

# Fast extraction of winter wheat planting area in Huang-Huai-Hai Plain using high-resolution satellite imagery on a cloud computing platform

Dongyan Zhang<sup>1</sup>, Mengru Zhang<sup>1</sup>, Fenfang Lin<sup>1,2</sup>, Zhenggao Pan<sup>1,3</sup>,  
Fei Jiang<sup>1,3</sup>, Liang He<sup>1,4</sup>, Hang Yang<sup>1,5</sup>, Ning Jin<sup>1\*</sup>

(1. National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230601, China;

2. Key Laboratory of Geospatial Technology for Middle and Lower Yellow River Regions (Henan University), Ministry of Education, Kaifeng 475004, Henan, China;

3. School of Informatics and Engineering, Suzhou University, Suzhou 234000, Anhui, China;

4. National Meteorological Center, Beijing 100081, China;

5. Aerospace Information Research Institute, Chinese Academy of Science, Beijing 100101, China)

**Abstract:** To extract regional winter wheat planting area using higher-resolution satellite imagery still faces many challenges due to large data size and long processing time in traditional remote sensing classification. Google Earth Engine (GEE), a cloud computing analysis platform based on global geospatial analysis, provides a new opportunity for rapid analysis of remote sensing data. In this study, high-quality Landsat-8 imagery was used to extract the winter wheat planting area from the Huang-Huai-Hai Plain in China. The random forest algorithm was used to identify and map the winter wheat sown in 2019 and harvested in 2020, and Sentinel-2 imagery was used to verify the results. The spectral indices, texture, and terrain features of the image were derived, and their contribution to the classification accuracy of winter wheat was evaluated by scoring. Then the top nine features were selected to form an optimal feature subset. Comparing the set of thirty-four features and the optimized feature subset as the input variables of the random forest classifier, the results show that the accuracy difference between the two feature classification schemes is small, but the classification effect of all feature sets is slightly better than the optimal feature subset. The overall classification accuracy of sample plots verification was 86%-95%, the Kappa coefficient was between 0.70 and 0.85, and the percentage error of the total area was 5.42%. The research demonstrates a reliable method for mapping a wide range of winter wheat planting area, and provides a good prospect for exploring the precise mapping of other crops, which is of great significance to crop monitoring and agricultural development.

**Keywords:** Google Earth Engine, regional scale, winter wheat, rapid mapping

**DOI:** 10.25165/j.ijabe.20221501.6917

**Citation:** Zhang D Y, Zhang M R, Lin F F, Pan Z G, Jiang F, He L, et al. Fast extraction of winter wheat planting area in Huang-Huai-Hai Plain using high-resolution satellite imagery on a cloud computing platform. *Int J Agric & Biol Eng*, 2022; 15(1): 241–250.

## 1 Introduction

Wheat is an important food crop for mankind, and it is widely planted around the world. Its production is of great significance to global food security. In recent years, China's wheat planting area has been ranked third in the world<sup>[1]</sup>, and the Huang-Huai-Hai Plain (HHHP) is the largest wheat planting area in China. According to statistics, the planting area of winter wheat in the HHHP has accounted for an average of 67.8% of the country in the past 10 years<sup>[2]</sup>. The planting and production of winter wheat in

this area are closely related to national food security. Therefore, the large-scale extraction of the spatial planting distribution of winter wheat is essential for government departments to macro-control food production information, and is essential for guiding agricultural production and formulating relevant agricultural policies<sup>[3]</sup>. At present, the widely used monitoring method by remote sensing technology is mainly single-machine classification<sup>[4]</sup>, that is, the experimental steps are completely carried out in a local operating program. Generally, when single-machine classification of remote sensing images is applied to a large area, it is necessary to download the data covering the research area to a local folder, and the amount of image data is as high as hundreds or thousands of scenes. A program runs locally to preprocess the imagery and perform other operations. The processing of images is complicated and cumbersome, the workload is large and consumes a lot of storage space and time cost<sup>[5]</sup>. Moreover, the classification speed is relatively slow, and it requires high computing performance and storage capacity of the computer<sup>[6-8]</sup>. Local computing cannot meet the identification and mapping of large-scale regional crops<sup>[4]</sup>. Therefore, this research was based on the Google Earth Engine (GEE) cloud computing platform and selected the HHHP as an example area to realize the regional-scale higher-resolution winter wheat identification.

There have been many studies on crop identification and classification using remote sensing technology<sup>[9,10]</sup>, among which

**Received date:** 2021-07-25 **Accepted date:** 2021-12-13

**Biographies:** **Dongyan Zhang**, PhD, Professor, research interest: agricultural remote sensing and information technology, Email: zhangdy@ahu.cn; **Mengru Zhang**, MS candidate, research interest: remote sensing information processing, Email: p19201088@stu.ahu.edu.cn; **Fenfang Lin**, PhD, Associate Professor, research interest: agricultural remote sensing, Email: linfenfang@126.com; **Zhenggao Pan**, MS, Professor, research interest: big data processing, Email: szxypzg@ahszu.edu.cn; **Fei Jiang**, MS, Professor, research interest: intelligent information processing, Email: feiwuhan@126.com; **Liang He**, PhD, Associate Research Fellow, research interest: crop growth models and climate change impacts in agricultural system, Email: heliang@cma.gov.cn; **Hang Yang**, PhD, Associate Professor, research interest: remote sensing information processing, yanghang@aircas.ac.cn.

\***Corresponding author:** **Ning Jin**, PhD, Associate Professor, research interest: agricultural remote sensing. Anhui University, Hefei 230601, China. Tel: +86-551-63861275, Email: jinn.13b@ignrr.ac.cn.

there was a large of survey experiments to identify large-scale wheat areas. However, most of them use the imagery collected by detection instruments such as MODIS and AVHRR with lower spatial resolutions (The spatial resolution of MODIS data is 250 m, 500 m, and 1 km, and that of AVHRR data is 1.1 km)<sup>[11-13]</sup>. Most regions of China belong to the cultivation mode of individual farmers, which leads to the phenomenon of field landscape fragmentation, complex planting structure and diverse crop types. There is a seriously mixed pixel phenomenon in low-resolution images<sup>[14,15]</sup>, which cannot provide fine spatial distribution of winter wheat. The multi-spectral images collected by Landsat series satellites and Sentinel-2 satellites have higher spectral resolution (The highest spatial resolutions of their bands are 30 m and 10 m respectively), and can better extract spatial information of ground objects. They have been widely used in crop identification and monitoring<sup>[16-18]</sup>, but their coverage is relatively small compared to large-scale crop area extraction studies, and many studies are limited to administrative divisions at or below the provincial level<sup>[19-21]</sup>. For example, Zhang et al.<sup>[21]</sup> used Sentinel-2 imagery to extract winter wheat planting areas in four counties in China using a random forest algorithm. Ali et al.<sup>[22]</sup> mapped the winter wheat areas in the Bekaa plain, the primary wheat production area in Lebanon. These wheat area mapping studies using high-resolution images have achieved good extraction results in a small area, but there were few studies to achieve fine spatial mapping of winter wheat on a large scale. Currently, many global and national-scale mapping products for land cover and land use are available, such as the 30 m fine resolution global land cover type map from Landsat satellite (FROM-GLC)<sup>[23]</sup>, and based on Sentinel-2 data the 10 m finer resolution global land cover observation map (FROM-GLC10)<sup>[24]</sup>, these remote sensing data sets provide useful information for studying the spatial distribution and management of land cover and land use on a global scale<sup>[22]</sup>, but their main purpose is to list land cover types (such as farmland, forest, grassland, etc.) extensively, and cannot provide detailed information about specific crops.

In recent years, the emergence of the GEE cloud computing platform has made it a reality to monitor crop planting information in large-scale areas. GEE provides a concurrent processing method, has the most advanced cloud computing capabilities, and can quickly process a large number of images<sup>[5,25]</sup>. Li et al.<sup>[26]</sup> combined Landsat and Sentinel imagery to draw a 30 m planting intensity map for provincial administrative areas in 2018. The recognition of wheat, rice, soybean and other crops based on the GEE cloud platform<sup>[4,5,26]</sup>. The rice map generated by Dong et al.<sup>[27]</sup> used phenological information has achieved high accuracy, the accuracy of producer (user) is 73% (92%), and they provide a reference for the study of large-scale regional winter wheat identification using Landsat-8 image data; Xiong et al.<sup>[28]</sup> combined the 2016 Sentinel-2 and Landsat-8 data to draw an annual farmland area map with a spatial resolution of 30 m in Africa, and the prospect of their later work is to draw the growth distribution map of a specific crop; You et al.<sup>[26]</sup> drew an annual map of maize, soybeans, and rice in 2017-2019, the three northeastern provinces of China, with an overall accuracy of 81% to 86%, which provides guidance for the study of crops in other large-scale regions; He et al.<sup>[4]</sup> extracted the area of winter wheat and rapeseed in Jiangsu province based on GEE. Xu et al.<sup>[5]</sup> drew the distribution map of winter wheat in Shandong province based on the GEE platform, and the random forest classifier has been proven to have significant advantages in large-scale crop classification, but in their study, the

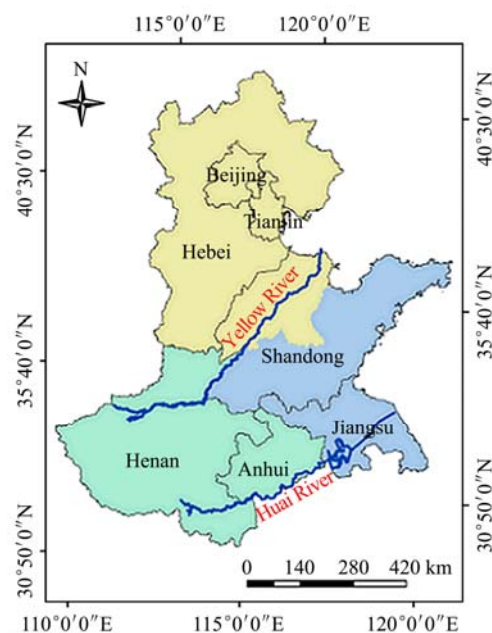
area of interest was limited to provincial administrative units. At present, studies related to rapid mapping of wheat planting area in the Huang-Huai-Hai Plain are relatively lacking, and experiments and research are urgently needed.

Therefore, the purpose of this study was to make use of GEE platform's cloud computing capability and random forest algorithm to achieve rapid identification and mapping of winter wheat in large-scale regions on higher-resolution imagery. Spectral bands and indices, texture, and topographic feature sets of Landsat-8 imagery were constructed to evaluate the contribution of each feature to classification accuracy. The Sentinel-2 data were used to verify the effectiveness of two classification schemes, classification with all feature sets and classification with preferred feature subsets, in a sample-square manner.

## 2 Materials

### 2.1 Study area

The Huang-Huai-Hai Plain is in the northern part of China (113°04'E-122°43'E, 31°23'N-42°37'N) and composed of large alluvial plains of the lower Yellow River and the Huai River, and the terrain is flat. It spans seven provinces including Beijing, Tianjin, Hebei, Shandong, Henan, Anhui, and Jiangsu. It has jurisdiction over 57 prefectures and 361 counties (cities, districts), with a total land area of approximately 620 000 km<sup>2</sup>. With reference to "Natural Agricultural Zoning of China" and "Climate Zoning of China" (Resources and Environment Science and Data Center: <http://www.resdc.cn/>), the HHHHP is divided into three sub-regions: north, west, and east of the HHHHP (Figure 1). The HHHHP is located in a warm temperate continental monsoon climate zone with a large latitude span and uneven distribution of climate resources. The annual accumulated temperature ( $\geq 10^{\circ}\text{C}$ ) is 3600°C-4900°C, the annual accumulated sunshine hours are 2300-2800 h, and the annual precipitation is 600-800 mm<sup>[29,30]</sup>, with more in the south than in the north<sup>[31]</sup>. The four seasons change significantly. HHHHP is a double-cropping system of winter wheat (October 2019-June 2020) and summer maize (June 2020-October 2020). Winter wheat is usually sown in mid-to-late October and harvested in mid-to-early June in China.



Note: Three colors represent three sub-areas, and the provincial borders are marked by black vector borders

Figure 1 Map of the study area and the results of the divisions

## 2.2 Datasets

The Landsat-8 satellite was launched by NASA in 2013. The Operational Land Imager (OLI) it carries has 9 bands. Except for band 8 which is a panchromatic band with a spatial resolution of 15 m, the other bands have a spatial resolution of 30 m (Table 1). The revisit period of the satellite is 16 d, and the imaging width is 185 km×185 km<sup>[32]</sup>. Due to the large span of the study area, the phenological periods of winter wheat in the three sub-study areas were significantly different. Therefore, all alternative images from January 30, 2020, to May 10, 2020, were selected for subsequent extraction research of winter wheat planting areas through GEE cloud platform.

**Table 1 Parameter information of Landsat-8 OLI and Sentinel-2 MSI**

Sensor	Band number	Band name	Center wavelength /nm	Band Width /nm	Spatial resolution /m
Landsat-8	1	Coastal blue	443	20	30
	2	Blue	482.5	65	30
	3	Green	562.5	75	30
	4	Red	655	50	30
	5	NIR	865	40	30
	6	SWIR1	1610	100	30
	7	SWIR2	2200	200	30
	8	Pan	640	180	15
	9	Cirrus	1375	30	30
Sentinel-2	1	Coastal aerosol	443	20	60
	2	Blue	490	65	10
	3	Green	560	35	10
	4	Red	665	30	10
	5	Vegetation red edge	705	15	20
	6	Vegetation red edge	740	15	20
	7	Vegetation red edge	783	20	20
	8	NIR	842	115	10
	8A	Narrow-NIR	865	20	20
	9	Water Vapor	945	20	60
	10	SWIR-Cirrus	1375	30	60
	11	SWIR	1610	90	20
12	SWIR	2190	180	20	

The multispectral imager (MSI) carried by the Sentinel-2 satellite is the only satellite with three red-side bands, which is very effective for vegetation observation<sup>[33]</sup>. The atmospherically corrected surface reflectance (SR) product of Landsat-8 OLI sensor, and the Level-2A product of Sentinel-2 data were archived in the GEE data catalog. The imagery has been orthorectified and atmospheric correction preprocessing<sup>[22,34]</sup>. In this study, sentinel-2 images from March 25 to April 15, 2020, were selected to verify the spatial distribution of winter wheat area extraction.

There are many mountains and forests in the northeastern part of HHHP. Digital elevation data (SRTMGL-003, spatial resolution is 30 m) was used for terrain features<sup>[35]</sup>. Elevation data from STRMGL-003 were produced by the Shuttle Radar Topography Mission (STRM) and provided by NASA Jet Propulsion Laboratory (JPL).

Due to the wide geographical area covered by the HHHP, the latitudinal span is large and the climate varies greatly, which leads to the great variation of winter wheat growth in the study area. Therefore, this study divided the study area into three sub-regions based on the "Natural Agricultural Zoning of China" and "Climate Zoning of China" (Resources and Environment Science and Data Center, <http://www.resdc.cn/>). The "Natural Agricultural Zoning of China" divides China into thirty-eight natural agricultural zones

based on temperature and humidity zones, and the temperature zones reflect the sum of daily average temperature accumulation during the crop growing period, that is, the distribution of cumulative temperature, which varies in different latitudes and can significantly affect the crop growth. The "Climate Zoning of China" is based on the heat and moisture indexes. The two zoning maps are combined to subdivide the HHHP, and then the area of winter wheat in the sub-regions is extracted and mapped separately.

## 3 Methods

### 3.1 Data preprocessing

Both Landsat-8 and Sentinel-2 images need to undergo pre-processing steps such as cloud removal, cloud-free image synthesis, mosaic, and clipping. A total of 1046 Landsat-8 SR images were covered in the study area and growth period, of which 325 images were selected for this study, and their cloud coverage was less than 20%. The quality assessment band (QA) of Landsat-8 data was used to mask cloud, shadow, snow, and ice, so as to reduce the pollution of cloud, shadow and snow to images. Finally, the values of each pixel were arranged in order and the median value was taken as the final value. The vector boundary of the study area was used for clipping to obtain a high-quality Landsat image covering the study area.

The level-2A product of Sentinel-2 contains three QA bands, of which the QA60 band contains cloud mask information, which can be used to remove cirrus and thick clouds in the image<sup>[36]</sup>. In this study, the imagery in the key phenological period is acquired, and the cloud amount is controlled within 20%, and the median value of all values is taken to constitute a higher-quality image. The spatial resolution of Sentinel-2 is 10 m, 20 m, and 60 m, respectively, for different bands. Due to the red edge band that is more effective for vegetation identification, this study uses the nearest neighbor resampling algorithm to resample it to 10 m, and the resolution is 10 m wavelength band (of 9 spectral bands) as the input of the classifier to identify winter wheat.

In this study, imagery retrieval, image preprocessing, band resampling, feature index calculation, image classification and accuracy verification were all realized by writing code online in GEE.

### 3.2 Candidate feature description

Thirty-four features were used to identify winter wheat, including seven spectral band reflectance features (30 m spatial resolution of the original spectral band B1-B7), five commonly used vegetation indexes (Table 2), eighteen texture features, and four terrain features, to explore the influence of feature variables and their linear combination on classification accuracy. Texture is an important attribute of an image and a structural feature<sup>[37]</sup>. The near-infrared band B8, which is more effective for vegetation monitoring, was selected to calculate texture features, and the glmTexture (size, kernel, Average) function provided by GEE could be used to achieve fast calculation<sup>[4]</sup>. In addition, there is a large area of mountains in the north and southeast of the study area, and the use of terrain features can improve the classification accuracy. Therefore, this study used SRTMGL1\_003 data to construct elevation, slope, aspect, and hillshade features.

### 3.3 Random Forest Algorithm for selecting feature and identifying winter wheat

Random Forest (RF) is a machine learning algorithm. Due to its stable performance, high efficiency, strong adaptability and good anti-noise ability, it has been widely used in remote sensing<sup>[43,44]</sup>

and has also been proved by many studies to be suitable for crop classification. RF algorithm uses an integrated learning method to integrate many independent decision trees into a forest. After random sampling of training samples, the category of unlabeled samples and the weight of each tree vote are finally determined through the voting results of each decision tree in the forest are the same<sup>[23]</sup>. In addition, many studies have shown that the random forest algorithm not only performs well in classification but also has advantages in selecting optimal features<sup>[45,46]</sup>.

**Table 2 Indices and candidate feature variables**

Feature	Band (Index name)	Expression	Reference
Indices	Normalized Difference Vegetation Index (NDVI)	$(B5-B4)/(B5+B4)$	Rouse et al., 1973 <sup>[38]</sup>
	Enhanced Vegetation Index (EVI)	$2.5(B5-B4)/(B5+6B4-7.5B2+1)$	Huete et al., 2002 <sup>[39]</sup>
	Soil Adjusted Vegetation Index (SAVI)	$1.5(B5-B4)/(B5+B4+0.5)$	Huete et al., 1988 <sup>[40]</sup>
	Modified Normalized Difference Water Index (MNDWI)	$(B3-B6)/(B3+B6)$	McFeeters, 1996 <sup>[41]</sup>
	Normalized Difference Building Index (NDBI)	$(B6-B5)/(B6+B5)$	Zha, 2003 <sup>[42]</sup>
Texture feature	Angular second moment (asm), contrast, correlation (corr), variance (var), inverse moment (idm), sum average (savg), sum variance (svar), sum entropy (sent), entropy (ent), difference variance (dvar), difference entropy (dent), related information measure 1 (imcorr1), related information measure 2 (imcorr2), maximum correlation coefficient (maxcorr), difference (diss), moment of inertia (inertia), Cluster cute (shade), cluster protrusion (prom)		
Terrain feature	Elevation, slope, aspect, hillshade		

Note: B2, B3, B4, B5, B6 represent the reflectance values in the spectral band of Landsat-8 OLI imagery.

In order to explore the contribution of spectral bands, indices, textures and terrain features to classification accuracy, this study evaluated the importance of feature variables by scoring the sample set using random forest algorithm. Combining the results of the visual interpretation of Google earth high-definition images and Sentinel-2 satellite images, the Landsat OLI data samples are marked online in the GEE program code, from a total of seven types of continental cover (winter wheat, water, bare land, urban, forest, other vegetation, and others) selected 13 000 samples evenly distributed in the study area. The samples used by the RF classifier for training are two-thirds of the samples randomly selected from the sample set, and the remaining one-third of the samples not involved in training are called out-of-bag samples (OOB)<sup>[47]</sup>, generated by OOB data Out of bag error is expressed by the unbiased estimation of the generalization error of the random forest classifier, which can be used to evaluate the classification ability of the RF classifier and the importance of feature variables<sup>[48]</sup>. The importance score of the evaluation feature variable  $j$  ( $VI_j$ ) is calculated as follows<sup>[49]</sup>:

$$VI_j = \frac{1}{N} \sum_{i=1}^N (OOB_{Ni}^j - OOB_{Oi}^j) \quad (1)$$

where,  $N$  represents the tree of the decision tree;  $OOB_{Ni}^j$  represents the OOB error of the decision tree  $i$  without the participation of noise. When random noise is added to feature  $j$ , if the OOB error ( $OOB_{Ni}^j$ ) changes greatly, it means that the feature variable  $j$  has a great influence on the classification result and is more important.

The sequence forward selection method was used to score all the features using the RF algorithm, and the feature variables

were input to the classifier in descending order. Each time a feature was added, the OOB sample was used to calculate the classification accuracy, and finally, a subset of different features was obtained. The classification accuracy of the constructed classification model and the dimension of the best feature subset were determined.

### 3.4 Accuracy assessment

This study verified the results from two aspects, spatial distribution verification and area extraction verification of winter wheat, using confusion matrix and area percentage respectively to achieve the results. The winter wheat distribution results obtained from Sentinel-2 imagery with higher spatial resolution were used as references, and the winter wheat distribution results obtained based on Landsat-8 remote sensing images were compared with the reference data, and a confusion matrix was generated based on this to evaluate the extraction effect. On the other hand, based on the winter wheat sown area data in the same growing season obtained by the National Bureau of Statistics, the area percentage was calculated for accuracy evaluation.

#### 3.4.1 Confusion matrix

The most accurate method to evaluate the classification results of remote sensing images is the confusion matrix<sup>[50]</sup>, also known as the error matrix, represented by a matrix with  $n$  rows and  $n$  columns. Evaluation indicators include overall accuracy (OA), user's accuracy (UA), Producer's accuracy (PA), and Kappa coefficient (Kappa). These accuracy indicators evaluate the classification accuracy of images from different perspectives. The calculation equation of Kappa coefficient is as follows<sup>[51,52]</sup>:

$$Kappa = \frac{M \sum_{i=1}^m A_{ii} - \sum_{i=1}^m A_{i+} A_{+i}}{M^2 - \sum_{i=1}^m A_{i+} A_{+i}} \quad (2)$$

where,  $M$  represents the total number of pixels;  $m$  is the number of categories;  $A_{ii}$  is the number of pixels on the diagonal in the confusion matrix;  $A_{i+}$  and  $A_{+i}$  represent the sum of the number of pixels in row  $i$ .

#### 3.4.2 Validation of statistical data

Estimate the winter wheat planting area based on the winter wheat distribution results extracted in the study area, and compare it with the winter wheat planting area data provided in the statistical yearbook data of each province or city. The area percentage error (PE) was used to evaluate the remote sensing recognition accuracy of winter wheat in the provinces and cities in the study area. The area percentage calculation equation is as follows:

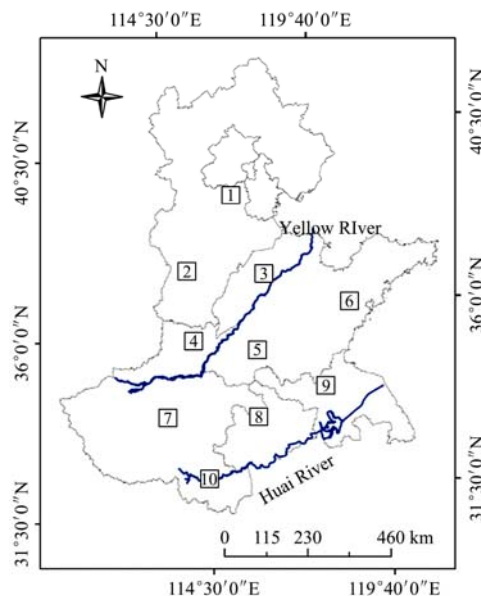
$$PE = \frac{|s - c|}{s} \times 100\% \quad (3)$$

where,  $s$  represents the statistical area;  $c$  represents the estimated area.

This study uses Sentinel-2 data and Landsat-8 images classification results to calculate a confusion matrix to verify the accuracy of the spatial distribution extracted by winter wheat. Follow the principle of uniformity and use the online editing function of GEE to plan ten sample plots (45 km×45 km). The distribution was shown in Figure 2. Export the sample plot to the GEE cloud disk, and crop the Sentinel-2 image and Landsat-8 classification map according to the vector boundary of the study area. According to previous studies, near the jointing stage, winter wheat is most separable from other land cover types, which is the best time to identify winter wheat<sup>[18]</sup>. Therefore, this study obtained the distribution of winter wheat in the plots extracted from

Sentinel-2 satellite images from March 25, 2020, to April 15, 2020, and used the results as reference data to verify its accuracy. Since the spatial resolution of Sentinel-2 is 10 m and the spatial resolution of Landsat-8 is 30 m, in order to match the spatial resolution of the two sensors, the nearest neighbor sampling algorithm is used to resample the Sentinel-2 image to 30 m, to achieve the spatial correspondence of each pixel in order to calculate the confusion matrix.

The process of identifying winter wheat planting distribution in this study was shown in Figure 3. All Landsat-8 images (325 scenes in total) of this period were acquired in GEE, using global 30 m land cover products, after removing clouds, missing value processing, stitching, and cropping, a high-quality cloud-free image were finally obtained. The distribution information of cultivated land in 2017 shields other types of land except cultivated land to reduce errors<sup>[21]</sup>. Secondly, construct classification features, including four types of features: spectral bands, spectral indices, texture, and terrain, and score all features to screen the best features. Analyze the impact of four types of features and their combinations, and priority features on classification accuracy. Finally, two methods of spatial distribution and area percentage were used to evaluate the classification effect.



Note: The numbers on the map are ten sample plots (45 km×45 km) that were planned by the online editing function of GEE. The picture shows the county-level division of Huang-Huai-Hai Plain.

Figure 2 Spatial distribution of ten sample plots with labels

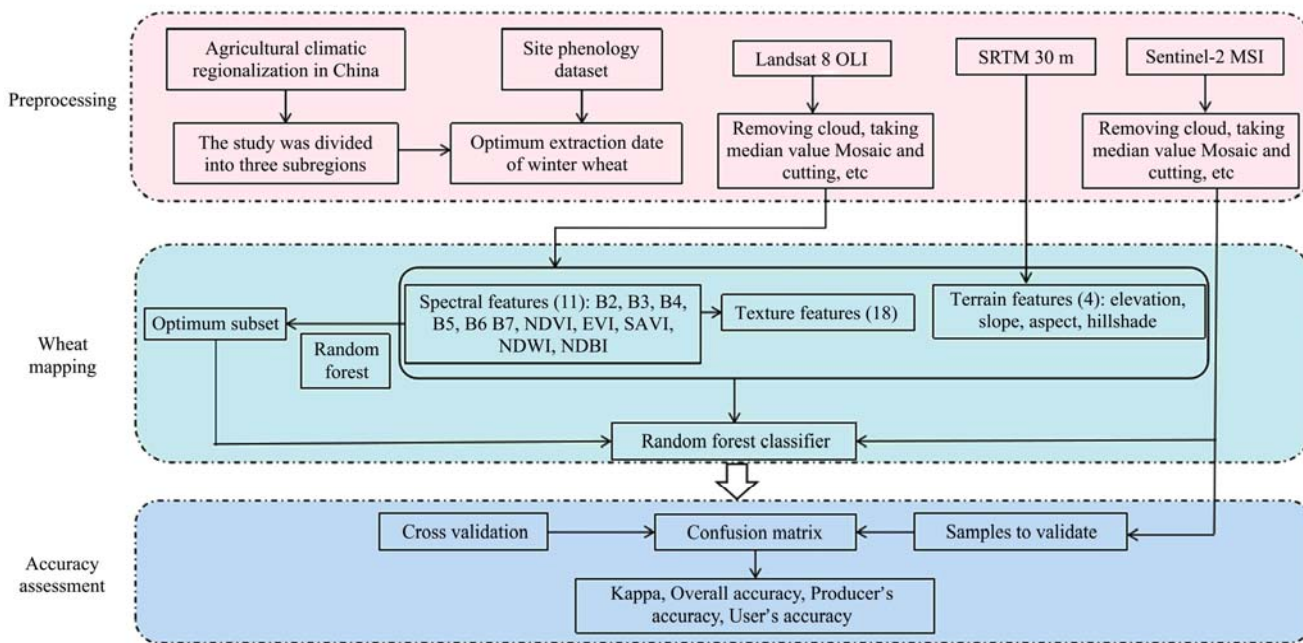


Figure 3 Workflow of winter wheat extraction

## 4 Results and analysis

### 4.1 Selection of optimal classification features

This study scores 34 features generated from the 325 high-quality images covering the study area. Random forest algorithm was used to score and sort all characteristic variables (Figure 4). The results showed that the slope had the highest score (5.18), and was the most critical characteristic variable for identifying winter wheat and other ground-object types, followed by altitude, B6, and B5 (short-wave infrared and near-infrared bands). The score of the moment of inertia (B8\_inertia) was the lowest (0.11) and was a relatively insignificant characteristic variable in winter wheat extraction.

In order to select the optimal dimension of the feature subset, according to the importance score of the feature, the feature

variable with the highest score was input into the classifier in descending order, and the classification accuracy was calculated using OOB samples. The results showed that after adding the features with the highest scores in sequence, the accuracy increases with the increase in the number of features, until the 10th feature was input, the accuracy no longer changed significantly. Therefore, it can be explained that using the top nine features of the score as the input of the classifier to build the model can achieve the highest accuracy, but this result can only explain the situation of the sample set, and the spatial distribution of winter wheat planting still needs to use Sentinel-2 data of sample plots is evaluated by calculating the Kappa coefficient. The top 9 features with importance scores (slope, elevation, Band 6, Band 5, sum average, NDVI, difference, Band 1, Band 2) in Figure 4 were selected as the best subset of features in the study area.



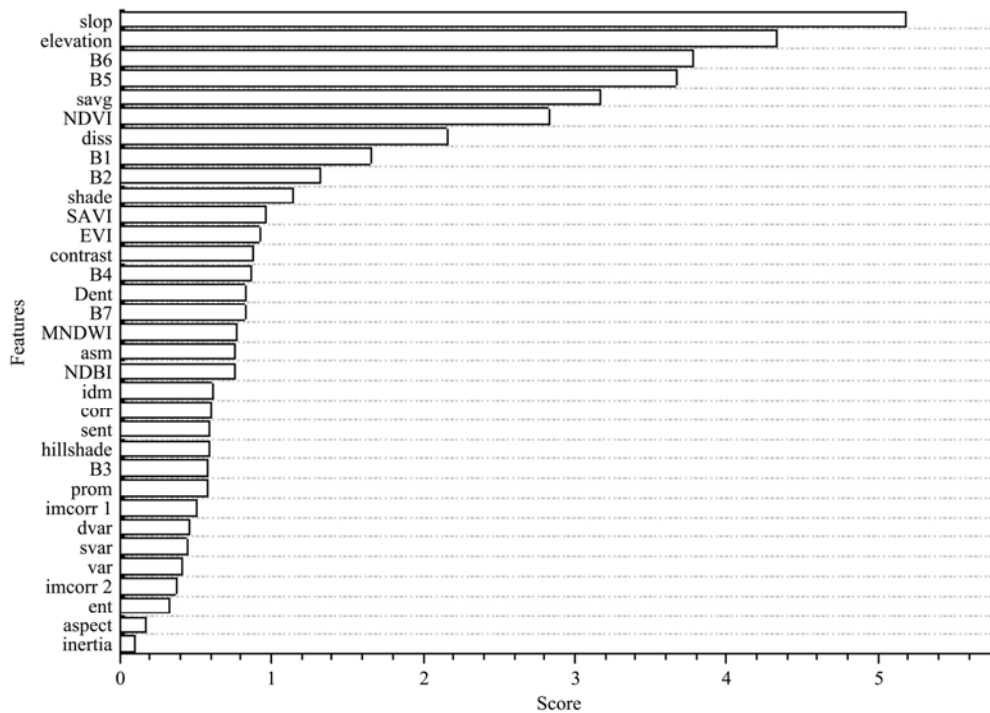


Figure 4 Score of the importance of each candidate feature ranked in descending order

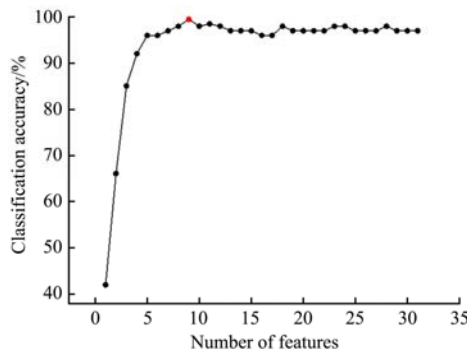


Figure 5 The relationship between the number of features and classification accuracy

**4.2 Results of sample plot validation**

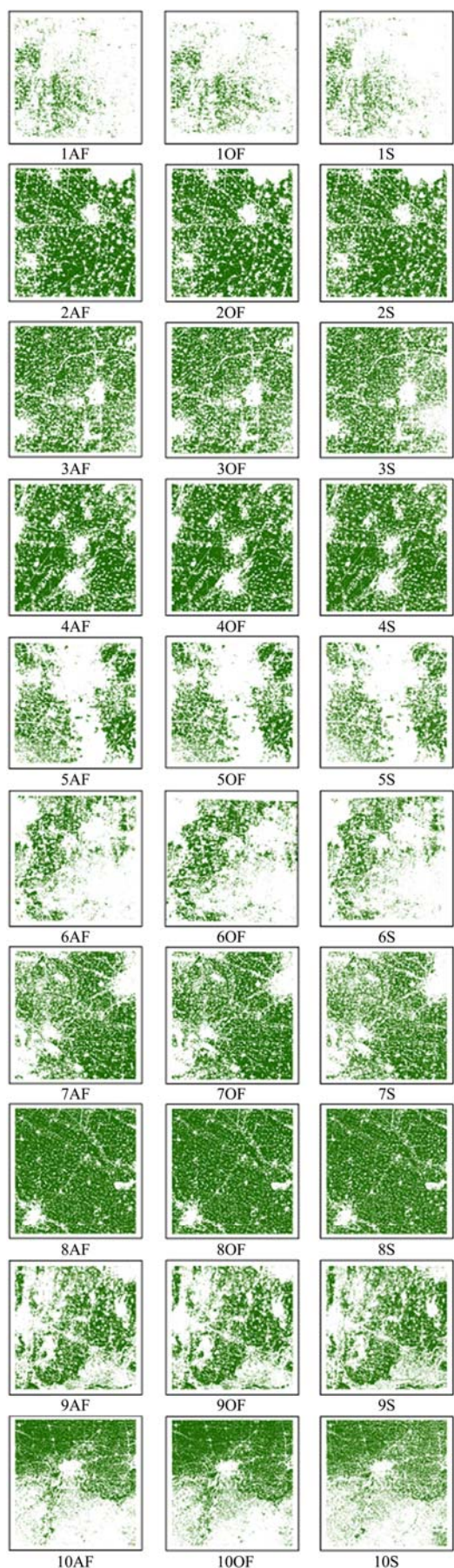
The distribution of winter wheat based on Sentinel-2 images was regarded as the true distribution on the ground. It used all feature set and the first 9 optimum feature subset respectively. Referred to the scheme all of features collection (AF) and scheme optimum features subset (OF) respectively, which was used as the input variable of the RF classifier for training and classification, and the confusion matrix was calculated with Sentinel-2 data. The overall accuracy of the scheme AF was between 86% and 95%, the Kappa coefficient was distributed between 0.70 and 0.84, the overall accuracy of the scheme OF was between 83% and 95%, and the range of the Kappa coefficient was between 0.67 and 0.85. It can be seen from the plots map that the spatial planting of winter wheat in the five wheat fields of plots 1, 5, 6, 9, and 10 was discontinuous, and the degree of fragmentation was high. Correspondingly, their Kappa coefficients were all lower than 0.8 except for plot 6. The wheat fields of plots 2, 4, 6, and 8 were planted densely and continuously, and the Kappa values were all above 0.80. It can be seen from Table 6 that in the two classification schemes, the overall accuracy of the continuous plots in the wheat field was above 92%, and the Kappa value was at least 0.81. The recognition effect of winter wheat was slightly better than other plots. In addition, among the results of classification

using the scheme AF, the mapping accuracy of 6 plots was higher than 90%, and the mapping accuracy of the scheme OF classification was higher than 90%, and there were 7 plots, indicating that the mapping accuracy of the scheme OF was better. However, its average overall accuracy was 0.4% lower than the scheme AF, and the Kappa value was 0.01 lower. The results showed that the classification scheme AF did not show obvious advantages in the recognition ability of winter wheat compared with the scheme OF. Both schemes can provide valuable information for future large-scale crop classification and extraction research.

**Table 3 Extraction accuracy of sample verification**

Scheme	Ground sample	OA/%	PA/%	UA/%	Kappa
All features	Plot 1	94.24	76.41	70.29	0.70
	Plot 2	92.35	97.02	90.41	0.84
	Plot 3	86.02	91.63	80.74	0.74
	Plot 4	91.96	92.44	94.46	0.83
	Plot 5	91.00	93.79	73.66	0.77
	Plot 6	93.88	86.43	84.20	0.81
	Plot 7	89.12	93.24	83.79	0.78
	Plot 8	93.14	97.00	93.77	0.82
	Plot 9	88.66	81.64	89.14	0.76
	Plot 10	86.50	86.47	79.00	0.72
	Mean	90.68	89.61	83.95	0.78
Optimized features	Plot 1	94.34	73.62	72.06	0.70
	Plot 2	92.67	96.21	91.52	0.83
	Plot 3	86.47	90.03	82.36	0.73
	Plot 4	91.97	95.9	91.54	0.82
	Plot 5	91.22	92.56	74.90	0.77
	Plot 6	93.58	89.00	81.46	0.81
	Plot 7	87.67	96.04	79.92	0.76
	Plot 8	92.75	96.99	93.35	0.81
	Plot 9	88.42	83.16	87.32	0.76
	Plot 10	83.77	90.13	72.55	0.67
	Mean	90.28	90.36	82.70	0.77

Note: OA: Evaluation indicators include overall accuracy; UA: User's accuracy; PA: Producer's accuracy; Kappa: Kappa coefficient.



Note: Green represents winter wheat; 1-10 represents plot number; AF represents the spatial distribution of Landsat winter wheat based on all candidate features; OF represents the distribution result of Landsat winter wheat classified using optimum features; S represents Sentinel-2 data Winter wheat distribution results.

Figure 6 Comparison of plot verification

### 4.3 Mapping the spatial distribution of winter wheat

In this study, 34 feature variables and preferred feature subsets were used to compare and study the effect of the construction of different feature spaces on the extraction of winter wheat area. The results showed that the difference in the recognition effect of the two methods for winter wheat was small. In this case, it is more convenient to extract winter wheat by using the optimized feature subset. The spatial distribution of winter wheat planting in the HHP area was extracted from 34 feature variables using a random forest classifier, as shown in Figure 7. The planting scale of winter wheat in the study area is relatively large, but the ground features are complex and diverse, covering a large area of mountains and forests in the northern part of the HHP, and no winter wheat is planted. Winter wheat is sparsely distributed in the areas north of the Yellow River and adjacent to the east of the Yellow River, and the planting structure is relatively complex and the field landscape is fragmented. The spatial distribution of winter wheat plots in the south of the Yellow River and the Huai River area is relatively continuous.

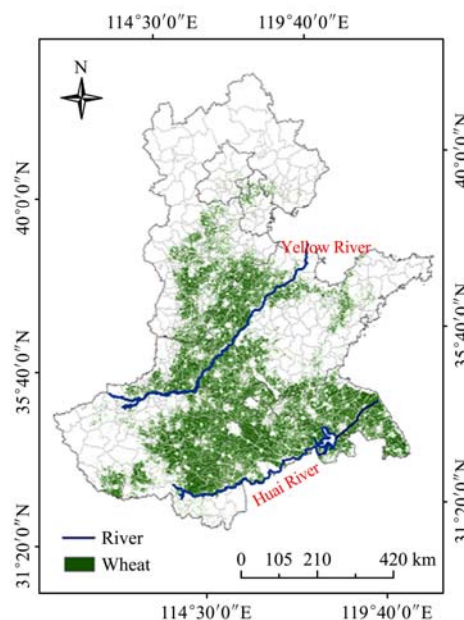


Figure 7 Spatial distribution map of winter wheat in the study area

### 4.4 Verification of the area percentage of winter wheat extraction

The winter wheat statistical data of seven provinces were used for data verification (Table 4), and the area percentage error (PE) of the total area was only 5.42%. Among them, the area percentage of four provinces was less than 10%, and only Hebei Province's winter wheat area extraction was underestimated, while the other provinces had slight overestimation, especially Beijing and Tianjin, with percentage errors of 26.36% and 17.22%, respectively.

Table 4 Statistical verification of winter wheat planting area in 2019

Province	Statistical area/hm <sup>2</sup>	Estimated area/hm <sup>2</sup>	PE/%
Beijing	8000	10 109	26.36
Tianjin	10 100	11 869	17.22
Hebei	2 322 500	2 176 158	6.30
Henan	5 706 650	6 266 558	8.20
Anhui	1 801 019	2 028 396	12.62
Shandong	4 001 753	4 068 260	1.66
Jiangsu	1 602 830	1 729 770	7.92
Sum	15 452 852	16 291 090	5.42

Source: Data from the Statistical Yearbook of 2021, <http://www.stats.gov.cn/>.

## 5 Discussion

In predecessors' research on crop recognition based on remote sensing images, a large number of winter wheat extraction and mapping methods have been developed<sup>[53-55,14]</sup>. However, their research left many problems to be solved in the future. For example, Xu et al.<sup>[5]</sup> mapped the distribution of winter wheat in Shandong Province, and realizing the map of winter wheat area in a larger area was the focus of their later work; Zhang et al.<sup>[47]</sup> focused on identifying winter wheat in the early season from the perspective of phenology, but lacked discussion on the characteristics that affect the identification of winter wheat; You et al.<sup>[29]</sup> used random forest algorithm to draw 10 m crop type maps of annual maize, soybean and rice in the three northeastern provinces, which provided powerful support for the mapping prospects of winter wheat. This study used the GEE platform to realize the rapid mapping of winter wheat on a large regional scale, which makes up for the deficiencies in the above research and provides references and case support for the remote sensing monitoring of large-scale crops.

### 5.1 Winter wheat mapping using all feature sets

This study first scored all the feature variables, selected the top 9 features to form an optimum feature subset, and evaluated the feature set and its combination using a cross-validation method. It is concluded that the spectral bands and indices are the main factors affecting the recognition accuracy of winter wheat. Sentinel-2 data were used for the final accuracy evaluation through the sample plots verification method. The average Kappa value of the two classification schemes differed only by 0.01, and the average overall accuracy differed only by 0.4%. Considering the computing performance of the computer and the large workload, the optimum feature subset can be selected first for classification mapping, but this study used the supercomputing power of GEE to select the scheme with higher precision, that is, all feature sets were classified.

### 5.2 Factors influencing the accuracy of winter wheat mapping

From the results of the confusion matrix calculated from the sample data of Sentinel-2 and Landsat-8, it can be seen that the producer accuracy of plot 2 and plot 8 is relatively high, while plot 1 and plot 9 are relatively low, and the producer accuracy is 76.4% and 81.6%, respectively. This may be caused by two factors, firstly, China's individual planting patterns dominate, leading to fragmentation of the distribution of cultivated land<sup>[55]</sup>; secondly, due to the differences in climatic conditions and wheat phenology, there are many phenomena of "same substance with the different spectrum, and different substance with same spectrum" when visually interpreting winter wheat in this region. It can also be observed in the plot distribution map that the producer accuracy is higher in areas where the planting distribution is continuously concentrated (Plots 2, 3, 4, 8), while the producer accuracy is lower in the areas where the planting distribution is fragmented (Plot 1, 5, 10).

### 5.3 Effects of rapeseed on winter wheat mapping

In winter crops, the spectra and phenological periods of rapeseed and winter wheat are similar, and they are more prone to misclassification. According to official statistics, the sown area of rapeseed accounts for only 0.7% of the total sown area of crops in the study area<sup>[2]</sup>, but the influence of rapeseed on the extraction accuracy of winter wheat cannot be completely ruled out. According to the related research shows that when the winter wheat at the jointing stage rapeseed at flowering period, the true color

(red, green, and blue wave period of combination shows) image on rapeseed presented is bright yellow, and winter wheat is full green, from the standard false-color (near red, red, and green band combination display) on the images of the same can distinguish, pink purple is rapeseed, The red is winter wheat<sup>[56]</sup>, which can be combined with the above two different image display methods for visual interpretation to distinguish rape and winter wheat. In this study, winter wheat was selected from the greening stage to the jointing stage to extract winter wheat<sup>[4]</sup>, and the pink-purple ground features were classified as "other vegetation" as a ground feature type, thereby reducing the influence of rape on the extraction results.

### 5.4 Uncertainty and limitation analysis

This study explored the identification and extraction of large-scale wheat fields and obtained preliminary results. However, this study still has some uncertainties and limitations that need to be further resolved by follow-up work. First of all, this study divided the study area into 3 sub-areas according to the agroclimatic divisions, and only one winter wheat growing season from 2019 to 2020 was selected for the study. From the perspective of obvious differences in winter wheat phenology, a good result has been achieved in identifying winter wheat. The use of time series or phenology knowledge to achieve multi-year regional-scale spatial mapping of winter wheat is the focus of our later work. Secondly, this study used Sentinel-2 data to verify the results in space. Due to the time limitation and actual conditions, a comprehensive field survey was not done. Only detailed field interpreting was made from the high-definition images of Google Earth and GEE platforms. There are certain deviations in the interpretation of highly mixed and broken areas of crop planting. Therefore, there is an urgent need for field survey data to improve the verification part of the results of this study.

## 6 Conclusions

The results of the study showed that the GEE cloud platform has great potential in large-scale identification and mapping of winter wheat. The short-wave infrared band B6 (1610 nm) and the near-infrared band B5 (865 nm) of Landsat-8 satellite imagery, as well as the slope and altitude of the elevation data, are the most effective features for distinguishing winter wheat from other ground features, and the texture features sum average and the vegetation index NDVI was more advantageous than other features in the identification of winter wheat. The optimum feature subset can also be applied to the extraction and mapping of winter wheat and other crops when computer computing performance is limited. The overall validation accuracy was in the range of 86% to 95%, and the Kappa coefficient was between 0.70 and 0.84, which accurately describes the spatial distribution of winter wheat. The results also indicated that random forest algorithm performs well in winter wheat identification. In this study, before the winter wheat harvest, the data resources of the GEE platform and the RF classification algorithm can be used to identify and map the main wheat-producing areas in a large-scale area of the HHHP, and obtained relatively reliable spatial distribution results of winter wheat, which made up for the shortage of research on winter wheat extraction under the condition of large-scale, complex climate type, limited storage space and operational memory. The current research is the basic work of crop phenological period observation and crop yield estimation. It also provides a valuable reference for government departments for agricultural decision-making and food security issues.



## Acknowledgements

This study was financially supported by the National Key Research and Development Project of China (Grant No. 2018YFC1507802); The Outstanding Young Talents Program in Colleges and Universities in Anhui Province (Grant No. GXYQ2020001); the National Natural Science Foundation of China (Grant No. 41871352); the Open Fund of Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions (Henan University), Ministry of Education (Grant No. GTYR202104). The authors acknowledge the data support from the National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China (<http://www.geodata.cn>).

## References

- [1] Food and Agriculture Organization of the United Nations. Available: <http://www.fao.org/faostat/en/#data>. Accessed on [2021-04-16].
- [2] National Bureau of Statistics. Available: <http://www.stats.gov.cn/>. Accessed on [2021-04-16].
- [3] Qiu B W, Luo Y H, Tang Z H, Chen C C, Lu D F, Huang H Y, et al. Winter wheat mapping combining variations before and after estimated heading dates. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017; 123: 35–46.
- [4] He Z X, Zhang M, Wu B F, Xing Q. Extraction of summer crop in Jiangsu based on Google Earth Engine. *Journal of Geo-Information Science*, 2019; 21(5): 752–766. (in Chinese)
- [5] Xu F, Li Z, Zhang S, Huang N, Prishchepov A V. Mapping winter wheat with combinations of temporally aggregated sentinel-2 and Landsat-8 data in Shandong Province, China. *Remote Sens*, 2020; 12: 2065. doi: 10.3390/rs12122065.
- [6] Guan S, Han P, Wang Y, Han Y, Lin Y I, Zhou T, et al. Study on the classification of typical plantations in south China. *Journal of Geo-information Science*, 2017; 19(11): 1538–1546. (in Chinese)
- [7] Zhang J, Zhang X, Tian L, Zhang Q F. The support vector machine method for RS images' classification in northwest arid area. *Science of Surveying and Mapping*, 2017; 42(1): 49–52, 58. (in Chinese)
- [8] Yang Y K, Jiang P A, Wu H Q, Zhu L. Recognition research of Korla fragrant pear planting area based on satellite images of GF-1 and Landsat 8. *Shandong Agricultural Sciences*, 2018; 50(1): 153–157. (in Chinese)
- [9] Wang C W, Chen Q, Fan H S, Yao C L, Sun X, Chan J, et al. Evaluating satellite hyperspectral (Orbita) and multispectral (Landsat 8 and Sentinel-2) imagery for identifying cotton acreage. *International Journal of Remote Sensing*, 2021; 42(11): 4042–4063.
- [10] Zhang S, Zhang J H, Bai Y, Yao F M. Extracting winter wheat area in Huanghuaihai Plain using MODIS-EVI data and phenology difference avoiding threshold. *Transactions of the CSAE*, 2018; 34(11): 150–158. (in Chinese)
- [11] Qiu B W, Luo Y H, Tang Z H, Chen C C, Lu D F, Huang H Y, et al. Winter wheat mapping combining variations before and after estimated heading dates. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017; 123: 35–46.
- [12] He T L, Xie C J, Liu Q S, Guan S Y, Liu G H. Evaluation and comparison of random forest and A-LSTM networks for large-scale winter wheat identification. *Remote Sensing*, 2019; 11(14): 1665. doi: 10.3390/rs11141665.
- [13] Yang Y J, Tao B, Ren W, Zourarakis D P, Masri B E, Sun Z G, et al. An improved approach considering intraclass variability for mapping winter wheat using multitemporal MODIS EVI images. *Remote Sensing*, 2019; 11(10): 1191. doi: 10.3390/rs11101191.
- [14] Tao J B, Wu W B, Zhou Y, Wang Y, Jiang Y. Mapping winter wheat based on time-series MODIS data. *Journal of Integrative Agriculture*, 2017; 16(2): 348–359.
- [15] Pan Y Z, Li L, Zhang J S, Liang S L, Zhu X F, Sulla-Menasha D. Winter wheat area estimation from MODIS-EVI time series data using the Crop Proportion Phenology Index. *Remote Sensing of Environment*, 2012; 119: 232–242.
- [16] Khan A, Hansen M C, Potapov P V, Adusei B, Pickens A, Krylov A, et al. Evaluating Landsat and RapidEye data for winter wheat mapping and area estimation in Punjab, Pakistan. *Remote Sensing*, 2018; 10(4): 489. doi: 10.3390/rs10040489.
- [17] Li R Y, Xu M Q, Chen Z Y, Gao B B, Cai J, Shen F X, et al. Phenology-based classification of crop species and rotation types using fused MODIS and Landsat data: The comparison of a random-forest-based model and a decision-rule-based model. *Soil and Tillage Research*, 2021; 206: 104838. doi: 10.1016/j.still.2020.104838.
- [18] Shoko C, Mutanga O. Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017; 129: 32–40.
- [19] Delegido J, Verrelst J, Alonso L, Moreno J. Evaluation of Sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors*, 2011; 11(7): 7063–7081.
- [20] Korhonen L, Saputra D H, Packalen P, Rautiainen M. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sens. Environ*, 2017; 195: 259–274.
- [21] Zhang D Y, Fang S M, She B, Zhang H H, Jin N, Xia H M, et al. Winter wheat mapping based on Sentinel-2 data in heterogeneous planting conditions. *Remote Sensing*, 2019; 11(22): 2647. doi: 10.3390/rs11222647.
- [22] Nasrallah A, Baghdadi N, Mhawej M, Faour G, Darwish T, Belhouchette H, et al. A novel approach for mapping wheat areas using high resolution Sentinel-2 images. *Sensors*, 2018; 18(7): 2089. doi: 10.3390/s18072089.
- [23] National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China. Available: <http://www.geodata.cn>. Accessed on [2021-03-08].
- [24] Gong P. Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Chinese Science Bulletin*, 2019; 64(6): 370–373.
- [25] Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 2017; 202: 18–27.
- [26] Pan L, Xia H M, Yang J, Niu W H, Wang R M, Song H Q, et al. Mapping cropping intensity in Huaihe basin using phenology algorithm, all Sentinel-2 and Landsat images in Google Earth Engine. *International Journal of Applied Earth Observations and Geoinformation*, 2021; 102: 102376. doi: 10.1016/j.jag.2021.102376.
- [27] You N, Dong J, Huang J, Du G, Xiao X. The 10-m crop type maps in Northeast China during 2017–2019. *Scientific Data*, 2021; 8: 41. doi: 10.1038/s41597-021-00827-9.
- [28] Dong J W, Xiao X M, Menarguez M A, Zhang G L, Qi Y W, Tnau D, et al. Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine. *Remote Sensing of Environment*, 2016; 185: 142–154.
- [29] Xiong J, Thenkabail P S, Tilton J C, Gumma M K, Teluguntla P, Oliphant A, et al. Nominal 30-m cropland extent map of continental africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on Google Earth Engine. *Remote Sensing*, 2017; 9(10): 1065. doi: 10.3390/rs9101065.
- [30] Zhou B Y, Ge J Z, Sun X F, Han Y L, Ma W, Ding Z S, et al. Research advance on optimizing annual distribution of solar and heat resources for double cropping system in the Yellow-Huaihe-Haihe Rivers plain. *Acta Agronomica Sinica*, 2021; 47(10): 1843–1853. (in Chinese)
- [31] Yang J Y, Mei X Y, Yan C R, Liu Q. Study on spatial pattern of climatic resources in North China. *Chinese Journal of Agrometeorology*, 2010; 31(S1): 1–5. (in Chinese)
- [32] Tian H F, Wang L, Niu Z, Qin Y C. Winter wheat planting area extraction based on new remote sensing data at county level. *Chinese Agricultural Science Bulletin*, 2015; 31(5): 220–227. (in Chinese)
- [33] Liu H. Extraction of crop planting structure in Hetao irrigation area based on Sentinel-2 image. *Arid Land Resources and Environment*, 2021; 35(2): 88–95. (in Chinese)
- [34] Wang Y M, Chen C. Remote sensing extraction of wheat planting area in northern Jiangsu based on time series Sentinel-2 satellite image data. *Modern Surveying and Mapping*. Jiangsu Mapping and Geographic Information Society, 2019; pp.86–89. (in Chinese)
- [35] Farr T G, Rosen P A, Caro E. The shuttle radar topography mission. In: *Space-Based Observation Technology*, 2007; 45(2): RG2004.
- [36] Wang S L. Rape and the same phenological crops area extraction based on multi-source remote sensing data in Jiangsu Province. Master dissertation. Yangtze University, 2015; 75p. (in Chinese)

- [37] Liu X Y, Zhong X C, Chen C, Liu T, Sun C M, Li D S, et al. Prediction of wheat yield using color and texture feature data of UAV image at early growth stage. *Journal of Triticeae Crops*, 2020; 40(8): 1002–1007. (in Chinese)
- [38] Rouse J W, Hass R H, Scheel J A, Deering D W. Monitoring vegetation systems in the Great Plain with ERTS. In: *Proceedings of the 3rd Earth Resources Technology Satellite-1 Symposium*, Washington, DC, USA, 1973; pp.309–317.
- [39] Huete A R, Didan K, Miura T, Rodriguez E P, Gao X, Ferreira L G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing Environment*, 2002; 83(1-2): 195–213.
- [40] Huete A R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 1988; 25(3): 295–309.
- [41] Mcfeeters S K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 1996; 17(7): 1425–1432.
- [42] Zha Y, Ni S X, Yang S. An effective approach to automatically extract urban land-use from TM imagery. *National Remote Sensing Bulletin*, 2003; 7(1): 38–40, 82. (in Chinese)
- [43] Wang N, Li Q Z, Du X, Zhang Y, Zhao L C, Wang H Y. Identification of main crops based on the univariate feature selection in Subei. *National Remote Sensing Bulletin*, 2017; 21(4): 519–530. (in Chinese)
- [44] Chan J C W, Paelinckx D. Evaluation of random forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment*, 2008; 112(6): 2999–3011.
- [45] Immitzer M, Atzberger C, Koukal T. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, 2012; 4(9): 2661–2693.
- [46] Immitzer M, Vuolo F, Atzberger C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sensing*, 2016; 8(3): 166. doi: 10.3390/rs8030166.
- [47] Zhang H Y, Du H Y, Zhang C K, Zhang L P. An automated early-season method to map winter wheat using time-series Sentinel-2 data: A case study of Shandong, China. *Computers and Electronics in Agriculture*, 2021; 182: 105962. doi: 10.1016/j.compag.2020.105962.
- [48] Rodriguez-Galiano V F, Olmo-Chica M, Abarca-Hernandez F, Atkinson P M, Jeganathan C. Random forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment*, 2012; 121: 93–107.
- [49] Genuer R, Poggi J M, Tuleau-Malot C. Variable selection using random forests. *Pattern Recognition Letters*, 2010; 31(14): 2225–2236.
- [50] Congalton R, Green K. *Assessing the accuracy of remotely sensed data*. Boca Raton, FL, USA: CRC Press, 2019; 200p.
- [51] Zhao L C, Shi Y, Liu B, Hovis C. Finer classification of crops by fusing UAV images and Sentinel-2A data. *Remote Sensing*, 2019; 11(24): 3012. doi: 10.3390/rs11243012.
- [52] Congalton R G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 1991; 37(1): 35–46.
- [53] Zheng Y, Zhang M, Zhang X, Zeng H W, Wu B. Mapping winter wheat biomass and yield using time series data blended from PROBA-V 100- and 300-m S1 products. *Remote Sensing*, 2016; 8(10): 824. doi: 10.3390/rs8100824.
- [54] Khan A, Hansen M C, Potapov P V, Adusei B, Pickens A, Krylov A, et al. Evaluating landsat and rapidEye data for winter wheat mapping and area estimation in Punjab, Pakistan. *Remote Sens*, 2018; 10: 489. doi: 10.3390/rs10040489.
- [55] Zhang Q S, Chu Y Y, Xue Y F, Ying H, Chen X H, Zhao Y J, et al. Outlook of China's agriculture transforming from smallholder operation to sustainable production. *Global Food Security*, 2020; 26: 100444. doi: 10.1016/j.gfs.2020.100444.
- [56] Yang Y J, Tao B, Ren W, Zourarakis D P, Masri B E, Sun Z G, et al. An improved approach considering intraclass variability for mapping winter wheat using multitemporal MODIS EVI images. *Remote Sensing*, 2019; 11: 1191. doi: 10.3390/rs11101191.