Detecting Forest Disturbance and Recovery in Primorsky Krai, Russia, Using Annual Landsat Time Series and Multi–Source Land Cover Products

Yang Hu 1,2 and Yunfeng Hu 1,2,*

1 State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; huy.17s@igsnrr.ac.cn
2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: huyf@lreis.ac.cn; Tel.: +86-10-6488-8020

Received: 26 November 2019; Accepted: 21 December 2019; Published: 1 January 2020

Abstract: The spatial distribution and dynamic changes of the forests in Primorsky Krai, Russia, are of great significance for regional ecological security and sustainable economic and societal development. With the support of the Google Earth Engine cloud computing platform, we first synthesized yearly Landsat surface reflectance images of the best quality of the research area and then used the random forest method to calculate the forest classification probability of the study area year by year from 1998 to 2015. Furthermore, we used a time–series segmentation algorithm to perform temporal trajectory segmentation for forest classification probability estimation, and determined the spatial and temporal distribution characteristics and change laws of the forest. We extended the existing algorithms and parameters of forest classification probability trajectory analysis, achieving a high overall accuracy (86.2%) in forest change detection in the study area. The extended method can accurately capture the time node information of the changes. In the present research we observed: (1) that from 1998 to 2015, the forest area of the whole district showed a net loss state, with a loss area of $0.56 \times 10^6$ ha, of which the cumulative forest disturbance area reached $1.12 \times 10^6$ ha, and the cumulative forest recovery area reached $0.55 \times 10^6$ ha; and (2) that more than 90% of the forest change occurred in areas with a slope of less than 18°, at a distance of less than 20 km from settlements, and at a distance of less than 10 km from roads. The forest disturbance monitoring results are consistent with the changes in official statistical results over time, but there was a 20% overestimation. The technical method we extended in this study can be used as a reference for large–scale and high–precision dynamic monitoring of the forests in Russia’s Far East and other regions of the world; it also provides a basis for estimating illegal timber harvesting and determining the appropriate amount of forest harvested.

Keywords: temporal segmentation; change detection; landsat; time series; forest disturbance; forest recovery

1. Introduction

Forest ecosystems are an important component of terrestrial ecosystems. According to the 2015 Global Forest Resources Assessment Report issued by the Food and Agriculture Organization (FAO) of the United Nations, the total area of global forest ecosystems is 3.99 billion hectares, accounting for 30.6% of the total land area. The forest ecosystem is also the largest carbon stock of the global terrestrial ecosystem; its stored carbon accounts for approximately 45% of the global terrestrial ecosystem carbon pool [1,2]. Therefore, research on the spatial distribution and monitoring of forest ecosystems is an important part of global change research. Forest changes occur in different forms and intensities, including natural disturbances (forest fires, pests, etc.) and human disturbances (harvesting, building...
construction, etc.), and the reverse process of the above two disturbances, namely, the forest recovery process, will have important impacts on forest ecosystems on multiple levels, such as forest cover, forest components, biodiversity, and ecosystem services. These impacts further affect the global or local climate, regional economy, and sustainable ecosystem development [3–6].

The forest is the largest ecosystem type in Primorsky Krai of the Russian Far East. The forest area of the whole region is approximately 12.3 million hectares, accounting for more than 75% of the total land area. The total output value of the forest industry accounted for approximately 8.5% of the GDP of the region. Approximately 20,000 people in the region were engaged in the forest industry, accounting for 1% of the total population in the region [7]. In addition, illegal logging is severe in the study area. In the Russian Far East, about 50% of illegal logging occurred in Primorsky Krai [8]. Therefore, the timely and accurate monitoring of the status, changes (disturbance and recovery processes), and change–driven mechanisms of the forest ecosystems in Primorsky Krai, Russia, is of great significance for assessing the regional ecological environment and enhancing the sustainable development capacity [9]. Most existing studies have been carried out around the objectives of regional forestry economic development [7,10], forestry resource management [11], and biodiversity conservation [12,13]. In the existing research on forest change in the study area, Loboda et al. [9] used Landsat data from 1972 to 2002 to obtain the forest disturbance dataset in the Far East for the past 30 years. Then, based on the Landsat archive, the real forest disturbance data for the four time periods, i.e., 2000–2002, 1990–1999, 1980–1989, and 1972–1979 were collected to verify the obtained forest disturbance dataset; the accuracy was 70%. The results showed that forest disturbances occurred mainly in the northern mountainous areas of the study area before 2000, while forest disturbances occurred mainly in the mountainous areas in the middle of the study area and in the southwestern area bordering China after 2000. In a recent study of forest change throughout Russia, Vivchar [14] analyzed the forest fire activities in Russia from 2000 to 2008 based on MODIS (Moderate Resolution Imaging Spectroradiometer) MCD 45 data. Similarly, Alexander et al. [15] estimated the area of forest fires in Russia from 2002 to 2011 based on several sets of existing global forest change datasets. With the opening of a large number of satellite remote sensing images, there is currently a great deal of global forest change data, including MODIS Burned Area products [16,17], MODIS forest disturbance mapping [18,19], SPOT–Vegetation circumpolar burned area mapping [20], and Hansen dataset (High–Resolution Global Maps of 21st–Century Forest Cover Change) [21]. To sum up, most existing studies on forest change in this study area have focused on forest disturbance events, but there is almost no research on forest recovery in the study area. Furthermore, most studies were based on existing global forest change data; however, all the data mentioned above cannot satisfy both high time and high spatial resolution, nor can it meet the data needs for the monitoring of time–space changes in forest resources and analyses of change–driven mechanisms on longer–term scales (such as since the collapse of the Soviet Union in the 1990s and after 2015). As a result, scholars have an insufficient understanding of the scale of forest disturbances (mainly deforestation and forest fires) and forest recovery in the study area and their significant impact on regional ecology, trade, and sustainable development.

In recent years, with the emergence and rapid development of remote sensing cloud platforms, the rapid processing and storage of massive–scale, long–term sequence satellite remote sensing images are no longer problematic [22]. Among the many remote sensing cloud service platforms, the Google Earth Engine (GEE) integrates massive global multisource satellite remote sensing data and ground observation data [22–25]. The GEE is the most advanced and convenient online cloud computing platform for remote sensing and geographic information processing in the world [22]; it not only provides basic computing functions for raster and vector data, but also provides application programming interfaces (APIs) based on the JavaScript and Python programming languages for developers to customize functions. Based on the GEE platform, researchers can easily conduct high–efficiency research on large–scale regions (such as provinces, countries, continents, and even at planetary scales) [23–25].
In the field of satellite remote sensing land change detection, there are two traditional methods: direct comparison (preclassification comparison) and PCC (postclassification comparison). The former method directly compares and detects the differences between two remote sensing images, such as the ratio method, the difference method, and the variation vector method. For example, Hu et al. [26] developed a change–vector analysis in posterior probability space (CVAPS) change monitoring method that the coupled normalized difference vegetation index (NDVI) time–series characteristics during the year and achieved good results in land use mapping in western China in 2009 and 2014. PCC first classifies the satellite images of two specific phases and then extracts the change map based on the land classification map. For example, Hu et al. [27] used the random forest (RF) method to conduct land mapping using satellite imagery of the Central Asian region from 2001 to 2017, and then analyzed the land change law and its driving mechanisms in the region. In general, the direct comparison method requires remote sensing images to be compared to the same satellite payload and sensor, with the same or similar phenological period [28]. The PCC method relies on two interpretations with the same land classification system and the same or similar classification accuracy [29]. However, in the above two methods, change detection based on two–two comparison makes the analysis process extremely complicated and cumbersome. Moreover, the accumulation of single–change detection errors will cause accumulated errors to exceed the allowable limit, and eventually, the drawing result will become unavailable [30].

To avoid the above problems, scholars have proposed a change detection method based on the time trajectory of feature parameters. Kennedy et al. [31] proposed the LandTrendr algorithm for the inflection point characteristics of time–series characteristic parameters (normalized burn ratio (NBR), NDVI, normalized difference moisture index (NDMI), etc.); it has been successfully used for forest disturbance monitoring in the Pacific Northwest. In more Landtrendr–based applications, Bost [32] used annual Landsat data to assess temporal and spatial patterns of the extent and magnitude of forest decline in California and the Pacific Northwest. Zhu et al. [33] proposed a land cover continuous change detection and classification (CCDC) algorithm which can continuously detect multiple land cover changes at any time point. In the New England land cover mapping test, the overall accuracy of the results obtained by the CCDC method reached 90%. Based on the principle of piecewise linear regression, and with the NDVI as the predictor variable, Verbesselt et al. [34] proposed the Breaks For Additive Season and Trend (BFAST) algorithm, which can detect trends and seasonal variations in time–series satellite images. The overall accuracy of the BFAST method in a southern Ethiopian mountain mapping test reached 78%. Continuous change detection at longer time scales, regional dynamic mapping on a larger scale, selecting more targeted surface features, optimizing the relevant parameters and thresholds of the model algorithm, and improving the automation of continuous change detection and the accuracy of the mapping results are the main directions of the current research on change detection algorithms [28,30]. In order to make full use of all the spectral information of remote sensing images, Yin et al. [35,36] used the LandTrendr algorithm to sequence the land cover probability derived from the RF classifier based on MODIS image data in Inner Mongolia. The resulting segmentation results were used to characterize changes in various land cover categories during the annual interval.

In this study, the author extended the above method, and selected Primorsky Krai of the Russian Far East as the study area, integrated the 1998–2015 long–term Landsat SR dataset, and applied the RF classification algorithm to conduct land classification. Then, using the time–series land classification probability dataset as input, the LandTrendr algorithm was used to perform time–series segmentation, and forest disturbance and forest recovery information was obtained. Combined with the environmental background of the research area and industrial development, we also examined the reliability of the forest change detection results in the study area and the advantages and disadvantages of the proposed method in this study. We attempted to answer the following three questions:

1. Is it possible to extend a forest change detection algorithm and parameter system to make them suitable for the Russian Far East based on the LandTrendr algorithm?
2. What is the forest change trend in Primorsky Krai, Russia, over the past 18 years?
3. What is the relationship between forest change (especially forest disturbance) and regional physical geography and economic and societal development?

2. Data and Methods

2.1. Research Area

Primorsky Krai (Figure 1) is located in the Asian part of Russia, between 130°–139° east longitude and 42°–48° north latitude. The area faces the Sea of Japan to the east and south, Khabarovsk Krai to the north, China to the west, and North Korea to the southwest.

![Figure 1. Topographic map of Primorsky Krai.](image)

Figure 1. Topographic map of Primorsky Krai. Primorsky Krai covers an area of 15.87 million hectares, and the forest area is approximately 12.3 million hectares, accounting for 77.5% of the total area. The shapefiles data used in the basemap is from the Resource and Environmental Science Data Center of the Chinese Academy of Sciences (http://www.resdc.cn). The digital elevation model data used in the topographic map comes from http://www2.jpl.nasa.gov/srtm.

Sihot Alin from the northeast to the southwest through the eastern and central parts of the Primorsky Territory and the west of the mountains is a lowland plain. The overall terrain of the study area is high in the northeast and low in the southwest. The highest peak of the study area is located in the Pozharsk district, with an elevation of 1993 m; the lowest point is located in the Horoli district, with an elevation of −104 m. There are 12 cities and 25 districts in Primorsky Krai, with a total area of 15.87 million hectares. The Primorsky Krai belongs to the temperate maritime monsoon climate, which is conducive to the development of forest vegetation. The total forest area is approximately 12.3 million hectares, accounting for 77.5% of the total area. The forest area mainly includes the following three forest types and tree species: Coniferous forests (cedar, fir, spruce, and larch), coniferous and broad-leaved forests (white birch, aspen, eucalyptus, and maple), and broad-leaved forests (oak, ash, eucalyptus, and yellow birch).
The total population of Primorsky Krai in Russia was approximately 1.93 million, of which approximately 20,000 (1% of the total population) were engaged in the forest industry [7]. From the official statistics of Primorsky Krai, the output value of the forest industry accounts for approximately 8.5% of the total industrial production value of the coastal area. Judging from the foreign trade reports of the Primorsky Krai in the past 10 years, China is a major timber exporter in Primorsky Krai. The main forest industrial area is Chuguyevkalin (Russia), and the main timber export port is Suifenhe (China). Among them, in 2015, the timber exports from Primorsky Krai to China accounted for 75.9% of the total timber exports in the region. The main form of timber exports is logs; the export volume of logs accounts for more than 85% of the total timber exports [37].

2.2. Basic Data

In the study, a variety of satellite imagery and geographic environment background data were used. See Table 1 for details.

<table>
<thead>
<tr>
<th>Data</th>
<th>Year(s)</th>
<th>Temporal Resolution</th>
<th>Spatial Resolution</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat 5 *</td>
<td>1990–2012</td>
<td>16 days</td>
<td>30 m</td>
<td><a href="http://landsat.usgs.gov/">http://landsat.usgs.gov/</a></td>
</tr>
<tr>
<td>Landsat 7 *</td>
<td>1999–2015</td>
<td>16 days</td>
<td>30 m</td>
<td><a href="http://landsat.usgs.gov/">http://landsat.usgs.gov/</a></td>
</tr>
<tr>
<td>Landsat 8 *</td>
<td>2013–2015</td>
<td>16 days</td>
<td>30 m</td>
<td><a href="http://landsat.usgs.gov/">http://landsat.usgs.gov/</a></td>
</tr>
<tr>
<td>SRTM3 *</td>
<td>2000</td>
<td>–</td>
<td>30 m</td>
<td><a href="http://www2.jpl.nasa.gov/srtm">http://www2.jpl.nasa.gov/srtm</a></td>
</tr>
<tr>
<td>MCD12Q1.006 *</td>
<td>2001–2015</td>
<td>1 year</td>
<td>500 m</td>
<td><a href="https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1">https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mcd12q1</a></td>
</tr>
<tr>
<td>GlobeLand30</td>
<td>2000 and 2010</td>
<td>–</td>
<td>30 m</td>
<td><a href="http://www.globeland30.com">www.globeland30.com</a></td>
</tr>
<tr>
<td>CCI</td>
<td>1998–2010</td>
<td>1 year</td>
<td>300 m</td>
<td><a href="https://www.esa-landcover-cci.org/">https://www.esa-landcover-cci.org/</a></td>
</tr>
<tr>
<td>OSM</td>
<td>Up to date</td>
<td>–</td>
<td>–</td>
<td><a href="http://www.openstreetmap.org">www.openstreetmap.org</a></td>
</tr>
</tbody>
</table>

In the GEE, the Landsat 5/7/8 SR data from 1998 to 2015 can be directly called. In this study, we also used the NDVI products calculated by the above SR products and the SRTM3 developed by the JPL laboratory in the United States. The NDVI products are introduced to strengthen the characteristics of the regional terrestrial vegetation cover [38]. SRTM3 products are intended to reduce the impact of mountains on the land cover vegetation data [39].

To construct a training sample set for automatic forest classification mapping, we introduced several sets of well–recognized land cover products in the field, including 2001–2015 year–on–year MCD12Q1 (MODIS land cover) products [40], 1998–2015 European Space Agency CCI (Climate Change Initiative) land cover products (https://www.esa-landcover-cci.org/), 2000 and 2010 global surface coverage data developed by the China National Basic Geographic Information Center (GlobeLand30) [41], and 2015 global surface coverage data product FROM–GLC, developed by Tsinghua University, China [42]. Among these land cover products, the MCD12Q1 product spatial resolution is 500 m, the CCI product spatial resolution is 300 m, and the GlobeLand30 and FROM–GLC data spatial resolutions are 30 m. The GEE platform has built–in MCD12Q1 products, which the application can use directly. The other three sets of land cover products can be called using GEE–related APIs after uploading to the GEE.
platform. All work was performed at 30 m resolution. The MCD12Q1 and CCI products were resampled in the GEE platform.

To evaluate the background characteristics of the natural geography and economic and societal development of forest change, we calculated the slopes, residential buffer zones, and road traffic network buffers. These indicators are calculated by SRTM3 V3 products [43] and OSM data products (https://www.openstreetmap.org).

To evaluate the forest disturbance data detected by the study, we used official statistics, including logging data and forest fire area. The logging data are from the official website of Primorsky Krai (http://primstat.gks.ru/wps/wcm/connect/rosstat_ts/primstat/ru). The data used for the conversion of the amount of logging (cubic meters) into a forest area (ha) are that in 2010, the forest area of Primorsky Krai was 11.833 million hectares, and the forest reserve was 1.799 billion cubic meters; that is, the amount of logging per cubic meter corresponded to a deforestation area of 65.78 square meters. The forest fire data come from the Russian official statistics website (http://www.gks.ru), and the data are as of 2012. The forest fire and logging data are tallied at the end of each year.

2.3. Forest Change Mapping

2.3.1. Overall Technical Process

Forest change detection mapping can be divided into the following four steps (Figure 2):

1. The best pixel synthesis method was applied, and Landsat SR synthetic images with no clouds and no shadows from 1998 to 2015 in the research area were constructed from the massive image library provided by the GEE. See Section 2.3.2 for the specific methods.
2. Based on four classic LULC products, i.e., MCD12Q1, CCI, GlobalLand30, and FROM–GLC, training and calibration samples were determined according to the principles of complete consistency and temporal stability of the pixel attributes (i.e., land type attributes) [27,44]. The specific method is described in Section 2.3.3.
3. The RF classification method was applied to classify the yearly images from 1998 to 2015 and to obtain yearly and pixel–by–pixel forest–type probability values. The specific method is provided in Section 2.3.4.
4. The LandTrendr algorithm was used to perform time–series segmentation on the yearly and pixel–by–pixel forest–type probability sequences obtained in the previous step. Through repeated comparison experiments, the key time–series segmentation parameters were determined. The specific method is described in Section 2.3.5.

2.3.2. Preparation of Time–Series Satellite Imagery

We extracted the SR Tier 1 datasets of the Landsat 5/7 B1–B5, B7 band and Landsat 8 B2–B7 band in the summers (June–September) of the target year and the years before and after. In view of the higher spectral and radiation resolutions of the Landsat 8 OLI sensor compared to Landsat 7 ETM+, to improve the consistency between Landsat 8 OLI (after 2013) and Landsat 7 ETM+ (before 2013), we applied the least squares regression (OLS) method proposed by Roy et al. [45]. The basic principle of the OLS method is that the straight line obtained by the best fit should yield the sum of the distances from each point to the line minimum [46]. The relevant band values of Landsat 8 are recalibrated to the level of the Landsat 7 ETM+ correlation band values by the best fit model parameters determined by Roy et al. [45]. Then, according to the CFMASK algorithm provided by the GEE, the pixel quality score (pixel_qa) is obtained, thereby eliminating pixels such as clouds, cloud shadows, and snow [47]. Finally, the cloudless, unshaded SR synthetic dataset (no pixel area in the whole area only accounts for 0.004%) of the summers from 1998–2015 in the study area was obtained by median synthesis.
**Figure 2.** Overall technical process. Forest change detection consists of four key steps (the gray background boxes in the flow chart): optimal image synthesis, filtering the training sample based on the principles of complete consistency and temporal stability, using the random forest algorithm to classify the yearly images, and using LandTrendr algorithm to perform time-series segmentation. PCC means postclassification comparison.

### 2.3.2. Preparation of Time-Series Satellite Imagery

We extracted the SR Tier 1 datasets of the Landsat 5/7 B1–B5, B7 band and Landsat 8 B2–B7 band in the summers (June–September) of the target year and the years before and after. In view of the higher spectral and radiation resolutions of the Landsat 8 OLI sensor compared to Landsat 7 ETM+, to improve the consistency between Landsat 8 OLI (after 2013) and Landsat 7 ETM+ (before 2013), we applied the least squares regression (OLS) method proposed by Roy et al. [45]. The basic principle of the OLS method is that the straight line obtained by the best fit should yield the sum of the distances from each point to the line minimum [46]. The relevant band values of Landsat 8 are recalibrated to the level of the Landsat 7 ETM+ correlation band values by the best fit model parameters determined by Roy et al. [45]. Then, according to the CFMASK algorithm provided by the GEE, the pixel quality score (pixel_qa) is obtained, thereby eliminating pixels such as clouds, cloud shadows, and snow [47]. Finally, the cloudless, unshaded SR synthetic dataset (no pixel area in the whole area only accounts for 0.004%) of the summers from 1998–2015 in the study area was obtained by median synthesis.

### 2.3.3. Selecting the Training and Validation Samples

The supervised classification of satellite remote sensing images usually requires a certain number of training and validation samples. The traditional method is to obtain training and validation sample points by referring to higher-resolution images, using manual visual interpretation methods, field survey marking ground validation points, etc. [48]. The above methods are difficult for large-scale land cover mapping studies such as countries and continents, and are not feasible in practice. Therefore, Hu et al. [27,44] proposed the principles of complete consistency and temporal stability based on multisource land cover products to achieve a highly reliable and automated method for the selection, training, and validation of sample points. The specific steps are described below.

First, the land classification systems of the multisource land cover products such as MCD12Q1, CCI, GlobalLand30, and FROM–GLC were unified, and the land cover types in the study area were divided into two categories (forest and nonforest). Then, for the annual 2001–2015 MCD12Q1 products, 1998–2015 CCI land cover products, 2000 and 2010 GlobeLand30 products, and 2015 FROM–GLC products, we conducted overlay analyses and selected pixels with completely consistent land cover types that had not changed from 1998 to 2015. MCD12Q1 and CCI were reduced to 30-m resolutions in the GEE using the `reduceResolution` function. GEE performs nearest neighbor resampling by default. In this study, we randomly selected pixels that met the aforementioned requirements and acquired
a total of 7664 points, including 6844 stable forest sample points and 820 nonforest sample points. According to the ratio of forest to nonforest area in the study area, we randomly selected 1750 stable forest sample points and 200 nonforest sample points as training samples from the sample library, and then randomly selected 50 stable forest samples and 50 nonforest samples as validation samples in the remaining sample points. In addition, we randomly selected 100 sample points from the above two types of sample points, and assessed the land cover classes through a visual interpretation of the high-resolution imagery available from Google Earth. We found that the classification accuracy of the selected sample points was 99.0%. The sample library we obtained can be used for classification training and validation.

In the same way that Loboda et al. [9] selected change samples, we selected forest disturbance and recovery samples over several time periods. Based on reference Google Earth historical images and Landsat synthetic images, through manual visual interpretation, we identified forest disturbance and recovery samples during the periods of 1998–2004, 2004–2010, and 2010–2015. In this study, we obtained a total of 320 change pixels. Among them, there were 60 forest disturbance and 50 forest recovery samples from 1998 to 2004, 52 forest disturbance and 53 forest recovery samples from 2004 to 2010, and 55 forest disturbance and 52 forest recovery samples from 2010 to 2015.

To validate the forest disturbance and recovery maps, we used two validation methods. The first was classification validation, which mainly includes four classifications: forest disturbance, forest recovery, stable forest, and nonforest. Based on the stable forest, nonforest validation samples, and forest change samples obtained above, the validation was performed every six years (a single cell has only one possible classification within this length of time). Among them, the stable forest samples and nonforest samples used in the three time periods are the same set of samples. The second method was to use statistical data for validation, to calculate the annual forest disturbance data, and compare the difference with the official data.

2.3.4. Random Forest Classification

The RF algorithm is a method to improve the accuracy of a single decision tree by combining multiple decision tree discrimination results [49]. Compared with traditional automatic classification algorithms such as MDC (minimum distance classification), MLC (maximum likelihood classification), and SVM (support vector machine), the information dimension of RF processing can be increasingly complex, so the classification effect is better and more stable [50,51]. In the RF model, the decision forest consists of multiple decision trees which are independent and unrelated. When a new sample is entered, each decision tree is judged separately to estimate which category the sample should belong to. By collecting and counting the output of all decision trees, it is possible to obtain the classification probability that a particular sample belongs to a certain category.

In the RF model used in this study, the number of variables per split was set to the square root of the number of variables, the number of decision trees was set to 1000, and the minimum number of terminal nodes was set to 10. Considering that there may be systematic differences in the satellite remote sensing images of the different years, we used the same training samples to train the images for each year from 1998 to 2015. The images data input into the RF model included the full range of Landsat bands (bands B1–B7 for Landsat 5/7 and B1–B7, B10, and B11 for Landsat 8), and NDVI and SRTM products. Among them, the NDVI were calculated using the bands of each image before synthesis. Then, applying the training result parameters of the current year, the images of the same year were classified and mapped. Finally, the probability map of the yearly image was obtained for the period of 1998 to 2015 (a total of 18 years).

2.3.5. Determining the Forest Classification and Change Thresholds

Based on the aforementioned yearly forest classification probability data, the LandTrendr algorithm was used for the periods of 1998–2004, 2004–2010, and 2010–2015 to perform time-series segmentation. Then, a series of characteristic parameters in regard to forest classification probability time-series
changes could be obtained, including the change magnitude of the classification probability, change
time duration, change start year, change end year, change start classification probability, and change end
classification probability; see Supplementary Material Figure S1 for specific characteristic parameters
definitions. In this study, we refer to the parameter scheme used by Kennedy et al. [52] in relevant
experiments in the United States. For the definition of the algorithm parameters and the parameters
used in the experiment [52,53], see Supplementary Material Table S1.

To determine the optimal characteristic parameters, we applied the best probability variation
amplitude calibration method proposed by [36]. We tested different characteristic parameters, and
combined stable forest samples, nonforest samples, and forest change samples from Section 2.3.3
as calibration samples to calculate forest disturbance and recovery user accuracy and producer
accuracy, which were further used to calculate the \( F_{1\text{max}} \) function. The relevant parameter threshold
corresponding to when \( F_{1\text{max}} \) reaches the maximum is the optimal threshold.

\[
F_{1\text{max}} = \max \{2 \times \left( \frac{UA_k \times PA_k}{UA_k + PA_k} \right) \}
\]

where \( UA \) is the user accuracy, \( PA \) is the producer accuracy, and \( k \) is the probability growth step of 0.01.
In this study, the forest classification probability test interval is 0.45–0.95, the forest change probability
test interval is 0.1–0.9, and the test interval for the change start classification probability and the change
end classification probability is 0.3–0.7.

The test showed that for forest type detection, when the forest classification probability was 0.91,
\( F_{1\text{max}} \) was the largest, and the forest classification accuracy was the best. For forest change detection,
we selected different discriminant parameters for different types of changes (Table 2), including the
change start year, the change end year, the change start classification probability, and the change
end classification probability. The discriminant parameter of the forest disturbance process is the
magnitude of change in the classification probability.

### Table 2. Optimal thresholds for forest recovery and disturbance in the different time periods.

<table>
<thead>
<tr>
<th>Period</th>
<th>Forest Recovery Conditions</th>
<th>Forest Disturbance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998–2004</td>
<td>StartYear ≥ 1990; EndYear &gt; 1998; StartVal &lt; 0.56; EndVal ≥ 0.56</td>
<td>Mag &gt; 0.36</td>
</tr>
<tr>
<td>2004–2010</td>
<td>StartYear ≥ 1994; EndYear &gt; 2004; StartVal &lt; 0.65; EndVal ≥ 0.65</td>
<td>Mag &gt; 0.42</td>
</tr>
<tr>
<td>2010–2015</td>
<td>StartYear ≥ 2003; EndYear &gt; 2010; StartVal &lt; 0.60; EndVal ≥ 0.60</td>
<td>Mag &gt; 0.40</td>
</tr>
</tbody>
</table>

Note: StartYear: change start year; EndYear: change end year; StartVal: change start classification probability; EndVal: change end classification probability; Mag: change magnitude of the classification probability. Forest recovery is a long-term process. In order not to miss the forest recovery process in the early part of each study period, we started 7–10 years ahead of the change in the detection process of each forest recovery process.

### 3. Results

#### 3.1. Quantitative Characteristics of Forest Change

From 1998 to 2015, the cumulative forest disturbance area reached \( 1.12 \times 10^6 \) ha, and the cumulative
forest recovery area reached \( 0.55 \times 10^6 \) ha. The cumulative forest disturbance area was 2.04 times the
cumulative forest recovery area, and the net loss in forest area in the whole district was \( 0.56 \times 10^6 \) ha
(Figure 3).
Figure 3. Interannual area change in forest disturbance and recovery. The area of forest disturbance in the study area was significantly larger than that of forest recovery. The forest area in the whole district showed a net loss trend, and the total forest loss area was $0.56 \times 10^6$ ha. The dotted line is used to divide the time period.

From 1998–2004, the area of forest disturbance fluctuated. The average annual forest disturbance area was $6.6 \times 10^4$ ha, and the rate of change (the coefficient of variation (CV) is defined as the ratio of the standard deviation to the mean) was 34%. In 2007, the area of forest disturbance was the smallest, i.e., only $3.5 \times 10^4$ ha. In 2010, the area of forest disturbance was the largest, reaching $11.3 \times 10^4$ ha. In the same period, the forest recovery area (average annual recovery area was $3.3 \times 10^4$ ha) was significantly smaller than the area of forest disturbance, but the CV of forest recovery was 51%, which was notably higher than the variation in forest disturbance. However, the study also showed that in 2000 and 2006, the forest recovery area ($5.0 \times 10^4$ ha and $7.4 \times 10^4$ ha, respectively) exceeded the forest disturbance area for those years.

We compared the relationship between the forest disturbance area and forest fire statistics and logging volume (Figure 4). According to official statistics, the amount of forest harvesting increased year by year before 2008; after 2008, forest harvesting declined slightly. The average annual forest harvest was $2.2 \times 10^4$ ha and the coefficient of variation was 28.6%. The annual fire fluctuation of the forest fire area was large. The forest fire area in most years was smaller than $1.1 \times 10^4$ ha, but in 1999, 2004, 2009, and 2010, it notably exceeded the historical average level, reaching $3 \times 10^4$ ha. The multiyear average forest loss caused by forest fires was $2.22 \times 10^4$ ha, but the coefficient of variation was as high as 125.6%. The total loss due to forest harvesting and forest fires was $4.42 \times 10^4$ ha, and the coefficient of variation was 62.3%. Correspondingly, the annual average forest loss based on time-series satellite remote sensing image detection was $5.3 \times 10^4$ ha, and the coefficient of variation was 55.6%. The remote sensing value was of the same order of magnitude as the statistical value, but the value was slightly higher (+20.1%). The time series of forest disturbance obtained by remote sensing monitoring has the same trend as the time series of the sum of forest harvesting and forest fires obtained by official statistics. The remote sensing monitoring results captured the statistically high value points of 1999, 2004, and 2010. The remote sensing monitoring results also reflected the troughs of several years, i.e., those in 2000, 2003, 2007, and 2013 well.
The facts showed that the algorithm constructed in this study can reproduce the temporal and
quantitative characteristics of forest disturbance (forest disturbances in Primorsky Krai were mainly
due to forest fires and logging [9,54]) in the study area. In terms of the time–series characteristics,
the remote sensing monitoring results can capture several characteristic points (peaks and valleys)
of forest disturbance. In terms of the quantity, the remote sensing monitoring results are also of
the same order of magnitude, but the remote sensing monitoring results are generally 20% higher.
The reasons for the differences may be as follows: 1. the images used in this study were collected from
the summer (June to September) in the study area, while forest fires occur mostly in spring, autumn,
and winter [9]. Due to the phase mismatch, it is reasonable for differences to occur between the
remote sensing monitoring results and the statistical data [55]; and 2. Since official statistics fail to
account for illegal logging, the official statistics are necessarily lower than the remote sensing data.

3.2. Accuracy Assessment

The overall accuracy of forest disturbance and recovery mapping in the study area from 1998
to 2015 was 86.17% (Table 3). Among these mapping types, the stable forest mapping accuracy in
1998–2004, 2004–2010, and 2010–2015 was higher than 96%, and the user accuracy was higher than
80%. For change detection, the forest disturbance detection accuracy was better, the user and producer
accuracies were higher than 83%, the forest recovery detection accuracy was slightly higher than 83%,
and the drawing accuracy was 80%.

3.3. Spatial Characteristics of Forest Disturbance and Recovery

Forest disturbance and recovery occurred mainly in the lowlands and plains west of Sihot Alin.
In the eastern and northern parts of Sihot Alin, there were also a small number of scattered forest
disturbance and recovery patches. There were slight differences in spatial distribution between forest
disturbance and recovery at different times. From 1998–2004, the forest disturbance process was mainly
distributed in the west and southwest of the study area, and sporadic forest recovery occurred in the
east and north of the study area (Figure 5a). From 2004–2010, in the context of the spatial distribution
pattern described above, significant forest recovery occurred in the surrounding area of Vladivostok in
the southwestern part of the study area (Figure 5b). The spatial distributions of forest disturbance

![Figure 4. Forest disturbance and forest fires and logging in the study area from 1999 to 2013. The main causes of the forest disturbance areas with a duration of one year are deforestation and forest fires. In the figure, the remote sensing monitoring data of forest disturbance in the current year are compared with the forest fire and logging data of the previous year.]
and recovery from 2010–2015 were similar to those from 1998–2004, but the forest recovery process in northeastern Sihot Alin increased (Figure 5c).

### Table 3. Accuracy Assessment Results.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>S</th>
<th>D</th>
<th>R</th>
<th>UA (%)</th>
<th>PA (%)</th>
<th>OA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998–2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>36</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>81.82</td>
<td>72.00</td>
<td>84.29</td>
</tr>
<tr>
<td>S</td>
<td>1</td>
<td>49</td>
<td>0</td>
<td>6</td>
<td>87.50</td>
<td>98.00</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>9</td>
<td>0</td>
<td>52</td>
<td>1</td>
<td>83.87</td>
<td>86.67</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>40</td>
<td>83.33</td>
<td>80.00</td>
<td></td>
</tr>
<tr>
<td>2004–2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>95.24</td>
<td>80.00</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1</td>
<td>48</td>
<td>4</td>
<td>7</td>
<td>80.00</td>
<td>96.00</td>
<td>87.80</td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>1</td>
<td>48</td>
<td>0</td>
<td>88.89</td>
<td>92.31</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>44</td>
<td>89.80</td>
<td>83.02</td>
<td></td>
</tr>
<tr>
<td>2010–2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>42</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>87.50</td>
<td>84.00</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>1</td>
<td>49</td>
<td>4</td>
<td>6</td>
<td>81.67</td>
<td>98.00</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>5</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>90.38</td>
<td>85.45</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>41</td>
<td>87.23</td>
<td>78.85</td>
<td></td>
</tr>
<tr>
<td>1998–2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>86.17</td>
</tr>
</tbody>
</table>

Note: N: Nonforest, S: Stable forest, D: Forest disturbance, R: Forest recovery, UA: User accuracy, PA: Producer accuracy, OA: Overall accuracy. The horizontal rows in the table represent actual data, and the vertical rows represent reference data. The kappa coefficient was 0.82 ± 0.02.

Figure 5. Cont.
Figure 5. Forest disturbance and recovery from 1998 to 2015. (a) Forest change from 1998 to 2004; (b) forest change from 2004 to 2010; (c) forest change from 2010 to 2015. Forest disturbance and recovery occurred mainly in the lowlands and plains west of Sihot Alin.
Based on the natural and human factor partitions in the study area, we analyzed the relationship between the forest change area and natural and human factors. The research showed that with increasing slope, the total area of forest disturbance and recovery showed a monotonous and rapid reduction trend. More than 90% of the forest disturbance and recovery processes occurred in areas with a slope of 0–18° (Figure 6a,b). In contrast to the above, the relationship between forest disturbance, recovery, and settlement distance was not monotonous. In the range of 0–3 km, the total area of forest disturbance and recovery increased rapidly with the distance from settlements. However, beyond 4 km, the total area of forest disturbance and recovery decreases rapidly with the distance from settlements. More than 90% of the forest disturbance and recovery processes occurred within 20 km of settlements (Figure 6c,d). Similar to the slope, as the distance from roads increased, the area of forest disturbance and recovery also decreased monotonously and rapidly. More than 90% of the forest disturbance and recovery processes occurred within 10 km of roads (Figure 6e,f).
This research showed that areas with a slope of less than 18°, a distance of less than 20 km from settlements, and a distance of less than 10 km from roads were more prone to forest change. These results revealed the key areas for the further use of high spatial resolution satellites and drone monitoring in the dynamic monitoring of forest resources.

3.4. Spatial Distribution Pattern of the Net Forest Change

From 1998 to 2015, the forest area in all administrative districts in Primorsky Krai was reduced (Figure 7). In the areas of Lesozavodsk, Hankaysk, Oktyabrsik, Horoli, and Nakhodka, the forest area was greatly reduced, and the net reduction rate of the forest area was reduced by more than 20%. Among these areas, the proportions of forest loss in Oktyabrsik, Horoli and Nakhodka City in the southwest of Sihot Alin were larger than 50%. In addition to the above administrative districts, the net forest change was represented by a smaller proportion of forest loss. Among the Chuguevskiy, Terney, Lazovsky, Kavalerovsk, Krasnoarmeyskyskiy, and Shkotovskyskiy areas in the eastern and central parts of Sihot Alin, the proportion of forest loss was no more than 3%.

![Figure 7](image-url) Net change rate of the forest area in the study area from 1998 to 2015. The forest area in each administrative district was reduced. In Oktyabrsik, Khorol, and Nakhodka City in the southwest of Sihot Alin, the reduction in forest area was the most severe.

The spatial combination pattern of the forest disturbance and recovery intensities in each administrative district (Figure 8) showed that the forest change intensity in the eastern administrative regions was low, while that in the other administrative regions in the west was relatively high. Among these regions, the total area of forest disturbance and recovery in all administrative districts in the east was smaller than 10% of the forest area corresponding to the administrative area in 1998. In the areas of Khorol, Oktyabrsik, Nakhodka City, and Khasansky, the degree of forest change was relatively high, and the cumulative area of forest disturbance and recovery exceeded 30% of the forest area of the corresponding administrative area in 1998. Among these regions, the forest disturbance area in Khorol,
Oktyabrsk, and Nakhodka City exceeded 130% of the forest area in the corresponding administrative area in 1998.

Figure 8. Forest change intensity in each administrative region of the study area. Low–intensity forest recovery and disturbance means that the cumulative change area accounts for less than 10% of the regional forest area in 1998. Medium–intensity forest recovery and forest disturbance means that the cumulative change area accounts for between 10% and 30% of the regional forest area in 1998. High–intensity forest recovery and disturbance means that the cumulative change area accounts for more than 30% of the regional forest area in 1998.

4. Discussion

4.1. New Discoveries and Their Potential Applications

The research results show that our improved forest change detection algorithm has higher mapping accuracy (86.2%), compared with the forest disturbance mapping accuracy (70.0%) of Lobodaa et al. [9]. In the study of forest classification probability time–series segmentation, scholars mainly detect changes based on MODIS images. Due to the limitation of image resolution, no matter how accurate the algorithm is, it cannot detect fine–scale forest change processes [36,56]. As for the technical route we have proposed, it can effectively detect more fine–scale forest disturbances and recovery in the Russian Far East. Specifically, the forest disturbance events in the study area are mainly distributed in the neighboring regions of Russia and China and the mountainous regions in the northeast (Figures 5, 7 and 8). The spatial distribution pattern is consistent with the forest disturbance results from 2000–2008 monitored by Lobodaa et al. [9]. Lobodaa et al. [9] also pointed out that a large number of temperate forests are distributed in the southern area of the study area, and that the pattern of regeneration after disturbance is more obvious than that in the northern area. This is similar to the monitoring results of the restoration of a large number of forest disturbances in Khasansky in the south of the study area during 1998–2010 (Figure 5a,b). More than 90% of the forest disturbance and recovery in the study area occurred in areas with a slope of less than 18°, a distance of less than 20 km from the settlements, and...
within 10 km from a road. This is similar to the result found by Kuemmerle et al. [57], that the higher degree of forest disturbance occurs on slopes of less than 20° in the Ukrainian Carpathian Mountains at the same latitude as the study area. It also shows that forest disturbances do not occur uniformly within a region, but are often concentrated in a relatively small area [58]. In addition, the results of forest disturbances we detected are consistent with the interannual trend of forest disturbance statistics (Figure 4). By analyzing the reasons for the differences between the monitoring results and the statistical values (see Section 3.1), we found that the forest disturbance monitoring data obtained from the study can be used to monitor illegal logging events in the study area. There is also research showing that by comparing forest disturbance monitoring data with statistical data, regional illegal logging data can be obtained [56].

On the one hand, we found that by comparing the results of forest disturbances obtained by this study with official statistics, it can be used to estimate the area of illegally–logged forests in the Primorsky Krai, Russian. On the other hand, because the forest harvesting results obtained in this study are higher than the official statistical results, this also provides a basis for Russian local authorities and forest authorities to formulate a reasonable forest harvesting volume and to compile forestry and reforestation planning. The improved technical methods proposed in this study can be used to develop forest resource management systems. Combining historical remote sensing images and real–time monitoring images (such as on–site monitoring, drone aerial surveys, etc.), it can be used in Russia to carry out forest monitoring dynamic changes [9], illegal logging monitoring [57,59], and reasonable forest use estimation [60]. These applications are exactly what the Russian administration urgently needs; there is currently a crisis of local forestry management in Russia [61].

4.2. Uncertainties of the Forest Change Detection Algorithm

In this study, our extended forest change detection algorithm also has many uncertainties. The spatial and temporal resolution and quality of remote sensing data are important factors affecting forest change detection. For this study, due to the cloud problem in high latitudes [3], Landsat cloudless pixels available in a year will not be sufficient to synthesize a complete image (for example, the cloudless pixels from June to September 2015 accounted for 97.2% of the entire region). This is also the main reason why this study used Landsat images of the target year and the years before and after. In addition, since the yearly composite image of the study area is synthesized from the median surface reflectance values of multiple images, this will affect the true value of the surface emissivity to a certain extent.

In addition, the technical route adopted by this research is different from those described in previous studies. In almost all recent, related studies, the researchers time–segmented their research period as a whole [3,31,35,36,56,62–64]. For this study area, due to the high degree of forest coverage (accounting for more than 75% of the area), if the entire study period were used for time series segmentation, as in the above study, the LandTrendr algorithm would remove many short–term drastic changes (especially severe forest disturbances, such as those due to forest fires and forest harvesting) as noise, and the monitoring results would greatly underestimate forest changes. In order to avoid the above problems, we divided the observation period into three subperiods (1998–2004, 2004–2010, 2010–2015), so that we could capture short–term, sharp changes well. But at the same time, as forest recovery is usually a long and ongoing process, in order not to omit the early forest recovery process of each study period, we advance the start year of change (see Section 2.3.5) in forest recovery detection at each time period by 7–10 years. However, for the forest recovery that occurred within the three periods, monitoring may cause some pixels to be detected as forest recovery in all three periods. As a result, the number and area of forest recovery tests will be exaggerated.

In the study of change detection based on the LandTrendr algorithm, the algorithm parameters and characteristic parameters obtained by time–segment segmentation (for specific definitions, see Supplementary material) are the focus of research. For algorithm parameters, we used the same parameters as in previous studies [35,36,63,64]. However, since the study area is located in a high latitude
region, most of its forest recovery process occurs over long periods. In this study, the recoveryThreshold (see Supplementary Material Table S1) value of the forest recovery test is 0.25. This parameter can eliminate a large number of pixels that quickly recover after forest disturbances caused by noise. This set of localized algorithm parameters is suitable for the same climate environment (the forest recovery process is longer), but for other regions with better climate environments (the forest recovery process is shorter), our algorithm parameters may no longer be applicable. For the characteristic parameters obtained by time–segment segmentation, the characteristic parameter combinations used in different study areas are often different. Algorithmic model corrections based on real sample data are the key to obtaining the best detection results [3]. For the study area, only a small amount of Google Earth historical images and year–to–year Landsat composite images were used to select the changing sample points. The skill level of the researchers and problems with the images themselves will affect the accuracies of the sample points. In future research, we need to consider introducing more high–resolution, historical images for real sample selection, such as QuickBird, Pleiades, Worldview, and other images.

5. Conclusions

In this study, research was carried out on the forest changes in Primorsky Krai, Russia, from 1998 to 2015. The study was performed on the GEE cloud computing platform based on massive satellite remote sensing image data. We constructed a high–confidence sample library and applied the RF classification algorithm to generate a forest classification probability dataset on an interannual pixel level. We then conducted time–series segmentation on the interannual forest classification probability, and finally, completed the identification, spatial mapping, and analysis of the forest disturbance and recovery processes in the study area. The forest change detection scheme based on the LandTrendr algorithm has better robustness than PCC. We extended the existing technical route of forest classification probability trajectory analysis, yielding an overall accuracy of 86.2% in forest disturbance and recovery mapping in the study area. The study found that the forest area in the region showed a net loss, and more than 90% of the forest changes occurred in areas with a slope of less than 18°, at a distance of less than 20 km from the settlement, and within 10 km of a road. The technical method extended by the authors and corresponding localized parameter schemes will serve as important reference values for forest change monitoring and forest management in the Russian Far East and in other regions of the world.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/12/1/129/s1, Figure S1: Trajectory segment feature parameters of forest classification for a forest loss pixels (Pixel with green dot) from 1998 to 2004, Table S1: Definition of time-series segmentation algorithm parameters.

Author Contributions: Y.H. (Yunfeng Hu) designed and supervised the research. Y.H. (Yang Hu) performed the experiments and preliminarily analyzed the results. All of the authors drafted and revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: National Key Research and Development Plan Program in China, Grant Number: 2016YFB0501502, 2016YFC0503701, Strategic Priority Research Program of Chinese Academy of Sciences, Grant Number: XDA20010202, XDA19040301, XDA23100201 and Key Project of High Resolution Earth Observation System in China, Grant Number: 00–Y30B14–9001–14/16.

Acknowledgments: We are grateful to Oregon State University’s Environmental Monitoring, Analysis and Process Recognition Lab (http://emapr.ceoas.oregonstate.edu) for providing the GEE–based LandTrendr algorithm. We sincerely thank the editors and three reviewers for their constructive comments and suggestions.

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

References


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).