LETTER

Identifying the spatiotemporal changes of annual harvesting areas for three staple crops in China by integrating multi-data sources

Yuchuan Luo, Zhao Zhang, Ziyue Li, Yi Chen, Liangliang Zhang, Juan Cao and Fulu Tao

State Key Laboratory of Earth Surface Processes and Resource Ecology & MOE Key Laboratory of Environmental Change and Natural Hazards, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, People’s Republic of China
Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, People’s Republic of China
College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, People’s Republic of China

Keywords: agricultural system, cropland, land use/cover change, rice, maize, wheat

Abstract

Reliable and continuous information on major crop harvesting areas is fundamental to investigate land surface dynamics and make policies affecting agricultural production, land use, and sustainable development. However, there is currently no spatially explicit and time-continuous crop harvesting area information with a high resolution for China. The spatiotemporal patterns of major crop harvesting areas at a national scale have rarely been investigated. In this study, we proposed a new crop phenology-based crop mapping approach to generate a 1 km harvesting area dataset for three staple crops (i.e. rice, wheat, and maize) in China from 2000 to 2015 based on GLASS leaf area index (LAI) products. First, we retrieved key phenological dates of the three staple crops by combining the inflexion- and threshold-based methods. Then, we determined the grids cultivated for a certain crop if its three key phenological dates could be simultaneously identified. Finally, we developed crop classification maps and a dataset of annual harvesting areas (ChinaCropArea1 km), comprehensively considering the characteristics of crop phenology and the references of drylands and paddy fields. Compared with the county-level agricultural statistical data, the crop classification had a high accuracy, with $R^2$ values consistently greater than 0.8. The spatiotemporal patterns of major crop harvesting areas during the period were further analyzed. The results showed that paddy rice harvesting areas had expanded aggressively in northeastern China but decreased in southern China. Maize harvesting areas expanded substantially in major maize cultivation areas across China. Wheat harvesting areas declined overall, although they increased notably in their major production areas. The spatiotemporal patterns could be ascribed to various anthropogenic, biophysical, and social-economic drivers, including urbanization, reduced cropping intensity in southern China, frequent disasters from climate change, and large areas of abandoned farmland in northern and southwestern China. The resultant dataset can be applied for many purposes, including land surface modeling, agro-ecosystem modeling, agricultural production and land use policy-making.

1. Introduction

Paddy rice, maize and wheat provide the most important staple foods, accounting for approximately 42.5% of the food calorie supply and 79.2% of the total harvested areas of cereals (FAO 2016). However, climate change, urbanization, economy and policy, and cropland degradation have led to significant changes in cropland area and threaten global food security (Liu et al 2014, Asseng et al 2015, Song et al 2018). Therefore, accurate and timely information on crop spatial distribution is essential, especially for land surface modeling, agro-ecosystem modeling, agricultural production prediction, policy-making, and environmental sustainability (Liu et al 2005, Gumma et al 2014, Zhang et al 2014a).
Field surveys and remote sensing are the main ways to obtain the spatial distributions of crop planting/harvesting areas. Compared to field surveys, remote sensing has the advantages of being cost-effective, having a wide spatial coverage and having a high revisiting frequency. Thus, combining remote sensing with ground truth from field surveys has been a powerful tool to detect crop planting/harvesting areas at regional or global scales (Xiao et al 2006, Chang et al 2007, Bouvet and Thuy 2011, Kuenzer and Knauer 2013). Previous efforts to map crop spatial distribution were mainly based on the normalized difference vegetation index (NDVI) time series data or enhanced vegetation index (EVI) time series data derived from the moderate resolution imaging spectroradiometer (MODIS), advanced very high resolution radiometer (AVHRR), and Landsat data (Ozdogan 2010, Pan et al 2012, Neeti et al 2012, Zhang et al 2014a, Gumma et al 2016, Zhong et al 2016). Since different crop types may have similar spectral and textural features within the same growing season in remote sensing images at a large scale, phenological differences play an important role in crop classification at a large scale (Wardlow et al 2007, Geerken 2009). Previous studies on crop classification using phenology-based methods could be mainly grouped into (1) detecting the seasonal dynamics of each crop and utilizing the image dates corresponding to key phenological dates of different crops for classification (Conrad et al 2010, Son et al 2014); (2) determining phenological metrics (e.g. starting and end dates of the whole growing season) and establishing classification rules based on crop calendar and stage-specific crop growth characteristics (Dong et al 2015, Qiu et al 2017a). The latter usually requires supervised classification methods (e.g. random forest, decision tree classification and support vector machine) with appropriate training samples for different land-cover classes (Brown et al 2013, Hao et al 2015). However, the performance of these supervised classification methods generally depends on the selection of appropriate training samples that are difficult to collect at a large scale. Additionally, the training samples used in previous studies were typically collected for a specific year, which could lead to unreliable mapping results when applying them for long-term classification due to the changes in land cover over time. Thus, it is essential to develop an automatic and rapid approach to classify and map major crops over the long term at a large scale.

In addition, further improvements in crop harvesting area mapping are necessary because: (1) no spatially explicit dataset on the harvesting areas of the three staple crops across China is available due to high diversity and complexity of agricultural cropping systems, which hinders related studies on global food security (Piao et al 2010, Ray et al 2015); (2) the vegetation index (VI) is less sensitive to a large amount of vegetation than is the leaf area index (LAI). The phenological information derived from the LAI is more reliable than that from VIs for classification because the seasonal patterns of phenology are better represented by leaf development, canopy structure, and the biochemical properties of leaves (Baret and Guyot 1991, Myneni and Williams 1994, Richardson et al 2009); (3) many previous studies have focused on mapping single-crop types without considering different cropping patterns or multi-crop types (Dong et al 2015, Zhang et al 2017, Qiu et al 2017b). Currently, there are a lack of annual and spatially explicit crop classification maps over a long-term period, especially for maize and wheat as many previous studies have focused on paddy rice (Dong et al 2015, Qin et al 2015, Zhang et al 2017, Qiu et al 2017a). Crop type maps at the national scale are available only in the United States and Canada, examples of which include the Cropland Data Layer (CDL) generated by the United States Department of Agriculture (USDA) and the Crop Inventory (CI) dataset generated by the Agriculture and Agri-Food Canada (AAFC) (Johnson and Mueller 2010). Thus, it is urgent to acquire spatially explicit crop harvesting area datasets over a long-term period in China. Detailed information on the harvesting area could be incorporated into agroecosystem models or environmental models to accelerate agricultural production estimates and policy-making (Gilmanov et al 2013, Sakamoto et al 2014, Lobell et al 2015, Azzari et al 2017). Moreover, this information can help us investigate the impacts of climatic, socioeconomic, or other factors on cropland land area change and reveal the drivers of agricultural land or bioenergy land use expansion (Wang et al 2014, Howard and Wylie 2014, Zhang et al 2017). Therefore, an accurate and timely crop harvesting area dataset is valuable for related scientific research and policy-making.

The main objectives of this study are three-fold: (1) to develop a new crop mapping method to obtain a spatially explicit harvesting area dataset of three staple crops (i.e. rice, wheat and maize) in China from 2000 to 2015 using the remotely sensed Global Land Surface Satellite (GLASS) LAI product (Xiao et al 2014); (2) to evaluate the accuracy of the crop phenology-based approach in mapping crop harvesting areas by comparing the obtained dataset with county-level agricultural statistical data; and (3) to investigate the spatial patterns and temporal changes in crop harvesting areas. To our knowledge, we provide the first dataset of annual harvesting areas for three staple crops across China from 2000 to 2015 at a resolution of 1 km (ChinaCropArea1 km).

2. Data and methods
2.1. Study area
The study area is mainland China, with a complex agricultural planting structure, complicated geographic environments, diverse cropping patterns and
The studied areas and the spatial distribution of Agricultural Meteorological Stations (AMS). The study areas include seven sub-regions: Northeastern China (NE), the Middle-Lower Yangtze plain (YZ), Southern China (ST), Northwestern China (NW), North China Plain (NCP), and Southwestern China (SW) [include Sichuan Basin (SC)].

2.2. Data sources

2.2.1. The GLASS LAI data
The improved MODIS-based LAI products (GLASS LAI) from 2000 to 2015 with a 1-km spatial resolution and an 8-day composite were used in this study. The GLASS LAI products, obtained from the Beijing Normal University Global Change Data Processing and Analysis Centre (http://glass-product.bnu.edu.cn/?pid=protect$\relax=38c$&protect$\relax=1$), were generated with general regression neural networks (GRNNs) approach based on time-series reflectance data (Liang et al 2013). Compared to other LAI products, the GLASS LAI is more temporally continuous and spatially complete (Xiao et al 2014), and it has been applied to global land cover monitoring and crop model assimilation (Chen et al 2018, Liu et al 2019).

2.2.2. Other datasets
The observed phenology of rice, wheat, and maize from 2000 to 2013 were collected from Agricultural Meteorological Stations (AMS) in China, which are governed by the China Meteorological Administration (CMA) (https://data.cma.cn/). Crop phenology was observed and recorded by well-trained agricultural technicians in experimental fields (the average field size was 0.15 ha) and then checked and managed by the Chinese Agricultural Meteorological Monitoring System (CAMMS) (Tao et al 2014). In this study, the observed key phenological dates for the three staple crops at the agrometeorological stations for more than 10 years were applied. The key phenological dates included the green-up date, emergence date, transplanting date, V3 stage (i.e. early vegetative stage of maize when the third leaf is fully expanded), heading date, and maturity date.

The National Land Cover Dataset (NLCD) of China, with a spatial resolution of 1-km, was provided by the Data Centre for Resources and Environmental Sciences, Chinese Academy of Sciences (http://www.resdc.cn/Default.aspx), and these data included the land use datasets for 2000, 2005, 2010, and 2015. These datasets were generated by visual interpretation.
and digitization of Landsat TM/ETM + images with an overall classification accuracy above 90% (Liu et al 2005, 2014). Agricultural census data were derived from the statistical yearbooks of counties in China from 2000 to 2008 (http://data.cnki.net/).

2.3. Methods
The flow chart applied in this study for mapping the three staple crops is presented in figure 2, including the data preprocessing, phenology-based crop mapping method, and accuracy assessment processes.

2.3.1. Data preprocessing
First, we projected or reprojected all raster data to the ‘Asia North Albers Equal Area Conic’ projection using the nearest-neighbor resampling approach because this projection had the minimum bias on the area calculation and was therefore suitable for comparison with the actual harvesting areas. Then, the cultivated land derived from the NLCD was used as a cropland mask. Finally, we combined 46 annual GLASS LAI images and obtained an LAI time series for each cropland pixel and further applied the commonly used Savitzky-Golay (S-G) filter method to reduce the noise of the GLASS LAI time series.

2.3.2. Phenology-based crop mapping method
Each crop has a specific cropping pattern and growth curve characteristics, which can be used to identify its specific type using remote sensing time-series imagery. Figure 3 illustrates the LAI temporal profiles of different cropping patterns in each subregion defined in figure 1. Apparently, the image dates corresponding to the key phenological stages help us distinguish crop type. For example, spring maize usually reaches the V3 stage 30 ~ 60 d earlier than summer maize, and therefore, this advanced V3 phase could be used for crop type classification. Likewise, winter wheat and spring wheat can be easily distinguished in each phenological period. Compared with maize, the key phenological period of spring wheat is generally earlier, especially in terms of the maturity date, which is nearly 60 d earlier. In addition, the number of peaks during a growing season can be used to distinguish a single cropping pattern from a double- or triple-cropping pattern, such as double cropping rice and single cropping rice, winter wheat-summer maize rotation, and spring maize and other single cropping types. Therefore, integrated and comprehensive information will greatly improve our ability to map the main crops accurately. The phenology-based crop mapping approach for a long-term period includes two processes: retrieving phenological stages and classifying crop type.

2.3.2.1. Retrieval of phenological stages
We used both the inflexion- and the threshold-based methods together to detect crop phenology (Luo et al 2020). First, we defined the inflection and maximum points of the LAI time series as the specific timing of key phenological stages for different crops (Sakamoto et al 2005, 2010, Sakamoto 2018). Specifically, we defined the date of inflection point (the point after
Figure 3. Temporal profiles of LAI for different crop cropping systems for each sub-region.

which the first derivative increases continuously or the second derivative equals 0) of the LAI time-series curves as the green-up date of winter wheat, emergence date of spring wheat, transplanting date of rice and V3 stage of maize. Moreover, we regarded the maximum LAI point in the LAI time-series curve as the heading date and considered another inflection point (the first derivative is negative, and the absolute value is the largest) as the maturity date. Then, we determined the restricted time windows corresponding to each key phenological stage for different crops as a 24-day period, with the average dates observed at the nearest AMS as the medians. Finally, we detected the key phenological dates based on the identification of the inflection points and the maximum value point of each LAI time-series curve at each grid within the restricted time windows. Additionally, we restricted the phenology detection for dryland crops (i.e., maize and wheat) and paddy rice on the dryland and paddy field layers derived from the NLCD, respectively.

2.3.2.2. Crop classification

After accomplishing the retrieval of key phenological stages for different crops, we determined the crop-cultivated grids for each crop on the basis of three key phenological stages (mentioned in section 2.3.2.1) that could be identified simultaneously. For example, if the green-up date, heading date and maturity date of winter wheat could be detected simultaneously for a specific grid, then the grid was regarded as a grid with winter wheat cultivation. Then, in some
provinces with complex cropping patterns, spring maize in northern China was easily confused with other spring vegetation. In such cases, we applied a threshold of the maximum LAI based on the characteristic of crop phenology to eliminate the pseudo cultivated grids with a maximum below the threshold for each crop (the bottom right corner of figure 2). Ultimately, we obtained the classification maps for each crop after the quality control that eliminated non-crop grids based on crop-cultivated grids.

2.3.3. Accuracy assessment
The statistical data were widely used for the accuracy assessment of satellite-derived crop planting/harvesting areas at large scales (Bouvet and Thuy 2011, Zhang et al 2014a, Qiu et al 2017b). The county-level statistical data in China were used to evaluate the consistency of the satellite-estimated crop area. First, area estimates were obtained by pixel counting. Then, we calculated the coefficient of determination ($R^2$), the root mean square error (RMSE), and the relative root mean square error (RRMSE) for each item in each year to describe the goodness of the linear relationship and the correspondence between classification maps and statistical data. The computational formulas were expressed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (O_i - E_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (O_i - E_i)^2}{n}} \quad (2)$$

$$RRMSE = \sqrt{\frac{\sum_{i=1}^{n} ((O_i - E_i)/O_i)^2}{n} \times 100} \quad (3)$$

where $O_i$ and $E_i$ are the statistical harvesting area and estimated harvesting area in county $i$, respectively. $n$ represents the total number of counties.

In addition, we spatialized the error to investigate the locations where classification was difficult, using the RRMSE as an indicator to measure error. First, we calculated the RRMSE of the harvesting areas between the GLASS LAI estimates and the statistical data during 2000–2008 for 1269 counties. Next, the previously calculated RRMSE value was assigned to the centroid of each county. Finally, we applied the kriging interpolation method to the RRMSE values of all the centroids to obtain the spatial distribution map of the error using a mask of crop-cultivated grids. The value of the error ranged from 0 to 100, where larger values indicated larger errors.

2.3.4. Spatiotemporal analysis
To investigate the spatiotemporal patterns of the three staple crops in China, we firstly upScaled the grid from 1 km to 5 km. We then calculated the linear trend in crop area percentage per grid by the Theil-Sen estimator. The significance of the trend in each grid was tested using the Mann-Kendall test. Finally, we summarized the area changes with a significant trend ($p < 0.1$) at the pixel and provincial scales to investigate the corresponding characteristics from 2000 to 2015. We also computed the gains and losses in area based on the 5-km grids that had a significant trend.

3. Results

3.1. Validation of the dataset products
The GLASS LAI-derived harvesting areas were compared with the agricultural statistical data of China at the county level. The comparisons illustrated the reliability of the phenology-based crop mapping approach in identifying crop spatial distribution. Overall, the estimated and statistical areas were evenly distributed around the 1:1 line for the three crops. The slopes were close to 1 (figure 4 and figures S1–S3 (stacks.iop.org/ERL/15/074003/mmedia)). In addition, the $R^2$ values were consistently greater than 0.8, with the RMSE and RRMSE values of the estimated areas consistently less than 12 Kha and 45%, respectively (table 1). The average $R^2$ was 0.86,
0.85, and 0.82 for rice, wheat, and maize, respectively, corresponding to an RMSE (RRMSE) of 8.16 Kha (34.81%), 10.02 Kha (36.78%), and 6.9 Kha (42.82%), respectively. Therefore, the crop harvesting area maps derived in this study had a high consistency with the agricultural census data, indicating they were appropriate for monitoring crop dynamics in China.

For the differences in accuracy among crops, the accuracy was the highest for paddy rice, with the largest $R^2$ and the lowest RRMSE. The reasons may be that rice was mainly distributed in paddy fields and paddy fields were not suitable for dryland crops such as maize and wheat. In contrast, the accuracy was the lowest for maize, mainly due to its wider spatial distribution and more heterogeneous planting environment.

Generally, the errors in most regions were lower, with an RRMSE less than 50% in 77% of the grids, while the RMSE was greater than 60% in only 1.5% of the grids for the three crops (figures S4–S6). Additionally, there were larger errors in northwestern China (NW) and in coastal areas than in central and northeastern China (NE).

### 3.2. Spatial characteristics of crop distribution across China

The geographic distributions of the three staple crops across China in 2000, 2005, 2010 and 2015 are presented in figures 5–7. For a certain crop, the interannual variations in the spatial patterns during the 16 years were similar. Paddy rice was more widely distributed, mostly in the Northeast Plain, southwestern China (SW) and southeastern China (figure 5). In the Northeast Plain (14.7% of total harvesting area), the distribution of paddy rice showed a dendritic shape alongside the rivers, with latitudes ranging from 44–48°N and longitudes ranging from 122–128°E and 130–134°E (figure 5 and figures S7(a), (d)). In SW and southeastern China (5.4 times the proportion of harvesting area in the Northeast Plain), paddy rice was mainly concentrated in the Middle-Lower Yangtze plain (YZ), Sichuan Basin (SC), the terraced fields of the Yunnan-Guiizhou Plateau and the Pearl River Delta, with latitudes ranging from 22–33°N and longitudes ranging from 103–122°E (figure 5 and figures S7(a), (d)). The distribution of wheat was concentrated, mainly in the NCP and SC, with latitudes ranging from 30–38°N and longitudes ranging from 111–120°E and 103–105°E, accounting for 85.8% of the total harvesting areas (figure 6 and figures S7(b), (e)). Compared with rice and wheat, maize was most widely distributed throughout China, including the NE, NCP, and SW, with latitudes ranging from 23–49°N and longitudes ranging from 100–127°E and 129–133°E, amounting for 86.7% of the total harvesting areas (figure 7 and figures S7(c), (f)).

### 3.3. Temporal trends in crop harvesting areas in China

We further analyzed the trends in crop harvesting areas at the 5-km grid scale (figures 8(a1)–(a3)) and summarized the grids with a significant trend ($p < 0.1$) according to crops and provincial administrative boundaries (figures 8(b1)–(b3), figure S8, table 2 and tables S1–S3). The positive trends in crop harvesting areas dominated, with an average of 0.38%/year in 64% of the pixels, indicating that crop harvesting area generally increased during the period. The negative trends were weak, with an average of $-0.41%$/year in 36% of pixels. Specifically, the proportion of pixels that had a positive trend in crop harvesting area was 84.4% for maize, 58.7% for rice, and 48% for wheat (table 2).

For paddy rice, the harvesting area increased obviously in northern China (north of 40°N and east of 122°E) but decreased in ST (south of 25°N) (figure 8(a1) and figures S8(a), (c)). The harvesting areas increased in 42% of provinces but decreased in 50% of provinces; nevertheless, there was an increasing trend overall because the provinces with reduced acreage had a lower proportion of harvesting area (i.e. 30%) (table S1). The largest area expansion occurred in Heilongjiang Province at a speed of

| Year | Paddy rice | Wheat | Maize |
|------|------------|-------|-------|
|      | $R^2$ | RMSE (Kha) | RRMSE (%) | $R^2$ | RMSE (Kha) | RRMSE (%) | $R^2$ | RMSE (Kha) | RRMSE (%) |
| 2000 | 0.86 | 8.38 | 35.15 | 0.88 | 8.96 | 37.19 | 0.81 | 7.27 | 43.73 |
| 2001 | 0.83 | 8.42 | 34.70 | 0.85 | 9.74 | 37.81 | 0.81 | 7.72 | 44.19 |
| 2002 | 0.88 | 7.00 | 33.92 | 0.89 | 8.73 | 32.37 | 0.85 | 5.68 | 43.25 |
| 2003 | 0.87 | 7.08 | 33.49 | 0.84 | 10.30 | 37.69 | 0.84 | 6.92 | 43.96 |
| 2004 | 0.85 | 7.98 | 35.19 | 0.86 | 9.86 | 37.38 | 0.82 | 7.10 | 43.57 |
| 2005 | 0.85 | 8.47 | 36.33 | 0.83 | 11.40 | 40.82 | 0.82 | 7.26 | 42.07 |
| 2006 | 0.86 | 8.48 | 33.55 | 0.82 | 10.90 | 37.21 | 0.80 | 6.54 | 42.77 |
| 2007 | 0.84 | 9.05 | 36.37 | 0.83 | 11.10 | 36.39 | 0.82 | 6.57 | 40.25 |
| 2008 | 0.86 | 8.56 | 34.59 | 0.88 | 9.22 | 34.18 | 0.85 | 7.12 | 41.60 |
| Mean | 0.86 | 8.16 | 34.81 | 0.85 | 10.02 | 36.78 | 0.82 | 6.90 | 42.82 |

Note: $R^2$ means coefficients of determination; RMSE means root mean square error; RRMSE means relative root mean square error; Kha, kilo hectare.
70.3 Kha/year, and this province includes the Sanjiang Plain and the Songnen Plain (figure 8(b1), table S1). The area decreased mainly in the coastal economically developed provinces (i.e. Fujian, Shanghai, Zhejiang, Guangxi, and Guangdong provinces) and SW (i.e. Sichuan, Guizhou and Chongqing provinces) with an average decline rate of −11.5 Kha/year, nearly one-sixth of that of Heilongjiang Province (figure 8(b1), table S1).

Contrary to the overall upward trend in the paddy rice harvesting area, the wheat harvesting area generally experienced a downward trend. It had an evident decline in 58% of provinces, while it increased in 29% of provinces, with an overall decreasing trend (table S2). Compared with paddy rice, the trends in wheat harvesting area did not have a spatially explicit pattern along latitudinal gradients (figure S8(b), (c)). Wheat harvesting areas declined significantly in SW, NW, and NE and increased significantly in the south and east of the NCP (figures 8(a2), (b2) and table S2). They declined most in SC and NW (i.e. Gansu, Shaanxi, and Shanxi provinces), with an average rate of −17.5 Kha/year (figure 8(b2), table S2). In contrast, in the NCP (except for Hebei Province with an insignificant decreasing trend) and the Yangtze plain (i.e. Jiangsu and Hubei provinces), accounting for 62% of the total harvesting area in 2015, they increased significantly, with an average rate of 19.4 Kha/year, approximately equal to that of decline (figure 8(b2), table S2).

Completely different from rice and wheat, the maize harvesting area increased significantly in most regions of China, accounting for 76% of all the provinces (table S3). The increasing rate of maize area declined gradually from north to south. The NCP, NW, and NE had an increasing rate that was 13 times that in SW and the YZ (figures 8(a3), (b3) and figure S8(c), (f), and table S3).

The interannual variations in total harvesting areas differed among crops and subregions from
Figure 6. Spatial patterns of wheat cultivation in China for 2000 (a), 2005 (b), 2010 (c) and 2015 (d).

2000 to 2015 (figure 9). The trends in crop harvesting areas based on the crop classification in this study were similar to those based on the official statistics (figures 9(a1)-(a3)). The paddy rice harvesting areas increased at a relatively low rate of 153.8 Kha/year. The harvesting areas of maize increased significantly, at 7 times the rate of rice (figures 9(a1), (a3)). Additionally, the paddy rice areas declined first, reaching a minimum of 25 227.7 Kha in 2003, then gradually increased until 2015 at a rate of 215.6 Kha/year (figure 9(a1)). In contrast, the wheat harvesting areas remained relatively stable over time. It declined sharply from 2000 to 2003 at a rate of −1092.4 Kha/year and then fluctuated slightly at approximately 22 436.8 Kha (figure 9(a2)). Paddy rice harvesting areas expanded aggressively in NE (with an increased rate of 156.9 Kha/year, 6 times the rate of other subregions) and YZ, while they contracted in SC and ST, indicating that the trends in harvesting areas in the north and south were opposite (figure 9(b1)). Wheat harvesting areas declined in most subregions (i.e. NW and SW) but increased significantly in the major wheat production region (i.e. portions of the NCP) (figure 9(b2)). Compared to paddy rice and wheat, the maize harvesting areas in all the subregions increased consistently, especially in NE, NW, and the NCP, with a dramatic increase rate that was 6 times that in SW (figure 9(b3)). For wheat and paddy rice, we separately summarized the total harvesting areas for the provinces where the harvesting areas increased and decreased significantly (p < 0.1), respectively (figures 9(a1), (a2)). Maize was not included because most of the provinces had experienced expansions in harvesting areas, and only a few provinces differed from this trend (figure 8(b3)). The results showed that the increased paddy rice harvesting areas were twice as large as the decreased areas, resulting in a small net
increase (figure 9(a1)). In contrast, the decrease and increase in wheat harvesting areas were roughly offset, resulting in a small change in wheat harvesting areas over the study period (figure 9(a2)).

4. Discussions

4.1. Advantages of the phenology-based crop mapping approach and dataset

In this study, we proposed a phenology-based approach to map the three staple crops in China from 2000 to 2015. The crop phenology-based approach had the following advantages: (1) it could be operated repeatedly over a large area without the need for region- or spectral-dependent classifier rules; (2) it was based on GLASS LAI products, which were more temporally continuous and spatially complete than other remote sensing products. This approach avoid the flooding signals of paddy rice fields that cannot be obtained due to poor data quality; additionally, these flooding signals cause the transplanting-based algorithm to malfunction (Zhou et al 2016). The resultant datasets have been well validated using county-level statistical data, which are promising in terms of providing valuable information for issues including food security, environmental sustainability, climate change, and agricultural policy-making.

4.2. Spatially explicit datasets of crop harvesting areas at a 1 km spatial resolution in China during 2000–2015

The comparisons between the ChinaCropArea1 km and agricultural census data demonstrated a high consistency and credibility. Additionally, the accuracy of the harvesting area maps was the highest for paddy rice, followed by that of wheat and maize. One possible reason for the difference is that paddy rice is mainly planted in paddy fields, which are unfit for dryland crops, resulting in less disturbance by the noise of other vegetation (Dong et al 2015,
Figure 8. Spatial distribution of the trends in crop harvesting area ($p < 0.1$) at 5 km grid scale (a1)–(a3), and the gain, loss areas and slope from summarized results at the provincial scale (b1)–(b3), which are both arranged in the order of paddy rice, wheat, and maize.

Qin et al (2015). The lowest accuracy of maize classification is attributed to the following two reasons: (1) compared with paddy rice and wheat, maize is more widely distributed across the whole country, including NE, the NCP, and SW, with latitudes ranging from 23–49°N and longitudes ranging from 100–127°E and 129–133°E, with complex topography and highly heterogeneous environmental
Table 2. Trends in crop harvesting area percentages from 2000 to 2015.

| Slope | Mean %/a | Percentage % | Mean %/a | Percentage % | Mean %/a |
|-------|----------|--------------|----------|--------------|----------|
| < 0   |          |              | > 0      |              |          |
| Rice  | 32.8     | 41.3         | 0.22     | 58.7         | 0.02     |
| Wheat | 39.5     | 52.0         | 0.46     | 48.0         | −0.02    |
| Maize | 46.8     | 15.6         | 0.46     | 84.4         | 0.32     |

*aPsig means the percent of grids with a significant trend at the level of p < 0.1.

Figure 9. Total Glass LAI-based crop harvesting area in China from 2000 to 2015 (a1)–(a3), and summarized results by sub-regions (b1)–(b3). For (a1)–(a3), the red line represents the total harvesting area, while the orange and blue lines represent the total harvesting area of the provinces where the harvesting area is significantly increased and decreased (p < 0.1), respectively (solid line for estimated area, dashed line for official statistics).

conditions (figure 7 and figures S7(c), (f)). (2) The planting structure of maize is the most complex, such as the single cropping of spring maize and the rotation of summer maize and other crops (e.g. winter wheat) (Qiu et al 2018).

Based on the spatiotemporal analysis of the datasets produced in this study, we found that paddy rice areas expanded significantly in NE during 2000–2015 but decreased significantly in ST. The maize harvesting area increased significantly in most of
the country. For wheat, its harvesting area decreased mainly in SW, NW, and NE but increased significantly in the main wheat production areas (i.e., the NCP and Yangtze plain). These findings are supported by some previous studies (Liu et al 2013, Wei et al 2015, Qiu et al 2017b, 2018). The possible reasons for the large decrease in paddy rice harvesting areas in ST might be the decline in the agricultural labor force, lower net benefit for farmers, urbanization, industrialization, and reduced cropping intensity (from double cropping to single cropping) (Cheng et al 2012, Dong et al 2015, Ning et al 2018). Another factor that cannot be ignored is climate change, which increases the frequency of droughts, flooding, and heat stresses that threaten rice agriculture (Huang et al 2014, Zhang et al 2014c). In contrast, the paddy rice harvesting areas increased apparently in NE due to the conversion of drylands into paddy fields and high subsidies for farmers planting paddy rice (Cheng et al 2012, Ning et al 2018). The overall decline in wheat harvesting areas is ascribed to the dramatic increases in the maize harvesting area, especially in northern China and NE, where maize substantially substitutes wheat. In addition, due to rapid urbanization and the continuous loss of rural labor, large areas of farmland have been abandoned in some regions of China (e.g., SW and the NCP), which is also a vital factor causing the decline of wheat harvesting areas (Zhang et al 2014b, Xiao et al 2019). The main reasons for the substitution of wheat by maize are the exacerbation of water deficits, increasing irrigation costs and the incentives of national agricultural policies (Wei et al 2015, Huang and Yang 2017). Maize is a crop with a high water use efficiency. Maize needs less water for irrigation and can provide the same production amount of wheat (Zwart and Bastiaanssen 2004). Moreover, technological advances and unlimited purchases from the government at the protected price (price intervention policy) lead to sprawling cultivation patterns of maize (Huang and Yang 2017). Although agricultural incentive policies were highly effective in ensuring food security, they broke the balances in planting structures (i.e., the overall decline in the wheat harvesting area) and aggravated the burden on stocks. Thus, in recent years, the government in China has adjusted intervention policies for crop planting structures, aiming to optimize crop planting structure, encourage the cultivation of other minor crops and improve crop production sustainability (Huang and Yang 2017).

4.3. Uncertainty analysis
The ChinaCropArea1 km dataset has been proven to be highly credible and accurate; nevertheless, there are some deficiencies in the study. First, the coarse spatial resolution of 1 km can lead to uncertainties in the results, as several land cover types or crops can be contained within a 1 km pixel. The complex terrain and subpixel composition resulted in a high uncertainty in mapping paddy rice in ST. In contrast, in some areas (e.g., the NCP), the cultivation patterns are complex, but almost all fields follow similar cropping practices and are cultivated with a few main crops. Thus, although the areas of farming fields are small, they have similar spectral signatures and behave like a ‘large field’, which mitigates the mixed pixel effects. Second, fewer grids were identified as crop-cultivated regions in several provinces, with less cultivation of a certain crop due to the lack of crop phenological observations. For instance, phenology observations of paddy rice in Inner Mongolia and Xinjiang are lacking. However, this lack of data might not cause large uncertainties because these areas are not the main production areas of these crops. Finally, we used only crop phenology information as the classification criteria in this study. Recent studies have successfully combined phenological indicators with additional information, such as temperature conditions, to improve the accuracy and efficiency of paddy rice classification (Qin et al 2015, Dong et al 2015). Nevertheless, unique phenomena, such as the requirement of transplanting and the inundation environment of paddy rice, are rare for dryland crops such as wheat and maize, which can be used as additional information for classification (Qiu et al 2018). Moreover, maize is more widely distributed nationwide than is paddy rice, with a higher heterogeneity of the planting environment and a more complex planting structure. This pattern will add complexity to the threshold setting of additional information. Thus, monitoring multiple crops at large scales using a phenology-based approach with additional information is still challenging.

5. Conclusion
To obtain the dynamic changes in the three staple crops in China for a long time, this study presented a new phenology-based crop mapping approach to generate a 1-km crop harvesting area dataset for three staple crops from 2000 to 2015 based on GLASS LAI products, i.e., ChinaCropArea1 km. In comparison with the county-level agricultural statistical data, ChinaCropArea1 km is confirmed to be highly consistent, and the phenology-based approach is robust and reliable. We also quantified the errors (i.e., RRMSE) at a national scale based on the crop classification map, which showed that the errors were low in most regions. Further spatiotemporal analyses revealed several remarkable patterns as follows: (1) paddy rice areas expanded significantly in NE but decreased in ST; (2) wheat harvesting area declined generally although increased significantly in the main wheat production areas; and (3) maize harvesting area increased remarkably in most regions of the country. The drivers underlying the spatiotemporal patterns
included urbanization, industrialization, reduced cropping intensity in ST, frequent disasters driven by climate change, large areas of abandoned farmland in the north and SW, and the incentive of national agricultural policies.

Compared to previous studies, the current research had additional advantages. First, it considered several major crop planting patterns. Second, the phenology-based approach could be applied repeatedly over a large area. Finally, it was based on GLASS LAI products, which have the advantages of high spatiotemporal continuity and completeness. In addition, for the first time, we provided annual and spatially explicit crop cultivation maps for three main crops across China with a resolution of 1 km. The datasets can be used for many purposes, including land surface modeling, agro-ecosystem modeling, agricultural production prediction, and agricultural policy-making.

Acknowledgments

This study was supported by the National Key Research and Development Program of China (project nos. 2019YFA0607401, 2017YFA0604703, and 2017YFD0300301) and the National Basic Research Program of China (Grant Nos. 41977405, 31561143003, 31761143006).

Data availability

The datasets can be used for many purposes, including land surface modeling, agro-ecosystem modeling, agricultural production prediction, and agricultural policy-making.

References

Asseng S et al 2015 Rising temperatures reduce global wheat production Nat. Clim. Change 5 143–7
Azzari G, Jain M and Lobell D B 2017 Towards fine resolution maps of crop yield: testing multiple methods and satellites in three countries Remote Sens. Environ. 202 129–41
Baret F and Guyot G 1991 Potentials and limits of vegetation indexes for LAI and APAR assessment Remote Sens. Environ. 35 161–73
Bouvet A and Thuy L T 2011 Use of ENVISAT/ASAR wide-swath data for timely rice fields mapping in the Mekong river Delta Remote Sens. Environ. 115 1090–101
Brown J C, Kastens J H, Coutinho A C, Victoria D D and Bishop C R 2013 Classifying multiyear agricultural land use data from Mato Grosso using time-series MODIS vegetation index data Remote Sens. Environ. 130 39–50
Chang J, Hansen M C, Pittman K, Carroll M and Dimiceli C 2007 Corn and soybean mapping in the united states using MODN time-series data sets Agron. J. 99 1654–64
Chen Y, Zhang X and Tao F L 2018 Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data Eur. J. Agron. 101 163–73
Cheng Y, Wang X, Guo J, Zhao Y and Huang J 2012 The temporal-spatial dynamic analysis of China rice production Sci. Agric. Sinica 45 3473–85
Conrad C, Fritsch S, Zeidler J, Rucker G and Dech S 2010 Per-field irrigated crop classification in arid central Asia using SPOT and ASTER data Remote Sens. Based 2 1035–56
Dong J W et al 2015 Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms Remote Sens. Environ. 160 99–113
FAO 2016 FAOSTAT Online Database http://faostat.fao.org/ (accessed April 2020)
Frolking S, Qui J J, Boles S, Xiao X M, Liu J Y, Zhuang Y H, Li C S and Qin X G 2002 Combining remote sensing and ground census data to develop new maps of the distribution of rice agriculture in China Glob. Biogeochern. Cycles 16
Geerken R A 2009 An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change ISPRS J. Photogramm. 64 422–31
Gilmanov T G et al 2013 CO₂ uptake and ecophysiological parameters of the grain crops of midcontinent North America: estimates from flux tower measurements Agric. Ecosyst. Environ. 164 162–75
Gumma M K, Thenkabail P S, Maunahan A, Islam S and Nelson A 2014 Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500 m data for the year 2010 ISPRS J. Photogramm. 91 98–111
Gumma M K, Thenkabail P S, Teluguntla P, Rao M N., Mohammed I A and Whitbread A M 2016 Mapping rice-fallow cropland areas for short-season grain legumes intensification in South Asia using MODIS 250 m time-series data Int. J. Digit Earth 9 981–1003
Hao P Y, Zhan Y L, Wang L, Niu Z and Shakir M 2015 Feature selection of time series MODIS data for early crop classification using random forest: a case study in Kansas, USA Remote Sens. Basel 7 5347–69
Howard D M and Wylie B K 2014 Annual crop type classification of the US great plains for 2000 to 2011 Photogramm. Eng. Remote Sens. 80 537–49
Huang J K and Yang G L 2017 Understanding recent challenges and new food policy in China Glob. Food Secur. Agric. 12 119–26
Huang W, Sui Y, Yang X, Dai S, Qu H and Li M 2014 Spatio-temporal characteristics of crop drought in southern China based on drought index of continuous days without available precipitation Trans. of the CSAFE 30 125–35
Johnson D M and Mueller R 2010 The 2009 cropland data layer Photogramm. Eng. Remote Sens. 76 1201–5
Kuenzer C and Knauer K 2013 Remote sensing of rice crop areas Int. J. Remote Sens. 34 2101–39
Li L, Friedli M A, Xin Q C, Gray J, Pan Y Z and Frolking S 2014 Mapping crop cycles in China using MODIS-EVI time series Remote Sens. Basel 6 2473–93
Li S L et al 2013 A long-term global land surface satellite (GLASS) data-set for environmental studies Int. J. Digit Earth 6 53
Liu H, Gong P, Wang J, Clinton N, Bai Y and Liang S 2019 Annual dynamics of global land cover and its long-term changes from 1982 to 2015 Earth Syst. Sci. Data Discuss. 2019 1–54
Liu J Y et al 2014 Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980 J. Geogr. Sci. 24 195–210
Liu J Y, Liu M L, Tian H Q, Zhang D F, Zhang Z X, Zhang W, Tang X M and Deng X Z 2005 Spatial and temporal patterns of China's cropland during 1990-2000: an analysis based on Landsat TM data Remote Sens. Environ. 98 442–56
Liu Z H, Li Z G, Tang P Q, Li Z P, Wu W B, Yang P, You L Z and Tang H J 2013 Change analysis of rice area and production in China during the past three decades J. Geogr. Sci. 23 1005–18
Lobell D B, Thau D, Seifert C, Engle E and Little B 2015 A scalable satellite-based crop yield mapper Remote Sens. Environ. 164 324–33
Luo Y C, Zhang Z, Chen Y, Li Z Y and Tao F L 2020 ChinacropPhen1km: a high-resolution crop phenological

14
dataset for three staple crops in China during 2000–2015 based on leaf area index (LAI) products Earth Syst. Sci. Data 12 197–214
Myneni R B and Williams D L 1994 On the relationship between Fapar and Ndvi Remote Sens. Environ. 49 200–11
Neeti N, Rogan J, Christman Z, Eastman J R, Milliones M, Schneider L, Nickl E, Schmoo B, Turner B L and Ghimire B 2012 Mapping seasonal trends in vegetation using AVHRR-NDVI time series in the Yucatan Peninsula, Mexico Remote Sens. Lett. 3 433–42
Ning J et al 2018 Spatiotemporal patterns and characteristics of land-use change in China during 2010–2015 J. Geogr. Sci. 28 547–62
Ozdogan M 2010 The spatial distribution of crop types from MODIS data: temporal unmixing using independent component analysis Remote Sens. Environ. 114 1190–204
Pan Y Z, Li L, Zhang J S, Liang S L, Zhu X F and Sulla-Menashe D 2012 Winter wheat area estimation from MODIS-EVI time series data using the crop proportion phenology index Remote Sens. Environ. 119 232–42
Piao S L et al 2010 The impacts of climate change on water resources and agriculture in China Nature 467 43–51
Qin Y W et al 2015 Mapping paddy rice planting area in cold temperate climate region through analysis of time series Landsat 8 (OLI), Landsat 7 (ETM+) and MODIS imagery ISPRS J. Photogramm. 105 220–33
Qiu B W, Huang Y Z, Chen C C, Tang Z H and Zou F L 2018 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. 153 82–93
Qiu B W, Lu D F, Tang Z H, Chen C C and Zou F L 2017a Automatic and adaptive paddy rice mapping using Landsat images: case study in Songyuan Plain in Northeast China Sci. Total Environ. 598 581–92
Qiu B W, Luo Y H, Tang Z H, Chen C C, Lu D F, Huang H Y, Chen Y Z, Chen N and Xu W M 2017b Winter wheat mapping combining variations before and after estimated heading dates ISPRS J. Photogramm. 123 35–46
Ray D K, Gerber J S, Macdonald G K and West P C 2015 Climate variation explains a third of global crop yield variation Nat. Commun. 6
Richardson A D, Braswell B H, Hollinger D Y, Jenkins J P and Ollinger S V 2009 Near-surface remote sensing of spatial and temporal variation in canopy phenology Ecol. Appl. 19 1417–28
Sakamoto T 2018 Refined shape model fitting methods for detecting various types of phenological information on major US crops ISPRS J. Photogramm. 138 176–92
Sakamoto T, Gitelson A A and Arkebauer T J 2014 Near real-time prediction of U.S. corn yields based on time-series MODIS data Remote Sens. Environ. 147 219–31
Sakamoto T, Wardlow B D, Gitelson A A, Verma S B, Suyker A E and Arkebauer T J 2010 A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data Remote Sens. Environ. 114 2146–59
Sakamoto T, Yokozawa M, Toritani H, Shibayama M, Ishitsuka N and Ohno H 2005 A crop phenology detection method using time-series MODIS data Remote Sens. Environ. 96 366–74
Son N T, Chen C F, Chen C R, Duc H N and Chang L Y 2014 A phenology-based classification of Time-series MODIS data for rice crop monitoring in Mekong Delta, Vietnam Remote Senses. Baseline 6 135–56
Song X P, Hansen M C, Stehman S V, Potapov P V, Tyukavina A, Vermote E J and Townshend J R 2018 Global land change from 1982 to 2016 Nature 560 639–+ Taro F L, Zhang S, Zhang Z and Botter P P 2014 Maize growing duration was prolonged across China in the past three decades under the combined effects of temperature, agronomic management, and cultivar shift Glob. Change Biol. 20 3686–99
Wang C Z, Zhong C and Yang Z W 2014 Assessing bioenergy-driven agricultural land use change and biomass quantities in the US Midwest with MODIS time series J. Appl. Remote Senses. 8 Wardlow B D, Eghbert S L, and Kastens J H 2007 Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US central great plains Remote Sens. Environ. 108 290–310
Wei X, Zhang Z, Shi P J, Wang P, Chen Y, Song X and Tao F L 2015 Is yield increase sufficient to achieve food security in China? PLoS One 10 e0221167
Xiao G F, Zhu X F, Hou C Y and Xia X S 2019 Extraction and analysis of abandoned farmland: a case study of Qingyun and Wudi counties in shandong province J. Geogr. Sci. 29 581–97
Xiao X M, Boles S, Frolking S, Li C S, Babu Y, Salas W and Moore B 2006 Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images Remote Sens. Environ. 109 95–113
Xiao Z Q, Liang S L, Wang J D, Chen P, Yin X J, Zhang L Q and Song J L 2014 Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance IEEE Trans. Geosci. Remote Sens. 52 209–23
Zhang G L et al 2017 Spatiotemporal patterns of paddy rice croplands in China and India from 2000 to 2015 Sci. Total Environ. 579 82–92
Zhang J H, Feng L L and Yao F M 2014a Improved maize cultivated area estimation over a large scale combining MODIS-EVI time series data and crop phenological information ISPRS J. Photogramm. 94 102–13
Zhang Y, Li X and Song W 2014b Determinants of cropland abandonment at the parcel, household and village levels in mountain areas of China: a multi-level analysis Land Use Policy 41 186–92
Zhang Z, Wang P, Chen Y, Song X, Wei X and Shi P J 2014c Global warming over 1960–2009 did increase heat stress and reduce cold stress in the major rice-planting areas across China Eur. J. Agron. 59 49–56
Zhong L H, Yu L, Li X C, Hu L N and Gong P 2016 Rapid corn water productivity values for irrigated wheat, rice, cotton and maize Agric. Water Manage. 69 115–33
