

Article

Large-Scale and High-Resolution Crop Mapping in China Using Sentinel-2 Satellite Imagery

Yulin Jiang ¹, Zhou Lu ², Shuo Li ¹, Yongdeng Lei ¹, Qingquan Chu ¹, Xiaogang Yin ¹
and Fu Chen ^{1,*}

¹ College of Agronomy and Biotechnology, China Agricultural University, Key Laboratory of Farming System, Ministry of Agriculture and Rural Affairs, Beijing 100193, China; jiangyulin93@163.com (Y.J.); lis@cau.edu.cn (S.L.); leiyongdeng56@163.com (Y.L.); cauchu@cau.edu.cn (Q.C.); xiaogangyin@cau.edu.cn (X.Y.)

² Institute of Geographic Sciences and National Resources Research, Chinese Academy of Sciences, Beijing 100193, China; luz@lreis.ac.cn

* Correspondence: chenfu@cau.edu.cn

Received: 24 August 2020; Accepted: 25 September 2020; Published: 26 September 2020



Abstract: Large-scale, high-resolution mapping of crop patterns is useful for the assessment of food security and agricultural sustainability but is still limited. This study attempted to establish remote sensing-based crop classification models for specific cropping systems using the decision trees method and monitored the distribution of the major crop species using Sentinel-2 satellites (10 m) in 2017. The results showed that the cropping areas of maize, rice, and soybean on the Northeast China Plain were approximately 12.1, 6.2, and 7.4 million ha, respectively. The cropping areas of winter wheat and summer maize on the North China Plain were 13.4 and 16.9 million ha, respectively. The cropping areas of wheat, rice, and rape on the middle-lower Yangtze River plain were 2.2, 6.4 and 1.3 million ha, respectively. Estimated images agreed well with field survey data (average overall accuracy = 94%) and the national agricultural census data ($R^2 = 0.78$). This indicated the applicability of the Sentinel-2 satellite data for large-scale, high-resolution crop mapping in China. We intend to update the crop mapping datasets annually and hope to guide the adjustment and optimization of the national agricultural structure.

Keywords: crop pattern; multiple indicator; classification model; cropping system; Sentinel-2

1. Introduction

Ensuring global food security is one of the greatest challenges for scientists around the world [1]. As the country with the largest population in the world, China's food security is not only important to economic development and social stability but also to global food patterns [2,3]. China has enacted a series of measures to ensure food security during recent decades, including food self-sufficiency policies and environmentally friendly agricultural development strategies [4–6]. These policies limit the scale of water, pesticide, fertilizer, and arable land use, which increases the difficulty of new round adjustment strategies for agricultural structure [7–9]. Updated information on crop pattern mapping can provide significant scientific evidence for estimating agricultural production and food security. Mapping major crop patterns in China has great implications for policymaking in agricultural sustainability [10,11]. Despite the development of global land cover products, the specific information on crop types at large scales is still limited [12]. Additional studies should be performed to enrich crop mapping data to assist in the optimization of agricultural distribution.

Earth observation satellites can be effective tools for land use mapping and crop classification due to their ability to quickly and efficiently collect information in the field [13,14]. The mapping

methods have been rapidly developed over the past 20 years [15,16]. From the methods based on single imagery [17,18] to time-series imagery [19–21], satellite mapping can fully capture temporal features. Comparative studies have proven that the methods that use multiple features are more suitable for regional mapping of areas with complicated planting structures [22]. Furthermore, the combination of remote sensing data and agricultural statistical data can realize large-scale crop pattern mapping with a relatively coarse resolution [23,24]. Future crop mapping aims to create “a map of crops” that covers more crop types and enlarged mapping areas [15,25]. On the other hand, many studies have proven that it is difficult to extract each crop type across large areas simply by applying the temporal profiles of the vegetation indices [12,26]. The combination of the multiple spectral indicators and unique signals of the target crops at specific times is necessary for large-scale, high-resolution crop pattern mapping.

At present, moderate-resolution imaging spectroradiometer (MODIS) data are widely used in crop mapping because of their high revisit frequency. However, the coarse spatial resolution (250 m–1 km) of the data limits their suitability for monitoring patchy and fragmented cropland, especially in the middle and south of China [27,28]. Landsat imagery with a higher spatial resolution (30 m) has also been widely used in recent years. However, it is difficult to collect enough cloud-free images for time-series analysis because of the low temporal resolution (16 days) of the data [29]. The GF-1 satellite can provide high-spatial-resolution (16 m) data and dense time-series (4 days) to cover the key growth phases of different crops. However, the limited spectral bands of the GF-1 data cannot provide enough information for multiple classification indices because different crops in the same region may share similar spectral characteristics [30]. Since 2017, with the combination of the Sentinel-2A and 2B satellites, the ESA has provided high-resolution (10 m) images with worldwide coverage and a 5-day revisit interval. These data make it possible to significantly advance routine global monitoring of crop patterns at no cost, especially in the field of vegetation research, through enhanced definition in the red-edge spectral reflectance [31–33]. There have been some successful applications in crop classification using Sentinel-2 in recent years. For example, Immitzer et al. used Sentinel-2 images to map six summer crop species in Lower Austria and to map seven different deciduous and coniferous tree species in Germany [34]. Hegarty-Craver et al. used imagery acquired from unmanned aerial vehicles to create a high-fidelity ground-truth dataset, and then implemented the dataset for classifying cropped land using Sentinel-1 and Sentinel-2 data [35]. Zhu et al. developed a method for tea plantation detection and mapped the tea planting area in the Henan Province in China using Sentinel-2 images [36].

Despite the progress of the satellite and monitoring methods, little emphasis has been placed on large-scale crop mapping. Most efforts have focused on the classification of land cover types, such as farmland, forest, and grassland. Few large-scale mapping studies have attempted to map specific crop types and the associated land use practices at the national scale, especially for the rapid land cover changes that commonly occur in annual crop rotations [27]. Since 1997, the National Agricultural Statistics Services (NASS) of the US Department of Agriculture (USDA) has produced the annual cropland data layer (CDL), which is a raster-format, georeferenced, and crop-specific land cover map [37]. Wardlow and Egbert used time-series MODIS 250 m data for large-area crop-related LULC mapping over the US Central Great Plains [27]. However, large-area crop monitoring is difficult in cropland with multiple cropping systems or small plot sizes [10,13]. Extreme variation of cropping system occurred in China from the north to the south. The single-cropping system was widely adopted in the Northeast Plain, while the triple-cropping system was adopted in the south of China. The degree of fragmentation in agricultural fields is also totally different. Many crop pattern mapping studies in China occurred in recent years [15]. These studies tried to perform large-scale crop mapping using moderate resolution images [12] or perform high-resolution crop mapping in the typical area using different satellite or monitoring methods [11,13]. However, few studies have focused on both high-resolution and large-scale crop pattern mapping.

The objective of this study was to investigate the applicability of the Sentinel-2 satellite data for large-scale, high-resolution crop mapping in the main grain-producing areas of China, to build different remote sensing-based models for crop classification in typical cropping systems that could be applied

to large areas, to map the distribution of the major crops in the main grain-producing regions from 2017 to 2018, and to evaluate the classification accuracy of the crop mapping. The primary research question in this study was as follows: What thematic accuracy can be achieved for large-scale and high-resolution crop pattern classification in different regions with the use of Sentinel-2 images?

2. Materials and Methods

2.1. Study Area

Considering the first-order agroecological zones in China, which were defined based on the climatic, soil, and landform characteristics [38], we chose the Northeast Plain (NEP), North China Plain (NCP), and middle-lower Yangtze River plain (MYRP) as the study area (Figure 1). This region, which has an area of approximately 2 million km², is the main grain-producing area in China, covers almost 60% of the total farmland in China, and produces more than 60% of the national grain yield. Considering the terrain complexity on the MYRP, we chose the plain area as the study area but monitored the entirety of the NEP and NCP. The cropping system in each zone was totally different. The single-cropping system was widely adopted in the NEP, while the double-cropping system was widely adopted in the NCP. Both the double-cropping system and the triple-cropping system existed in the MYRP.

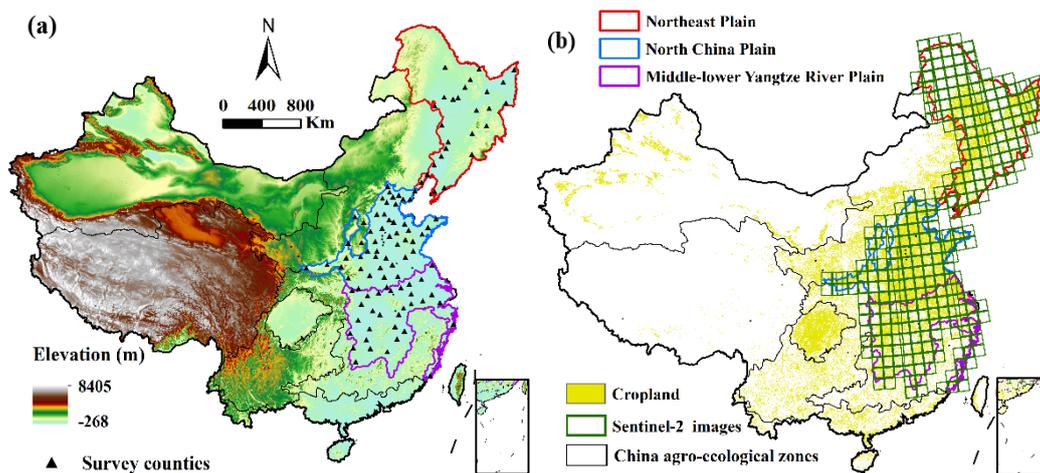


Figure 1. The spatial distribution of the elevation and survey sites (a), and the cropland and Sentinel-2 images (b) in the main grain-producing area of China, including the Northeast Plain (red frame), North China Plain (blue frame), and middle-lower Yangtze River plain (purple frame). The rectangle in the right-bottom corner represents China's Nansha Islands and the nine-dash line.

2.2. Datasets

In this study, we collected a total of 6,266 Sentinel-2 images (almost 2.3 TB of computational storage) from February to October of 2017 and 2018, which covered the whole study area and the key phenological periods of the target crop species (Figure 1). All images were collected from the open-source data of the European Space Agency (<http://scihub.copernicus.eu>) and we set a 20% standard to control the cloud cover. The Sentinel-2 sensor provides 13 bands at a 10-m resolution and contains three bands at the red edge, which can better classify different types of vegetation. All data were acquired under low cloud conditions and were georeferenced to the UTM-WGS84 projection system. In total, we chose more than 100 actual sampling points in the study area (Figure 1a). These sampling points were selected by local agricultural experiment stations and agricultural observation stations. The training dataset was used to determine the general decision trees and to adjust the thresholds in subregions. We also established another dataset for accuracy evaluation. We randomly chose a total of 113 test counties and 34,995 sample points within the study area under the principle of uniform spatial distribution.

We also used the 90-m resolution SRTM DEM and SRTM SLOPE data, which were obtained from the Geospatial Data Cloud (<http://www.gscloud.cn>), as the terrain layer to assist the crop mapping research. We used 91Weitu software to choose 0.5-m resolution Google Map temporal images from February to October of 2017 and 2018 during crop-specific growing periods (crop jointing stage or mature stage), which were used as the data source for the accuracy test. We chose national statistical data obtained from the National Statistical Bureau of China (NSBC) (<http://www.stats.gov.cn>) to establish the target crop species and to evaluate the classification accuracy.

2.3. Methodology

We chose multiple indicators for the crop mapping and built remote sensing-based models for crop classification in the different cropping systems. The general technical process included the following procedures (Figure 2): The selection of the target crop species, data collection and preprocessing, design of the phenology-based indicators, crop area mapping, and accuracy evaluation.

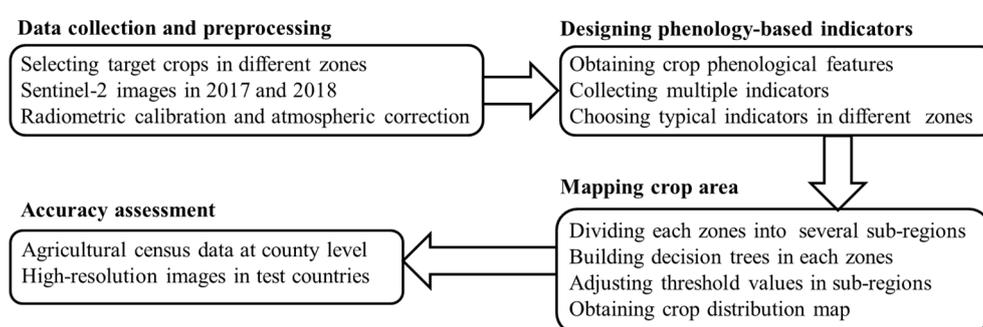


Figure 2. Overview of the methodology in the crop pattern mapping research.

2.3.1. Selecting Target Crop Species

The cropping systems in the different zones were totally different. We used national statistical data from 2015 to calculate the ratio of the areas of the major crops that were sown in the different regions (Table 1). We calculated nine major kinds of crop area (rice, maize, wheat, soybean, rape, peanut, cotton, potato, and coarse cereals) and selected targeted crops in each study area. On the Northeast Plain, a single-cropping system is used due to the limited temperatures and rainfall resources. Maize, soybean, and paddy rice were the main crops in this area, accounting for 98.11% of the total planting area. On the North China Plain, most farms chose a winter wheat-summer maize double-cropping system. The area of the wheat and maize can cover 90.26% of the farmland in this zone. On the middle-lower Yangtze River plain, both double-cropping systems and multiple cropping systems exist because of the sufficient climatic resources. Paddy rice, wheat and rape were major crops in this zone, accounting for 93.03% of the total planting area. Thus, we chose maize, soybean, and paddy rice as the target crop species on the NEP; wheat and maize as target crop species on the NCP; and paddy rice, wheat, and rape as target crop species on the MYRP.

Table 1. Percentage of areas of target crop species sown in different zones in 2015.

%	Maize	Soybean	Paddy Rice	Wheat	Rape	Total
Northeast Plain	64.95	14.43	18.73	-	-	98.11
North China Plain	40.63	-	-	49.63	-	90.26
Middle-lower Yangtze River plain	-	-	61.15	15.85	16.03	93.03

2.3.2. Designing Phenology-Based Indicators

A combination of multiple spectral indicators and unique signals of targeted crops at specific times was necessary for large-scale high-resolution crop pattern mapping. Many indicators have

been established in the application of agricultural remote sensing [15]. Considering the different cropping systems and the mixed crop density in different regions, we chose a total of seven common indicators and designed a new indicator to better classify the crop patterns by investigating the sampling points. The normalized difference vegetation index (NDVI) is widely used in to monitor vegetation coverage [39]. The NDVI time-series (Δ NDVI) for the crop growing period can be used in the classification of crops and other land cover types. The normalized difference water index (NDWI) and normalized difference building index (NDBI) can distinguish between water and buildings in remote sensing images [40]. The land surface water index (LSWI) more accurately classify different crops, especially paddy rice [41]. Other indicators used in our research included the red-edge parameter (REP) and the blue-green (B-G) ratio [42,43]. We also designed a new indicator called the normalized difference rice index (NDRI), which can better identify the paddy rice on the Northeast Plain. The calculation formula of NDRI was as follows:

$$NDRI = \frac{\rho_{swir} - \rho_{red}}{\rho_{swir} + \rho_{red}} \quad (1)$$

where ρ_{swir} and ρ_{red} represent the shortwave infrared band and red band, respectively. The indicators used in the different regions are shown in Table 2.

Table 2. Different indicators used for crop classification in the different study regions.

Region	Crops	Indicators
Northeast Plain (NEP)	Maize, Soybean, Rice	NDVI, REP, NDRI
North China Plain (NCP)	Maize, Wheat	NDVI, Δ NDVI
Middle-lower Yangtze River plain (MYRP)	Wheat, Rice, Rape	NDVI, Δ NDVI, LDSI, NDBI, NDWI, B-G

2.3.3. Mapping Crop Area

The choice of classification algorithm is usually based upon a series of factors, such as the availability of software, ease of use, performance, and expected overall accuracy. The decision trees (DTs) method has been widely used in land cover classification during past years, considered the advantages of being computationally fast, making no statistical assumptions, and the ability to handle data from different measurement scales [15,16]. On the other hand, software to implement DTs is readily available over the Internet. DTs construction involves the recursive partitioning of a set of training data, which is split into increasingly homogenous subsets on the basis of tests applied to one or more of the feature values. Thus, we chose the decision tree classifier for this task. We chose the univariate decision tree algorithm and built trees in each target study region considering different cropping systems.

For each study region, we labelled the training sample, calculated the value of spectral indicators in specific phenological period, and compared the differences of spectral characteristics occurring in different crop species. We tried to determine the threshold manually by maximizing the dissimilarity or minimizing the similarity of the descendant nodes, and obtained a preliminary decision tree of classifications in three agroecological zones. However, according to our investigation, we found that the crop phenological features could be different even in the different provinces. The time difference of the Sentinel-2 images causes different threshold values. Building a decision tree based on a single threshold value for each study region could make a certain error. Therefore, we set 10% of value difference as the standard of subregion division and divided the Northeast Plain (NEP), North China Plain (NCP), and middle-lower Yangtze River plain (MYRP) into 13, 16, and 10 subregions, respectively, according to the differences in crop phenological features and image times (Figure 3). The goal of subdivision was to better identify the threshold of decision trees and improve the overall accuracy. We built different decision trees for the crop classification of different cropping systems and slightly adjusted the threshold values in different subregions because of the varying image times and management practices. Meanwhile, for each pixel of cropland, a time series dataset of the targeted spectral indicator

was developed. We implemented the specific decision tree in specific subregions and obtained the large-scale crop map by combining of all the subregions.

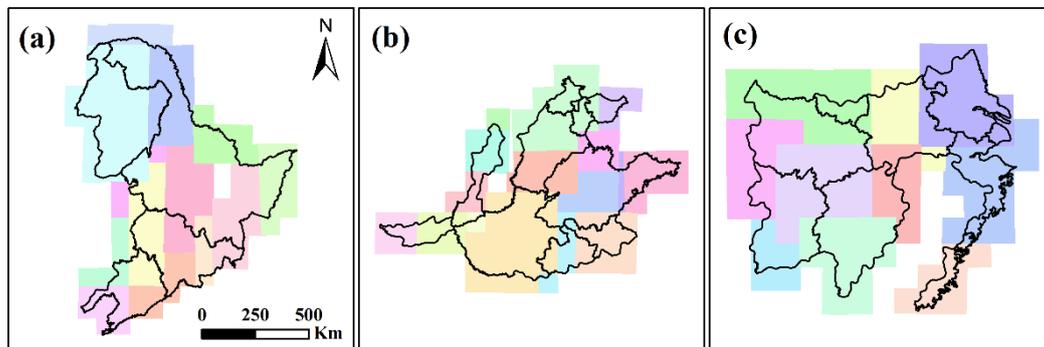


Figure 3. Distribution of the subregions on the Northeast Plain (a), North China Plain (b), and middle-lower Yangtze River plain (c). Different colors represent different subregions.

2.3.4. Accuracy Assessment

In this study, we calculated the whole crop area directly by remote sensing. We evaluated the Sentinel-2-derived crop maps based on ground truth data and agricultural census data. We referenced the ‘National specifications of inspection, acceptance and quality assessment of digital surveying and mapping products’ (GB-T18316-2001) and selected more than 10% of total counties as evaluation targets. Data from a total 113 survey counties with around 400 sample points of each were collected. We randomly chose these test counties and sample points under the principle of uniform distribution, which covered Northeast, North, and South China. We identified the crop pattern through visual interpretation using 0.5 m-resolution Google Maps temporal images. The spatial consistency of the crop pattern maps with the Google Maps image results were assessed for the test counties. We calculated the user’s, producer’s, and overall accuracies. The kappa index was also computed [44].

Due to the limitation of field survey data and error of manual visual interpretation, an accuracy assessment was also conducted with the national agricultural census dataset. A dataset of crop areas at the county level was calculated based on remote sensing derived crop maps. Estimated sown areas were then compared to those from the national agricultural census datasets. We conducted regression analysis for each crop and chose the r-squared value to evaluate the accuracy of results. We set 95% as the confidence coefficient and then tested the *p*-value.

3. Results

3.1. Remote Sense-Based Crop Classification Models in the Typical Cropping Systems

3.1.1. Classification Model on the Northeast Plain

Considering the limited temperature and precipitation resources, a single-cropping system was widely adapted on the NEP. The crop growing season usually starts in April and ends in October. Similar growth periods for different crops increased the difficulty of the crop classification with the use of a single indicator. We chose the NDVI, REP, and NDRI for this area, which can better distinguish the different crop types. We found that using the value of the NDVI in May and July could easily suppress the interference of buildings, water, and forest. The value of the REP in the maize sampling sites during August was higher than those of other crops, which could be used to identify the maize planting areas. We also found that the NDRI in the rice sampling sites during August were lower than those in the maize and soybean sampling sites, which could be used to identify the paddy rice planting areas (Figure 4). The general decision tree on the NEP is shown in Figure 5.

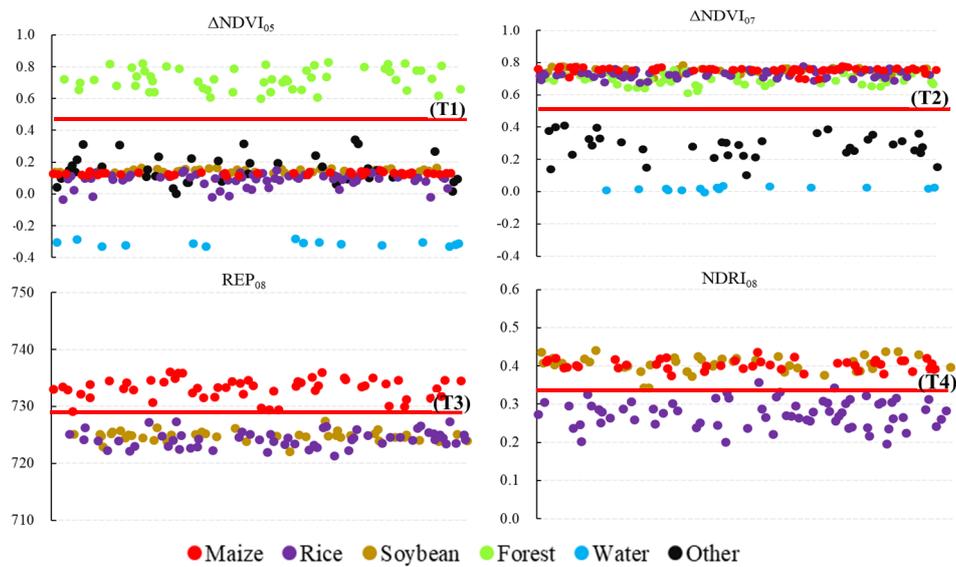


Figure 4. Comparison between the temporal profiles of the normalized difference vegetation (NDVI), red-edge parameter (REP), and normalized difference rice index (NDRI) for the different land cover types on the Northeast Plain. Red lines represent the classification threshold.

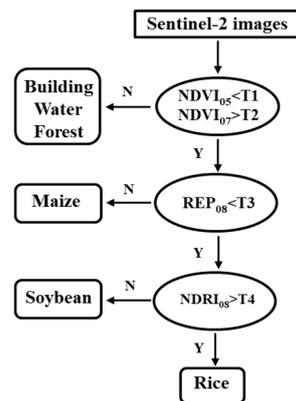


Figure 5. Decision tree for crop classification on the Northeast Plain. T1-4 represent threshold values of each procedure, with slight differences under subregions.

3.1.2. Classification Model for the North China Plain

The NCP is a major winter wheat-summer maize double-cropping system in China. Farmers usually plant winter wheat from October to June and plant summer maize from June to late September. The value of the NDVI was highest in the crop heading period, which can be used to distinguish the crops from other land cover types, such as buildings and water. To differentiate between the crops and other types of vegetation, such as forest and grass, we used the Δ NDVI between the crop heading and harvest period, which was significantly higher than those at other times (Figure 6). The general decision tree for the NEP is shown in Figure 7.

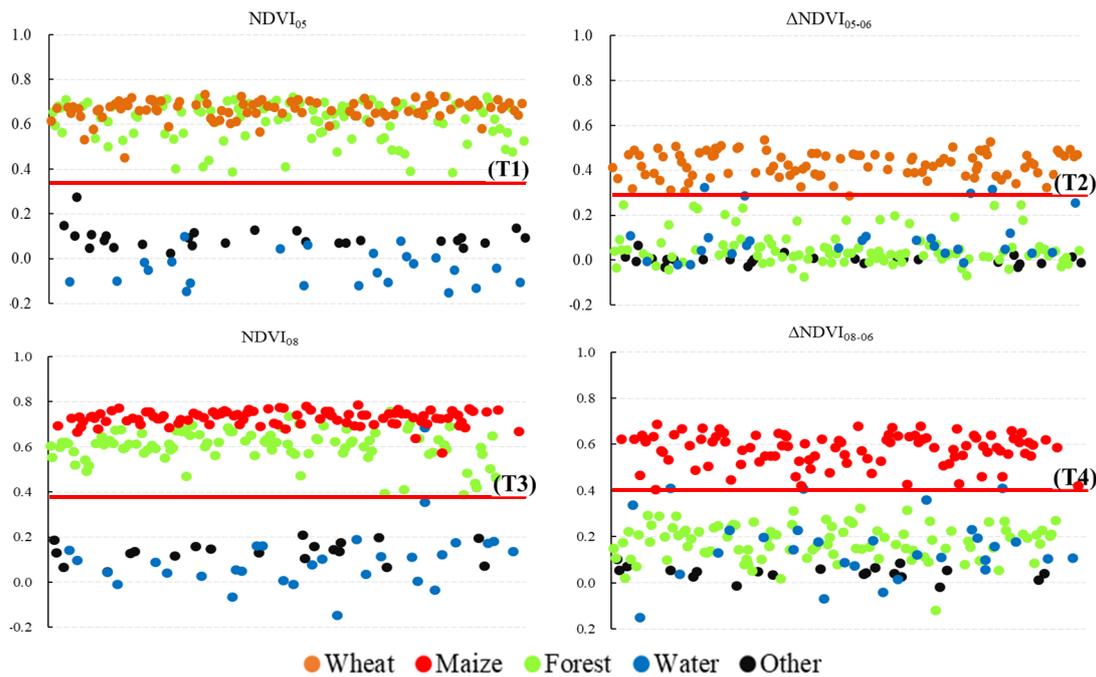


Figure 6. Comparison between the temporal profiles of normalized difference vegetation (NDVI) and NDVI time series (Δ NDVI) for the different land cover types on the North China Plain. Red lines represent the classification threshold.

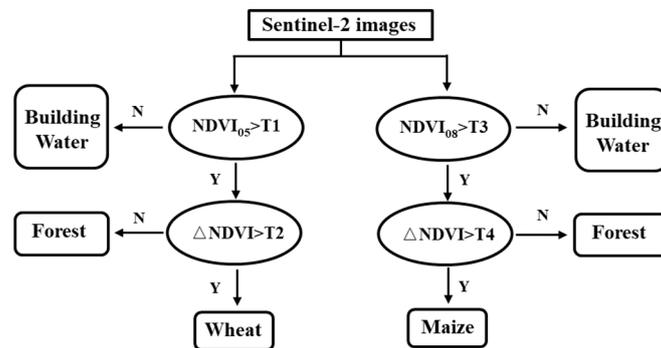


Figure 7. Decision tree for crop classification on the North China Plain. T1-4 represent the threshold values of each procedure with slight differences among the subregions.

3.1.3. Classification Model for the Middle-Lower Yangtze River Plain

Both the double-cropping system and the triple-cropping system occurred on the MYRP. Farmers usually plant single rice (from April to October), early rice (from April to July), or late rice (from July to October) in the summer and plant winter wheat or rape in the winter. We chose the Sentinel-2 images for late April or early May that were taken during the rice transplanting period to identify the rice cropping area. We found that the value of the NDVI in April between T1 and T2 could remove the influence of the water, wheat, and forest. The characteristics of the NDBI and NDWI in April were used to distinguish the rice from other types, such as buildings and rape (Figure 8). For the classification of winter wheat and rape, the images from middle of late March that were taken during the jointing period were chosen in our research. The value of the NDVI in March was totally different between the vegetation and the other land cover types (Figure 8). We used this characteristic to extract vegetation. Considering the color variation in the rape flowers and the other crops, we found that the value of the B-G in the rape sampling sites was significantly lower than those in wheat and forest. This feature could be used to identify the rape cropping areas. We also found that the value of the

LSWI in the wheat sampling sites was higher than that in forest, which was used to identify the wheat cropping areas. The general decision tree for the MYRP is shown in Figure 9.

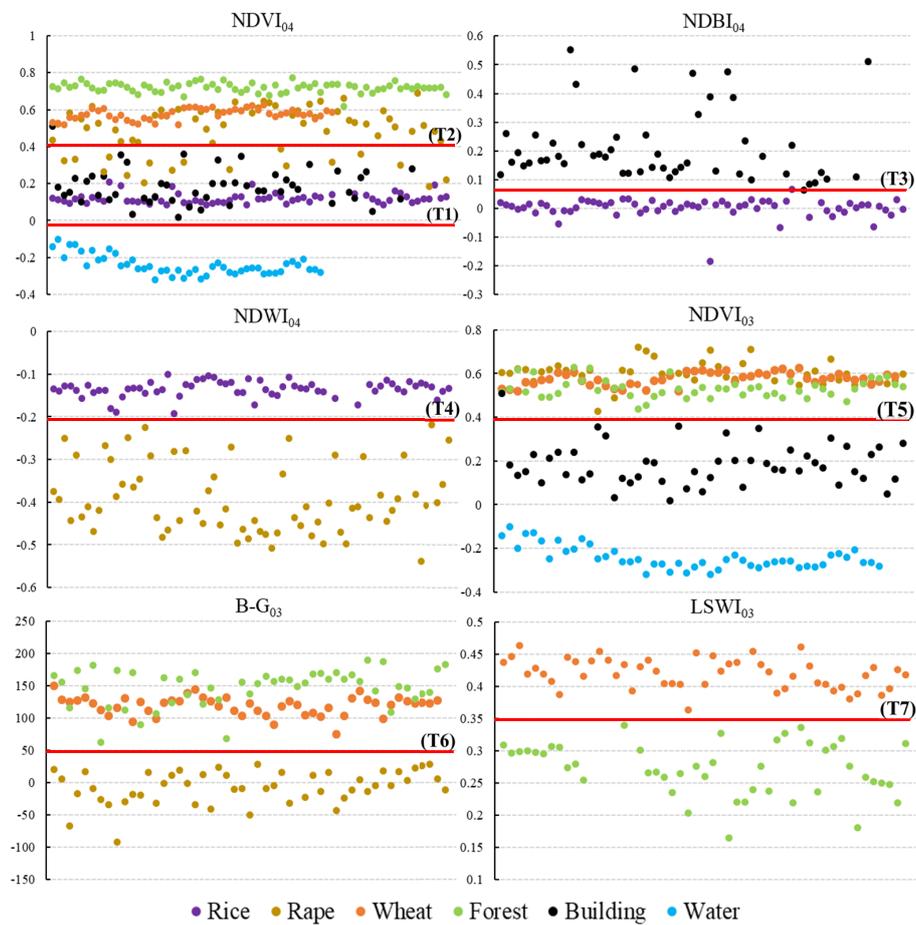


Figure 8. Comparison between the temporal profiles of the normalized difference vegetation (NDVI), normalized difference water index (NDWI), land surface water index (LSWI) and blue-green ratio (B-G) for the different land cover types on the middle-lower Yangtze River plain. Red lines represent the classification threshold.

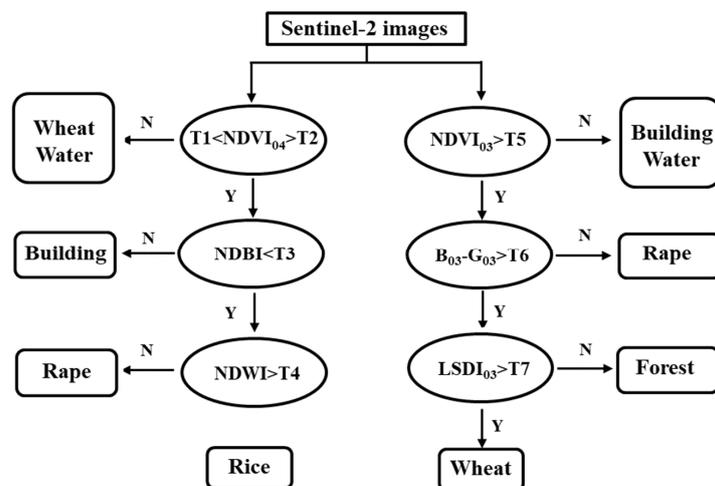


Figure 9. Decision tree for crop classification on the middle-lower Yangtze River plain. T1-7 represent the threshold values of each procedure with slight differences among the subregions.

3.2. Spatial Distribution of the Major Crops in the Main Grain-Producing Regions

We monitored the areas of the major crop sown in the three grain-producing areas from 2017 to 2018. The results showed that the crop planting area on the NEP was mostly concentrated on the Songnen Plain, Liaodong Plain, and Sanjiang Plain. We estimated that the total areas sown with maize, rice, and soybeans on the NEP were 12.1, 6.2, and 7.4 million ha, respectively. The maize cropping area was mainly distributed in the central-western portion of the NEP, while a few areas were located on the Sanjiang Plain. The rice cropping areas were mainly concentrated on the Sanjiang Plain in the Heilongjiang Province, while a few areas were located on the Liaodong Plain. The soybean cropping area was primarily distributed in the Jilin and Heilongjiang provinces in the northern-central region of the NEP (Figure 10a–c).

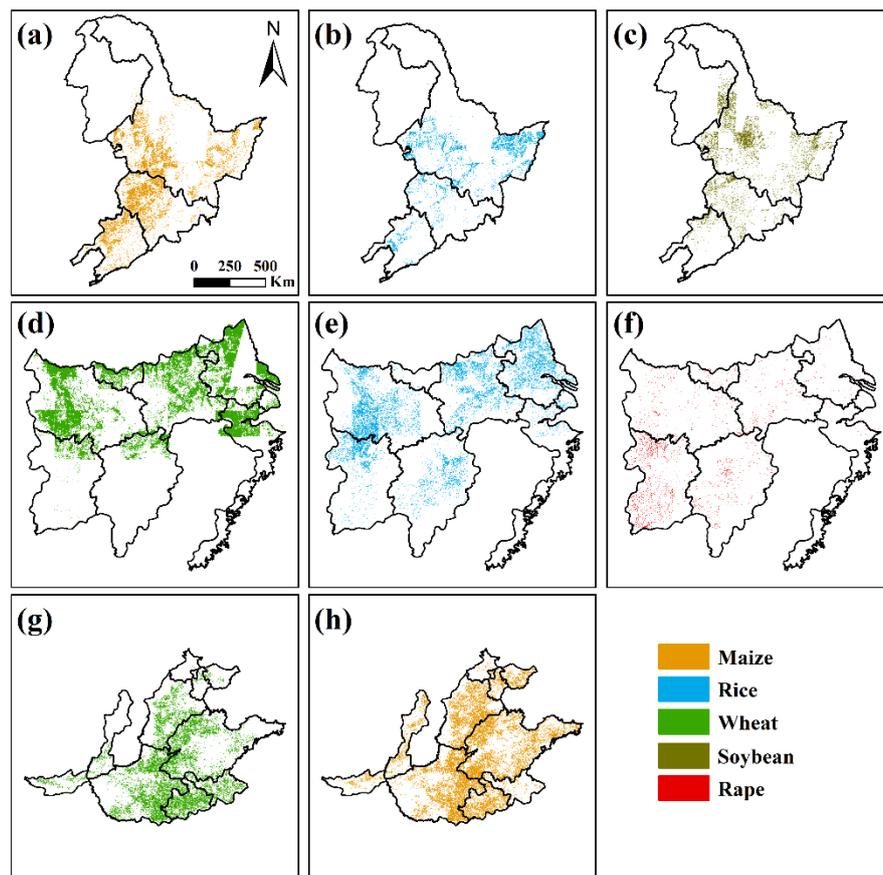


Figure 10. Spatial distribution maps of the major crops on the Northeast Plain (a–c), middle-lower Yangtze River plain (d–f), and North China Plain (g–h) during 2017 and 2018.

The crop planting area on the NCP covered almost the whole region except the mountains and urban areas. We estimated that the total areas sown with winter wheat and summer maize on the NCP were 13.4 and 16.9 million ha, respectively. The winter wheat cropping area was mainly distributed in the piedmont region of the Taihang Mountain, Luxi Plain, and Huang-Huai Plain. Few areas of winter wheat were planted in the Fenhe Valley in the western part of the NCP. The maize cropping area on the NCP covered almost the whole area, including the Jiadong Peninsula and some mountain areas (Figure 10g–h).

Rice is a major crop on the MYRP, including early rice, single rice, and late rice. We estimated that the total areas sown with rice, winter wheat, and winter rape on the MYRP were 6.4, 2.2, and 1.3 million ha, respectively. The rice cropping areas were mostly distributed in the northern part of the MYRP and the plains areas in the south. The winter wheat planting area was mostly concentrated

in the northern part of the MYRP, while the rape planting area was distributed in the western and southern areas of the MYRP (Figure 10d–f).

3.3. Spatial Agreement with the Google Maps Image Results

We evaluated the crop map estimated from the Sentinel-2 imagery based on the 0.5-m resolution Google Maps image results for the test counties. The results showed that the Sentinel-estimated crop map was consistent with the interpreted images from Google Maps, especially on the NCP. The overall accuracy on the NEP was 0.93 with a kappa index of 0.90 (Table 3). The identification accuracy for maize was relatively high in comparison with those of the other crops. The high accuracy occurred on the NCP with an overall accuracy of 0.96 and a kappa index of 0.92 because of the low interference from the other crop types, especially in the winter wheat cropping areas (Table 4). We achieved a relatively high accuracy on the MYRP considering the patchy and fragmented cropland in this area. The overall accuracy on the MYRP was 0.93, with a kappa index of 0.90 (Table 5). The results fully demonstrate the advantage of high-resolution images in crop mapping research in areas with complex land features.

Table 3. Accuracy assessment using the interpreted images from Google Maps for the NEP.

	Maize	Soybean	Rice	Others	User Accuracy
Maize	4557	103	24	180	0.94
Soybean	92	2800	69	148	0.90
Rice	73	37	1269	40	0.89
Others	159	116	14	4506	0.94
Producer accuracy	0.93	0.92	0.92	0.92	
Overall accuracy	0.93				
Kappa	0.90				

Table 4. Accuracy assessment using the interpreted images from Google Maps for the NCP.

	Winter Wheat	Summer Maize	Others	User Accuracy
Winter wheat	4635	-	125	0.97
Summer maize	-	5670	305	0.95
Others	356	464	18557	0.96
Producer accuracy	0.93	0.92	0.98	
Overall accuracy	0.96			
Kappa	0.92			

Table 5. Accuracy assessment using the interpreted images from Google Maps for the MYRP.

	Wheat	Rape	Rice	Others	User Accuracy
Wheat	2686	24	-	203	0.92
Rape	45	2961	-	81	0.96
Rice	-	-	3661	175	0.95
Others	193	150	294	5110	0.89
Producer accuracy	0.92	0.94	0.93	0.92	
Overall accuracy	0.93				
Kappa	0.90				

3.4. Evaluation with Agricultural Statistical Data

Due to the limitations of the field survey data, we also compared the remote sensing-estimated crop areas with those from the national agricultural census data at the county level. These two datasets

agreed reasonably well, with an overall r-squared value of 0.78. The confidence coefficient of all models could reach the 99% level. The coefficients of determination between the Sentinel-2-estimated dataset and the agricultural census data at the county level on the NEP, NCP, and MYRP were 0.83, 0.81, and 0.71, respectively (Figure 11). Considerable agreement was obtained for the maize and rice on the NEP and the wheat in the NCP, with r-squared values higher than 0.85. The accuracy of the estimated crop area on the MYRP was relatively low because of the interference of cloud and mountain elements. Fragmented cropland on the MYRP also increased the difficulty of the crop mapping, especially in the areas sown with rape.

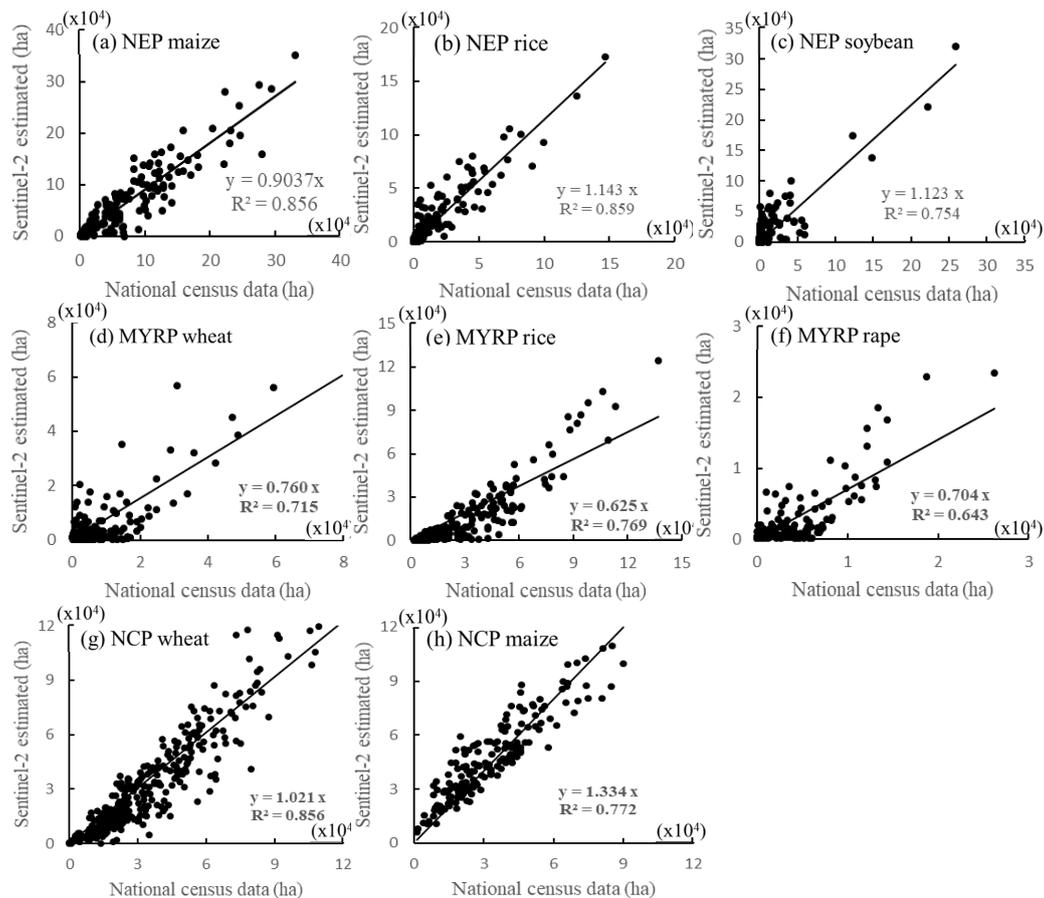


Figure 11. Correlations of areas sown with different crop species between the Sentinel-2-estimated and national census data on the Northeast Plain (a–c), middle-lower Yangtze River plain (d–f), and North China Plain (g–h).

4. Discussion

This study contributed to large-scale crop mapping using high-resolution satellite images and multiple indicators in China. Previous research has mapped the spatiotemporal dynamics of the maize and cropping intensity trends in China at the national scale by using MODIS (500 m) images [12,45]. Moderate-resolution images can achieve coverage of large areas and make large-scale monitoring possible. However, limited resolution and indicator systems could cause errors in crop distribution mapping research [30]. Multiple indicator systems have been widely used in local-scale studies [46]. It is possible to achieve almost completely accurate results for small areas with the development of quantitative monitoring technologies for remote sensing. However, few studies have focused on the extensive use of multiple indicator systems in large areas. On the basis of previous studies, we created a practical application for crop mapping research at the national scale with a combination of multiple indicator systems and high-resolution satellite images.

High-resolution satellite data have played an important role in remote sensing-based agricultural monitoring. The cost of large-scale use of high-resolution imagery is prohibitive. Some recent studies have monitored crop distributions in China with data from the GF-1 satellite, which was launched by China to obtain high-resolution imagery (i.e., 2 m/8 m/16 m) and a short revisit cycle (i.e., 4 days). However, limited wave bands (four bands) increase the difficulty of crop classification, especially in ferruginous plots [28]. This study is also a pioneering research to prove that the Sentinel-2 satellites could be useful for large-scale, high-resolution crop mapping research in China. We realized relatively high accuracy on the middle-lower Yangtze River plain, which has complex land cover features and fragmented cropland. The results fully demonstrated that the Sentinel-2 satellites, which have high resolutions, low revisit periods, multiple wave bands, and the advantage of low cost, can easily provide images for specific times and multiple indicator systems, which could be widely used in large-scale classification of land cover types.

We achieved relatively high classification accuracy compared with the interpreted Google Maps images. Though the temporal difference between Google Maps and Sentinel-2 images may bring some errors, cropping system in the study area would not change frequently. Thus, we did not take this kind of error into consideration. Visual interpretation of Google Maps would also cause some mistakes considering the different individual standard, especially in some blurry areas. We marked these challenge sample points and made the final decision through team discussion in order to reduce the human error in this part. However, the accuracy assessment with the county-level national census data was unsatisfactory. Previous studies of large-scale crop mapping have usually conducted an accuracy assessment with census data at the provincial level [45]. The use of provincial data could ignore some errors in local areas and reach relatively high accuracy levels. We conducted an accuracy assessment at the county level, which could enlarge the discrete degree to a certain extent. On the other hand, national census data at the county level have hysteresis characteristics and are easily affected by human factors, which could cause errors in the accuracy assessment. Cloud cover is also an uncertain factor for monitoring studies. A higher cloud cover ratio indicated higher uncertainties in the estimated results. The lack of images for specific times during the crop growing period increases the difficulty of crop classification. For example, entire crop cycles might be missed due to persistent cloud cover [47]. In this study, there is limited influence caused by cloud cover because of the short revisit period of Sentinel-2 satellite. We could almost obtain more than three available images in a single month, especially in NEP and NCP, while we could almost obtain at least one available image in a single month in MYRP. On the other hand, although we set a 20% standard to control cloud cover when downloading the images, missing images in some areas invalidated our decision model, especially in the critical crop growing period. We calculated the whole cloud area in our research and found that cloud area occurring in the specific growing period was less than 5% of the whole area, which could cause certain errors for the detection of the true situation.

Large-scale and high-resolution remote sensing monitoring is significant for the adjustment and optimization of agricultural production. Understanding the spatial and temporal variations of crop patterns is important for the creation of policy, assessment of food security, and development of sustainable agricultural practices [8,48]. We found that the sown area of the winter wheat on the NCP was much lower than that of the summer maize, especially in the Hebei Province. This result was also much lower than previous crop mapping research in 2009 [49]. The results showed that the crop rotation and fallow policy have been properly implemented because of groundwater overexploitation on the NCP [50]. Building remote sensing-based models in China for multiple cropping systems is also meaningful for nationwide crop mapping research [49]. Creating and periodically updating different crop distribution datasets is meaningful for estimating crop yield, analyzing changes in cropping intensity and adapting agriculture to climate change [45].

We estimated the main crop distribution in the major grain-producing areas of China and achieved relatively high accuracy in the plain areas. However, it is difficult to ensure accuracy for crop monitoring in mountain areas. We used 90-m resolution SRTM DEM and SRTM SLOPE data to remove

the mountain areas and assumed that few crops were distributed in the mountains. Some crops could be planted in the mountain areas with relatively low yields due to limitations of the landscape, especially on the middle-lower Yangtze River plain. We could focus on the classification of crops in mountain areas and build a reasonable model for improving the accuracy of crop mapping in the future. We could also refine the crop types (i.e., spring maize/summer maize) and build a classification method for other field crops in the future, such as cotton and potato. On the other hand, the decision trees method is an ordinary but efficient method in crop classification research [15,16]. There are also some alternative methods based on machine learning or deep learning in this field. For example, the random forest algorithm has been widely used in recent years. It is an ensemble learning algorithm that consists of multiple decision trees or classified regression trees and can reduce error caused by artificial subjective factor to a certain extent [34,36]. In the future, we will conduct comparative studies between the different classification methods and optimize our algorithm in different cropping systems. We built an open-access dataset on the internet (www.cross.ac.cn) to share the crop monitor results at county level and intend to update it annually. We hope our results can guide the adjustment and optimization of the national agricultural structure in the future.

5. Conclusions

This study built three remote sensing-based models for typical cropping systems using multiple indicators and monitored the distribution of the major crops in the main grain-producing area of China using Sentinel-2 satellite data. We found that the areas sown with maize, rice, and soybean on the NEP were 12.1, 6.2 and 7.4 million ha, respectively. The maize and wheat planting areas on the NCP were 16.9 and 13.4 million ha, respectively. The total sown areas of the rice, wheat, and rape on the MYRP were 6.4, 2.2, and 1.3 million ha, respectively. The estimated images agreed well with field survey data (average overall accuracy = 94%) and the national agricultural census data ($R^2 = 0.78$). This proves the applicability of the Sentinel-2 satellite data for large-scale, high-resolution crop mapping in China. We intend to update the crop mapping datasets annually and hope to guide the adjustment and optimization of the national agricultural structure.

Author Contributions: Conceptualization and design, Y.J. and Z.L.; methodology, Y.J.; software, Z.L.; formal analysis, Y.J.; resources, S.L.; writing—original draft preparation, Y.J.; writing—review and editing, Y.L. and Q.C.; supervision, X.Y.; project administration, F.C.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number (2016YFD0300201).

Acknowledgments: We thank the research team at the Institute of Geographic Sciences and National Resources Research for their careful guidance and thank all reviewers and editors for their suggestions, which substantially improved the manuscript.

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

References

1. He, G.; Zhao, Y.; Wang, L.; Jiang, S.; Zhu, Y. China's food security challenge: Effects of food habit changes on requirements for arable land and water. *J. Clean. Prod.* **2019**, *229*, 739–750. [[CrossRef](#)]
2. Wang, Q.; Liu, X.; Yue, T.; Wang, C.; Wilson, J.P. Using models and spatial analysis to analyze spatio-temporal variations of food provision and food potential across China's agro-ecosystems. *Ecol. Model.* **2015**, *306*, 152–159. [[CrossRef](#)]
3. Wang, Y.; Wu, P.; Engel, B.; Sun, S. Comparison of volumetric and stress-weighted water footprint of grain products in China. *Ecol. Indic.* **2015**, *48*, 324–333. [[CrossRef](#)]
4. Kang, Y. Food safety governance in China: Change and continuity. *Food Control.* **2019**, *106*, 106752. [[CrossRef](#)]
5. Yang, H.; Liu, J.; Savenije, H. China's move to higher-meat diet hits water security. *Nature* **2008**, *454*, 397.
6. Zang, Z.; Zou, X.; Xi, X.; Zhang, Y.; Zheng, D.; Sun, C. Quantitative characterization and comprehensive evaluation of regional water resources using the three red lines method. *J. Geogr. Sci.* **2016**, *26*, 397–414. [[CrossRef](#)]

7. Feng, Z.; Yang, Y.; Zhang, Y.; Zhang, P.; Li, Y. Grain-for-green policy and its impacts on grain supply in West China. *Land Use Policy* **2005**, *22*, 301–312. [[CrossRef](#)]
8. Liu, Z.; Yang, P.; Wu, W.; You, L. Spatiotemporal changes of cropping structure in China during 1980–2011. *J. Geogr. Sci.* **2018**, *28*, 1659–1671. [[CrossRef](#)]
9. Xie, H.; Cheng, L.; Lu, H. Farmers' responses to the winter wheat fallow policy in the groundwater funnel area of China. *Land Use Policy* **2018**, *73*, 195–204. [[CrossRef](#)]
10. Xiao, X.; Boles, S.; Liu, J.; Zhuang, D.; Froking, S.; Li, C.; Salas, W.; Moore, B. Mapping paddy rice agriculture in Southern China using multi-temporal MODIS images. *Remote Sens. Environ.* **2005**, *95*, 480–492. [[CrossRef](#)]
11. Yin, Q.; Liu, M.; Cheng, J.; Ke, Y.; Chen, X. Mapping paddy rice planting area in Northeastern China using spatiotemporal data fusion and phenology-based method. *Remote Sens.* **2019**, *11*, 1699. [[CrossRef](#)]
12. Qiu, B.; Huang, Y.; Chen, C.; Tang, Z.; Zou, F. Mapping spatiotemporal dynamics of maize in China from 2005 to 2017 through designing leaf moisture-based indicator from normalized multi-band drought index. *Comput. Electron. Agric.* **2018**, *153*, 82–93. [[CrossRef](#)]
13. Jia, K.; Wu, B.; Li, Q. Crop classification using HJ satellite multispectral data in the North China plain. *J. Appl. Remote Sens.* **2013**, *7*, 73576. [[CrossRef](#)]
14. Vintrou, E.; Desbrosse, A.; Bégué, A.; Traoré, S.; Baron, C.; Seen, D. Crop area mapping in West Africa using landscape stratification of MODIS time series and comparison with existing global land products. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14*, 83–93. [[CrossRef](#)]
15. Hu, Q.; Wu, W.; Song, Q.; Yu, Q.; Yang, P.; Tang, H. Recent progress in research of crop patterns mapping by using remote sensing. *Sci. Agric. Sin.* **2015**, *48*, 1900–1914.
16. Tang, H.; Wenbin, W.; Yang, P. Recent progresses in monitoring crop spatial patterns by using remote sensing technologies. *Sci. Agric. Sin.* **2010**, *43*, 2879–2888.
17. Mathur, A.; Foody, G. Crop classification by support vector machine with intelligently selected training data for an operational application. *Int. J. Remote Sens.* **2008**, *29*, 2227–2240. [[CrossRef](#)]
18. Yang, C.; Everitt, J.; Murden, D. Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Comput. Electron. Agric.* **2011**, *75*, 347–354. [[CrossRef](#)]
19. Cai, Y.; Guan, K.; Peng, J.; Wang, S.; Seifert, C.; Wardlow, B.; Li, Z. A high-performance and in-season classification system of field-level crop types using time-series landsat data and a machine learning approach. *Remote Sens. Environ.* **2018**, *210*, 35–47. [[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. [[CrossRef](#)]
21. Thenkabail, P.; Biradar, C.; Noojipady, P.; Dheeravath, V.; Li, Y.; Velpuri, M.; Gumma, M.; Gangalakunta, O.; Turrall, H.; Cai, X.; et al. Global irrigated area map (GIAM), derived from remote sensing, for the end of the last millennium. *Int. J. Remote Sens.* **2009**, *30*, 3679–3733. [[CrossRef](#)]
22. Biradar, C.; Thenkabail, P.; Noojipady, P.; Li, Y.; Dheeravath, V.; Turrall, H.; Velpuri, M.; Gumma, M.; Gangalakunta, O.; Cai, X.; et al. A global map of rainfed cropland areas (GMRCA) at the end of last millennium using remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 114–129. [[CrossRef](#)]
23. Leff, B.; Ramankutty, N.; Foley, J.A. Geographic distribution of major crops across the world. *Glob. Biogeochem. Cycles* **2004**, *18*, GB1009. [[CrossRef](#)]
24. Monfreda, C.; Ramankutty, N.; Foley, J.A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* **2008**, *22*, GB1022. [[CrossRef](#)]
25. Wu, B.; Li, Q. Crop planting and type proportion method for crop acreage estimation of complex agricultural landscapes. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *16*, 101–112. [[CrossRef](#)]
26. Qiu, B.; Luo, Y.; Tang, Z.; Chen, C.; Lu, D.; Huang, H.; Chen, Y.; Chen, N.; Xu, W. Winter wheat mapping combining variations before and after estimated heading dates. *ISPRS J. Photogramm. Remote Sens.* **2017**, *123*, 35–46. [[CrossRef](#)]
27. Wardlow, B.; Egbert, S. Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sens. Environ.* **2008**, *112*, 1096–1116. [[CrossRef](#)]
28. Zhou, Q.; Yu, Q.; Liu, J.; Wu, W.; Tang, H. Perspective of Chinese GF-1 high-resolution satellite data in agricultural remote sensing monitoring. *J. Integr. Agric.* **2017**, *16*, 242–251. [[CrossRef](#)]
29. Song, Q.; Zhou, Q.; Wu, W.; Hu, Q.; Lu, M.; Liu, S. Mapping regional cropping patterns by using GF-1 WVF sensor data. *J. Integr. Agric.* **2017**, *16*, 337–347. [[CrossRef](#)]

30. Li, H.; Chen, Z.; Jiang, Z.; Wu, W.; Ren, J.; Liu, B.; Tuya, H. Comparative analysis of GF-1, HJ-1, and landsat-8 data for estimating the leaf area index of winter wheat. *J. Integr. Agric.* **2017**, *16*, 266–285. [[CrossRef](#)]
31. Cabezas-Rabadán, C.; Pardo-Pascual, J.E.; Palomar-Vázquez, J.; Fernández-Sarría, A. Characterizing beach changes using high-frequency sentinel-2 derived shorelines on the valencian coast (Spanish mediterranean). *Sci. Total Environ.* **2019**, *691*, 216–231. [[CrossRef](#)] [[PubMed](#)]
32. Griffiths, P.; Nendel, C.; Hostert, P. Intra-annual reflectance composites from sentinel-2 and landsat for National-scale crop and land cover mapping. *Remote Sens. Environ.* **2019**, *220*, 135–151. [[CrossRef](#)]
33. Sánchez-Espinosa, A.; Schröder, C. Land use and land cover mapping in wetlands one step closer to the ground: Sentinel-2 versus landsat 8. *J. Environ. Manag.* **2019**, *247*, 484–498. [[CrossRef](#)] [[PubMed](#)]
34. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classification in Central Europe. *Remote Sens.* **2016**, *8*, 166. [[CrossRef](#)]
35. Hegarty-Craver, M.; Polly, J.; O’Neil, M.; Ujeneza, N.; Rineer, J.; Beach, R.; Lapidus, D.; Temple, D. Remote crop mapping at scale: Using satellite imagery and UAV-acquired data as ground truth. *Remote Sens.* **2020**, *12*, 1984. [[CrossRef](#)]
36. Zhu, J.; Pan, Z.; Wang, H.; Huang, P.; Sun, J.; Qin, F.; Liu, Z. An improved multi-temporal and multi-feature tea plantation identification method using Sentinel-2 imagery. *Sensors* **2019**, *19*, 2087. [[CrossRef](#)]
37. Boryan, C.; Yang, Z.; Mueller, R.; Craig, M. Monitoring US agriculture: The US department of agriculture, National agricultural statistics service, cropland data layer program. *Geocarto Int.* **2011**, *26*, 341–358. [[CrossRef](#)]
38. Liu, X.; Chen, F. *Farming Systems in China*; China Agriculture Press: Beijing, China, 2005.
39. Wu, W.; Yang, P.; Tang, H.; Zhou, Q.; Chen, Z.; Shibasaki, R. Characterizing spatial patterns of phenology in cropland of China based on remotely sensed data. *Agric. Sci. China* **2010**, *9*, 101–112. [[CrossRef](#)]
40. Firozjaei, K.; Alavipanah, K.; Liu, H.; Sedighi, A.; Mijani, N.; Kiavarz, M.; Weng, Q. A PCA–OLS model for assessing the impact of surface biophysical parameters on land surface temperature variations. *Remote Sens.* **2019**, *11*, 2094. [[CrossRef](#)]
41. Chang, Q.; Xiao, X.; Jiao, W.; Wu, X.; Doughty, R.; Wang, J.; Du, L.; Zou, Z.; Qin, Y. Assessing consistency of spring phenology of snow-covered forests as estimated by vegetation indices, gross primary production, and solar-induced chlorophyll fluorescence. *Agric. For. Meteorol.* **2019**, *275*, 305–316. [[CrossRef](#)]
42. Carbas, B.; Machado, N.; Oppolzer, D.; Ferreira, L.; Brites, C.; Rosa, E.; Barros, A. Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for the determination of nutritional and antinutritional parameters in common beans. *Food Chem.* **2020**, *306*, 125509. [[CrossRef](#)] [[PubMed](#)]
43. Zhang, M.; Su, W.; Fu, Y.; Zhu, D.; Xue, J.; Huang, J.; Wang, W.; Wu, J.; Yao, C. Super-resolution enhancement of sentinel-2 image for retrieving LAI and chlorophyll content of summer corn. *Eur. J. Agron.* **2019**, *111*, 125938. [[CrossRef](#)]
44. Congalton, R. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 270–279. [[CrossRef](#)]
45. Qiu, B.; Lu, D.; Tang, Z.; Song, D.; Zeng, Y.; Wang, Z.; Chen, C.; Chen, N.; Huang, H.; Xu, W. Mapping cropping intensity trends in China during 1982–2013. *Appl. Geogr.* **2017**, *79*, 212–222. [[CrossRef](#)]
46. Qiu, B.; Lu, D.; Tang, Z.; Chen, C.; Zou, F. Automatic and adaptive paddy rice mapping using landsat images: Case study in songnen plain in Northeast China. *Sci. Total Environ.* **2017**, *598*, 581–592. [[CrossRef](#)]
47. Gray, J.; Friedl, M.; Frolking, S.; Ramankutty, N.; Nelson, A.; Gumma, M. Mapping Asian cropping intensity with MODIS. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 3373–3379. [[CrossRef](#)]
48. Tang, H.; Wu, W.; Yang, P.; Li, Z. Systematic synthesis of impacts of climate change on China’s crop production system. *J. Integr. Agric.* **2014**, *13*, 1413–1417. [[CrossRef](#)]
49. Tao, J.; Wu, W.; Zhou, Y.; Wang, Y.; Jiang, Y. Mapping winter wheat using phenological feature of peak before winter on the North China plain based on time-series MODIS data. *J. Integr. Agric.* **2017**, *16*, 348–359. [[CrossRef](#)]
50. Jiang, Y.; Wang, X.; Ti, J.; Yin, X.; Lei, Y.; Chu, Q.; Chen, F. Winter wheat water-saving potential in groundwater overexploitation district of North China plain. *Agron. J.* **2020**, *112*, 44–55. [[CrossRef](#)]

