Article

Automatic Crop Classification in Northeastern China by Improved Nonlinear Dimensionality Reduction for Satellite Image Time Series

Yongguang Zhai ¹ , Nan Wang ^{2,*}, Lifu Zhang ³ , Lei Hao ⁴ and Caihong Hao ⁵

¹ College of Water Conservancy and Civil Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China; ychia@imau.edu.cn

² Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

³ State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; zhanglf@radi.ac.cn

⁴ College of Resources and Environmental Economics, Inner Mongolia University of Finance and Economics, Hohhot 010070, China; haolei@imufe.edu.cn

⁵ Branch of Animal Husbandry and Veterinary of Heilongjiang Academy of Agricultural Sciences, Qiqihar 161005, China; haocaihong@hljnykxy.haasep.cn

* Correspondence: wangnan@radi.ac.cn

Received: 9 June 2020; Accepted: 19 August 2020; Published: 24 August 2020

Abstract: Accurate and timely information on the spatial distribution of crops is of great significance to precision agriculture and food security. Many cropland mapping methods using satellite image time series are based on expert knowledge to extract phenological features to identify crops. It is still a challenge to automatically obtain meaningful features from time-series data for crop classification. In this study, we developed an automated method based on satellite image time series to map the spatial distribution of three major crops including maize, rice, and soybean in northeastern China. The core method used is the nonlinear dimensionality reduction technique. However, the existing nonlinear dimensionality reduction technique cannot handle missing data, and it is not designed for subsequent classification tasks. Therefore, the nonlinear dimensionality reduction algorithm Landmark–Isometric feature mapping (L–ISOMAP) is improved. The advantage of the improved L–ISOMAP is that it does not need to reconstruct time series for missing data, and it can automatically obtain meaningful featured metrics for classification. The improved L–ISOMAP was applied to Landsat 8 full-band time-series data during the crop-growing season in the three northeastern provinces of China; then, the dimensionality reduction bands were inputted into a random forest classifier to complete a crop distribution map. The results show that the area of crops mapped is consistent with official statistics. The 2015 crop distribution map was evaluated through the collected reference dataset, and the overall classification accuracy and Kappa index were 83.68% and 0.7519, respectively. The geographical characteristics of major crops in three provinces in northeast China were analyzed. This study demonstrated that the improved L–ISOMAP method can be used to automatically extract features for crop classification. For future work, there is great potential for applying automatic mapping algorithms to other data or classification tasks.

Keywords: crop classification; nonlinear dimensionality reduction; northeastern China; time series; L–ISOMAP

1. Introduction

A spatial distribution map of crops provides fundamental information for agriculture-related research, such as crop growth monitoring, yield estimation, planting structure optimization, etc.,

which are of great significance for national food security [1–3]. The conventional methods of crop statistics or monitoring are not only labor-intensive and time-consuming, but they are also not precise enough. Nowadays, remote sensing has become the main means to obtain the spatial distribution map of crops at the regional or global scale [4–6]. However, using remote sensing technology to obtain the accurate and timely spatial distribution map is also a challenge, because many crop types have similar spectral characteristics at specific phenological stages. This is especially obvious on multi-spectral images with a limited number of spectral bands [7,8]. If only a single multi-spectral image within an inappropriate time is used, the accuracy of mapping may be severely limited [9–12].

To overcome this shortcoming, a sequence of satellite images taken from different dates throughout the growing season for crops have been used in many studies, namely satellite image time series (SITS) [13–18]. The possibility of crops being successfully classified will increase by using SITS because they may exhibit different temporal characteristics at distinct phenological stages, which will translate into spectral separability to facilitate their discrimination on the images. In multi-temporal remote sensing, research studies based on vegetation index (VI) time-series data are impressive [19,20]. Some studies established decision tree models that directly used original VI values from different periods of the growing season [21–23]. This method is effective for vegetation types with unique temporal characteristics. However, some useful information in time series may be ignored because the ordered relationship of different temporal observation was not explicitly considered [24]. A more prevalent method for processing VI time series is to extract feature metrics from the time series and then input them into a supervised classifier [25]. Commonly used feature metrics include phenological and related statistical features, such as the maximum or minimum value of VI, as well as the quantile, mean, standard deviation, slope, and gradient of VI time-series curves [26–28]. Compared to using original VI values, this may improve accuracy in classification tasks [29,30]. Predefined mathematical models as more complex temporal feature extractors have been broadly adopted in vegetation phenology research and SITS classification [31]. VI time series were described as a function or a batch of functions such as Savitzky–Golay filter [32,33], Kalman Filter [34], Fourier transform [35,36], wavelet transform [37,38], spline fitting [39,40], linear regression [41,42], Hidden Markov Model [43,44], or a merger of function with various forms [45].

It is worth noting that VI is often developed by two spectral bands, which means that the remaining bands in SITS are not utilized. It can be reasonably supposed that if all spectral bands in SITS are used as the feature set instead of VI, better classification accuracy can be achieved [46,47]. This is because the remaining bands may also have obvious responses to vegetation dynamics. For example, short-wave infrared (SWIR) bands have been verified to be sensitive to vegetation moisture [48,49]. The use of full-band SITS increases the possibility of discriminating different crop types [50]. Nonetheless, the resulting high-dimensional SITS dataset also brings challenges to data processing and classification. In hyperspectral remote sensing, machine learning methods show good performance in high-dimensional data processing, especially when the number of training samples is limited [51,52]. SITS have a very similar data structure to hyperspectral images, and they contain a large number of related spectral bands. Obviously, the similarity of the data structure allows machine learning techniques used in hyperspectral images to also be used for SITS [53]. Dimensionality reduction (DR) is a machine learning method that is widely used to extract spectral features from hyperspectral images. There is great potential for applying this method to automatically extract spectral–temporal features from SITS [54].

DR is a data compression technique that projects high-dimensional data onto low-dimensional space. It can reduce noise while retaining as much useful information as possible [55,56]. DR can be divided into linear and nonlinear methods. Linear DR methods may not be suitable for optical SITS because multiple scattering occurs between different scene components during the imaging process, so that the contribution of each scene component is not linearly proportional to its surface area. This makes optical remote sensing inherently show nonlinear characteristics [57]. For SITS,

the reflectance contribution of most scene components will change temporally, which further enhances its nonlinearity.

At present, only a few studies have applied nonlinear DR methods to SITS. Yan and Roy (2015) used the nonlinear DR algorithm Laplacian Eigenmap (LE) for time-series land-cover classification and achieved higher accuracy than the traditional metrics method. Zhai et al. (2018) also proposed a time-series land-cover classification method that integrates spectral-temporal and spatial features obtained from the LE algorithm. The above methods are only tested on the SITS of a single scene, and it may have problems when used in large-scale areas. This is because SITS frequently suffer from missing data due to cloud and snow coverage, and nonlinear DR algorithms have the limitation that they cannot directly handle missing data. It is necessary to reconstruct the missing data for SITS before completing classification [58]. However, reconstructing the missing data does not increase the real information, but it brings uncertainty to the subsequent classification because of prediction errors. Therefore, in order to obtain true results from DR, a more reasonable way should be to delete the missing data directly while retaining all remaining data from SITS before DR. This requires a nonlinear DR method that can handle SITS with unequal lengths in the same periods.

In this paper, an automated crop classification method based on SITS is established. A nonlinear DR method that is commonly used for hyperspectral data has been improved for application to SITS with missing observations. The methodology is applicable to any SITS but takes Landsat data as the example in this study. The study area and the data used are first described in Section 2. Then, the automatic method for crop classification is proposed in Section 3. The isometric feature mapping (ISOMAP) nonlinear DR method, and how to improve it to deal with missing data in SITS, are also introduced in Section 3. Section 4 shows the result of applying the proposed method to three provinces in northeastern China for automatic crop classification. The performance of the improved DR method, and the effect of different proportions of training data on DR and classification, are also analyzed in this section. Section 5 discusses the reliability and uncertainty of applying the proposed method to crop classification in large-scale areas. Finally, Section 6 is the conclusion.

2. Study Areas and Data

2.1. Study Area

Northeastern China consists of Heilongjiang, Jilin, and Liaoning Provinces, which have a total area of 787,100 km², accounting for 8.2% of the land area in China. Northeastern China is located in the high latitude of China (38°43' N–53°33' N, 118°53'–135°05'), and it has a temperate monsoon climate. Summers are short and warm. Winters are long and cold. The main crops are planted once per year, including soybean, maize, and rice. The main crop growth season is from April to September with mean temperatures of 14°C to 19°C and mean precipitation of 400–700 mm. The study area is the core area of “Golden Maize Belt” in China. The Golden Maize Belt of China is at the same latitude as the American maize belt and the Ukrainian maize belt. The three maize belts together are called the “Three Major Golden Maize Belts”. The geographical position of the study area is shown in Figure 1.

2.2. Data

2.2.1. Landsat Data

In this study, Landsat 8 OLI (Operational Land Imager) Level 1 terrain-corrected (L1T) products during the crop growth season in 2015 were used. All the data were acquired from the USGS website (<http://earthexplorer.usgs.gov/>). Sixty-eight scenes with different paths/rows of Landsat 8 data are needed for covering three provinces of northeastern China. Each scene corresponds to 11 or 12 different temporal images during the crop-growing season from April 1 to September 30, with a total of 775 images for this study. For each image, we used the Landsat 8 OLI bands with 30 m spatial resolution, which included band 1 (coastal), 2 (blue), 3 (green), 4 (red), 5 (nir), 6 (swir 1), 7 (swir 2),

and two cloud mask files. Since it is sensitive to water absorption, band 9 (cirrus) was not used [59]. The path/row of Landsat data was used, and the acquisition date and cloud coverage of all temporal images for each path/row are shown in Figure 2. The raw DN values were converted into surface reflectance using an atmospheric correction tool from the Landsat Ecosystem Disturbance Adaptive Processing System [60].

Figure 1. The geographical location of the study area.

Figure 2. Cont.

Figure 2. The Landsat path/row for the whole study area and the acquisition date and cloud coverage of all temporal images for each scene with different paths/rows in the crop-growing season. The cloud coverage percentage is calculated based on all the temporal images of each scene. (a) The acquisition date and cloud coverage for the paths/rows with 11 different temporal images during the crop-growing season. (b) The acquisition date and cloud coverage for the paths/rows with 12 different temporal images during the crop-growing season. The legend shows the acquisition date (month.day) of images.

2.2.2. Land-Cover Map

A land-cover map provides a theoretical reference for cultivated land distribution. The 2015 land-cover product, with 30 m spatial resolution provided by geographical information monitoring cloud platform (<http://www.dsac.cn/>), was adopted in this study to generate an initial cultivated land mask for subsequent use. This product series was first released in 1995 and updated every few years. They are achieved by visual interpretation and human–computer interaction methods using Landsat satellite image data, which mainly includes six first-level categories: cultivated land, woodland, grassland, water bodies, built-up land, and unused land. The validation procedure shows that the overall accuracy of this product series is above 90% [61].

2.2.3. Field Data

Training and testing data were obtained from a field survey and Google Earth high-resolution images. During the entire growth period of the crop from April to September in 2015, the relevant researchers divided three groups into three provinces of the study area to carry out the field survey. As a result of the large scope of the study area, we choose two main agricultural cities in each province, including Qitaihe and Mudanjiang in Heilongjiang Province, Siping and Liaoyuan in Jilin Province, and Chaoyang and Panjin in Liaoning Province. The main content of the survey includes the latitude, longitude, and elevation information of the survey point, the soil type, the crop planting type, the crop growth period, and the crop growth condition. The sampling interval of the survey point is about 0.4 km, and the area of the survey sample is about 20 m × 20 m. One photo of each orientation survey point, including north, south, east, and west, was taken to assist in the identification of the crop category. Chaoyang in Liaoning Province has the largest number of field survey data, which is 436, and Panjin has the lowest number, which is 318. Besides the 2165 ground truth images from the field survey, the identification of crops was also replenished from Google Earth high-resolution images from May to August in 2015 through visual interpretation. Finally, a total of 3512 reference sites in 16 cities

were obtained. The difference of geographic locations for these cities ensure the spatial heterogeneity of the training and testing data [62,63]. The number of samples for each crop type is shown in Table 1. The cities where the training and testing samples were collected are shown in Figure 3.

Table 1. Number of crop reference sites.

Crop	Number of Field Survey	Number of Google Earth
Maize	1028	657
Rice	543	503
Soybean	389	114
Others	205	73

Figure 3. The cities where the training and testing samples were collected. Photos of maize taken by the field survey group in Chaoyang and Liaoyuan city.

2.2.4. Statistical Data

In addition to the field survey, we collected statistical yearbooks for three provinces. The China Statistical Yearbook 2016, published by the National Bureau of Statistics (NBS), systematically collects statistical data on all aspects of economic and social development at the country and province level for 2015 (National Statistical Bureau 2016). The relevant statistical data of agriculture (Chapter 12) in the yearbook are used as the partial validation data for this study. Statistical data for maize, rice, and soybean of northeastern China in 2015 are given in Table 2.

Table 2. Statistical data of planting acreage for three major crops in northeastern China, 2015 (1000 ha).

Province	Maize	Rice	Soybean
Heilongjiang	5821.1	3147.8	2476.1
Jilin	3800.0	761.7	284.6
Liaoning	2416.8	544.9	114.5

3. Methods

The overall procedure for mapping the distribution of maize, rice, and soybean in northeastern China is illustrated in Figure 4. The main procedure includes (i) time series construction, (ii) training samples learning, (iii) nonlinear dimensionality reduction, (iv) automatic crop mapping, and (v) accuracy assessment.

Figure 4. The flowchart to extract the distribution of crops.

3.1. Time-Series Construction

In order to effectively utilize the temporal characteristics of crops, we rearranged the satellite image time series according to the spectral–temporal mode. All of the temporal data for the same band were stacked together to construct the time series. A 2015 land-cover map was used to build a cultivated land mask. Then, we applied the mask to the spectral–temporal time series for obtaining a spectral–temporal time series of the cultivated land. For Landsat 8 satellite images with a low acquisition period, we often encounter some regions on the image that are covered by clouds. In this

study, the cfmask bands are used to remove the missing data covered by clouds and snow in SITS, while retaining all available data. As a result, the lengths of time series for data points may not be equal.

3.2. Nonlinear Dimensionality Reduction

3.2.1. ISOMAP

The ISOMAP DR method was selected for this study for the following reasons. First, it is a nonlinear DR algorithm that has been demonstrated to have better performance than linear DR algorithms when applied to hyperspectral data. Second, ISOMAP is a global optimization algorithm that uses geodesic distance to powerfully represent the true geometric structure of data. This feature is particularly attractive for remote sensing image classification because it can provide the greatest separation between classes to project the high-dimensional data into a low-dimensional space. Third, it is easy to incorporate different distance metrics because ISOMAP only uses the distance relationship between the data. This provides a possible way to handle SITS with missing data. The ISOMAP algorithm can be briefly described as follows:

- (a) Using the K nearest neighbor method to construct a neighborhood graph that can express the neighborhood structure of samples.
- (b) The geodesic distances between all samples are calculated using Dijkstra's or Floyd's shortest path algorithm.
- (c) Using these geodesic distances as input, the classical multidimensional scaling (MDS) algorithm is run to reconstruct the new sample points in a low-dimensional space.

3.2.2. L-ISOMAP

When the number of sample points increases, the computational cost of the ISOMAP algorithm will be very expensive. First, it requires obtaining the shortest paths from all the root nodes in the neighborhood graph with an order of $O(kN^2 \log N)$ using Dijkstra's algorithm, where N is the number of sample points and k is the size of the neighborhood. Second, it depends on the eigen-decomposition for the geodesic distances matrix, with a computational complexity of $O(N^3)$.

To overcome this drawback, a modified version of ISOMAP, called Landmark ISOMAP (L-ISOMAP), was developed by Silva and Tenenbaum [64].

Let vector $X = [x_1, x_2, \dots, x_N]^T \in \mathbb{R}^D$ and $Y = [y_1, y_2, \dots, y_N]^T \in \mathbb{R}^d$ represent the original input dataset and the corresponding low-dimensional embedding, where D and d denote input and embedding dimensions ($d \ll D$). The landmark dataset is $X_P = \{x_p^l\}_{l=1}^n \in X$, $n \leq N$, and the low-dimensional embedding for the landmark dataset and non-landmark dataset is Y_L and $Y_Q = Y/Y_L$, respectively. The L-ISOMAP algorithm includes four main steps. First, the neighborhood graph is constructed using K to the nearest neighbor method. If x_i is a neighbor of x_j defined by KNN, the weight of the edge between them is the Euclidean distance; otherwise, the weight is 0. Second, the geodesic distance matrix $M_{n \times N}$ between landmark dataset X_P and the original input data X are calculated using the neighborhood graph based on the shortest path distance algorithm. Clearly, only the geodesic distance matrix $M_{n \times N}$ is calculated when the computational complexity is reduced from $O(N^3)$ to $O(nN \log N)$. Third, the d -dimensional embedding of landmark dataset Y_L is obtained by MDS. The solutions of Y_L are the eigenvectors of the first d -largest eigenvalues of the inner-product matrix $\Delta_{n \times n} = -HM_GH/2$, where M_G is the squared distance matrix between pairwise landmarks, and H is the mean-centering matrix. Finally, the embedding of non-landmarks Y_Q are found by the transformation via $Y_Q = \frac{1}{2}Y_L^+(\bar{M} - M_{n \times N})$, where Y_L^+ is the pseudo inverse of landmark Y_L , and $\bar{M}$ is the column mean of the geodesic distance matrix $M_{n \times N}$. In order to project non-landmark points uniquely in the embedded space, the number of landmarks n is larger than the dimension of embeddings d .

3.2.3. TL-ISOMAP-DTW

In this study, an improved version of L-ISOMAP, called Training Landmark-ISOMAP-Dynamic Time Warping (TL-ISOMAP-DTW), was proposed for crop classification using time-series remote sensing data in a large area. We have made two improvements to the original L-ISOMAP algorithm.

(1) The first improvement is to use a dynamic time warping (DTW) metric instead of Euclidean distance to construct the nearest neighbor graph and calculate the weight of the edge. Since the missing data was eliminated, and all the available data were retained during the construction of time series data in Section 3.1, the final TS data for subsequent DR and classification is not equal in length for data points. The original L-ISOMAP algorithm cannot handle unequal time series; therefore, we use the DTW metric instead of Euclidean distance to calculate the distance between points. DTW has been proven to be capable of comparing time series with unequal lengths and handling the temporal distortion and extension [65,66].

The methodology of DTW for time series is as follows. Assume a time series M of length t , and a time series N of length s .

$$M = m_1, m_2, \dots, m_i, \dots, m_t \quad (1)$$

$$N = n_1, n_2, \dots, n_j, \dots, n_s \quad (2)$$

We construct a t -by- s path matrix based on time series M and N where the $(i$ th, j th) element of the matrix corresponds to a distance between the points m_i and n_j . The warping path is subject to several constraints such as:

Endpoint constraint: the starting and ending points of the warping path must be the first point (m_1, n_1) and the last point (m_t, n_s) of the path matrix.

Continuity constraint: in the path matrix, each step forward of the path is confined to neighboring points, $m_i - m_{i-1} \leq 1$ and $n_j - n_{j-1} \leq 1$.

Monotonicity: the warping path must move forward monotonically with respect to time, $m_{i-1} \leq m_i$ and $n_{j-1} \leq n_j$.

The cumulative distance $\gamma(i, j)$ between two time series is calculated by DTW based on the following dynamic programming formulation:

$$\gamma(i, j) = d(i, j) + \min[\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)] \quad (3)$$

where $d(i, j)$ denotes the distance between m_i and n_j ; for example, the square of the difference or the magnitude of the difference.

$$d(i, j) = (t_i - s_j)^2 \quad (4)$$

$$d(i, j) = |t_i - s_j| \quad (5)$$

Among all possible paths in the path matrix, the cumulative distance of the warping path found by DTW is the minimum. The path matrix and the DTW metric (red squares) between time series M and N is shown in Figure 5.

(2) The second improvement is the use of training samples as landmark points. In the original L-ISOMAP, the DR result mainly depends on the quantity and quality of the selected landmark points. However, how to determine landmark points lacks a common standard. Choosing the training samples as landmark points not only meets the random principle of the landmark selection, but training samples are also strongly representative of what is beneficial to subsequent classification. In this study, training samples as landmark points of the SITS of each scene with different paths/rows for DR are obtained by an automatic learning method (details in Section 3.3). In addition, the parameter (k) in TL-ISOMAP-DTW needs to be set. The “graph growth” strategy [67] is used to automatically determine the value of k . This strategy adaptively selects the neighbor size for each point instead of selecting the same number of neighbors for all points.

Figure 5. The dynamic time warping (DTW) metric (red squares) between time series M and N.

In addition, after applying the cultivated land mask, the number of data points retained in each scene is greatly reduced, which allows most scenes to be directly applied to improve the DR algorithm. However, there are still several scenes where the number of data points retained is too much to directly apply the DR algorithm. For this problem, we adopt the solution of block processing; that is, the scene is further divided into subsets according to the spatial range, and the time series of each subset is independently processed for DR.

3.3. Training Samples Learning

In this study, the different paths/rows for the SITS of each scene require a certain number of training samples as landmark points for nonlinear DR. Due to the limited number of training samples collected for the study area, it is difficult to ensure that the SITS of each scene have enough training samples as landmark points. Therefore, a method of learning new training samples for the SITS of each scene is still needed. We have built an automatic sample learning method based on the dynamic time warping (DTW) metric, and a spectral–temporal curve database for the main crops in northeastern China.

3.3.1. Training Samples Learning Based on Spectral–Temporal Curve Database

The following procedures were adopted to learn of the new training samples, which serve as landmark points for the SITS of each scene with a different path/row:

(a) Sampling. A certain percentage of field survey points (10% in this study) were randomly selected for each crop type, and then the spectral–temporal curve corresponding to each point was extracted.

(b) Construction of spectral–temporal curve. Since the samples of each crop type may come from SITS of different scenes, the dates of acquiring observations may be different. In order to obtain a complete spectral–temporal curve during the whole growing season, all the dates of acquiring observations are used in temporal order. This may cause the temporal interval of the constructed spectral–temporal curve to be inconsistent, but it has little effect on the DTW algorithm. If there are multiple available data for the same day, the mean value is calculated. The spectral–temporal curve of maize, soybean, and rice in Northeastern China is shown in Figure 6.

(c) Finding similar pixels. For the SITS of each scene with a different path/row, we used the DTW metric to search for a specified number of pixels closest to the spectral curve of each crop.

(d) Sorting by DTW distance. All similar pixels for the three different crop types were sorted by the DTW distance calculated in the last step.

(e) Choosing training samples. According to the number of landmark points required, the same number of similar pixels with the smallest DTW distance were reserved for SITS in each scene, with a different path/row as training samples for DR and classification.

Figure 6. The spectral-temporal curve for maize, soybean, and rice in Northeastern China. (The horizontal labels Band 1, . . . , Band 7 is the abbreviation of all temporal data stacked together for Band 1, . . . , Band 7).

3.3.2. Determine Landmark Points Size

In order to determine the reasonable number of training samples as landmark points, we used 3512 verification points collected to test the effect of different sizes of landmark points in the TL-ISOMAP-DTW method on the classification accuracy. Different percentages (0.1%, 0.2%, 0.3%, 0.5%, 0.7%, 1%, 2%, 3%, and 5%) of samples were selected as landmark points, respectively. The results of DR using TL-ISOMAP-DTW were inputted into a random forest (RF) classifier. To reduce the possible errors caused by the classification process, the classifier also used different proportions of training samples (the proportion is the same as the percentage chosen for landmark points). The rest of the data were used as testing samples. Finally, the DR performance was evaluated by the classification accuracy. Each experiment was repeated 10 times, with the mean value as the final result. Among the results with the highest supervised classification accuracy, the minimum percentage was used in this study to determine the number of landmark points for the SITS of each scene with a different path/row.

3.4. Automatic Crop Mapping and Accuracy Assessment

We input the DR data of each scene with different paths/rows and the corresponding training samples from the last step into random forests (RFs) [68] classifiers to complete the classification for each scene, and we mosaic all of the classification results to obtain the final crop distribution map in Northeastern China. The RF classifier includes multiple tree classifiers, and each tree classifier has an equal vote to decide the class of the input vector [69]. To run an RF classifier, two parameters have to be set, including the number of classification trees and the number of input variables used at each node. In this paper, we chose settings of 500 trees and 2 variables at each node for the RF classifier. Finally, the statistical data and field data were used to assess the classification accuracy. Since 10% of the field survey data is used when constructing the spectral-temporal curve, all remaining field data are used for accuracy assessment.

4. Results

4.1. Algorithm Sensitivity to Landmark Points Size

The classification accuracy of landmark points with different percentages in TL-ISOMAP-DTW for 3512 verification points is shown in Figure 7. Meanwhile, the L-ISOMAP method is used for comparison. Since L-ISOMAP cannot handle SITS with missing data, we need to use temporal interpolation for these points before L-ISOMAP DR. All settings are the same with TL-ISOMAP-DTW

except for the landmark points, which are randomly selected in L-ISOMAP instead of the activated learning method in TL-ISOMAP-DTW.

It can be seen from Figure 7 that the overall accuracy increases with the rising of the percentage of landmark points in TL-ISOMAP-DTW, which is especially obvious when the percentage of landmark points is less than 0.7%. With the error bar repeated 10 times for each experiment and the percentage of landmark points less than 0.7%, we find that the classification accuracy changes sharply. This shows that the DR results have a great impact on the classification accuracy. In other words, when the percentage of landmark points is less than 0.7%, it is difficult to obtain stable DR results. However, when the percentage of landmark points is greater than 0.7%, the overall classification accuracy tends to be stable, and the highest classification accuracy is almost the same for different percentages of landmark points. The range of the error bar for these percentages also becomes significantly smaller, which means that a stable DR result was obtained. Although the overall classification fluctuates as the proportion of training samples in the RF classifier changes, it is obvious that this fluctuation is not due to the results of DR, but from the classifier itself. Therefore, for the SITS of each scene with different paths/rows, 0.7% of the data points are selected by training samples learning as landmark points for TL-ISOMAP-DTW DR in this study.

Figure 7. Classification accuracy and error bar of landmark points with different proportions for both Landmark–Isometric feature mapping (L-ISOMAP) and Training Landmark–ISOMAP–Dynamic Time Warping (TL-ISOMAP-DTW) applied to 3512 verification points. Top row: landmark points with a percentage of 0.1%, 0.2%, or 0.3%. Middle row: landmark points with a percentage of 0.5%, 0.7%, or 1%. Bottom row: landmark points with a percentage of 2%, 3%, or 5%. The horizontal axis represents the proportion of training samples selected in random forest (RF) classifier.

Figure 7 also shows the overall classification accuracy of the L-ISOMAP method using different percentages of landmark points. When the percentage of landmark points is less than 0.5%, the results of L-ISOMAP and TL-ISOMAP-DTW are very close. The error bar, after 10 repeated experiments for both methods, varies greatly. However, when the percentage of landmark points is greater than 0.5%,

the TL-ISOMAP-DTW method shows a higher overall classification accuracy than the L-ISOMAP method. In addition, as an advanced classifier, RF can achieve higher classification accuracy with fewer training samples. From this experiment, we can see that when the proportion of training samples reaches 1%, the classification results of the RF classifier tend to stabilize. As the proportion of training samples increases, the classification accuracy does not increase significantly. In each of the SITS scenes with different paths/rows, the number of data points is significantly higher than the magnitude of that used in this experiment. It can be predicted that the proportion of training samples required by the classifier for each SITS scene to achieve stable accuracy will be further reduced. Therefore, considering the efficiency and accuracy of the algorithms in this study, the proportion of training samples required in RF classifier for the SITS of each scene is also set to 0.7%; that is, the training samples used by the RF classifier are the same as the landmark points used for TL-ISOMAP-DTW.

4.2. Overall Classification Performance

The distribution of maize, rice, and soybeans in the study area is shown in Figure 8. The classification validation based on sample points from field surveys and Google Earth indicate that the crop map of northeastern China in 2015 had high accuracy, with an overall accuracy of 84.83% (Table 3). The producer accuracy and user accuracy of the three main crops including maize, rice, and soybeans are significantly higher than those of other crops. The producer accuracy of rice is the highest, which is 0.51% and 5.69% higher than that of maize and soybeans, respectively. In contrast, the user accuracy of maize is the highest, which is 4.33% and 15.58% higher than that of rice and soybeans, respectively.

Figure 8. The distribution of the main crops cultivated area in northeastern China in 2015.

Table 3. Confusion matrix of classification validation based on sample points from field survey and Google Earth.

Overall Accuracy = 84.83%						
Class	Reference				Producer Accuracy (%)	User Accuracy (%)
	Maize	Rice	Soybean	Others		
Maize	1381	98	70	33	87.29	89.56
Rice	72	871	36	13	87.80	85.23
Soybean	51	24	381	8	82.11	73.98
Others	38	29	28	163	63.18	75.12

The specific crop area produced by this study is shown in Table 4. The area of each type = the number of pixels $\times$ the square of spatial resolution. The comparison and percentage difference between the classification results of this study and the 2015 statistical yearbook data is shown in Figure 9. The percentage difference of each type in each province = |the area difference| / the mean of corresponding type in the statistical yearbook and classification results. From Figure 9, we find that the classified area of rice in all three provinces is higher than that in the statistical yearbook data, especially in Liaoning Province, where the difference is up to 150.5 (1000 ha). Rice has the greatest difference between the total classified area in northeastern China and the statistical yearbook data among these three main crops. The classified area of maize in Heilongjiang Province is higher than the statistical yearbook data, while in Jilin Province and Liaoning Province, it is lower than the yearbook data, which makes the total area difference of maize in northeastern China much smaller than that of rice. The classified area of soybeans is the closest to the statistical yearbook among the three crops, which is mainly because the planting area of soybeans is small. Therefore, it is necessary for us to further obtain the percentage difference based on the absolute value of difference and mean between statistic data and classification results. As can be seen from Figure 9, the percentage difference of maize is less than 10% in the three provinces, and it is also less than 5% in Heilongjiang Province. The percentage difference of rice and soybeans is the highest in Liaoning Province, with both of them exceeding 20%. For each province, the lowest percentage difference appears on the crop with the largest planting area, and the highest percentage difference occurs on the crop with the smallest planting area.

Figure 9. The area and percentage difference of three crops in three provinces from official statistics and classification produced by this study (HLJ: Heilongjiang province, JL: Jilin province, LN: Liaoning province).

Table 4. The crop area from classification results in this study in northeastern China, 2015 (1000 ha).

Province	Maize	Rice	Soybean
Heilongjiang	5962.5	3241.7	2408.2
Jilin	3655.3	846.2	329.8
Liaoning	2300.4	695.4	162.1

4.3. DR Performance

To verify the proposed method more intuitively, we chose 100 field samples for each of the three crops, i.e., maize, rice, and soybeans to complete DR. We chose the ISOMAP and L-ISOMAP algorithms to compare with TL-ISOMAP-DTW. Since ISOMAP and L-ISOMAP cannot handle time-series data with unequal length, we need to interpolate time-series data first for these two algorithms. Landmark points for L-ISOMAP and TL-ISOMAP-DTW were selected both at random and by the samples learning method designed in this study, respectively. For simplicity, the target dimensionality of all the DR algorithm is two.

Figure 10a displays the original distribution for bands 82 (red band in July) and 118 (NIR band in August) from time-series data constructed in Section 3.1. Figure 10b–d shows the projected data distribution using different DR algorithms. It is overly obvious that the results of ISOMAP and L-ISOMAP are very poor according to the separation of three classes. ISOMAP is a global algorithm, and although the global structure of the data can be retained, the DR results do little to the follow-up classification, since the various classes are still overlapped. In addition, the complexity of the ISOMAP algorithm is significantly higher than that of the other two algorithms. One possible reason for the poor performance of L-ISOMAP is that the landmark points are randomly selected and are not well representative of different crop classes. The TL-ISOMAP-DTW algorithm separates the three classes quite nicely. Mainly, this is not only because the selection of landmark points is very representative, but also because the DTW metric has a good adaptability to the loss and distortion of time-series data.

Figure 10. Projected data produced by different dimensionality reduction (DR) algorithms. (a) Original (Bands 104 and 106 from TS data). (b) ISOMAP. (c) L-ISOMAP. (d) TL-ISOMAP-DTW.

4.4. Geographical Characteristics of Crop Distribution in Northeastern China

The maize planting area of Heilongjiang Province accounts for 45.7% of the total cultivated land area of the whole province, which was mainly distributed in Heihe City, Qiqihar City, Daqing City, Suihua City, and Harbin City in the southwest of Heilongjiang Province. For Hegang, Jiamusi, Mudanjiang, Qitaihe, and other eastern cities, there is also some maize planting. The rice planting area accounts for 24.8% of the total cultivated land area of the province, which is mainly concentrated in the east-central part of Heilongjiang Province, including Harbin, Jiamusi, and Qiqihar. Soybeans are mainly distributed in Mudanjiang, Suihua, Daqing, and other places. In addition, a number of other crops have been planted in the east and midwest.

Cultivated land in Jilin Province is mainly concentrated in the northwest of Jilin Province, including Chanchun City, Jilin City, Liaoyuan City, Siping City, and Baicheng City. These areas are rich in annual precipitation, sustained periods of sunshine, and fertile black land; thus, they become a very suitable zone for maize planting and growth. This area includes the southern part of Heilongjiang Province, the main cultivated land areas in Jilin Province, and the northern part of Liaoning Province. The Golden Maize Belt in Jilin Province includes Changchun City, Siping City, Songyuan City, Baicheng City, Liaoyuan City, and Jilin City. It encompasses a total of 6 cities and 22 counties, with a total area of more than 60,000 square kilometers, and the cultivated land area here accounts for about 75% of the total cultivated land area of Jilin Province. This provides a unique resource advantage and brings much wealth to Jilin Province.

The total area of cultivated land in Liaoning Province is the smallest among the three provinces in northeastern China; it is mainly concentrated in the central and northern part of Liaoning Province. The maize-planting area accounts for more than 50% of the total cultivated land in Liaoning Province, which is mainly distributed in Tieling, Fuxin, Chaoyang, Jinzhou, Panjin, Shenyang, and Benxi city. Rice is mainly distributed in Panjin, Anshan, and Fushun, and soybeans are only planted sporadically.

5. Discussion

5.1. Reliability of Crop Mapping in Northeastern China

This study proposed the automated and reliable yearly 30-m crop mapping method in a large scale area and the improved DR method based on the L-ISOMAP algorithm and DTW metric for Landsat 8 time-series images. To our knowledge, this study was the first investigation to map the three main crops at 30-m spatial resolution in northeastern China. The implementation of this study can be attributed to the following factors: improved TS data, samples learning method, modified DR algorithm, and simple natural conditions.

First, the enhanced observation capability of Landsat 8 provides observation with high-temporal frequency, which has much higher advantages in terms of data availability than Landsat 5 and 7 in both the yearly scale and the transplanting phase [59,70]. Landsat 8 TS data contain abundant phenological information of crop growth. Although the average missing (cloud/snow cover) data accounted for 33.73% of the total Landsat 8 TS images used in this study during the crop-growing period in 2015 (Figure 2), we used all available observations of individual pixels to heighten the representation of phenological information [71], instead of choosing scenes with less cloud/shadow. In addition, we rearranged the Landsat 8 TS data to obtain information about the variation of each spectral band over the whole crop growth, so that the phenological characteristics of crops can be used more effectively and directly.

Second, in remote sensing, the method of labeling samples is mainly achieved by in situ surveys or image photointerpretation [72]. However, in the crop classification on a large regional scale, it is difficult to obtain a large number of samples, which leads to a large amount of human effort and time. Therefore, many studies adopted the samples learning method to label new training samples for supervised algorithms [73–75]. However, there is often missing data in TS images, and the performance is poor by directly using the common samples learning methods based on the uncertainty and the diversity

criteria [76]. This study demonstrates that samples learning based on a spectral–temporal curve database extracted from SITS can label effective samples for crop classification in a large-scale area.

Third, data redundancy exists in time-series data, and DR is often required before classification. The nonlinear DR algorithms have been successfully applied to time-series data [54,77]. However, because the complexity of the nonlinear DR algorithm is quite high, there is no research to apply the nonlinear DR algorithm to the time-series images of a large area. This study proved that the TL-ISOMAP-DTW method is reliable for a large area in northeastern China, which is helpful for applying the nonlinear DR algorithm to the time series images of a large area.

Fourth, compared with tropical regions, the simpler natural conditions in northeastern China are another important factor in the success of this study. Northeastern China is different from the natural conditions in tropical regions: there is lower planting density (generally single cropping), a longer crop growing season, and less cloud coverage in the study area. All of these variables provide opportunities that can aid in the completion of this study.

5.2. Uncertainty Analysis and Implications for Other Regions

This paper presents an automatic crop classification method that uses all available data and is not reliant on human intervention. The impact of cloud coverage on the accuracy of classification is difficult to quantify when using time-series data. Since the time-series data includes different regions, years, and frequency, the proportion of cloud coverage is quite different. It is difficult to show the complexity of cloud cover for satellite image time series, regardless of the total or average cloud cover of time-series data. This complexity is more evident in large-scale area studies, which include more time-series data with different paths and rows. In this study, the cloud cover of satellite image time series with different paths and rows varies greatly, and this brings some uncertainty to this method. In fact, this is also a problem that most image classification using satellite image time series will face [25]. This paper uses a series of robust strategies: using a DTW metric to improve the nonlinear dimensionality reduction algorithm, and building a spectral–temporal curve database and learning training samples to limit the uncertainty caused by cloud cover, to a certain extent.

Furthermore, due to the low frequency of observations, this method is not recommended for satellite image time series covering only the beginning of the crop-growing season. It has been pointed out that Landsat 7 data are insufficient for annual crop mapping due to the limited observation capabilities and SLC (Scan Line Corrector)-off gaps [71]. Consequently, many studies have focused on improving the temporal resolution of satellite image time series through data synthesis, which sheds light on classification methods that use temporal features because it can improve the availability of effective observations [78,79]. This study demonstrates the feasibility of mapping crop distribution using Landsat 8 data during the growing season in northeastern China. However, if this method is used in tropical areas with higher cloud cover, such as southeast China, where planting is more dispersed and intensive, there may be more challenges. With the implementation of the sharing policy for earth observation by many countries, more and more satellite image time series are available free of charge, such as Sentinel-2 of ESA (European Space Agency), whose temporal resolution has been improved to 5 days, and HLS (Harmonized Landsat Sentinel-2) data synthesized using Sentinel-2 and Landsat 8 data, which has a temporal resolution of 2–3 days [80,81]. Therefore, with the enrichment of time series, the wide application of the method proposed in this paper for other areas can be predicted and is encouraged.

6. Conclusions

In this study, we propose a new automated method based on Landsat 8 OLI time-series data to map the spatial distribution of three main crops including maize, rice, and soybeans in northeastern China. The most impressive feature of the automated method is that it reduces the need for reference data and expert knowledge in model training, thereby significantly lower the cost of crop mapping. This is especially important for high-resolution mapping at regional or larger scales. Considering the

data characteristics of SITS, a simple and robust nonlinear DR algorithm was improved, which not only automatically handles the missing data in time series but can provide meaningful feature metrics for classification. The final crop map shows an overall classification accuracy of 83.68%. Compared with official statistics, the planting areas of three crops in the classification map show good consistency. Since only single-sensor data are used, this may limit the accuracy of the mapping. It is recommended to apply the proposed method to higher temporal resolution data or multi-temporal hyperspectral data, which will lead to more accurate crop distribution and finer crop types. The successful application of the proposed method in the cropland mapping of maize, rice, and soybean in northeastern China indicates that the automated method has more potential applications.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z.; validation, N.W., and C.H.; formal analysis, L.H.; writing—original draft preparation, Y.Z.; writing—review and editing, N.W. and L.Z.; funding acquisition, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Open Fund of State Key Laboratory of Remote Sensing Science under grant number OFSLRSS201916 and Inner Mongolia Natural Science Foundation under grant number 2018BS04001.

Acknowledgments: The authors would like to thank Bart Widmer Leu for his support and assistance on English editing of this paper.

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

References

1. Kang, Y.; Khan, S.; Ma, X. Climate change impacts on crop yield, crop water productivity and food security—A review. *Prog. Nat. Sci.* **2009**, *19*, 1665–1674. [\[CrossRef\]](#)
2. Ramankutty, N.; Mehrabi, Z.; Waha, K.; Jarvis, L.; Kremen, C.; Herrero, M.; Rieseberg, L.H. Trends in Global Agricultural Land Use: Implications for Environmental Health and Food Security. *Annu. Rev. Plant Biol.* **2018**, *69*, 789–815. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Mingwei, Z.; Zhou, Q.; Chen, Z.; Jia, L.; Yong, Z.; Chongfa, C. Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 476–485. [\[CrossRef\]](#)
4. Immitzer, M.; Atzberger, C.; Koukal, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sens.* **2012**, *4*, 2661–2693. [\[CrossRef\]](#)
5. Immitzer, M.; Vuolo, F.; Atzberger, C. First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sens.* **2016**, *8*, 166. [\[CrossRef\]](#)
6. Atzberger, C. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sens.* **2013**, *5*, 949–981. [\[CrossRef\]](#)
7. Esch, T.; Metz, A.; Marconcini, M.; Keil, M. Combined use of multi-seasonal high and medium resolution satellite imagery for parcel-related mapping of cropland and grassland. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 230–237. [\[CrossRef\]](#)
8. Peña-Barragan, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316. [\[CrossRef\]](#)
9. Murthy, C.S.; Raju, P.V.; Badrinath, K.V.S. Classification of wheat crop with multi-temporal images: Performance of maximum likelihood and artificial neural networks. *Int. J. Remote Sens.* **2003**, *24*, 4871–4890. [\[CrossRef\]](#)
10. Shelestov, A.; Lavreniuk, M.; Kussul, N.; Novikov, A.; Skakun, S. Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. *Front. Earth Sci.* **2017**, *5*, 17. [\[CrossRef\]](#)
11. Yan, L.; Roy, D. Conterminous United States crop field size quantification from multi-temporal Landsat data. *Remote Sens. Environ.* **2016**, *172*, 67–86. [\[CrossRef\]](#)
12. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [\[CrossRef\]](#)
13. Bargiel, D. A new method for crop classification combining time series of radar images and crop phenology information. *Remote Sens. Environ.* **2017**, *198*, 369–383. [\[CrossRef\]](#)

14. Khan, A.; Hansen, M.; Potapov, P.V.; Stehman, S.V.; Chatta, A.A. Landsat-based wheat mapping in the heterogeneous cropping system of Punjab, Pakistan. *Int. J. Remote Sens.* **2016**, *37*, 1391–1410. [\[CrossRef\]](#)
15. Song, X.-P.; Potapov, P.V.; Krylov, A.; King, L.; Di Bella, C.M.; Hudson, A.; Khan, A.; Adusei, B.; Stehman, S.V.; Hansen, M. National-scale soybean mapping and area estimation in the United States using medium resolution satellite imagery and field survey. *Remote Sens. Environ.* **2017**, *190*, 383–395. [\[CrossRef\]](#)
16. Khan, A.; Hansen, M.; Potapov, P.; Adusei, B.; Pickens, A.; Krylov, A.; Stehman, S. Evaluating Landsat and RapidEye Data for Winter Wheat Mapping and Area Estimation in Punjab, Pakistan. *Remote Sens.* **2018**, *10*, 489. [\[CrossRef\]](#)
17. Inglada, J.; Arias, M.; Tardy, B.; Hagolle, O.; Valero, S.; Morin, D.; Dedieu, G.; Sepulcre, G.; Bontemps, S.; Defourny, P.; et al. Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.* **2015**, *7*, 12356–12379. [\[CrossRef\]](#)
18. Inglada, J.; Vincent, A.; Arias, M.; Tardy, B.; Morin, D.; Rodes, I. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sens.* **2017**, *9*, 95. [\[CrossRef\]](#)
19. Dempewolf, J.; Adusei, B.; Becker-Reshef, I.; Hansen, M.; Potapov, P.; Khan, A.; Barker, B. Wheat Yield Forecasting for Punjab Province from Vegetation Index Time Series and Historic Crop Statistics. *Remote Sens.* **2014**, *6*, 9653–9675. [\[CrossRef\]](#)
20. Chang, J.; Hansen, M.C.; Pittman, K.; Carroll, M.; DiMiceli, C. Corn and Soybean Mapping in the United States Using MODIS Time-Series Data Sets. *Agron. J.* **2007**, *99*, 1654–1664. [\[CrossRef\]](#)
21. Friedl, M.; Brodley, C.; Strahler, A. Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 969–977. [\[CrossRef\]](#)
22. Yang, C.-C.; Prasher, S.O.; Enright, P.; Madramootoo, C.; Burgess, M.; Goel, P.K.; Callum, I. Application of decision tree technology for image classification using remote sensing data. *Agric. Syst.* **2003**, *76*, 1101–1117. [\[CrossRef\]](#)
23. Kandrika, S.; Roy, P. Land use land cover classification of Orissa using multi-temporal IRS-P6 awifs data: A decision tree approach. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 186–193. [\[CrossRef\]](#)
24. Gao, B.-C. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [\[CrossRef\]](#)
25. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *116*, 55–72. [\[CrossRef\]](#)
26. Zhang, X.; Xiong, Q.; Di, L.; Tang, J.; Yang, J.; Wu, H.; Qin, Y.; Su, R.; Zhou, W. Phenological metrics-based crop classification using HJ-1 CCD images and Landsat 8 imagery. *Int. J. Digit. Earth* **2017**, *11*, 1–22. [\[CrossRef\]](#)
27. Al-Shammari, D.; Bishop, T.F.; Fuentes, I.; Filippi, P. Mapping cotton fields using phenology-based metrics derived from a time series of Landsat imagery. In Proceedings of the 2019 Agronomy Australia Conference, Wagga Wagga, Australia, 25–29 August 2019.
28. Zeng, L.; Wardlow, B.; Xiang, D.; Hu, S.; Li, D. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ.* **2020**, *237*, 111511. [\[CrossRef\]](#)
29. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [\[CrossRef\]](#)
30. Zhong, L.; Hu, L.; Yu, L.; Gong, P.; Biging, G.S. Automated mapping of soybean and corn using phenology. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 151–164. [\[CrossRef\]](#)
31. Zhong, L.; Hu, L.; Zhou, H. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* **2019**, *221*, 430–443. [\[CrossRef\]](#)
32. Cao, R.; Chen, Y.; Shen, M.; Chen, J.; Zhou, J.; Wang, C.; Yang, W. A simple method to improve the quality of NDVI time-series data by integrating spatiotemporal information with the Savitzky-Golay filter. *Remote Sens. Environ.* **2018**, *217*, 244–257. [\[CrossRef\]](#)
33. Badhwar, G. Classification of corn and soybeans using multitemporal thematic mapper data. *Remote Sens. Environ.* **1984**, *16*, 175–181. [\[CrossRef\]](#)
34. Vicente-Guijalba, F.; Martinez-Marin, T.; Lopez-Sanchez, J.M. Crop Phenology Estimation Using a Multitemporal Model and a Kalman Filtering Strategy. *IEEE Geosci. Remote Sens. Lett.* **2013**, *11*, 1081–1085. [\[CrossRef\]](#)
35. Xu, C.; Katchova, A.L. Predicting Soybean Yield with NDVI Using a Flexible Fourier Transform Model. *J. Agric. Appl. Econ.* **2019**, *51*, 402–416. [\[CrossRef\]](#)

36. Evans, J.P.; Geerken, R. Classifying rangeland vegetation type and coverage using a Fourier component based similarity measure. *Remote Sens. Environ.* **2006**, *105*, 1–8. [\[CrossRef\]](#)
37. Wang, L.; Song, W.; Liu, P. Link the remote sensing big data to the image features via wavelet transformation. *Clust. Comput.* **2016**, *19*, 793–810. [\[CrossRef\]](#)
38. Sakamoto, T.; Van Nguyen, N.; Ohno, H.; Ishitsuka, N.; Yokozawa, M. Spatio-temporal distribution of rice phenology and cropping systems in the Mekong Delta with special reference to the seasonal water flow of the Mekong and Bassac rivers. *Remote Sens. Environ.* **2006**, *100*, 1–16. [\[CrossRef\]](#)
39. Vorobiova, N.; Chernov, A. Curve fitting of MODIS NDVI time series in the task of early crops identification by satellite images. *Procedia Eng.* **2017**, *201*, 184–195. [\[CrossRef\]](#)
40. Bradley, B.A.; Jacob, R.W.; Hermance, J.F.; Mustard, J.F. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sens. Environ.* **2007**, *106*, 137–145. [\[CrossRef\]](#)
41. Johnson, M.; Caragea, P.C.; Meiring, W.; Jeganathan, C.; Atkinson, P.M. Bayesian Dynamic Linear Models for Estimation of Phenological Events from Remote Sensing Data. *J. Agric. Biol. Environ. Stat.* **2018**, *24*, 1–25. [\[CrossRef\]](#)
42. Funk, C.; Budde, M.E. Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sens. Environ.* **2009**, *113*, 115–125. [\[CrossRef\]](#)
43. Gong, W.; Fang, S.; Yang, Z.; Ge, M. Using a Hidden Markov Model for Improving the Spatial-Temporal Consistency of Time Series Land Cover Classification. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 292. [\[CrossRef\]](#)
44. Siachalou, S.; Mallinis, G.; Tsakiri-Strati, M. A Hidden Markov Models Approach for Crop Classification: Linking Crop Phenology to Time Series of Multi-Sensor Remote Sensing Data. *Remote Sens.* **2015**, *7*, 3633–3650. [\[CrossRef\]](#)
45. Jin, X.; Kumar, L.; Li, Z.; Feng, H.; Xu, X.; Yang, G.; Wang, J. A review of data assimilation of remote sensing and crop models. *Eur. J. Agron.* **2018**, *92*, 141–152. [\[CrossRef\]](#)
46. Peña, M.; Brenning, A. Assessing fruit-tree crop classification from Landsat-8 time series for the Maipo Valley, Chile. *Remote Sens. Environ.* **2015**, *171*, 234–244. [\[CrossRef\]](#)
47. Müller, H.; Rufin, P.; Griffiths, P.; Siqueira, A.J.B.; Hostert, P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. *Remote Sens. Environ.* **2015**, *156*, 490–499. [\[CrossRef\]](#)
48. Ghulam, A.; Li, Z.-L.; Qin, Q.; Yimit, H.; Wang, J. Estimating crop water stress with ETM+ NIR and SWIR data. *Agric. For. Meteorol.* **2008**, *148*, 1679–1695. [\[CrossRef\]](#)
49. Lobell, D.B.; Asner, G.P. Moisture Effects on Soil Reflectance. *Soil Sci. Soc. Am. J.* **2002**, *66*, 722. [\[CrossRef\]](#)
50. Peña, M.; Liao, R.; Brenning, A. Using spectrottemporal indices to improve the fruit-tree crop classification accuracy. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 158–169. [\[CrossRef\]](#)
51. Ghamisi, P.; Plaza, J.; Chen, Y.; Li, J.; Plaza, J. Advanced Spectral Classifiers for Hyperspectral Images: A review. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–32. [\[CrossRef\]](#)
52. Khan, M.J.; Khan, H.S.; Yousaf, A.; Khurshid, K.; Abbas, A. Modern Trends in Hyperspectral Image Analysis: A Review. *IEEE Access* **2018**, *6*, 14118–14129. [\[CrossRef\]](#)
53. Nitze, I.; Barrett, B.; Cawkwell, F. Temporal optimisation of image acquisition for land cover classification with Random Forest and MODIS time-series. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *34*, 136–146. [\[CrossRef\]](#)
54. Zhai, Y.; Qu, Z.; Hao, L. Land Cover Classification Using Integrated Spectral, Temporal, and Spatial Features Derived from Remotely Sensed Images. *Remote Sens.* **2018**, *10*, 383. [\[CrossRef\]](#)
55. Keogh, E.; Chakrabarti, K.; Pazzani, M.J.; Mehrotra, S. Dimensionality Reduction for Fast Similarity Search in Large Time Series Databases. *Knowl. Inf. Syst.* **2001**, *3*, 263–286. [\[CrossRef\]](#)
56. Shen, H.T.; Risch, T.; Canli, T.; Khokhar, A.; Yang, J.; Munagala, K.; Silberstein, A.; Chrysanthi, P.K.; Pitoura, E.; Ganti, V.; et al. Dimensionality Reduction. *Encycl. Database Syst.* **2009**, *10*, 843–846. [\[CrossRef\]](#)
57. Somers, B.; Cools, K.; Delalieux, S.; Stuckens, J.; Van Der Zande, D.; Verstraeten, W.W.; Coppin, P. Nonlinear Hyperspectral Mixture Analysis for tree cover estimates in orchards. *Remote Sens. Environ.* **2009**, *113*, 1183–1193. [\[CrossRef\]](#)
58. Shen, H.; Li, X.; Cheng, Q.; Zeng, C.; Yang, G.; Li, H.; Zhang, L. Missing Information Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geosci. Remote Sens. Mag.* **2015**, *3*, 61–85. [\[CrossRef\]](#)

59. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [\[CrossRef\]](#)
60. Masek, J.G.; Vermote, E.F.; Saleous, N.E.; Wolfe, R.; Hall, F.G.; Huemmrich, K.F.; Feng, G.; Kutler, J.; Teng-Kui, L. A Landsat Surface Reflectance Dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72. [\[CrossRef\]](#)
61. Zhang, Z.; Wang, X.; Zhao, X.; Liu, B.; Yi, L.; Zuo, L.; Wen, Q.; Liu, F.; Xu, J.; Hu, S. A 2010 update of National Land Use/Cover Database of China at 1:100000 scale using medium spatial resolution satellite images. *Remote Sens. Environ.* **2014**, *149*, 142–154. [\[CrossRef\]](#)
62. Olofsson, P.; Foody, G.; Herold, M.; Stehman, S.; Woodcock, C.; Wuldere, M. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **2014**, *148*, 42–57. [\[CrossRef\]](#)
63. Olofsson, P.; Arévalo, P.; Espejo, A.; Green, C.; Lindquist, E.; McRoberts, R.; Sanz, M. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.* **2020**, *236*, 111492. [\[CrossRef\]](#)
64. Silva, V.D.; Tenenbaum, J.B. Global versus local methods in nonlinear dimensionality reduction. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 8–13 December 2003; pp. 721–728.
65. Petitjean, F.; Inglada, J.; Gançarski, P. Satellite Image Time Series Analysis Under Time Warping. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3081–3095. [\[CrossRef\]](#)
66. Belgiu, M.; Csillik, O. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sens. Environ.* **2018**, *204*, 509–523. [\[CrossRef\]](#)
67. Zhai, Y.; Zhang, L.; Wang, N.; Guo, Y.; Cen, Y.; Wu, T.; Tong, Q. A Modified Locality-Preserving Projection Approach for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1059–1063. [\[CrossRef\]](#)
68. Breiman, L.; Last, M.; Rice, J. Random Forests: Finding Quasars. *Stat. Chall. Astron.* **2006**, *45*, 243–254. [\[CrossRef\]](#)
69. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [\[CrossRef\]](#)
70. Dong, J.; Xiao, X.; Menarguez, M.A.; Zhang, G.; Qin, Y.; Thau, D.; Biradar, C.; Moore, B. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sens. Environ.* **2016**, *185*, 142–154. [\[CrossRef\]](#)
71. Dong, J.W.; Xiao, X.M.; Kou, W.; Qin, Y.W.; Zhang, G.; Li, L.; Jin, C.; Zhou, Y.T.; Wang, J.; Biradar, C.; et al. Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. *Remote Sens. Environ.* **2015**, *160*, 99–113. [\[CrossRef\]](#)
72. Cheng, G.; Han, J.; Lu, X. Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proc. IEEE* **2017**, *105*, 1865–1883. [\[CrossRef\]](#)
73. Haut, J.M.; Paoletti, M.E.; Plaza, J.; Li, J.; Plaza, J. Active Learning With Convolutional Neural Networks for Hyperspectral Image Classification Using a New Bayesian Approach. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6440–6461. [\[CrossRef\]](#)
74. Ji, S.; Zhang, C.; Xu, A.; Shi, Y.; Duan, Y. 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images. *Remote Sens.* **2018**, *10*, 75. [\[CrossRef\]](#)
75. Shi, Q.; Du, B.; Zhang, L. Spatial Coherence-Based Batch-Mode Active Learning for Remote Sensing Image Classification. *IEEE Trans. Image Process.* **2015**, *24*, 2037–2050. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Demir, B.; Minello, L.; Bruzzone, L. An Effective Strategy to Reduce the Labeling Cost in the Definition of Training Sets by Active Learning. *IEEE Geosci. Remote Sens. Lett.* **2013**, *11*, 79–83. [\[CrossRef\]](#)
77. Yan, L.; Roy, D.P. Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction. *Remote Sens. Environ.* **2015**, *158*, 478–491. [\[CrossRef\]](#)
78. Zhu, Z.; Woodcock, C.E.; Holden, C.; Yang, Z. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. *Remote Sens. Environ.* **2015**, *162*, 67–83. [\[CrossRef\]](#)
79. Zhou, F.; Zhong, D. Kalman filter method for generating time-series synthetic Landsat images and their uncertainty from Landsat and MODIS observations. *Remote Sens. Environ.* **2020**, *239*, 111628. [\[CrossRef\]](#)

80. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascón, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; et al. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
81. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [[CrossRef](#)]

© 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/>).