in Bangladesh and implications. *Glob. Environ. Change Hum. Policy Dimens.* **2002**, *12*, 127–138. [CrossRef]
5. Seidl, R.; Thom, D.; Kautz, M.; Martin-Benito, D.; Peltoniemi, M.; Vacchiano, G.; Wild, J.; Ascoli, D.; Petr, M.; Honkaniemi, J.; et al. Forest disturbances under climate change. *Nat. Clim. Change* **2017**, *7*, 395–402. [CrossRef]
6. Grings, F.; Roitberg, E.; Barraza, V. EVI Time-Series Breakpoint Detection Using Convolutional Networks for Online Deforestation Monitoring in Chaco Forest. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 1303–1312. [CrossRef]
7. Hamunyela, E.; Verbesselt, J.; de Bruin, S.; Herold, M. Monitoring Deforestation at Sub-Annual Scales as Extreme Events in Landsat Data Cubes. *Remote Sens.* **2016**, *8*, 651. [CrossRef]
8. Kasischke, E.S.; Turetsky, M.R. Recent changes in the fire regime across the North American boreal region—Spatial and temporal patterns of burning across Canada and Alaska. *Geophys. Res. Lett.* **2006**, *33*, 213–222.
9. Westerling, A.L. Increasing western US forest wildfire activity: Sensitivity to changes in the timing of spring. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*, 20150178. [CrossRef] [PubMed]
10. Meigs, G.W.; Kennedy, R.E.; Gray, A.N.; Gregory, M.J. Spatiotemporal dynamics of recent mountain pine beetle and western spruce budworm outbreaks across the Pacific Northwest Region, USA. *For. Ecol. Manag.* **2015**, *339*, 71–86. [CrossRef]
11. Paritsis, J.; Veblen, T.T. Dendroecological analysis of defoliator outbreaks on Nothofagus pumilio and their relation to climate variability in the Patagonian Andes. *Glob. Change Biol.* **2015**, *17*, 239–253. [CrossRef]
12. Emanuel, K. Increasing destructiveness of tropical cyclones over the past 30 years. *Nature* **2005**, *436*, 686–688. [CrossRef]
13. Ye, S.; Rogan, J.; Zhu, Z.; Eastman, J.R. A near-real-time approach for monitoring forest disturbance using Landsat time series: Stochastic continuous change detection. *Remote Sens. Environ.* **2021**, *252*, 112167. [CrossRef]
14. Huo, L.; Persson, H.J.; Lindberg, E. Early detection of forest stress from European spruce bark beetle attack, and a new vegetation index: Normalized distance red & SWIR (NDRS). *Remote Sens. Environ.* **2021**, *255*, 112240.
15. Bolton, D.K.; Coops, N.C.; Wulder, M.A. Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne lidar data. *Remote Sens. Environ.* **2015**, *163*, 48–60. [CrossRef]
16. Miller, A.B.; Bryant, E.S.; Birnie, R.W. An analysis of land cover changes in the Northern Forest of New England using multitemporal Landsat MSS data. *Int. J. Remote Sens.* **1998**, *19*, 245–265. [CrossRef]
17. Huang, C.; Goward, S.N.; Masek, J.G.; Thomas, N.; Zhu, Z.; Vogelmann, J.E. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sens. Environ.* **2010**, *114*, 183–198. [CrossRef]
18. Woodcock, C.E.; Macomber, S.A.; Pax-Lenney, M.; Cohen, W.B. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sens. Environ.* **2001**, *78*, 194–203. [CrossRef]
19. Kennedy, R.E.; Yang, Z.; Cohen, W.B. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910. [CrossRef]
20. Jamali, S.; Jönsson, P.; Eklundh, L.; Ardö, J.; Seaquist, J. Detecting changes in vegetation trends using time series segmentation. *Remote Sens. Environ.* **2015**, *156*, 182–195. [CrossRef]
21. Verbesselt, J.; Hyndman, R.; Newnham, G.; Culvenor, D. Detecting trend and seasonal changes in satellite image time series. *Remote Sens. Environ.* **2010**, *114*, 106–115. [CrossRef]
22. Smith, V.; Portillo-Quintero, C.; Sanchez-Azofeifa, A.; Hernandez-Stefanoni, J.L. Assessing the accuracy of detected breaks in Landsat time series as predictors of small scale deforestation in tropical dry forests of Mexico and Costa Rica. *Remote Sens. Environ.* **2019**, *221*, 707–721. [CrossRef]
23. Zhao, K.; Wulder, M.A.; Hu, T.; Bright, R.; Wu, Q.; Qin, H.; Li, Y.; Toman, E.; Mallick, B.; Zhang, X.; et al. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sens. Environ.* **2019**, *232*, 111181. [CrossRef]
24. Rogan, J.; Mietkiewicz, N. Land cover change detection. In *Land Resources Monitoring, Modeling, and Mapping with Remote Sensing*, 1st ed.; Thenkabail, P.S., Ed.; CRC Press: Boca Raton, FL, USA, 2015; pp. 579–603.
25. Anees, A.; Aryal, J. Near-Real Time Detection of Beetle Infestation in Pine Forests Using MODIS Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 3713–3723. [CrossRef]
26. Decuyper, M.; Chávez, R.O.; Lohbeck, M.; Lastra, J.A.; Tsendbazar, N.; Hackländer, J.; Herold, M.; Vågen, T. Continuous monitoring of forest change dynamics with satellite time series. *Remote Sens. Environ.* **2022**, *269*, 112829. [CrossRef]

27. Olsson, P.; Kantola, T.; Lyytikäinen-Saarenmaa, P.; Jönsson, A.M.; Eklundh, L. Development of a method for monitoring of insect induced forest defoliation—Limitation of MODIS data in Fennoscandian forest landscapes. *Silva. Fenn.* **2016**, *50*, 1495. [CrossRef]
28. Pasquarella, V.; Bradley, B.; Woodcock, C. Near-Real-Time Monitoring of Insect Defoliation Using Landsat Time Series. *Forests* **2017**, *8*, 275. [CrossRef]
29. Shang, R.; Zhu, Z.; Zhang, J.; Qiu, S.; Yang, Z.; Li, T.; Yang, X. Near-real-time monitoring of land disturbance with harmonized Landsats 7–8 and Sentinel-2 data. *Remote Sens. Environ.* **2022**, *278*, 113073. [CrossRef]
30. Zhu, Z.; Zhang, J.; Yang, Z.; Aljaddani, A.H.; Cohen, W.B.; Qiu, S.; Zhou, C. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sens. Environ.* **2019**, *238*, 111116. [CrossRef]
31. Verbesselt, J.; Zeileis, A.; Herold, M. Near real-time disturbance detection using satellite image time series. *Remote Sens. Environ.* **2012**, *123*, 98–108. [CrossRef]
32. Zhu, Z.; Woodcock, C.E. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* **2014**, *144*, 152–171. [CrossRef]
33. Tang, X.; Bullock, E.L.; Olofsson, P.; Estel, S.; Woodcock, C.E. Near real-time monitoring of tropical forest disturbance: New algorithms and assessment framework. *Remote Sens. Environ.* **2019**, *224*, 202–218. [CrossRef]
34. Estay, S.A.; Chávez, R.O. npphen: An R-package for non-parametric reconstruction of vegetation phenology and anomaly detection using remote sensing. *BioRxiv* **2018**, *1*, 301143.
35. Estay, S.A.; Chávez, R.O.; Lastra, J.A.; Rocco, R.; Gutiérrez, A.G.; Decuyper, M. MODIS Time Series Reveal New Maximum Records of Defoliated Area by Ormiscodes amphimone in Deciduous Nothofagus Forests, Southern Chile. *Remote Sens.* **2023**, *15*, 3538. [CrossRef]
36. Estay, S.A.; Chávez, R.O.; Rocco, R.; Gutiérrez, A.G. Quantifying massive outbreaks of the defoliator moth Ormiscodes amphimone in deciduous Nothofagus-dominated southern forests using remote sensing time series analysis. *J. Appl. Entomol.* **2019**, *143*, 787–796. [CrossRef]
37. Chávez, R.O.; Castillo-Soto, M.E.; Traipe, K.; Olea, M.; Lastra, J.A.; Quinones, T. A Probabilistic Multi-Source Remote Sensing Approach to Evaluate Extreme Precursory Drought Conditions of a Wildfire Event in Central Chile. *Front. Environ. Sci.* **2022**, *10*, 865406. [CrossRef]
38. Broich, M.; Huete, A.; Paget, M.; Ma, X.; Tulbure, M.; Coupe, N.R.; Evans, B.; Beringer, J.; Devadas, R.; Davies, K.; et al. A spatially explicit land surface phenology data product for science, monitoring and natural resources management applications. *Environ. Modell. Softw.* **2015**, *64*, 191–204. [CrossRef]
39. Chávez, R.O.; Rocco, R.; Gutiérrez, A.G.; Dörner, M.; Estay, S.A. A Self-Calibrated Non-Parametric Time Series Analysis Approach for Assessing Insect Defoliation of Broad-Leaved Deciduous Nothofagus pumilio Forests. *Remote Sens.* **2019**, *11*, 204. [CrossRef]
40. Tang, X.; Bullock, E.L.; Olofsson, P.; Woodcock, C.E. Can VIIRS continue the legacy of MODIS for near real-time monitoring of tropical forest disturbance? *Remote Sens. Environ.* **2020**, *249*, 112024. [CrossRef]
41. Olsson, P.; Lindström, J.; Eklundh, L. Near real-time monitoring of insect induced defoliation in subalpine birch forests with MODIS derived NDVI. *Remote Sens. Environ.* **2016**, *181*, 42–53. [CrossRef]
42. Hamunyela, E.; Rosca, S.; Mitt, A.; Engle, E.; Herold, M.; Gieseke, F.; Verbesselt, J. Implementation of BFASTmonitor Algorithm on Google Earth Engine to Support Large-Area and Sub-Annual Change Monitoring Using Earth Observation Data. *Remote Sens.* **2020**, *12*, 2953. [CrossRef]
43. Cohen, W.; Healey, S.; Yang, Z.; Stehman, S.; Brewer, C.; Brooks, E.; Gorelick, N.; Huang, C.; Hughes, M.; Kennedy, R.; et al. How Similar Are Forest Disturbance Maps Derived from Different Landsat Time Series Algorithms? *Forests* **2017**, *8*, 98. [CrossRef]
44. Cohen, W.B.; Yang, Z.; Healey, S.P.; Kennedy, R.E.; Gorelick, N. A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sens. Environ.* **2018**, *205*, 131–140. [CrossRef]
45. Dixon, L. Guaviare: Colombia's Frontline in the Country's Battle to Stop Deforestation in the Amazon. Available online: https://news.mongabay.com/2017/05/guaviare-colombias-frontline-in-the-countrys-battle-to-stop-deforestation-in-the-amazon/ (accessed on 8 August 2023).
46. University of Maryland and World Resources Institute. Global Primary Forest Loss. Available online: https://www.globalforestwatch.org/dashboards/country/COL/16/ (accessed on 23 February 2024).
47. Wu, J.; Liu, Q.; Wang, L.; Chu, G.; Liu, J. Vegetation and climate change during the last deglaciation in the Great Khingan Mountain, Northeastern China. *PLoS ONE* **2016**, *11*, e146261. [CrossRef] [PubMed]
48. Li, Y.; Gao, C.; Liu, H.; Han, D.; Cong, J.; Li, X.; Wang, G. Distribution of phosphorus forms in surface soils of typical peatlands in northern Great Khingan Mountains and its potential to reconstruct paleo-vegetations. *J. Environ. Manage.* **2022**, *302*, 114033. [CrossRef] [PubMed]
49. Veblen, T.T.; Schlegel, F.M.; Oltremari, J.V. Temperate broad-leaved evergreen forests of South America. In *Ecosystems of the World*; Ovington, J.D., Ed.; Elsevier Science Publishers: Amsterdam, The Netherlands, 1983; pp. 5–31.
50. Veblen, T.T.; Read, J.; Hill, R.S. The Ecology and Biogeography of Nothofagus Forest. *Mt. Res. Dev.* **1996**, *6*, 403.

51. Masek, J.G.; Goward, S.N.; Kennedy, R.; Cohen, W.B.; Schleweiss, K. United States Forest Disturbance Trends Observed Using Landsat Time Series. *Ecosystems* **2013**, *16*, 1087–1104. [CrossRef]
52. Wulder, M.A.; Masek, J.G.; Cohen, W.B.; Loveland, T.R.; Woodcock, C.E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sens. Environ.* **2012**, *122*, 2–10. [CrossRef]
53. Ye, S.; Rogan, J.; Sangermano, F. Monitoring rubber plantation expansion using Landsat data time series and a Shapelet-based approach. *ISPRS-J. Photogramm. Remote Sens.* **2018**, *136*, 134–143. [CrossRef]
54. Baret, F.; Morissette, J.T.; Fernandes, R.A.; Champeaux, J.; Myneni, R.B.; Chen, J.; Plummer, S.; Weiss, M.; Bacour, C.; Garrigues, S. Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1794–1803. [CrossRef]
55. Xiao, Z.; Liang, S.; Wang, J.; Chen, P.; Yin, X.; Zhang, L.; Song, J. Use of General Regression Neural Networks for Generating the GLASS Leaf Area Index Product from Time-Series MODIS Surface Reflectance. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 209–223. [CrossRef]
56. Sellers, P.J.; Berry, J.A.; Collatz, G.J.; Field, C.B.; Hall, F.G. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using enzyme kinetics- electron transfer models of leaf physiology. *Remote Sens. Environ.* **1992**, *42*, 187–216. [CrossRef]
57. Key, C.H. Ecological and Sampling Constraints on Defining Landscape Fire Severity. *Fire Ecol.* **2006**, *2*, 34–59. [CrossRef]
58. Xiao, Z.; Liang, S.; Tian, X.; Jia, K.; Yao, Y.; Jiang, B. Reconstruction of Long-Term Temporally Continuous NDVI and Surface Reflectance from AVHRR Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2017**, *10*, 5551–5568. [CrossRef]
59. Scott, D.W. Multivariate Density Estimation: Theory, Practice, and Visualization, Second Edition. *Biometrics* **1994**, *383*, 425.
60. García-Portugués, E. *Notes for Nonparametric Statistics*; Carlos III University of Madrid: Madrid, Spain, 2023.
61. Olofsson, P.; Arevalo, P.; Espejo, A.B.; Green, C.; Lindquist, E.; McRoberts, R.E.; Sanz, M.J. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.* **2020**, *236*, 111492. [CrossRef]
62. Cochran, W.G. *Sampling Techniques*; John Wiley & Sons: New York, NY, USA, 1977.
63. Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [CrossRef]
64. Stehman, S.V.; Wickham, J.D.; Smith, J.H.; Yang, L. Thematic accuracy of the 1992 National Land-Cover Data for the eastern United States: Statistical methodology and regional results. *Remote Sens. Environ.* **2003**, *86*, 500–516. [CrossRef]
65. Hansen, M.C.; Krylov, A.; Tyukavina, A.; Potapov, P.V.; Turubanova, S.; Zutta, B.; Ifo, S.; Margono, B.; Stolle, F.; Moore, R. Humid tropical forest disturbance alerts using Landsat data. *Environ. Res. Lett.* **2016**, *11*, 34008–34019. [CrossRef]
66. Mongolian Forests. Available online: https://www.toursmongolia.com/about-mongolia/mongolian-geography/forest (accessed on 26 January 2025).
67. Yan, M.; Li, Z.Y.; Chen, E.X.; Tian, X.; Gu, C.Y.; Li, C.M.; Fan, W.W. Vegetation fractional coverage change in Daxinganling Genhe forest reserve of Inner Mongolia. *Chin. J. Ecol.* **2016**, *35*, 508–515.
68. Hao, X.Y.; Zhang, Z.; Zheng, H.; Gao, N.Z. Detection of spatial and temporal variation characteristics of vegetation cover in the economic zone of the northern slopes of Tianshan Mountains and the influencing factors. *China Environ. Sci.* **2023**, *2023*, 1–12.
69. Brooks, E.B.; Wynne, R.H.; Thomas, V.A.; Blinn, C.E.; Coulston, J.W. On-the-Fly Massively Multitemporal Change Detection Using Statistical Quality Control Charts and Landsat Data. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3316–3332. [CrossRef]
70. Reiche, J.; Hamunyela, E.; Verbesselt, J.; Hoekman, D.; Herold, M. Improving near-real time deforestation monitoring in tropical dry forests by combining dense Sentinel-1 time series with Landsat and ALOS-2 PALSAR-2. *Remote Sens. Environ.* **2018**, *204*, 147–161. [CrossRef]
71. Tyukavina, A.; Potapov, P.; Hansen, M.C.; Pickens, A.H.; Stehman, S.V.; Turubanova, S.; Parker, D.; Zalles, V.; Lima, A.; Kommareddy, I.; et al. Global Trends of Forest Loss Due to Fire From 2001 to 2019. *Front. Remote Sens.* **2022**, *3*, 825190. [CrossRef]
72. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [CrossRef] [PubMed]
73. Huo, L.; Persson, H.J.; Bohlin, J.; Lindberg, E. Green Attack or Overfitting? Comparing Machine-Learning- and Vegetation-Index-Based Methods to Early Detect European Spruce Bark Beetle Attacks Using Multispectral Drone Images. In Proceedings of the IGARSS 2023–2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023; pp. 546–549.
74. Jamali, S.; Olsson, P.O.; Müller, M.; Ghorbanian, A. Kernel-Based Early Detection of Forest Bark Beetle Attack Using Vegetation Indices Time Series of Sentinel-2. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2024**, *17*, 12868–12877. [CrossRef]
75. Ye, S.; Rogan, J.; Zhu, Z.; Hawbaker, T.J.; Hart, S.J.; Andrus, R.A.; Meddens, A.J.H.; Hicke, J.A.; Eastman, J.R.; Kulakowski, D. Detecting subtle change from dense Landsat time series: Case studies of mountain pine beetle and spruce beetle disturbance. *Remote Sens. Environ.* **2021**, *263*, 112560. [CrossRef]
76. Vargas, C.; Montalban, J.; Leon, A.A. Early warning tropical forest loss alerts in Peru using Landsat. *Environ. Res. Commun.* **2019**, *1*, 121002. [CrossRef]

77. Reiche, J.; de Bruin, S.; Hoekman, D.; Verbesselt, J.; Herold, M. A Bayesian Approach to Combine Landsat and ALOS PALSAR Time Series for Near Real-Time Deforestation Detection. *Remote Sens.* **2015**, *7*, 4973–4996. [CrossRef]
78. Mera, D.; Bolon-Canedo, V.; Cotos, J.M.; Alonso-Betanzos, A. On the use of feature selection to improve the detection of sea oil spills in SAR images. *Comput. Geosci.* **2017**, *100*, 166–178. [CrossRef]
79. Sarigai; Yang, J.; Zhou, A.; Han, L.; Xie, Y. Monitoring urban black-odorous water by using hyperspectral data and machine learning. *Environ. Pollut.* **2021**, *269*, 116166. [CrossRef] [PubMed]
80. Lei, L.; Liu, L. Assessment and analysis of collapsing houses by aerial images in the Wenchuan earthquake. *J. Remote Sens.* **2010**, *14*, 333–344.