remote sensing

MDPI

Article

A Cropland Disturbance Monitoring Method Based on Probabilistic Trajectories

Jiawei Jiang [1,2], Juanle Wang [2,3,4,*], Keming Yang [1], Denis Fetisov [5], Kai Li [2,3], Meng Liu [2,6] and Weihao Zou [1,2]

[1] College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing 100083, China
[2] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[3] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
[4] Jiangsu Centre for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China
[5] Institute for Complex Analysis of Regional Problems, Far Eastern Branch Russian Academy of Sciences, 679016 Birobidzhan, Russia; fetisov@icarp-febras.ru
[6] School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang 222005, China
* Correspondence: wangjl@igsnrr.ac.cn

Abstract: Acquiring the spatiotemporal patterns of cropland disturbance is of great significance for regional sustainable agricultural development and environmental protection. However, effective monitoring of cropland disturbances remains a challenge owing to the complexity of the terrain landscape and the reliability of the training samples. This study integrated automatic training sample generation, random forest classification, and the LandTrendr time-series segmentation algorithm to propose an efficient and reliable medium-resolution cropland disturbance monitoring scheme. Taking the Amur state of Russia in the Amur river basin, a transboundary region between Russia and China in east Asia with rich agriculture resources as research area, this approach was conducted on the Google Earth Engine cloud-computing platform using extensive remote-sensing image data. A high-confidence sample dataset was then created and a random forest classification algorithm was applied to generate the cropland classification probabilities. LandTrendr time-series segmentation was performed on the interannual cropland classification probabilities. Finally, the identification, spatial mapping, and analysis of cropland disturbances in Amur state were completed. Further cross-validation comparisons of the accuracy assessment and spatiotemporal distribution details demonstrated the high accuracy of the dataset, and the results indicated the applicability of the method. The study revealed that 2815.52 km^2 of cropland was disturbed between 1990 and 2021, primarily focusing on the southern edge of the Amur state. The most significant disturbance occurred in 1991, affecting 1431.48 km^2 and accounting for 50.84% of the total disturbed area. On average, 87.98 km^2 of croplands are disturbed annually. Additionally, 2495.4 km^2 of cropland was identified as having been disturbed at least once during the past 32 years, representing 83% of the total disturbed area. This study introduced a novel approach for identifying cropland disturbance information from long time-series probabilistic images. This methodology can also be extended to monitor the spatial and temporal dynamics of other land disturbances caused by natural and human activities.

Keywords: cropland disturbance; machine learning; LandTrendr; agriculture; Heilongjiang River basin; Amur state

check for updates

Citation: Jiang, J.; Wang, J.; Yang, K.; Fetisov, D.; Li, K.; Liu, M.; Zou, W. A Cropland Disturbance Monitoring Method Based on Probabilistic Trajectories. *Remote Sens.* **2024**, *16*, 4048. https://doi.org/10.3390/rs16214048

Academic Editor: Sandra Eckert

Received: 28 August 2024
Revised: 25 October 2024
Accepted: 28 October 2024
Published: 30 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Land use and land cover (LULC) is an important field for studying global environmental change, with far-reaching impacts on ecosystem services, climate change, and human survival and development [1]. As a major land-use type, cropland dynamics are closely related to food security, ecological stability, and socio-economic development [2,3]. With

the acceleration of population growth and urbanization [4], changes in croplands have become more frequent and drastic worldwide. Therefore, accurate access to information on cropland changes is important in understanding the driving mechanisms of agricultural resources [5].

Cropland disturbances involve various forms of influences, such as floods and droughts, urban expansion or Grain for Green projects [6], illegal reclamation, and so on. Traditional methods for monitoring croplands mainly rely on ground surveys and statistical data, which are time-consuming and laborious, despite their high accuracy. Realizing continuous monitoring on a large scale is difficult, leading to inconvenience and poor timeliness of information updates [7]. As remote-sensing technology advances rapidly, satellite image data have become pivotal tools for monitoring cropland changes on a large scale [8]. Medium- and high-spatial-resolution images, which provide finer details and extended observation durations, represent an effective approach for detecting fragmented croplands and monitoring cropland changes over many years [9,10]. By analyzing multitemporal remote-sensing images, spatial and temporal change information on croplands can be efficiently and accurately obtained.

Current research on cropland monitoring using medium- and high-spatial-resolution imagery has focused on two main approaches. The primary approach incorporates all the time-series information from remotely sensed images into a change-detection model, fitting time trajectories using the temporal features of spectral bands or indices to identify abrupt and continuous trends. Representative methods include the Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr) [11], Vegetation Change Tracker (VCT) [12], Breaks For Additive Seasonal and Trend (BFAST) [13], and Continuous Change Detection and Classification (CCDC) algorithms [14]. LandTrendr was initially developed for monitoring forest disturbances, tracking changes caused by deforestation, fires, and other factors, as well as their subsequent recovery processes. Recently, it has been applied to monitor agricultural land disturbances. The VCT algorithm is primarily used for forest management and large-scale change monitoring, providing information on forest dynamics by capturing growth and loss patterns. The BFAST algorithm is employed to monitor the impacts of various vegetation disturbances, separates trend and seasonal components through time series decomposition, and identifies the specific points in time when disturbance events occur. LandTrendr is straightforward to apply on a large scale to identify both abrupt and gradual changes and demands less computational resources than alternative algorithms.

Traditional algorithms for data preprocessing and execution on local servers are time-consuming. To address this issue, the LandTrendr algorithm based on the Google Earth Engine (LT-GEE) was developed for broader applications [15]. The LT-GEE algorithm can quickly process time-series data, which are useful for recognizing cropland disturbances over extended periods. Zhu et al. [16] utilized the LT-GEE algorithm with Landsat satellite data to effectively consider changes in cropland in Dongting Lake following the 1998–2018 fallow–return project. Similarly, Li [17] applied the LandTrendr algorithm to Landsat image data to identify and explore the spatial and temporal distribution changes in loess terraces in Guyuan City, north-east China. However, because cropland changes exhibit more diverse spectral characteristics in a time series, the results can exhibit significant variability and are frequently intermixed with other vegetation types like forests and grasslands.

The second approach considers croplands as a general land cover category and classifies them separately for each year using supervised classifiers. Gradinaru et al. [18] studied the abandonment of croplands in Bucharest, Romania using a Classification And Regression Tree (CART) model and validated it using Google Earth images, achieving an accuracy of over 80%. Zhang et al. [19] mapped the spatial distribution of abandoned croplands in China using the high-precision, long-time-series China Land Cover Dataset (CLCD). Data spanning 1990–2019 were analyzed using the time-sliding window method. This indicates that a method based on land-use classification can identify the change characteristics of abandoned land within a specific period. However, data quality significantly affects the accuracy of the classification results, and the final change map combines the

errors present in each map, potentially impairing its ability to capture real cropland use transitions [20]. Changes in croplands are usually gradual and occur over a long period. Therefore, relying solely on overlaying multi-year cropland distribution maps to identify changes is a challenge. Utilizing both methods could be an effective solution to overcome these shortcomings. First, cropland maps or cropland probability data at multiple time points were generated through supervised classification, and these mediated data were then used as inputs for change detection.

The key challenge in supervised classification is the acquisition of training samples. Although manual visual interpretation is highly accurate, its quick implementation is difficult on a large scale. Some researchers have proposed extracting training sample data from existing land cover datasets. Huang et al. [21] proposed a method to voluntarily extract land cover samples based on multi-source data to construct an automatic land cover classification applicable at large scales. It has a higher accuracy for global land cover products produced by manual sample selection methods. Zhang et al. [22] screened and incorporated irrigated cropland data from multiple land cover sources to develop a training sample library. They employed a random forest classifier, along with MODIS-derived spectral indices, climate data, and topographic variables, to create a map of irrigated cropland in China for the year 2015.

In summary, systematic monitoring offers a more reliable approach for studying cropland changes over time and mitigating the effects of data quality issues, complex spectral property features, and mixed pixels. This study proposes a cropland disturbance monitoring method based on probabilistic trajectories for long-time-series remote-sensing data. Training samples were automatically generated based on multi-source land cover datasets. The long-time-series Landsat images incorporating spectral features were then input into a random forest classifier to generate interannual cropland probabilities. The LandTrendr algorithm was used to obtain spatiotemporal information on cropland disturbances.

2. Study Area and Data Sources

2.1. Study Area

The Amur state is located in the Amur (Heilongjiang) River Basin, a transboundary area between China and Russia in Far East Asia. Its geographical position is advantageous, with the southern part near the Heilongjiang River on the Sino–Russian border, making it a key hub for border trade and transportation between China and Russia. This region occupies an area of 361,900 square kilometers; the region consists predominantly of plains and hills. It experiences a temperate continental climate with an annual precipitation of 800–900 mm, which is characterized by mostly sunny, windless weather throughout the year. The winters are cold and long, whereas summers are warm and humid. The state of Amur boasts large amounts of fertile black soil rich in organic matter and good moisture retention, making it an important agricultural and mineral region in Russia. The plains of Amur state are suitable for large-scale agricultural production, with the main crops being wheat, soybeans, corn, and potatoes. An overview of the study area is shown in Figure 1 (The land cover data used in the Figure 1 is the GlobeLand30 for 2020).

2.2. Basic Datasets

Various satellite images and geographical background data in this region were obtained. Detailed information is provided in Table 1.

Landsat 5/7/8 TM/ETM+/OLI Surface Reflectance (SR) multispectral imagery datasets from the U.S. Geological Survey (USGS) underwent basic preprocessing, including geometric, radiometric, and atmospheric corrections. All the Landsat image data were directly accessed on the Google Earth Engine cloud platform and were used to map the annual land cover probabilities. The Shuttle Radar Topographic Mission 3 (SRTM3) dataset provided by the Jet Propulsion Laboratory (JPL) was used in this study. The SRTM3 product was designed to minimize the impact of mountainous terrain on the classification of vegetation features [23]. The GlobeLand30 datasets were obtained from the China National Basic

Geographic Information Center. The FROM-GLC datasets were obtained from Tsinghua University. The GLC-FCS datasets were obtained from Zenodo. The land cover datasets were used to build the training sample set.

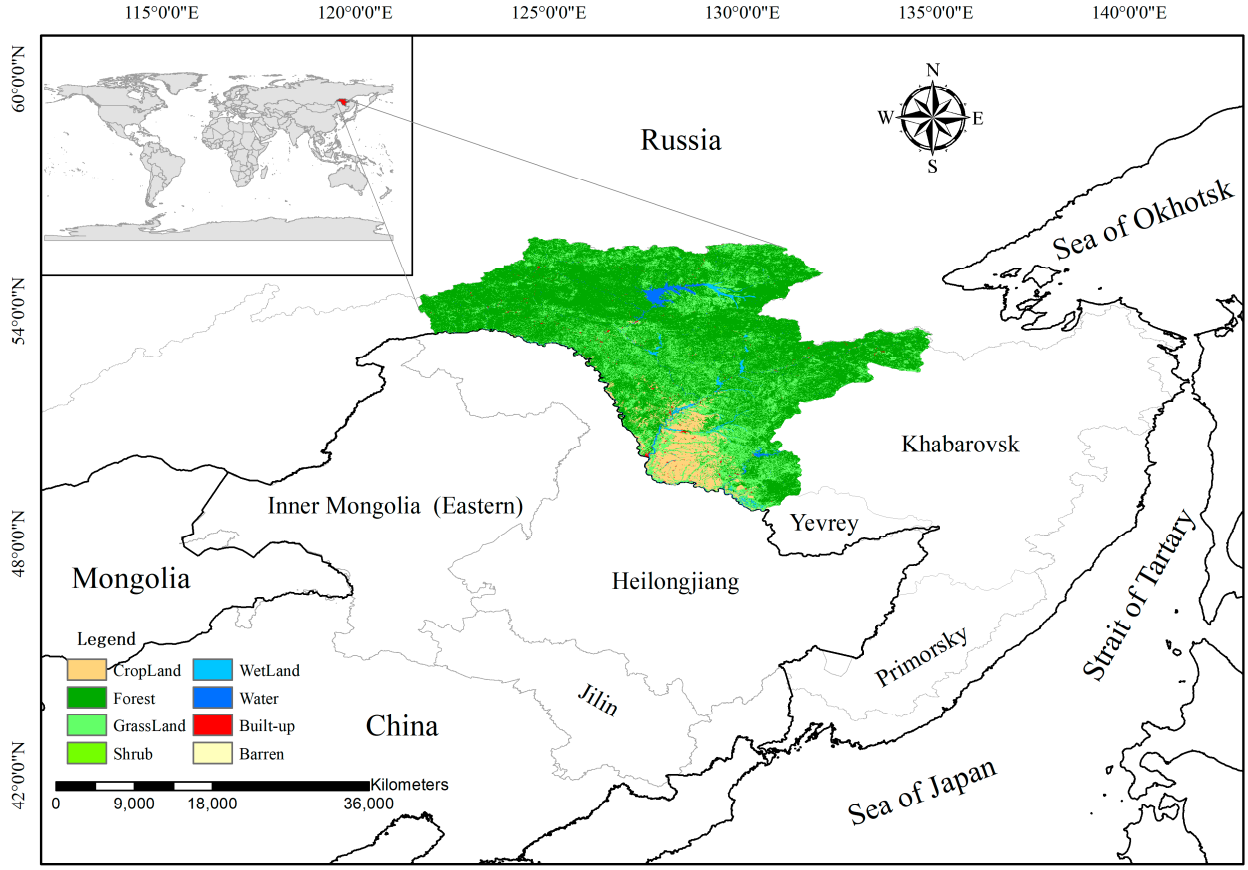


Figure 1. Situation of Amur state.

Table 1. Datasets used for this study.

Date	Year(s)	Temporal Resolution	Spatial Resolution	Data Sources
Landsat5-SR	1990–2012	16 d	30 m	http://landsat.usgs.gov/ (Accessed on 28 March 2024)
Landsat7-SR	1999–2021	16 d	30 m	http://landsat.usgs.gov/ (Accessed on 28 March 2024)
Landsat8-SR	2013–2021	16 d	30 m	http://landsat.usgs.gov/ (Accessed on 28 March 2024)
SRTM3	2000	-	30 m	http://www2.jpl.nasa.gov/srtm (Accessed on 28 March 2024)
GlobeLand30	2000, 2010, and 2020	-	30 m	http://www.globeland30.com (Accessed on 28 March 2024)
FROM-GLC	2015	-	30 m	http://data.ess.tsinghua.edu.cn /data/Simulation/ (Accessed on 28 March 2024)
GLC-FCS	1990 and 2015	-	30 m	https://zenodo.org/records/8239305 (Accessed on 28 March 2024)

3. Methods

This study proposes a probabilistic trajectory-based method for monitoring cropland disturbances. The method combines machine learning and change detection techniques, leveraging land cover probabilities derived from random forest classifiers to characterize

the year-to-year transitions between croplands and non-croplands. Change-detection algorithms were employed to capture both rapid and gradual changes, ultimately resulting in the mapping of annual cropland disturbances. The overall technical route is shown in Figure 2 and can be summarized in the following five parts:

1. Applying the optimal pixel synthesis method and constructing cloud-free and shadow-free Landsat SR synthetic images from 1990 to 2021;
2. Fusing multi-source land cover data products and automatically acquiring the training samples;
3. Generating year-by-year probabilistic maps of croplands using a random forest classification method on Landsat SR imagery from to 1990–2021;
4. Using the LandTrendr algorithm to segment the cropland probabilities to map the cropland disturbances annually;
5. Calculate the confusion matrix based on the validation samples and assess the accuracy of the maps in conjunction with the evaluation metrics.

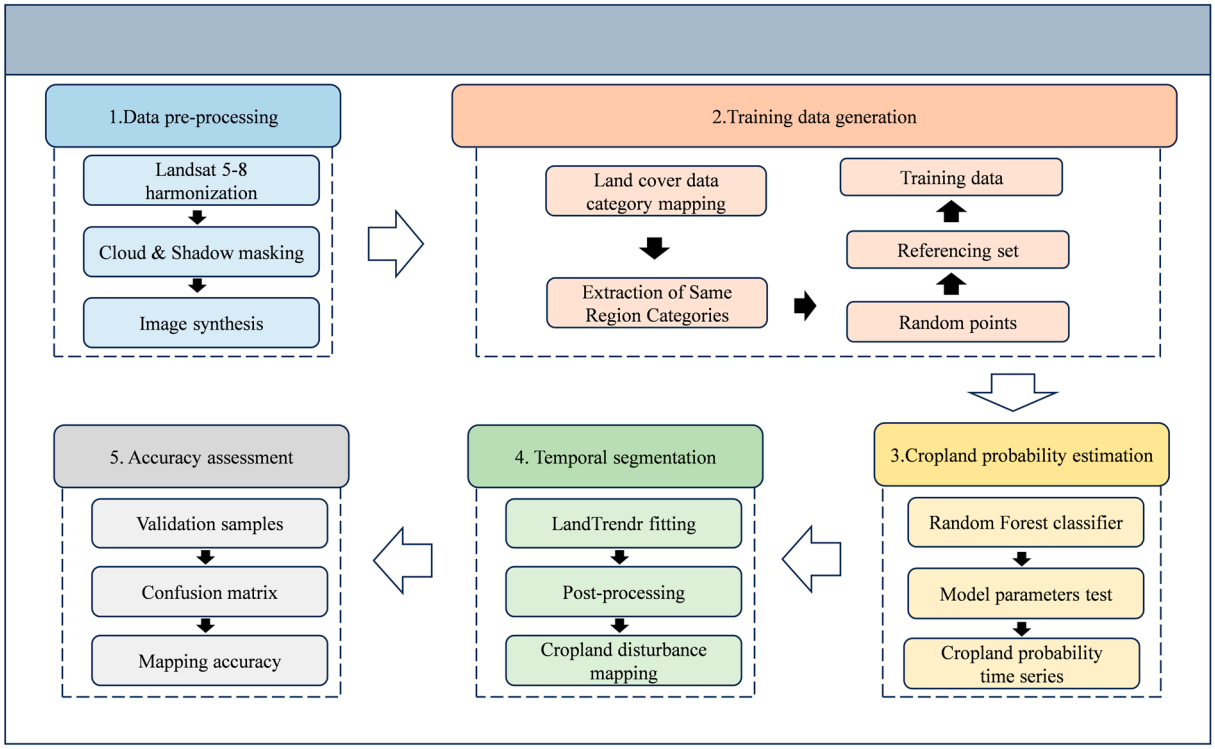


Figure 2. Overall technical process.

3.1. Data Preprocessing

This study employed annual synthetic data derived from Landsat as the principal satellite data for the time-series analysis. To reduce the annual differences associated with seasonal shifts, Landsat images from June to September of each year were used to establish the Landsat time series. A total of 8427 remote-sensing images were used in this study. Compared to the Landsat 7 ETM+, the Landsat 8 OLI has a higher spectral and radiometric resolution. To minimize the variability in the reflected wavelengths between different sensors, the data were processed for consistency by requesting the harmonic function from OLI to ETM+ using a code [24]. Clouds, snow, water, and shadows were removed using Landsat QA and CFMask bands for quality assessment. It utilized Landsat data from the target year, as well as the preceding and following years, to fill in any missing areas within the imagery. Annual synthetic images of the study area without clouds or shadows were obtained using the median synthesis method [25]. The number of available Landsat images from 1991 to 2020 for Amur state is shown in Figure 3.

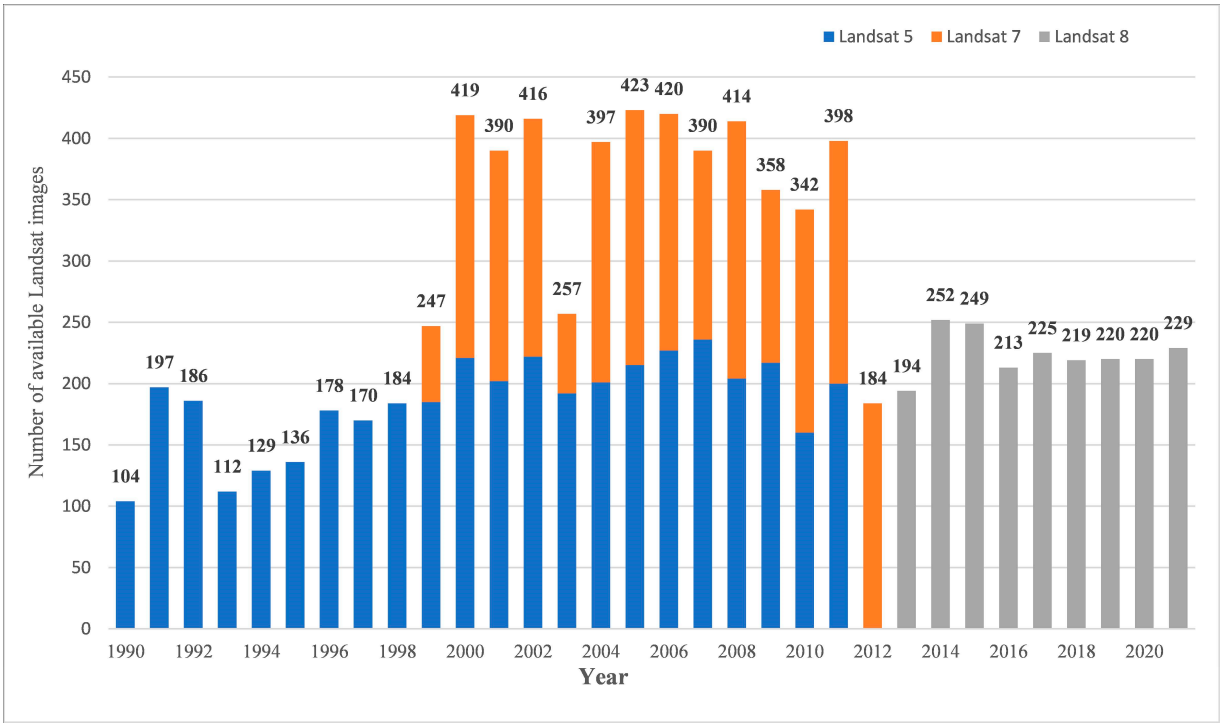


Figure 3. The number of available Landsat images during 1991–2020 in Amur state.

3.2. Training Samples Generation

By fusing multi-source land cover product in accordance with the principle of complete consistency in land cover categories and spatiotemporal alignment, training samples can be automatically generated with high accuracy and efficiency. The detailed steps are as follows: First, the classification systems of different land cover products were unified, and the land cover types were categorized into two groups: cropland and non-cropland. Next, the GlobeLand30 datasets for 2000, 2010, and 2020 were merged with FROM-GLC for 2015 and GLC-FCS for 1990 and 2015, based on the criteria described above, to identify regions where no changes occurred across the multi-source land cover products. Training samples were randomly generated from these stable regions and the sample size was determined by repeating the experiment. After manual visual interpretation of historical images from Google Earth Pro, 899 training samples were selected, including 209 stable cropland and 690 stable non-cropland samples.

3.3. Cropland Probability Estimation

The random forest algorithm improves the accuracy of a single decision tree by constructing multiple voting decision trees to determine the outcome [26]. Thus, multiple independent and unrelated decision trees form the decision forests. When a new sample is inputted, each decision tree votes separately to estimate the category to which it belongs. After tallying the outputs of all the decision trees, the classification probability that a particular sample belongs to a particular category can be obtained. Random forest classifiers are less prone to overfitting or overtraining than individual classification trees [27] and are particularly effective at modeling nonlinearities and interactions between predictor variables [28].

In this study, the number of decision trees was determined to be 1000, the number of variables selected for each split was calculated as the square root of the number of decision trees, and the minimum number of end nodes was set to 10. The class probability was defined as the ratio of the number of votes for a specific class to the total number of trees, estimating single pixel probability for cropland and non-cropland based on the proportion of tree voting for a given class. The image data input into the model included the bands

B1–B7 for Landsat5/7/8-SR. The training parameters were then applied to the SR images from the corresponding years for classification. Finally, annual cropland probability maps were obtained for the period of 1990–2021 [29].

3.4. LandTrendr Temporal Segmentation

Cropland disturbance is a complex long-term process. In the study for cropland disturbances, noise owing to image quality must be eliminated to accurately detect subtle changes. This is compatible with the advantages of traditional change detection methods that reduce interannual signal noise and can capture short-term dramatic changes by fitting the trend of the change to the long-time series [30]. The line segments in the LandTrendr fitting results consist of a series of breakpoints that represent the spectral trajectory of an image element. These segments were depicted as straight lines with a series of vertex divisions. Long- or short-term changes in features within an image can be inferred from these straight-line segments and breakpoints. The progressive and restoring trends are observed in an exponential sequence. The algorithm comprises the following processes: removal of noise-induced peaks, identification of potential vertices based on regression methods, fitting using point-to-point connections between vertices, and regression between two segments. The model was then simplified, the best model was selected by removing anomalous breakpoints, and the average probability value within each segment was calculated to obtain the modified probability, as shown in Figure 4.

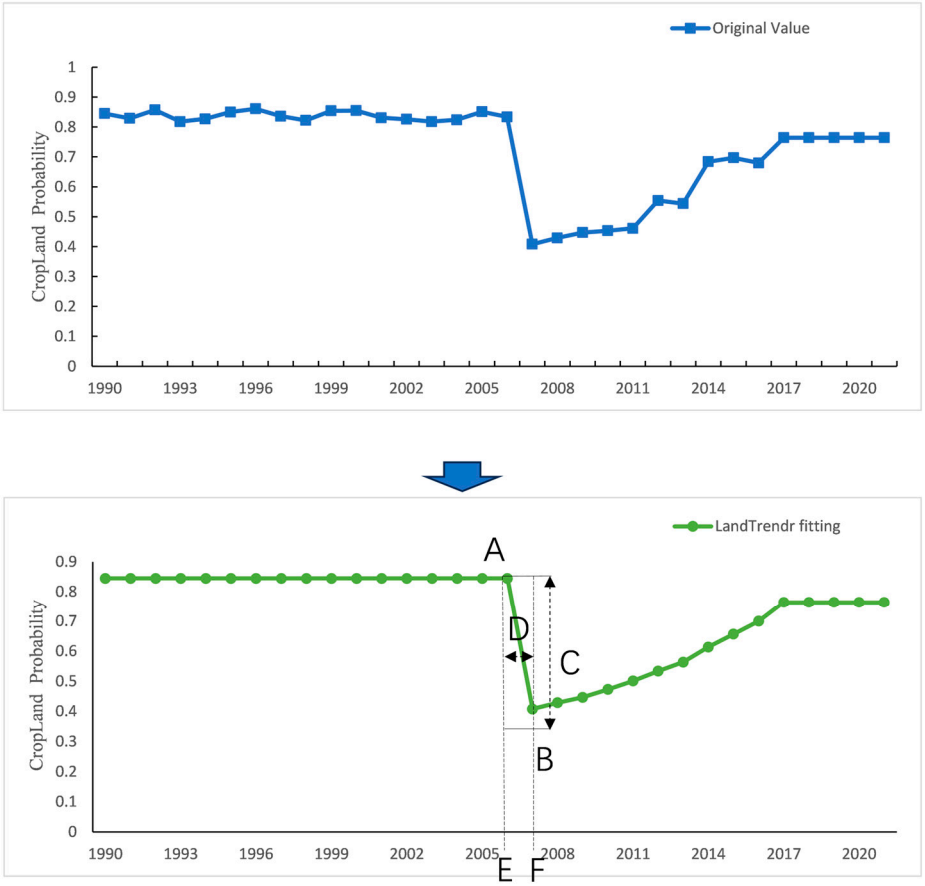


Figure 4. LandTrendr time-series segmentation of a cropland disturbance pixel. A: disturbance start cropland probability. B: disturbance stop cropland probability. C: C = A-B, disturbance magnitude of the cropland. E: disturbance start year. F: disturbance end year. D: D = F-E, disturbance time duration.

LandTrendr can only segment a single band or spectral index at a time, which limits its effectiveness in land-cover change detection [31]. Unlike most studies, this study did not use common vegetation indices, such as the normalized difference vegetation index

(NDVI) or normalized burn ratio (NBR). Instead, it utilizes land cover probability to describe the likelihood that a pixel belongs to a particular land cover class to represent actual changes in features, as demonstrated by Yin et al. [32]. Some studies have shown that probabilities can reflect the relationship between changes in different land cover categories more directly than spectral values or vegetation indices [33]. Treating land cover as a continuous process of change, rather than as a discrete single category, can better differentiate between interannual changes and those caused by classification uncertainty. The land cover probability also captures sub-pixel dissimilarities and the mixing of different types of low-resolution imagery more effectively than discrete classification [34].

The LT-GEE algorithm was applied to annual cropland probability maps to determine the year of occurrence and duration of disturbances, thereby obtaining information on cropland disturbances from 1990 to 2021. As the model parameter settings directly influence the accuracy of cropland disturbance detection, the most suitable LandTrendr parameters for the study area were identified through an experimental comparative analysis. Table 2 presents the parameters used in the study.

Table 2. Definition of LandTrendr algorithm parameters. In this study, the Max Segments parameter was set to eight to better identify cropland disturbances in the study, and the other parameters were adopted from the scheme used by Kennedy et al. [11].

Parameters	Type	This Work	Definition
Max Segments	Integer	8	Maximum number of segments to be used for time-series fitting
Spike Threshold	Float	0.9	Threshold for curbing peaks (1.0 means no curbing)
Vertex Count Overshoot	Integer	3	If the number of vertices in the initial model exceeds maxSegments + 1, change it to maxSegments + 1.
Prevent One Year Recovery	Boolean	true	Avoiding the appearance of segments representing one year of recovery
Recovery Threshold	Float	0.25	Eliminate segments with an annual recovery rate greater than 1
Pval Threshold	Float	0.05	If the *p*-value of the fitted model surpasses this threshold, the current model is rejected and another model is fitted using the optimizer
Best Model Proportion	Float	0.75	Take the model with the largest number of vertices if the *p* value of the model differs from the model with the smallest *p* value by up to this proportion
Min Observations Needed	Integer	6	Minimum number of observations to conduct the output fitting

Minimum Mapping Unit (MMU) filtering [35] was applied to eliminate potentially mixed pixels and enhance image quality, with the minimum unit set to 11 pixels (0.99 hectares) to filter preliminary cropland disturbances identified by LandTrendr. Croplands categories from GlobeLand30 for 2000, 2010, and 2020 were selected and merged to identify areas with potential cropland change. This merged area was then used to mask non-cropland regions in the LandTrendr results.

3.5. Accuracy Evaluation

This study evaluates the accuracy of the LandTrendr algorithm in recognizing cropland disturbances based on a confusion matrix. The evaluation metrics included user (UA), producer (PA), overall accuracies (OA), the kappa coefficient, and the F1 Score. The confusion matrix was calculated by comparing the position of each measured image element with that of the classified image. The overall accuracy was calculated based on the percentage of properly classified pixels relative to the total number of pixels. The kappa

coefficient is an evaluation metric of the classification results used for consistency testing and usually lies between 0 and 1, with higher values indicating higher accuracy. The F1 score serves as the primary metric for evaluating accuracy, balancing precision and recall effectively. Two hundred sample points were selected from the study area, of which 100 were sample points with disturbances and the remaining 100 were sample points without disturbances. Combined with high-spatial-resolution Google Earth imagery, each sample point was manually visually interpreted to record the sample point change and time of change. A total of 200 validation samples were used to generate a confusion matrix to assess the accuracy of the study.

$$PA = \frac{x_{ii}}{x_{+i}} \times 100\% \quad (1)$$

$$UA = \frac{x_{ii}}{x_{i+}} \times 100\% \quad (2)$$

$$OA = \frac{\sum_{i=1}^{r} x_{ii}}{N} \times 100\% \quad (3)$$

$$Kappa = \frac{N \cdot \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})} \quad (4)$$

$$\mathrm{F1} = 2 \times \left(\frac{PA \times UA}{PA + UA} \right) \quad (5)$$

where x_{ii} is the number of correctly classified pixels, x + i is the total number of pixels of type i in the reference data, x_{i+} is the total number of pixels of type i in the cropland disturbance product to be verified, r is the number of types, and N is the total number of pixels.

4. Results

4.1. Spatiotemporal Patterns of Cropland Disturbance

Figure 5 illustrates the annual trajectory of cropland disturbance in Amur state from 1990 to 2021. Considering the initial year of cropland disturbance, the spatial and temporal distributions of cropland disturbances over the 32-year period were determined using a gradient. Cropland disturbances of varying degrees occurred during the past three decades, and disturbance events were frequent, mainly in the southern part of Amur state, with a relatively concentrated spatial distribution and a high degree of patch fragmentation. Table 3 shows that the total area of disturbed cropland during 1990–2021 was 2815.52 km^2, with an average area of 87.98 km^2 per year, of which the cropland was disturbed the most in 1991, with an area of 1431.48 km^2, accounting for 50.84% of the total disturbance area. The disturbed cropland areas showed a fluctuating downward trend, with 2002, 2004, and 2017 being the main years of disturbance. The disturbed cropland areas were 85.41, 96.71, and 225.66 km^2, respectively. 2011 had the least disturbed area of cropland, with 8.14 km^2, which only accounted for 0.29% of the total disturbance area. Analysis of the time scale showed that the disturbance area of croplands exhibited an overall decreasing trend. The average annual disturbance area of cropland was 128.65 km^2 during the period of 1991–2006, and the average annual disturbance area of cropland was 50.47 km^2 during the period of 2007–2021. This indicates that the disturbance events in croplands in the area have gradually decreased.

The distribution of the duration of cropland disturbance between 1990 and 2021 is shown in Figure 6, and the area of cropland subjected to a disturbance duration between 1 and 31 years is illustrated in Figure 7. The duration of cropland disturbance events in the Amur state was dominated by 1–2 years, and the area of cropland with a disturbance duration was the highest, at 2072.60 km^2, accounting for 83% of the total area. The areas

with two and three years of disturbance duration were 120.95 and 15.84 km^2, accounting for 5% and 1% of the total area, respectively. The area of disturbance with a duration greater than three years was 541.26 km^2, accounting for 11% of the total area.

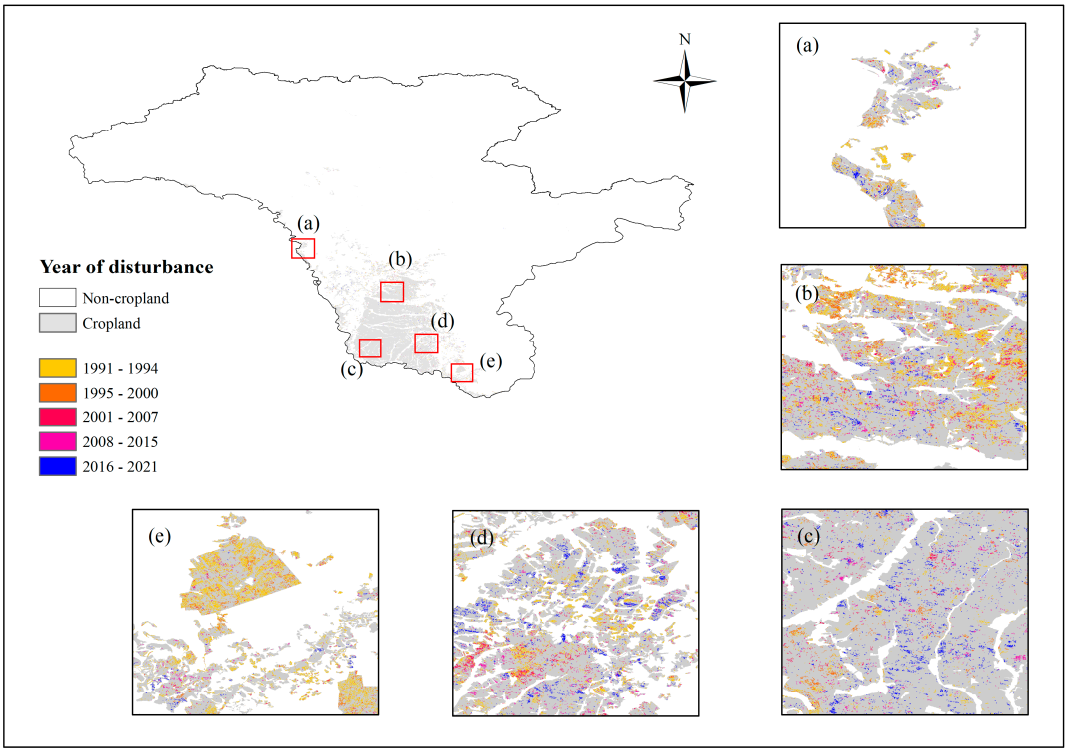


Figure 5. Spatial and temporal distribution of cropland disturbance.

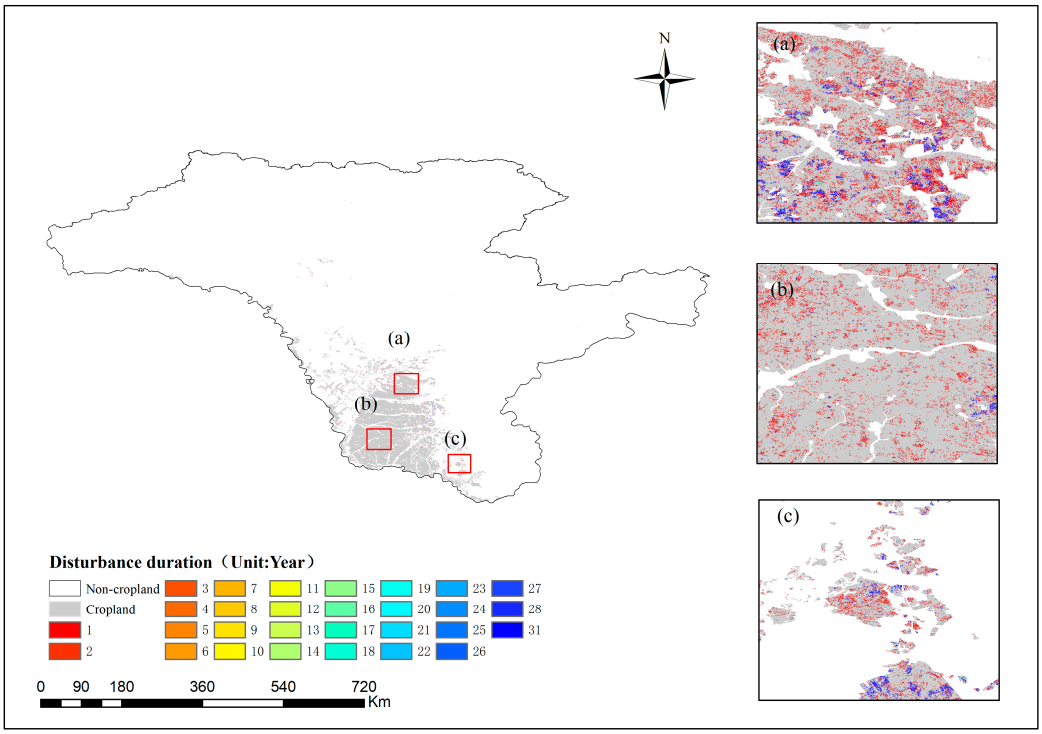


Figure 6. Duration of cropland disturbance.

Table 3. Area of cropland disturbance during 1990–2021 for each year.

Year	Area (km^2)	Proportion
1991	1431.48	50.84%
1992	1.49	0.05%
1993	8.17	0.29%
1994	17.14	0.61%
1995	59.88	2.13%
1996	32.72	1.16%
1997	24.91	0.88%
1998	51.48	1.83%
1999	74.17	2.63%
2000	31.99	1.14%
2001	38.87	1.38%
2002	85.51	3.04%
2003	70.37	2.50%
2004	96.71	3.43%
2005	12.82	0.46%
2006	20.72	0.74%
2007	23.17	0.82%
2008	8.28	0.29%
2009	23.14	0.82%
2010	17.21	0.61%
2011	8.14	0.29%
2012	32.37	1.15%
2013	70.99	2.52%
2014	32.19	1.14%
2015	40.89	1.45%
2016	71.76	2.55%
2017	225.66	8.01%
2018	24.57	0.87%
2019	25.78	0.92%
2020	77.21	2.74%
2021	75.71	2.69%
Total	2815.52	100%

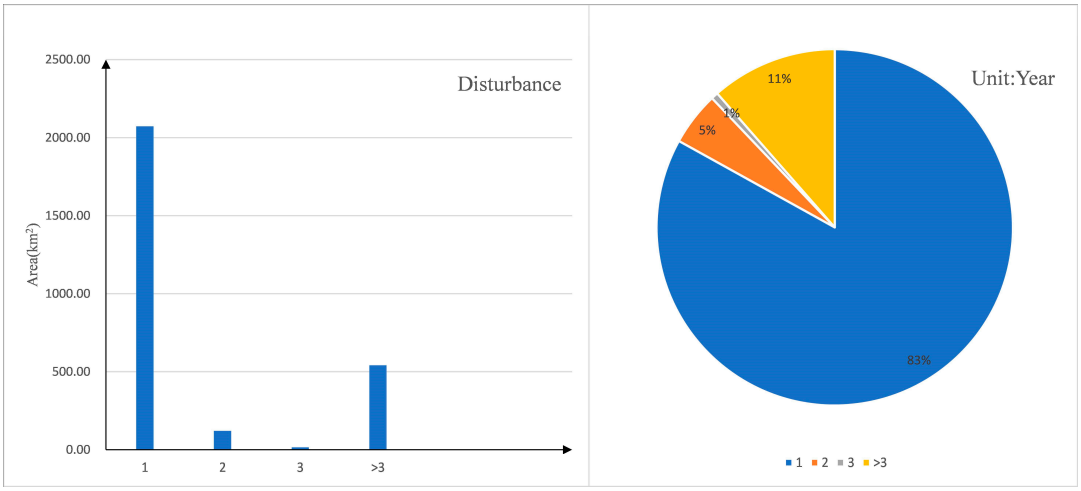


Figure 7. Area of the duration of cropland disturbance (unit = year). The legend indicates the duration of the disturbance. The left graph shows the area of cropland disturbance duration, and the right graph shows the percentage of area of cropland disturbance duration.

4.2. Accuracy Evaluation of Cropland Disturbance Maps

Table 4 shows the 91% overall accuracy of the LandTrendr algorithm in recognizing cropland disturbances and a kappa coefficient of 0.82. The F1 scores for cropland disturbance and no disturbance were 0.90 and 0.92, respectively. The producer and user accuracies of the cropland disturbance monitoring were 96.59% and 85%, respectively. This indicates that the LandTrendr algorithm, in combination with the machine-learning method, can effectively monitor cropland disturbances in Amur state.

Table 4. Accuracy assessment for cropland disturbance.

	Disturbance	No Disturbance	Total	User Accuracy
Disturbance	85	15	100	85%
No Disturbance	3	97	100	97%
Total	88	112	200	
Producer Accuracy	96.59%	86.61%		
Overall Accuracy				91%
Kappa Coefficient				0.82
F1 Scores	0.90	0.92		

The results of this study show that the LandTrendr algorithm combined with a machine-learning method is effective for detecting disturbances and capturing their occurrences and trajectories of disturbance events. To validate the recognition results of the temporal trajectory segmentation algorithm more intuitively and clearly from temporal and spatial perspectives, five typical sample areas were selected for validation. Figure 8 shows, from left to right, the Google satellite image before the disturbance, the Google satellite image monitoring in the year of the disturbance, the cropland disturbance monitoring results, and the cropland disturbance trajectory fitted by the LandTrendr algorithm. This demonstrates the changes in Google images before and after the disturbance and the matching degree of the monitored cropland disturbance information in this study. As shown in Figure 8, the LandTrendr algorithm could better monitor the years of sudden changes in cropland disturbance events and match the actual years of disturbance occurrence. The monitoring results showed good spatial consistency with the disturbance events. The results showed that the disturbance trajectories fitted by the LandTrendr algorithm could better monitor cropland disturbances in the study area.

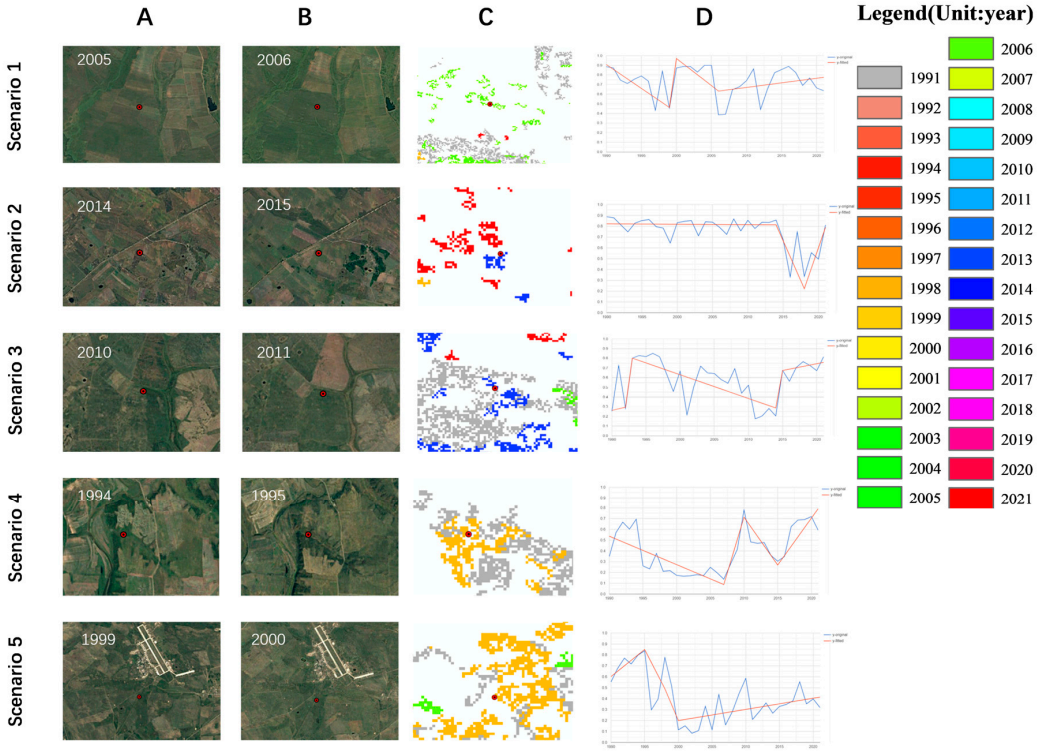


Figure 8. Comparison between cropland disturbance monitoring results with Google satellite images. (**A**,**B**) Google satellite images, (**C**) cropland disturbance monitoring results, and (**D**) disturbed trajectories of cropland based on LandTrendr algorithm fitting. The color gradient of the disturbance monitoring results represents different disturbance years.

5. Discussion

This study used the LandTrendr algorithm with machine learning to monitor cropland disturbances annually at a spatial resolution of 30 m. First, training samples were automatically generated based on multi-source land cover datasets. This saves time and effort compared with traditional methods for acquiring training samples, demonstrating significant cost-effectiveness in the large-scale mapping of cropland disturbances. Secondly, annual cropland probability estimations were performed using a random forest classifier. The time series of these probabilities were then integrated into the LandTrendr algorithm as input data for cropland trajectory fitting. The combination of change detection techniques and machine learning not only improved the accuracy of the disturbance results but also enhanced their spatiotemporal consistency. This study demonstrates that the probabilistic time series of croplands obtained from Landsat data can be used to map annual changes in croplands. The trajectory-based change-detection algorithm, LandTrendr, has been widely applied to discern forest disturbance and restoration. The algorithm was applied to map multiple LULC changes, followed by the masking of non-cropland areas in Amur state, resulting in an overall accuracy of 91.0%. The probability more effectively differentiates between cropland and non-cropland conversions compared to the results reported by Zhu et al. [16]. The boundary between croplands and non-croplands can be distinguished better from disturbance events. Navin et al. [36] extracted cropland in South America using multi-temporal Landsat images and a semi-automated field extraction method but were unable to capture dynamic change information. The method proposed in this study demonstrates significant advantages in processing long-term time-series data and in capturing cropland disturbances. In this study, it was found that cropland disturbance in the first year was significantly higher than that in subsequent years. LandTrendr usually performs poorly when detecting changes in the first and last years because of a lack of Landsat imagery data before the first year and after the last year [11]. Therefore, the confidence levels of the first and last years of the monitoring results in this study were lower than those of other years.

Cropland disturbances can be caused by a combination of factors. The driving factors of cropland disturbance are not identical in different countries. Prishchepov et al. [37] studied and summarized the reasons for cropland disturbances in Russia from to 1990–2000. It was agreed that the main factors were low food production and labor shortages due to low population density. Based on the results of this study, Benayas [38] concluded that the main factors affecting the occurrence of disturbances in croplands were related to socioeconomic development. A sharp decrease in the population of urban and rural migrants has led to the abandonment of large quantities of cropland. Urbanization and economic development have led to an increase in the employment transfer of rural labor. This may lead to labor shortages and an increase in the cost of rural labor, resulting in an annual decrease in cropland. Urban expansion and land degradation may also lead to the conversion of croplands to non-croplands [39]. Flooding in the state of Amur is primarily influenced by climate change and seasonal precipitation [40]. For example, the major floods of 2013 were due to extreme rainfall and increased river flow, which had a significant impact on agriculture, infrastructure, and the lives of the population. As documented by the Russian Federal State Statistics Service, cropland areas began to decrease rapidly in the early 1990s because of the economic turmoil and shifts in agricultural policies associated with the dissolution of the Soviet Union [41]. The rate of cropland decrease was lower in the period up to 2005 than in the period 2005–2015, which is largely in line with the findings of this study. Russia has undergone an economic transition with the dissolution of many collective farms and a decrease in agricultural inputs. Aging agricultural infrastructure in Amur state, along with the lack of modern technology and financial support, has prevented many farmers from continuing their agricultural practices on cropland. In addition, international market fluctuations and unstable prices of agricultural products have made it difficult for farmers to maintain their livelihoods, resulting in some croplands being abandoned or converted for other uses. Gellrich et al. [42] found that mountainous areas in Switzerland are disturbed by insufficient effective soil thickness, sloping terrain,

and inaccessible mountain roads. In addition, there are countries that have been forced to abandon croplands due to wars, such as Iraq, Colombia, and Yugoslavia, where portions of cropland were destroyed and could not be cultivated [43]. Disturbance duration can indicate varying levels of stress within cropland ecosystems. Short disturbances may suggest temporary disruptions due to factors such as seasonal weather patterns or pest outbreaks, while prolonged disturbances could signal more severe issues, including soil degradation or economic challenges faced by farmers. By examining these durations, we can gain insights into the adaptive capacity of agricultural practices and inform strategies for sustainable land management.

Errors in cropland disturbance maps may be related to poor estimates of cropland probability or some uncertainty in the LandTrendr fitting. This study set a relatively high change magnitude threshold to identify each transition between cropland and non-cropland to balance misclassification and omission errors. However, a reduction in false positives may result in an inability to monitor subtle categories of change, such as urban sprawl adjacent to croplands. Testing the sensitivity of the method in areas where these pixels are mixed relies heavily on high-quality imagery and requires more accurate reference data obtained from Landsat and Google. At the same time, LandTrendr is based on image-level algorithms that treat each image as a spatially isolated entity and does not utilize spatial information of neighboring images to improve the accuracy of cropland disturbance detection. This can lead to weak cropland disturbance signals received by images located at the edges of cropland disturbance events. Although the use of a filter function with the potential cropland extent eliminates the Pretzel noise present in the maps with some unreasonable results, it is also possible that the normal values in the results will be eliminated. The method proposed in this study demonstrates the potential of the LandTrendr algorithm for identifying dynamic changes across various land cover. It leverages the robust data storage and computational capabilities of the GEE platform, facilitating rapid and large-scale disturbance mapping. Furthermore, by adjusting parameters and developing tailored classification systems based on specific geographic environments, this approach can be adapted for monitoring other land types. In the future, with the continued advancement of deep learning technology, combining LandTrendr with high-resolution remote-sensing data will further enhance the accuracy and real-time performance of disturbance detection.

Varying the parameters of the LandTrendr algorithm altered the results [44]. Fixed parameters do not yield optimal results and must be combined with official verification procedures and field-collected samples. This study adopted an information extraction method that was able to obtain the spatial and temporal distribution of the disturbance tracks of croplands in Amur state. However, this approach cannot explore the causes of disturbance events, such as urban expansion, climate change, and geohazards. Only information such as the year of disturbance occurrence and duration can be detected. However, no information on the specific type of disturbance can be obtained. Therefore, in future studies, specific driver types will be identified in combination with spatiotemporal information on the identified disturbances in croplands.

6. Conclusions

This study developed a novel and cost-effective scheme for cropland disturbance monitoring that integrates Landsat time-series datasets, automated training of sample data generation, machine learning, and change-detection techniques. This approach supports the automatic generation of training data in stable and invariant regions while simplifying the cropland category and change analysis. This effectively demonstrates the combination of machine learning and change-detection methods for accurately identifying cropland disturbance information. Utilizing comprehensive Landsat data and GEE, this study mapped the distribution and duration of cropland disturbances at a 30-m resolution from 1990 to 2021 in Amur state. The results show the following: (1) The overall accuracy of the LandTrendr algorithm, based on probabilistic trajectories, in identifying cropland disturbances was 91% with a kappa coefficient of 0.82. (2) Cropland disturbances between

1990 and 2021 are primarily concentrated in the southern region of Amur state. Spatially, the distribution was relatively localized, with a total area of 2815.52 km^2 and an annual average of 87.98 km^2 of disturbed cropland. Temporally, the overall area of disturbed croplands showed a gradual decreasing trend over time. (3) The proportion of cropland disturbances lasting 1–2 years was higher from 1990 to 2021, with the largest area of disturbance occurring in those lasting one year, totaling 2072.60 km^2. In summary, this study presents a novel approach for extracting cropland disturbance information from long-term probabilistic time-series images. This aids decision-makers in improving their understanding of temporal and spatial variations in cropland disturbance. In addition, the proposed framework can be extended to monitor spatial and temporal information of other land disturbances caused by natural or human activities.

Author Contributions: Conceptualization, J.J. and J.W.; methodology, J.J.; software, D.F. and J.J.; validation, J.J., J.W. and K.Y.; formal analysis, W.Z.; investigation, M.L.; resources, J.J.; data curation, K.L. and J.J.; writing—original draft preparation, J.J.; writing—review and editing, J.W.; visualization, K.L.; supervision, K.Y.; project administration, M.L.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Alliance of International Science Organizations (Grant No. ANSO-CR-KP-2022-06), the Science & Technology Fundamental Resource Investigation Program of China (Grant No. 2022FY101902), and the Construction Project of China Knowledge Centre for Engineering Sciences and Technology (Grant No. CKCEST-2023-1-5).

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

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

References

1. Saxena, R.; Watson, L.T.; Wynne, R.H.; Brooks, E.B.; Thomas, V.A.; Zhiqiang, Y.; Kennedy, R.E. Towards a polyalgorithm for land use change detection. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 217–234. [CrossRef]
2. Liu, F. *Research on Land Use Change Detection Methods Based on Single-Phase Remote Sensing Images and GIS*; Chinese Academy of Surveying and Mapping: Beijing, China, 2006.
3. Xu, Z.; Guan, Y.; Wang, S. Detection of Cultivated Land Change by Integrating Object Image Analysis and OCNN. *Remote Sens. Inf.* **2022**, *37*, 15–22.
4. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food Security: The Challenge of Feeding 9 Billion People. *Science* **2010**, *327*, 812–818. [CrossRef] [PubMed]
5. Waldner, F.; Fritz, S.; Di Gregorio, A.; Defourny, P. Mapping Priorities to Focus Cropland Mapping Activities: Fitness Assessment of Existing Global, Regional and National Cropland Maps. *Remote Sens.* **2015**, *7*, 7959–7986. [CrossRef]
6. Fan, X.; Qu, Y.; Zhang, J.; Bai, E. China's vegetation restoration programs accelerated vegetation greening on the Loess Plateau. *Agric. For. Meteorol.* **2024**, *350*, 109994. [CrossRef]
7. Zhao, X.; Li, C.; Si, L.; Yuan, K.; Tian, Y. A New System for Forest Resources Survey Based on Information Technology. *J. Beijing For. Univ.* **2002**, *24*, 147–155.
8. Wessels, K.J.; Van den Bergh, F.; Roy, D.P.; Salmon, B.P.; Steenkamp, K.C.; MacAlister, B.; Swanepoel, D.; Jewitt, D. Rapid land cover map updates using change detection and robust random forest classifiers. *Remote Sens.* **2016**, *8*, 888. [CrossRef]
9. Woodcock, C.E.; Allen, R.; Anderson, M.; Belward, A.; Bindschadler, R.; Cohen, W.; Gao, F.; Goward, S.N.; Helder, D.; Helmer, E.; et al. Free access to Landsat imagery. *Science* **2008**, *320*, 1011. [CrossRef]
10. Wulder, M.A.; White, J.C.; Loveland, T.R.; Woodcock, C.E.; Belward, A.S.; Cohen, W.B.; Fosnight, E.A.; Shaw, J.; Masek, J.G.; Roy, D.P. The global landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* **2016**, *185*, 271–283. [CrossRef]
11. 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]
12. 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]
13. DeVries, B.; Verbesselt, J.; Kooistra, L.; Herold, M. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sens. Environ.* **2015**, *161*, 107–121. [CrossRef]
14. 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]
15. Kennedy, R.E.; Yang, Z.; Gorelick, N.; Braaten, J.; Cavalcante, L.; Cohen, W.B.; Healey, S. Implementation of the LandTrendr algorithm on google earth engine. *Remote Sens.* **2018**, *10*, 691. [CrossRef]

16. Zhu, L.; Liu, X.; Wu, L.; Tang, Y.; Meng, Y. Long-Term Monitoring of Cropland Change near Dongting Lake, China, Using the LandTrendr Algorithm with Landsat Imagery. *Remote Sens.* **2019**, *11*, 1234. [CrossRef]
17. Li, W. *Remote Sensing Monitoring of Temporal and Spatial Changes of Loess Terraces in Guyuan Based on GEE and Machine Learning*; Ningxia University: Yinchuan, China, 2022.
18. Grădinaru, S.R.; Kienast, F.; Psomas, A. Using multi-seasonal Landsat imagery for rapid identification of abandoned land in areas affected by urban sprawl. *Ecol. Indic.* **2019**, *96*, 79–86. [CrossRef]
19. Zhang, M.X.; Li, G.Y.; He, T.T.; Zhai, G.; Guo, A.D.; Chen, H.; Wu, C.F. Reveal the severe spatial and temporal patterns of abandoned cropland in China over the past 30 years. *Sci. Total Environ.* **2023**, *857*, 159591. [CrossRef]
20. Zhu, Z.; Chen, J.; Chen, J. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2013**, *34*, 2607–2654. [CrossRef]
21. Huang, Y.; Liao, S. Automatic Extraction of Land Cover Samples from Multi-Source Data. *J. Remote Sens.* **2017**, *21*, 757–766.
22. Zhang, C.; Dong, J.; Xie, Y.; Zhang, X.; Ge, Q. Mapping irrigated croplands in China using a synergetic training sample generating method, machine learning classifier, and Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *112*, 102888. [CrossRef]
23. Bai, J.; Chen, X.; Li, J.; Yang, L. Changes of inland lake area in arid Central Asia during 1975–2007: A remote-sensing analysis. *J. Lake Sci.* **2011**, *23*, 80–88.
24. Roy, D.P.; Kovalskyy, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* **2016**, *185*, 57–70. [CrossRef] [PubMed]
25. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]
26. Gislason, P.O.; Benediktsson, J.A.; Sveinsson, J.R. Random Forests for land cover classification. *Pattern Recognit. Lett.* **2006**, *27*, 294–300. [CrossRef]
27. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]
28. Cutler, D.R.; Edwards, T.C., Jr.; Beard, K.H.; Cutler, A.; Hess, J.G.; Lawler, J.T. Random forests for classification in ecology. *Ecology* **2007**, *88*, 2783–2792. [CrossRef]
29. Yin, H.; Pflugmacher, D.; Kennedy, R.E.; Sulla-Menashe, D.; Hostert, P. Mapping Annual Land Use and Land Cover Changes Using MODIS Time Series. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3421–3427. [CrossRef]
30. Yan, X.; Wang, J. The Forest Change Footprint of the Upper Indus Valley, from 1990 to 2020. *Remote Sens.* **2022**, *14*, 744. [CrossRef]
31. Kennedy, R.E.; Yang, Z.; Cohen, W.B.; Pfaff, E.; Braaten, J.; Nelson, P. Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sens. Environ.* **2012**, *122*, 117–133. [CrossRef]
32. Yin, H.; Pflugmacher, D.; Li, A.; Li, Z.; Hostert, P. Land use and land cover change in Inner Mongolia—Understanding the effects of China's re-vegetation programs. *Remote Sens. Environ.* **2018**, *204*, 918–930. [CrossRef]
33. Broich, M.; Hansen, M.C.; Potapov, P.; Adusei, B.; Lindquist, E.; Stehman, S.V. Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 277–291. [CrossRef]
34. Colditz, R.R.; Schmidt, M.; Conrad, C.; Hansen, M.C.; Dech, S. Land cover classification with coarse spatial resolution data to derive continuous and discrete maps for complex regions. *Remote Sens. Environ.* **2011**, *115*, 3264–3275. [CrossRef]
35. 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]
36. Graesser, J.; Ramankutty, N. Detection of cropland field parcels from Landsat imagery. *Remote Sens. Environ.* **2017**, *201*, 165–180. [CrossRef]
37. Prishchepov, A.V.; Mueller, D.; Dubinin, M.; Baumann, M.; Radeloff, V.C. Determinants of agricultural land abandonment in post-Soviet European Russia. *Land Use Policy* **2013**, *30*, 873–884. [CrossRef]
38. Rey Benayas, J. Abandonment of agricultural land: An overview of drivers and consequences. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* **2007**, *2*, 14. [CrossRef]
39. Yan, D.R.; Le, L. Forest carbon storage and its regional change feature in Inner Mongolia. *J. Inn. Mong. For. Sci. Technol.* **2010**, *36*, 19–22.
40. Zhang, W.X.; Wang, J.L. Flood monitoring of Heilongjiang River basin in China and Russia transboundary region based on SAR backscattering characteristics. *J. Geo-Inf. Sci.* **2022**, *24*, 802–813. [CrossRef]
41. Kerckhof, A.; Spalevic, V.; Van Eetvelde, V.; Nyssen, J. Factors of land abandonment in mountainous Mediterranean areas: The case of Montenegrin settlements. *Springerplus* **2016**, *5*, 485. [CrossRef]
42. Gellrich, M.; Zimmermann, N.E. Investigating the regional-scale pattern of agricultural land abandonment in the Swiss mountains: Aspatial statistical modelling approach. *Landsc. Urban Plan.* **2007**, *79*, 65–76. [CrossRef]
43. Witmer, F.D.W. Detecting war-induced abandoned agricultural land in northeast Bosnia using multispectral, multitemporal Landsat TM imagery. *Int. J. Remote Sens.* **2008**, *29*, 3805–3831. [CrossRef]
44. Hislop, S.; Jones, S.; Soto-Berelov, M.; Skidmore, A.; Haywood, A.; Nguyen, T.H. A fusion approach to forest disturbance mapping using time series ensemble techniques. *Remote Sens. Environ.* **2019**, *221*, 188–197. [CrossRef]

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