Patterns and causes of winter wheat and summer maize rotation area change over the North China Plain

Zhengjia Liu1,2,*, Yansui Liu1,*, Jinwei Dong1,2,*, Muhammad Hasan Ali Baig4, Wenhong Chi3, Liuying Peng5 and Jieyong Wang1,2,*

1 Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China
3 College of Resources and Environmental Economics, Inner Mongolia University of Finance and Economics, Inner Mongolia, Hohhot 010070, People's Republic of China
4 Institute of Geo-Information and Earth Observation (IGEO), Arid Agriculture University, Rawalpindi 46300, Pakistan
5 School of Economics, Shandong University of Finance and Economics, Shandong, Jinan 250014, People's Republic of China
* Authors to whom any correspondence should be addressed.
E-mail: liuzj@igsnrr.ac.cn and wjy@igsnrr.ac.cn

Keywords: crop phenology, MODIS NDVI anomaly, ensemble learning, cropping frequency, North China plain

Abstract
Cropland area and cropping frequency play very crucial roles in determining regional food production. However, rapid urbanization accompanied by declining surplus-agricultural labor force has greatly altered patterns of agriculture land use and cropping frequency. Due to lack of continuous cropland and cropland-use intensity maps, our knowledge is still limited to understand whether the urbanization process must have a negative effect for changes in cropland-use intensity. Herein, we took the North China Plain (NCP), both the largest winter wheat and summer maize rotation area and rapidly urbanized area in China, as the study area, and used 250 m moderate resolution imaging spectroradiometer NDVI anomaly data, the correlation of NDVI time series in two neighboring years and machine learning algorithms to investigate spatiotemporal patterns and trends of cropland area and cropping frequency change over the NCP from 2000 to 2019. Results showed a significantly decreased cropland area observed since 2004 (slope = 783.8 km² a⁻¹, p < 0.01), while area of double-season cropping presented a relatively steady trend (slope = 446.9 km² a⁻¹, p = 0.335). As expected, decreased croplands were mainly occupied by urban and built-up land expansion, however, existing cropland-use intensity was yet improved. Patterns and trends of double-season cropping types were varied spatially. Particularly, the area of winter wheat and summer maize rotation presented a significantly increasing trend (slope = 3423.3 km² a⁻¹, p < 0.01) in entire period. Land-use and grain subsidy policies are considered as largely responsible for this phenomenon. These satellite-observed findings highlight that positive land-use policies and managements will be helpful for profitably keeping/improving the harvest area.

1. Introduction
Population growth and rapid socio-economic development pose a tremendous challenge for global agricultural systems and food supply (Godfray et al 2010, Bryan et al 2018, Liu 2018, Abu Hatab et al 2019). Improving cropland-use intensity and expanding cropland area are regarded as the traditional and widely-used means to maximize crop yields in most regions of the globe (Lobell et al 2008, Ray and Foley 2013, Spera et al 2014, Cohn et al 2016). Double cropping plays a crucial role in meeting food demands of...
the temperate zones and it is also an important evaluation indicator of land-use intensification. However, rapid urbanization associated with declining surplus-agricultural labor force has greatly altered patterns of agriculture land use and cropping frequency. This process was usually regarded as having a negative effect for changes of cropland area and/or cropland-use intensity (Cui et al 2019). For example, some rural household-based surveys found that some farmers passively or actively give up farming due to rapid urbanization and sagging grain prices (Spera et al 2014, Liu et al 2014b, Liu 2018). Croplands in some regions were readily swallowed by urban and built-up lands (Liu et al 2014a, Ning et al 2018) or appeared abnormally with low-intensity use, such as single-season or double-season agricultural land abandonment instead of intensive cultivation (Prishchepov et al 2012, 2013, Spera et al 2014). Yet, it is worth noting that lack of available large-area maps could largely limit our systematic understanding for changes of cropland area and cropping frequency in response to rapid urbanization. Therefore, lots of communities, including remote sensing, agriculture, land system science and climate change, have paid more attention to studies of cropland area and cropping frequency changes in recent years (Ray and Foley 2013, Spera et al 2014, Cohn et al 2016, Wu et al 2018).

Satellite-based remote sensing data have provided great opportunities to monitor large-area changes of land-use types, given their features of open access, continuous time series and wide-coverage observations (Mueller-Warrant et al 2016, Leroux et al 2017, Waldhoff et al 2017, Lasko et al 2018, Weiss et al 2020). Among the satellite data, 30 m Landsat, 10 m Sentinel and 250 m moderate resolution imaging spectroradiometer (MODIS) are the most widely applied for the related studies of crop mapping and monitoring (Wardlow et al 2007, Spera et al 2014, Dong et al 2015, Veloso et al 2017, Chen et al 2018, Hao et al 2018). In comparison with the data of medium-high resolution Landsat and Sentinel satellites, MODIS data have a slightly coarse spatial resolution but have much better-quality time series information owing to relatively weaker impacts by the noise, e.g. cloud contamination. By contrast, MODIS data thus have greater potentials to map croplands and cropping frequency in view of the importance of time series information for agriculture (Wardlow et al 2007, Cohn et al 2016, Chen et al 2018). Some studies employed 250 m MODIS time series data to investigate the changes of cropping frequency in Mato Grosso state of Brazil and they also provided very valuable information for us to map cropland area and crop types by using time series data (Spera et al 2014, Cohn et al 2016). The automatic and continuous mapping, by directly using time series vegetation index, could be effective in relatively short-term mapping. Moreover, CO₂ fertilization and land management practices reportedly increase the vegetation biomass (Liu et al 2020). Our preliminary tests also found that the above-mentioned effects cannot be ignored for the long-term continuous mapping of crop types. Fortunately, annual vegetation index anomaly could minimize these impacts of CO₂ fertilization and land management as they are usually calculated by using time series values subtracting averaged value each year. To do so, we can better reserve the temporal behavior of each crop, which is important for improving specific-crop classification. On the other hand, proper approaches are also very crucial for cropland-use and cropping frequency mapping. Previously, machine learning approaches and Twi-difference approach are usually used to map cropping frequency (Panigrahy and Sharma 1997, Wardlow and Egbert 2008, Yang et al 2013, Spera et al 2014). Compared to Twi-difference approach, machine learning approaches are more promising. When we map land-use types in a specific region, a core limitation is that there are rarely fine-resolution maps of entire study area available to assess the robustness of inferred land-use maps (Lunetta et al 2010). One of the methods being practiced by researchers to tackle this challenge is to use results from multiple machine learning algorithms (or termed as ensemble learning). Using the base map derived from multiple machine learning algorithms as well as the image change detection analysis method, we can produce the continuous maps.

In this study, our aims are: (a) to produce annual map of cropland and double cropping area across the North China Plain (NCP) by using machine learning algorithms and corresponding time series of MODIS-NDVI anomaly data, (b) to investigate the spatiotemporal dynamics of cropland and double cropping in the context of rapid urbanization, and (c) to clarify the reasons behind changes in croplands and double cropping practices in 2000–2019.

2. Data and methods

2.1. Study area

The NCP (32°00’ N–40° 24’ N, and 112°48’ E–122°45’ E) is located in the eastern part of China (figure 1), covering southern Beijing municipality, Tianjin municipality, southeastern Hebei province, eastern Henan province, Shandong province, northeastern Anhui province and northern Jiangsu province. The overall area of this study is approximately 40 × 10⁶ km² with the average elevation of ~50 m above the sea level. It has a typical temperate monsoonal climate with annual precipitation of 500–1000 mm and annual averaged temperature of 8 °C–15 °C. The prevailing double cropping system is winter wheat and summer maize rotation. It is one of the most important granaries of China (Wang et al 2018).
2.2. Remote sensing data and preprocessing

The 250 m MODIS vegetation index collection 6 (C6) data with 16 d interval (MOD13Q1), covering the period of 2000–2019, were collected from the Land Processes Distributed Active Archive Center. An earlier study has evidently reported that multi-temporal vegetation index data track similar seasonal responses for the crop growth, and are highly correlated across the growing season (Wardlow et al. 2007). Therefore, this study used NDVI data in view of its higher values and more widely-used characteristics (Wardlow and Egbert 2008, Estel et al. 2015, Chen et al. 2018).

The noises induced by poor atmospheric conditions and cloud contamination are inevitable in satellite-based data. To minimize the impacts of noises in time series satellite-based data, previous studies suggested the advantages of using modified Savitzky–Golay (mSG) filter for reconstructing a high-quality 16 d NDVI data (Chen et al. 2004, Liu et al. 2017, Wang et al. 2017). Thus, this study first employed mSG filter with smoothing window size of $2m + 1$ ($m = 4$) and quartic polynomial to smooth all NDVI data.

Besides, some auxiliary data, e.g. existed 1 km land-use data and 300 m land cover data, were used for the cross-validation of typical-year cropland distributions. The 1 km spatial resolution land-use data in 2000, 2010 and 2018, derived from Landsat TM/ETM+ images with the help of human-computer interaction mapping technique, were collected from Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (CAS), thus were also known as CAS land use data (Liu et al. 2014a, Ning et al. 2018, Cui et al. 2019). The first-class land-use types include forests, grasslands, croplands, urban and built-up land, water body and unused land. The dataset has widely been used in previous studies of China’s land-use management and change monitoring (Liu et al. 2014a, Ning et al. 2018, Cui et al. 2019). The 300 m climate change initiative (CCI) land cover products were collected from the European Space Agency. Croplands (label codes of 10, 11, 12, 20 and 30) and urban areas (label codes of 190) of CCI data were used to help investigate the impacts of urbanization on cropland area.

2.3. Croplands and specific crops mapping

2.3.1. Cropland mapping

The primary land-use types in this study area included forests, grasslands, croplands, built-up lands, barren lands and water bodies (Liu et al. 2014a, Ning et al. 2018). In this study, forests, grasslands, built-up land, barren land and water body were also regarded as non-croplands. A random sample selection was conducted to collect training pixels for each land-use type. The total number of 2993 samples were manually labeled with the help of 2013–2019 fine-resolution images of Google Earth by considering the dominated land-use types of each 250 m pixel (i.e. over 90% homogeneity), including 1234 cropland samples and 1759 non-croplands samples. More
than 45% of samples were from the year of 2017. This study used the latitude, longitude and time information of samples to extract the corresponding time series of NDVI anomaly data

$$\text{ANDVI}_i = \text{NDVI}_i - \bar{\text{NDVI}}.$$  \hspace{1cm} (1)

where ANDVI$_i$ and NDVI$_i$ are the $i$th NDVI anomaly and NDVI values, and $\bar{\text{NDVI}}$ is the averaged NDVI during Julian days 65–289 of each year.

A hierarchical mapping strategy was used in this study, i.e. mapping annual cropland mask and then mapping annual crop types with the help of annual cropland mask (figure 2). To classify croplands and non-croplands, three simple but efficient machine learning algorithms, including a feedforward neural network (NN) algorithm, a decision tree (DT) algorithm and a support vector machine (SVM) algorithm, were used in this study. Specifically, the NN algorithm included 15 input nodes corresponding to 15 time series MODIS-NDVI anomaly images from Julian days 65–289 (i.e. the growing season from March to October). The remaining data were discarded in view of low information contents (Lunetta et al 2010). Reducing redundancy input nodes is also helpful for improving the operating efficiency of NN algorithm. In the hidden layer, this study set seven nodes. The output layer included two nodes, namely croplands and non-croplands. The NN classifier was trained by using 2/3 data of all samples collected from Google Earth images. The remaining 1/3 data were used for validating the performance of the NN classifier. Meanwhile, the DT and SVM algorithms were also employed to extract the cropland mask. Similarly, 2/3 data of all samples were trained to obtain the threshold values, and the remaining 1/3 data were used for validating the performance of the classifier. These trained NN-based, DT-based and SVM-based classifiers were subsequently employed to classify croplands and non-croplands of the entire study area in 2017 (considering the most training data from the year of 2017), respectively. A decision rule of cropland mask is defined as two croplands out of three predictions at least. A final cropland mask map in 2017 was obtained according to this decision rule of the ensemble learning.

Furthermore, based on the derived cropland map in 2017, we updated the cropland mask map in 2016 and 2018, respectively. Taking the cropland mapping in 2016 as a case, the specific updated method was detailed explained. This study firstly calculated the correlation and significance of NDVI of Julian days 65–289 and the bias values of maximum NDVI (NDVImax) between 2016 and 2017. For a specific pixel, if time-series NDVI data between 2016 and 2017 had a significant correlation ($p < 0.05$) and the bias absolute value of 2 year NDVImax was less than 0.1 (∼1 standard deviation of cropland NDVImax, figure S1 (available online at stacks.iop.org/ERL/17/044056/mmedia)), land-use type of the specific pixel was not updated. Otherwise, the pixel would be reclassified by using the best performance algorithm out of the three algorithms (e.g. NN algorithm, see the following section 3.1) as well as corresponding NDVI anomaly data in 2016. Using the similar method, we updated the cropland mask map in 2015 based on the derived cropland map in 2016. Similarly, we finally produced all cropland maps in 2000–2019.
2.3.2. Peak information extraction
A 6\(^{\circ}\) polynomial fitting algorithm was used to fit the curve of vegetation growth for Julian days 65–289 and reconstruct daily NDVI time series year by year (Liu and Wang 2018)

\[ y(\text{DOY}) = p_0 + p_1 \cdot X + p_2 \cdot X^2 + \cdots + p_n \cdot X^n. \]  

\[ y'(\text{DOY}) = p_1 + 2p_2 \cdot X + \cdots + np_n \cdot X^{n-1}. \]

where \( y(\text{DOY}) \) is the daily NDVI value at time DOY; \( Y \) is the normalized value of corresponding Julian (e.g. when \( \text{DOY} \) is 1, \( X \) is 1/365), \( p_0, p_1, \ldots, p_n \) are the fitting parameters; \( n \) is 6 in this study. The corrected Akaike Information Criterion (AICc) (Akaike 1974, Hurvich and Tsai 1989), root mean square error (RMSE) and correlation coefficient (\( R \)) were used to test potential overfitting risk of the six-degree polynomial fitting algorithm. To obtain the \( \text{DOY} \) and corresponding NDVI values of peaks and valleys in the curve of vegetation growth, the first derivative of function \( y \) was computed by using the following function

2.3.3. Crop-specific type mapping
Winter wheat and summer maize, two major crop types, were considered in this study. Referring to the NCP farming calendar, winter wheat is usually harvested in May and June and subsequently farmers grow summer maize. For ease of classification, this study is subjectively divided into two periods by using \( \text{DOY} 177 \). That is, data from \( \text{DOY} 65-177 \) (a total of eight images) were used for classification of winter wheat and non-wheat, and data from \( \text{DOY} 177-289 \) (a total of eight images) for classification of summer maize and non-maize.

This study collected 1767 winter wheat samples and 244 non-wheat samples in growing period of winter wheat, and 1396 summer maize samples and 655 non-maize samples in growing period of summer maize, from the field survey and high-resolution images (unmanned aerial vehicle and GF-2 images) in 2019. The NN algorithm was separately used for classifications of winter wheat and summer maize in view of the best performance in three algorithms (see the following section 3.1). This study took the classification of winter wheat in 2019 as a case to state the specific steps. This study first used the latitude and longitude information of samples to extract the corresponding time series of NDVI anomaly data in 2019. Then, this study employed 2/3 winter wheat and non-wheat samples to train the NN classifier, and the remaining 1/3 data were used for validating the performance of NN classifier. The NN classifier included eight input nodes corresponding to eight time-series based on NDVI anomaly images from \( \text{DOY} 65–177 \). In the hidden layer, this study set four nodes. The output layer included two nodes, namely winter wheat and non-wheat. Third, the trained NN classifier as well as 2019 cropland map and peak information was employed to map winter wheat and non-wheat across the entire study area in 2019. Fourth, based on the derived winter wheat map in 2019 as well as cropland map and peak information in 2018, we updated the winter wheat map mask in 2018. Similar to updating cropland map, this study calculated the correlation and significance of NDVI of \( \text{DOY} 65–177 \) and the bias absolute value of NDVImax between 2018 and 2019. For a specific pixel, if it has a significant correlation (\( p < 0.05 \)) and the bias absolute value of NDVImax less than 0.1, its land-use type will not be updated. Otherwise, the pixel will be reclassified by using the NN classifier as well as corresponding NDVI anomaly data in 2018. Repeating the fourth step, we finally produced all winter wheat and non-wheat maps in 2000–2019. Similar to winter wheat mapping, this study also produced all summer maize and non-maize maps from 2000 to 2019.

2.3.4. Validation of winter wheat and summer maize mapping
In this study, the city-level statistical planting area data of winter wheat and summer maize form statistical yearbooks during the year of 2000–2018 and mapping products of previous studies, i.e. 2016–2019 30 m winter wheat maps of Dong et al (2020) and 2000–2015 1 km winter wheat and summer maize maps of Luo et al (2020), were employed to validate the robustness of mapping results. In view of the continuity of integrity of statistical yearbook data, if data in a city were less than five years, the corresponding data were removed from the final validation dataset. Therefore, high-quality data from 32 cities of NCP were finally used in this study, covering four cities of Hebei province (2007–2011), five cities of Henan province (2000–2018), 17 counties of Shandong province (2003–2017), five cities of Anhui province (2000–2018) and one city of Jiangsu province (2000–2018). Besides, purchase prices (2001–2018), planting area (1999–2020), total yield (1999–2020) of dominant grain crops (including wheat, maize, soybean and rice) were collected from the statistical yearbook of Shandong, Henan and Hebei provinces. This study
further calculated three-province averaged crop purchase prices (yuan kg\(^{-1}\)), per-unit yield (kg ha\(^{-1}\)) and per-unit economic benefits (yuan ha\(^{-1}\)), respectively.

2.4. Statistical analyses

The linear relationship between estimated data and statistical data, determination coefficient (\(R^2\)) and statistical significance were used to assess the overall performance of satellite-based algorithm. Of them, the student’s \(t\)-test was used to calculate the significance at the 0.05 or 0.01 level. The user’s accuracy (UA), producer’s accuracy (PA), overall accuracy (OA) and Kappa coefficient \((k)\) were used to assess the algorithm performances.

\[
UA_i = \frac{X_{i,j}}{\sum_{j=1}^{n} X_{i,j}} \cdot 100\%
\]  
\[
PA_j = \frac{X_{i,j}}{\sum_{i=1}^{n} X_{i,j}} \cdot 100\%
\]  
\[
OA = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,j}}{\sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,j}} \cdot 100\%
\]

\[
k = \frac{n \cdot \sum_{i=1}^{n} X_{i,j} - \sum_{i=1}^{n} \left( \sum_{j=1}^{n} X_{i,j} \cdot \sum_{j=1}^{n} X_{i,j} \right)}{n^2 - \sum_{i=1}^{n} \left( \sum_{j=1}^{n} X_{i,j} \cdot \sum_{j=1}^{n} X_{i,j} \right)}
\]

where \(n\) is the total number of all classifications; \(X_{i,j}\) is an element, locating at \(i\)th row and \(j\)th column of the confusion matrix (or error matrix).

This study aggregated 250 m maps of croplands, double cropping area, and winter wheat and summer maize rotation area into corresponding fractional (%) maps within a \(40 \times 40\) pixel window (\(\sim 10 \times 10\) km\(^2\)) and then computed a map of change rates of croplands, double cropping area, and winter wheat and summer maize rotation area from 2000 to 2019 by using the linear regression method.

3. Results

3.1. Analyses of cropland classifications, crop classifications and cropping patterns

Three classifiers used in this study all showed good performances in the training phase. The overall accuracies were separately 90.9%, 88.0% 88.9% for NN-based, DT-based and SVM-based classifiers, while corresponding Kappa coefficients were 0.81, 0.75, 0.78, respectively. The NN-based croplands gave the best user’s accuracy (92.6%) and producer’s accuracy (86.0%) compared to those from DT-based and SVM-based classifiers. Given the best performance of NN-based classifier, it was thus used for updating cropland mask and producing crop-specific maps referring to the described steps of method section 2.2.3. Regarding NN-based winter wheat, the overall accuracy, user’s accuracy and producer’s accuracy were 94.5%, 97.3% and 96.4%, respectively, and the Kappa coefficient was 0.76. Regarding NN-based summer maize, the overall accuracy, user’s accuracy and producer’s accuracy were 92.2%, 93.1% and 95.7%, respectively, and the Kappa coefficient was 0.82. In addition, this study also validated the robustness of satellite-based predicted areas of winter wheat and summer maize against 32 city-level statistical planting area of winter wheat and summer maize from the NCP and compared with previous winter wheat and summer maize maps. Satellite-based predicted area of winter wheat and summer maize both performed good agreements against statistical data, as \(R^2\) showed high values with values of 0.89 and 0.81 (\(p < 0.01\)), and the linear regression-based slopes being close to line 1:1 with 0.84 and 0.69 (figure S2).

Figure 3(a) shows the NCP cropland mask map in 2017 according to the predetermined decision rule that at least two croplands out of three predictions from trained NN-based, DT-based and SVM-based classifiers. Figure 3(b) shows that double-season cropping patterns were more located in the western and southern NCP, while the single-season cropping patterns were mainly located in the mountainous regions of mid-eastern NCP and coastal regions. Regions with winter wheat and summer maize rotation were mostly concentrated upon western and southwestern NCP, namely southern Hebei province, western Shandong province, eastern Henan province, and northern Anhui and Jiangsu provinces.

3.2. Annual area changes of croplands, single-/double-season cropping, and winter wheat and summer maize rotation

Net cropland area in the NCP presented a significant decreasing trend with slope of \(-783.8 \text{ km}^2 \text{ a}^{-1}\) (\(p < 0.01\)) during 2004–2019, although it displayed a slightly increasing trend in the entire study period (\(p = 0.211\)) because of rapid cropland expansion before the year of 2004 (figure 4(a)). Figure 4(b) shows that single-season cropping area displayed a slightly decreasing trend, and that trend was much steeper after the year of 2004 (slope = \(-938.3 \text{ km}^2 \text{ a}^{-1}, p = 0.111\)) compared to that of the entire study period (slope = \(-61.0 \text{ km}^2 \text{ a}^{-1}, p = 0.893\)). In contrast, the cropland loss did not directly decrease double-season cropping area. Double-season cropping area always remained a slightly increasing trend during the entire study period (slope = \(446.9 \text{ km}^2 \text{ a}^{-1}, p = 0.335\)). Particularly, when considering double-season cropping types, this study, however, observed that the area of the winter wheat and summer maize rotation presented a significantly increasing trend (slope = \(3423.3 \text{ km}^2 \text{ a}^{-1}, p < 0.01\)). The area of other three double-season cropping types, including winter wheat and non-maize rotation, non-wheat and summer maize rotation and non-wheat and non-maize rotation, were yet all significantly decreasing trends (figure 4(c)).
other words, double-season cropping area remained relatively stable, but double-season cropping types had respective patterns and trends. This study also found that overall area of winter wheat and overall area of summer maize both exhibited significantly increasing trends with slope of $2953.8 \text{ km}^2 \text{ a}^{-1}$ and $2002.9 \text{ km}^2 \text{ a}^{-1}$ ($p < 0.01$) in the entire study period.

To further clarify the spatial changes of croplands (figures 5(a)–(d), double-season cropping (figures 5(b)–(e)), and winter wheat and summer maize rotation (figures 5(c)–(f)), this study also investigated their corresponding patterns and trends at a $10 \text{ km} \times 10 \text{ km}$ grid cell scale. This study observed that unlike patterns of cropland and double-season cropping with areas being more than $40 \text{ km}^2$, winter wheat and summer maize rotation area more focused on the southern and western NCP. In 2000–2019, areal trends of cropland and double-season cropping showed similar spatial distributions. Their increased regions were mainly concentrated on northeastern and southern NCP. Peri-urban regions however presented significantly decreased cropland and double cropping areal trends. In contrast, significantly increasing regions of winter wheat and summer maize rotation mostly covered plain regions with elevation being less than $50 \text{ m}$.

4. Discussion

4.1. Reliabilities and uncertainties of cropland and crop-specific type mapping
In this study, 250 m MODIS data and the best performed machine learning approach (i.e. NN algorithm) were used to derive annual cropland and
crop-specific type maps. We further proved the reliabilities of satellite-derived cropland and crop type mapping. The good performances could be more contributed to the following two reasons. On the one hand, different from some previous cropland mapping studies that only relied on one machine learning approach, this study employed three typical approaches (i.e. NN, DT and SVM algorithms), time series MODIS NDVI anomaly data and a new decision rule. Taking cropland mapping as an example, that is, cropland mask is defined as two croplands out of three predictions, at least. This kind of treatment declined the uncertainty of derived result that only relied on single machine learning approach to some extent. The methods used by previous studies that employed time series vegetation index values to map croplands or crop types, could be more effective in relatively short term (Wardlow et al 2007, Skakun et al 2017, Chen et al 2018). Because they could overlook the greening effects of CO₂ fertilization and land management for improving the vegetation biomass (Chen et al 2019, Liu et al 2020). The annual vegetation index anomaly could largely minimize impacts of CO₂ fertilization and land management, and thus likely help to improve accuracies of cropland and specific-crop classifications. At the same time, the best performance of machine learning approach was selected to apply for updating cropland and crop-specific type mapping. On the other hand, the update approaches of annual cropland and crop-specific type mapping could play an important role. For each pixel, the correlation and significance of NDVI time series in two neighboring years as well as the bias absolute value of 2 year NDVImax were calculated. The derived information is subsequently used for change detection analyses, that is, determining whether a specific pixel of land-use or crop type will be updated. If a specific pixel has a significant correlation ($p < 0.05$) and the bias absolute value of NDVImax less than 0.1, its corresponding land-use or crop type will not be updated, or vice versa. This simple method not only improves the efficiency of classifications but it also makes sure the consistency of areal trends of land-use or crop type. Particularly, the latter point is largely very important for investigating the areal trends of croplands and crop-specific types. In view of the open access 250 m MODIS data and robust algorithms, this study successfully obtained good performance cropland and crop-specific type maps with high classification accuracies (user’s accuracy, producer’s accuracy and overall accuracy being greater than 86%).

Some previous studies had completed winter wheat mapping of the NCP in several specific years (Wang et al 2015, Qiu et al 2017, Dong et al 2020). For example, a study from Wang et al (2015) used the traditional widely-used method of Twi-difference algorithm to derive three-year winter wheat areas of the NCP (2001, 2006 and 2011). They found that winter wheat areas of the NCP increased from 10.43 km$^2$ in 2001 to 11.10 km$^2$ in 2006 and to 11.99 km$^2$ in 2011. In view of using a similar study area in this study and Wang’s study, we compared the results of the two studies. The comparison of winter...

![Figure 5. Satellite-based averaged cropland area (a), double-season cropping area (b), winter wheat and summer maize rotation area (c) in the NCP, and their corresponding spatial trends and significance at the 0.05 level during 2000–2019 (d)–(f).](image-url)
wheat areas showed that both results were relatively consistent in 2011, but our derived areas in 2001 (9.69 km²) and 2006 (11.70 km²) were slightly lower or greater than Wang's results. Besides of the difference in methods used, one of the reasons for differences is probably that they used MODIS Collection 5 (C5) EVI time series data, which could be inevitably affected by relatively lower values and data quality of C5 EVI compared to C6 NDVI (Didan et al 2015, Zhang et al 2017). In comparison to C5 vegetation index, C6 products have gone through four efforts and improvements, including use of pre-composited data, modifying the constrained view angle maximum value composite composite method, updating the EVI backup algorithm from SAVI to a new robust two-band EVI, and adjusting the relative azimuth angle and using a new dynamic range from –180° to 180° (Didan et al 2015). Areal amounts had allowable differences (figures S3–S6), but winter wheat patterns in the NCP showed good spatial consistence with those of previous studies (Wang et al 2015, Qiu et al 2017, Dong et al 2020, Luo et al 2020). Yet different from above-mentioned studies, this study could provide annual ongoing winter wheat products (2000–2019) with the help of relationships between time series vegetation index information in two neighboring years and NN-based machine learning approach. Importantly, long-term maps take us to have the ability to estimate areal trends of the winter wheat. Besides, to investigate changes of double cropping, annual summer maize maps were produced that previous studies rarely reported (Luo et al 2020). This study also observed that the developed method had a relatively stronger ability to derived winter wheat area compared to summer maize area, such as, \( R^2 \) being 0.89 for winter wheat estimation and 0.81 for summer maize estimation at the validation stage, respectively. This is largely because that there are more various crop types in summer maize growing phase compared to winter wheat growing phase. In collected field training data, autumn-harvest crops cover summer maize, paddy rice, peanut, cotton and soybean. Complex crop types greatly increased the classification challenge of summer crops (Duveiller and Defourny 2010, Chen et al 2016, Bégue et al 2018). Besides, to improve the economic income, farmers usually use intercropping systems in the NCP, e.g. the summer maize and soybean intercropping system. Yet, we found that our summer maize maps well matched spatial distributions and inter-annual trends of Luo et al (2020), although total areas of summer maize were slightly higher than their results. This is likely because that 250 m MODIS NDVI data could more decrease the impacts of mixed pixels compared to 1 km LAI data used in their study. Although there are still some limitations and uncertainties, this study is of significance for understanding the changes of regional croplands, single-/double-cropping systems and specific crop types.

4.2. Reasons for areal changes of double cropping, and winter wheat and summer maize rotation
The entire NCP showed a relatively stable status of double cropping area during the years from 2000 to 2019. This stability appears as one of the important reasons that the decreased double cropping area in some regions were totally offset by increased double cropping area in some heavily agricultural counties (figure 5(e)). This study observed that the significant decreased cropland and double cropping areas were mainly located in regions around cities, in which cropland and double cropping area decreasing trends also gave similar spatial distributions (figures 5(d) and (e)). Rapid urbanization, decreasing rural active labor force and changing regional cropping systems are likely to be several important reasons of areal decrease. Previous studies have documented that urban expansion occupied a mass of high-quality croplands around cities (Liu et al 2014a, Ning et al 2018). The 1 km resolution NCP land-use data in 2000 and 2015 showed that cropland loss was mainly caused by urban expansion, accounting for 86.4% of cropland total loss (Liu et al 2014a, Ning et al 2018, Cui et al 2019). Another result based on 1992–2020 CCI data displays that 91.6% of decreased croplands are converted into urban and built-up land in the NCP (figure 6). Total area of converting from croplands to urban and built-up land accounts for 74.3% of all urban expansion area, and note that the rate has overpassed 80% since 2010. On the other hand, rural active labor forces, especially people of 18–50 years old, are inclined to work in cities in view of finer wage income, and therefore have little time to return form cities for being engaged in agriculture (Liu et al 2014b). Reports from China's National Bureau of Statistics present that China's urbanization rate has nearly doubled from 30.9% in the year of 1999 to 59.6% in the year of 2018 (Du and Lin 2019). This means that there are more than 10 million agricultural population entering cities every year in China. These farmers’ cultivated land could suffer from the single-season or double-season abandonment risk if not cultivated by other family members (Liu et al 2014b, Liu 2018). Besides, plastic greenhouse usage and cash crops planting expansion in some special regions likely affect the magnitude of double cropping area. For example, plastic greenhouse is widely used for vegetables or flowers planting in Shouguang county and Qingzhou county of Shandong province (Yang et al 2017), given that the economic income from vegetables or flowers can better support rural livelihoods compared to those from grain crops.

Excitingly, this study however finds that although cropland area has presented a significantly decreasing trend since 2004, double cropping area of the entire NCP performs a relatively stable status. Particularly, the area of winter wheat and summer maize rotation presents a significantly increasing trend. This phenomenon is largely contributed to the respective
area increases of winter wheat and summer maize, but is also greatly associated with relatively higher economic benefits of winter wheat and summer maize in the NCP (figure 7), including advantage of per-unit yield, the increasing purchase price of grain crops, as well as transitions beginning to subsidize rather than tax agriculture in China in 2004. Rice also has a relatively higher economic benefits, but available water resource of the NCP limits its planting area. Besides, this study observes a steep increasing areal trend of winter wheat and summer maize rotation in the NCP after the year of 2010. A potential reason

Figure 6. Percentage of C2U in total cropland loss and C2U in total urban expansion. C2U represents conversion from croplands to urban and built-up land.

Figure 7. Interannual changes of averaged economic benefits (a), crop purchase prices (c), the shaded area represents the corresponding standard deviation, pre-unit yield (d) and the relationships between averaged economic benefits and area (b). The symbol ** represents the significance at the 0.01 level.
is that national and some local land-use and subsidy policies of supporting agriculture and giving favorable treatments to farmers, e.g. improving protective grain purchase price, increasing grain subsidies and enhancing subsidies of purchasing agricultural machinery, have greatly promoted farmers’ enthusiasm since the beginning of the year of 2010. These findings suggest that positive local land-use policies and stabilized grain price are very important for promoting cropland intensive level, and maintaining the stability of the double cropping area and grain crop areas.

4.3. Potential applications and future development
The developed method in this study has a large potential to be used for continuous and large-scale cropping frequency mapping owing to its robust performances. In view of the purposes of this study, we only used the developed method and products to analyze patterns of croplands, double cropping, and winter wheat and summer maize rotation, and their areal changes. However, some potentially fruitful extensions of this developed method and products are being mentioned. For example, the developed method and products could have the large potential to monitor single cropping pattern and land abandonment by considering annual cropping frequency changes, which would be greatly helpful for land-use management and policy-making (Prishchepov et al 2013, Renwick et al 2013, Spera et al 2014, Estel et al 2015). Yet, it should be noted that land abandonment monitoring is challenging owing to its transient and fragmentized features.

Although the developed method has a robust performance and potentially fruitful extensions as AICc with averaged value and standard deviation of $6.76 \pm 1.20$, RMSE with $0.0135 \pm 0.0075$ and $R$ with $0.988 \pm 0.015$ (figure S7), its corresponding limitations also need to be noted. First, the developed method is currently appropriate for regions where cropping frequency is two crops per year or less. Because it uses a six-degree polynomial function to fit vegetation growth curve and the first derivative of fitting curve to obtain peaks and valleys in its core algorithm. If cropping frequency is more than two crops per year, this method could be overfitting crop growth curve. And then, it could estimate some incorrect peaks and valleys. Certainly, estimating cropping frequency of three crops or more per year could be, to our knowledge, a widespread challenge for current satellite-based cropping frequency extraction methods (Wardlow and Egbert 2008, Cohn et al 2016, Chen et al 2018). Second, cropping frequency in this study relies on peak number and their corresponding NDVI values. In normal years, it is relatively easy for accurately extracting cropping frequency for good crop phenological profiles. However, when crops suffer from extreme climatic events, e.g. severe drought event (Gao and Yang 2009, Wang et al 2016), crop phenological profiles are largely affected. Similar phenomena could lead to an underestimation of cropping frequency. Thus, our further efforts will focus on overcoming these above-mentioned limitations.

5. Conclusions
This study introduced an approach, i.e. employing 250 m time-series MODIS NDVI anomaly data, the correlation and significance of NDVI time series in two neighboring years, crop peak information and typical machine learning algorithms (e.g. a feedforward NN algorithm, a DT algorithm and a SVM algorithm), to derive and update annual croplands and specific crop type area maps. The introduced method gave good performances against validated data and statistical data. More importantly, it estimated long-term continuous dynamics of croplands and cropping frequency. We found that cropland area presented a significantly decreasing trend since 2004 in the context of rapid urbanization. However, the double cropping area of the NCP showed a balanced magnitude but divergent patterns. Particularly, areas of the winter wheat and summer maize rotation presented significant increasing trends. Areas of other three double-season cropping types, including winter wheat and non-maize rotation, non-wheat and summer maize rotation and non-wheat and non-maize rotation yet followed significantly decreasing trends. To our knowledge, this study is the first endeavor in mapping long-term and spatial distributions of winter wheat and summer maize rotation in the most important granaries of China. The satellite-based maps and findings in this study could potentially provide some valuable information for regional land-use management and agriculture policy-making.

Data availability statement
The data that support the findings of this study are openly available at the following URL/DOI: https://figshare.com/articles/dataset/Winter_wheat_and_summer_maize_datasets_over_the_North_China_Plain_in_2000-2019/19228224.

Acknowledgments
The special thanks are given to Master students, Huimin Zhong, Yuqing Jian, Ming Wang, Fangxin Chen and Wencun Zhou for their helps in the processing of training data sampling/labelling, and the grain price collection. This study is funded by the National Natural Science Foundation of China (Grant No. 41971218 and 42171266) and the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA23070302).
References

Abu Habat A, Cavinato M E R and Lagerkvist C J 2019
Urbanization, livestock systems and food security in
developing countries: a systematic review of the literature
Food Secur. 11 279–99

Akaike H 1974 A new look at the statistical model identification
IEEE Trans. Autom. Control 19 716–23

Bégout A et al 2018 Remote sensing and cropping practices: a
review Remote Sens. 10 99

Bryan B A et al 2018 China’s response to a national land-system
sustainability emergency Nature 559 193–204

Chen C et al 2019 China and India lead in growing of the world
through land-use management Nat. Sustain. 2 122–9

Chen J, Jönsson P, Tamura M, Gu Z, Matsushita B and Eklundh L
2004 A simple method for reconstructing a high-quality
NDVI time-series data set based on the Savitzky–Golay filter
Remote Sens. Environ. 91 332–44

Chen Y, Lu D, Moran E, Batistella M, Dutra L V, Sanches I D A, da
Silva R F B, Huang J, Luiz A J B and de Oliveira M A F 2018
Mapping croplands, cropping patterns, and crop types using
MODIS time-series data Int. J. Appl. Earth Obs. 69 133–47

Chen Y, Song X, Huang J, Luiz A J B and Mansaray L R 2016 Impacts
of spatial heterogeneity on crop area mapping in Canada
using MODIS data ISPRS J. Photogramm. Remote Sens.
119 451–61

Cohn A S, VanWey L K, Spera S A and Mustard J F 2016 Crop
frequencies and area response to climate variability can
exceed yield response Nat. Clim. Change 6 601

Cui Y, Liu J, Xu X, Dong J, Li N, Fu Y, Lu S, Xia H, Si B and Xiao X
2019 Accelerating cities in an unsustainable landscape:
urban expansion and cropland occupation in China,
1990–2030 Sustainability 11 2283

Didan K, Munoz A B, Solano R and Huette A 2015 MODIS
vegetation index user’s guide (MODI3 series) version 3.00
(collection 6) Vegetation Index and Phenology Lab. The
University of Arizona (available at: https://vip.arizona.edu/)

Dong J 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

Dong J, Fu Y, Wang J, Tian H, Fu S, Niu Z, Han W, Zheng Y, Huang J
and Yuan W 2020 Early-season mapping of winter
wheat in China based on landsat and sentinel images Earth
Syst. Sci. Data 12 3081–95

Du Z and Lin B 2019 Changes in automobile energy consumption
during urbanization: evidence from 279 cities in China
Energy Policy 132 309–17

Duveiller G and DeFourny P 2010 A conceptual framework to
define the spatial resolution requirements for agricultural
monitoring using remote sensing Remote Sens. Environ.
114 2637–50

Estel S, Kueemmerle T, Alcántara C, Levers C, Prischepov A
and Hostert P 2015 Mapping farmland abandonment and
recultivation across Europe using MODIS NDVI time series
Remote Sens. Environ. 163 312–25

Gao H and Yang S 2009 A severe drought event in northern China
in winter 2008–2009 and the possible influences of La Niña
and Tibetan Plateau J. Geophys. Res. 114 D24104

Godfrey H C J, Reddington J R, Crute I R, Haddad L, Lawrence D,
Muir J F, Pretty J, Robinson S, Thomas S M and Toloumin C
2010 Food security: the challenge of feeding 9 billion people
Science 327 812–8

Hao P, Tang H, Chen Z and Liu Z 2018 Early-season crop
mapping using improved artificial immune network (IAIN)
and sentinel data PeerJ 6 e5431

Hurvich C M and Tsai C-L 1989 Regression and time series model
selection in small samples Biometrika 76 297–307

Lasko K, Vadrevu K P, Tran V T and Justice C 2018 Mapping
double and single crop paddy rice with sentinel-1A at
varying spatial scales and polarizations in Hanoi, Vietnam
IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 11 498–512

Leroux L, Bégout A, Lo Seen D, Jolivot A and Kayitakire F 2017
Driving forces of recent vegetation changes in the Sahel:
lessons learned from regional and local level analyses Remote
Sens. Environ. 191 38–54

Liu J et al 2014a Spatiotemporal characteristics, patterns,
and causes of land-use changes in China since the late 1980s J.
Geogr. Sci. 24 195–210

Liu Y 2018 Introduction to land use and rural sustainability in
China Land Use Policy 74 1–4

Liu Y, Yang R, Long H, Gao J and Wang J 2014b Implications of
land-use change in rural China: a case study of Yucheng,
Shandong province Land Use Policy 40 111–8

Liu Z, Wang J, Wang X and Yang W 2020 Understanding the
impacts of ‘grain for green’ land management practice on
land greening dynamics over the Loess plateau of China
Land Use Policy 99 105017

Liu Z and Wang S 2018 Detecting changes of wheat vegetative
growth and their response to climate change over the North
China plain IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.
11 4630–6

Liu Z, Wu C, Peng D, Wang S, Gonsamo A, Fang B and Yuan W
2017 Improved modeling of gross primary production from
a better representation of photosynthetic components in
vegetation canopy Agric. For. Meteor. 233 222–34

Lobell D B, Burke M B, Tebaldi C, Mastrandrea M D, Falcon W P
and Naylor R L 2008 Prioritizing climate change
adaptation needs for food security in 2030 Science
319 607–10

Lunetta R S, Shao Y, Ediriwickrema J and Lyon J G 2010
Monitoring agricultural cropping patterns across the
Laurentian Great Lakes basin using MODIS-NDVI data Int.
J. Appl. Earth Obs. 12 81–88

Luo Y, Zhang Z, Chen Y, Li Z and Tao F 2020
ChinaCropPhen1km: a high-resolution crop phenological
dataset for three staple crops in China during 2000–2015
based on leaf area index (LAI) products Earth Syst. Sci. Data
12 197–214

Mueller-Warrant G W, Sullivan C, Anderson N and
Whittaker G W 2016 Detecting and correcting logically
inconsistent crop rotations and other land-use sequences
Int. J. Remote Sens. 37 29–59

Ning J et al 2018 Spatiotemporal patterns and characteristics of
land-use change in China during 2010–2015 J. Geogr. Sci.
28 547–62

Panigrahy S and Sharma S A 1997 Mapping of crop rotation using
multidate Indian remote sensing satellite digital data ISPRS
J. Photogramm. Remote Sens. 52 85–91

Prischepov A V, Müller D, Dubinin M, Baumann M and
Radeloff V C 2013 Determinants of agricultural land
abandonment in post-Soviet Eastern Europe Russia Land Use
Policy 30 873–84

Prischepov A V, Radeloff V C, Dubinin M and Alcantara C 2012
The effect of landsat ETM/ETM+ image acquisition dates
on the detection of agricultural land abandonment in
Eastern Europe Remote Sens. Environ. 126 195–209

Qiu B, Luo Y, Tang Z, Chen C, Lu D, Huang H, Chen Y, Chen N
and Xu W 2017 Winter wheat mapping combining
variations before and after estimated heading dates ISPRS J.
Photogramm. Remote Sens. 123 35–46

Ray D K and Foley J A 2013 Increasing global crop harvest
frequency: recent trends and future directions Environ. Res.
Lett. 8 044041

Renwick A, Jansson T, Verbarg P H, Revoredo-Giha C, Britz W,
Gocht A and McCracken D 2013 Policy reform and

ORCID iDs

Zhengjia Liu https://orcid.org/0000-0002-4577-446X
Jinwei Dong https://orcid.org/0000-0001-5687-803X
agricultural land abandonment in the EU Land Use Policy
30 446–57
Skakun S, Franch B, Vermote E, Roger J-C, Becker-Reshef I, Justice C and Kussul N 2017 Early season large-area winter crop mapping using MODIS NDVI data, growing degree days information and a Gaussian mixture model Remote Sens. Environ. 195 244–58
Spera S A, Cohn A S, VanWey L K, Mustard J F, Rudorff B E, Risso J and Adami M 2014 Recent cropping frequency, expansion, and abandonment in Mato Grosso, Brazil had selective land characteristics Environ. Res. Lett. 9 064010
Veloso A, Mermoz S, Bouvet A, Le Toan T, Planells M, Dejoux J-F and Ceschia E 2017 Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications Remote Sens. Environ. 199 415–26
Waldhoff G, Lussem U and Bareth G 2017 Multi-data approach for remote sensing-based regional crop rotation mapping: a case study for the Rur catchment, Germany Int. J. Appl. Earth Obs. 61 53–69
Wang S, Mo X, Hu S, Liu S and Liu Z 2018 Assessment of droughts and wheat yield loss on the North China plain with an aggregate drought index (ADI) approach Ecol. Indic. 87 107–16
Wang S, Mo X, Liu S, Lin Z and Hu S 2016 Validation and trend analysis of ECV soil moisture data on cropland in North China plain during 1981–2010 Int. J. Appl. Earth Obs. 48 110–21
Wang S, Mo X, Liu Z, Baig M H A and Chi W 2017 Understanding long-term (1982–2013) patterns and trends in winter wheat spring green-up date over the North China plain Int. J. Appl. Earth Obs. 57 235–44
Wang X, Li X, Tan M and Xin L 2015 Remote sensing monitoring of changes in winter wheat area in North China plain from 2001 to 2011 Trans. Chin. Soc. Agric. Eng. 31 190–9
Wardlow B D and Egbert S I 2008 Large-area crop mapping using time-series MODIS 250 m NDVI data: an assessment for the U.S. Central great plains Remote Sens. Environ. 112 1096–116
Wardlow B D, Egbert S I and Kastens J H 2007 Analysis of time-series MODIS 250 m vegetation index data for crop classification in the U.S. Central great plains Remote Sens. Environ. 108 290–310
Weiss M, Jacob F and Duveiller G 2020 Remote sensing for agricultural applications: a meta-review Remote Sens. Environ. 236 111402
Wu W, Yu Q, You L, Chen K, Tang H and Liu J 2018 Global cropping intensity gaps: increasing food production without cropland expansion Land Use Policy 76 515–25
Yang D, Chen J, Zhou Y, Chen X, Chen X and Cao X 2017 Mapping plastic greenhouse with medium spatial resolution satellite data: development of a new spectral index ISPRS J. Photogramm. 128 47–60
Yang R, Liu Y, Chen Y and Li T 2013 The remote sensing inversion for spatial and temporal changes of multiple cropping index and detection for influencing factors around Bohai Rim in China Sci. Geogr. Sin. 33 588–93
Zhang Y, Song C, Band L E, Sun G and Li J 2017 Reanalysis of global terrestrial vegetation trends from MODIS products: browning or greening? Remote Sens. Environ. 191 145–55