



Mapping rapeseed in China during 2017-2021 using Sentinel data: an automated approach integrating rule-based sample generation and a one-class classifier (RSG-OC)

Yunze Zang, Yuean Qiu, Xuehong Chen, Jin Chen, Wei Yang, Yifei Liu, Longkang Peng, Miaogen Shen & Xin Cao

To cite this article: Yunze Zang, Yuean Qiu, Xuehong Chen, Jin Chen, Wei Yang, Yifei Liu, Longkang Peng, Miaogen Shen & Xin Cao (2023) Mapping rapeseed in China during 2017-2021 using Sentinel data: an automated approach integrating rule-based sample generation and a one-class classifier (RSG-OC), *GIScience & Remote Sensing*, 60:1, 2163576, DOI: [10.1080/15481603.2022.2163576](https://doi.org/10.1080/15481603.2022.2163576)

To link to this article: <https://doi.org/10.1080/15481603.2022.2163576>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 10 Jan 2023.



[Submit your article to this journal](#)



[View related articles](#)



[View Crossmark data](#)

Mapping rapeseed in China during 2017-2021 using Sentinel data: an automated approach integrating rule-based sample generation and a one-class classifier (RSG-OC)

Yunze Zang^{a,b}, Yuean Qiu^{a,b}, Xuehong Chen^{a,b}, Jin Chen^{a,b}, Wei Yang^c, Yifei Liu^{a,b}, Longkang Peng^{a,b}, Miaogeng Shen^d and Xin Cao^{a,b}

^aState Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing, China; ^bBeijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing, China; ^cCenter for Environmental Remote Sensing, Chiba University, Chiba, Japan; ^dInstitute of Land Surface System and Sustainable Development, Faculty of Geographical Science, Beijing Normal University, Beijing, China

ABSTRACT

Rapeseed mapping is important for national food security and government regulation of land use. Various methods, including empirical index-based and machine learning-based methods, have been developed to identify rapeseed using remote sensing. Empirical index-based methods commonly employed empirically designed indices to enhance rapeseed's bright yellow spectral feature during the flowering period, which is straightforward to implement. Unfortunately, the heavy cloud cover in the flowering period of China would lead to serious omission errors; and the required flowering period varies spatially and yearly, which often cannot be acquired accurately. Machine learning-based methods mitigate the reliance on clear observations during the flowering period by inputting all-season imagery to adaptively learn features. However, it is difficult to collect sufficient samples across all of China considering the large intraclass variation in both land cover types of rapeseed and non-rapeseed. This study proposed an automated rapeseed mapping approach integrating rule-based sample generation and a one-class classifier (RSG-OC) to overcome the shortcomings of the two types of methods. First, a set of sample selection rules based on empirical indices of rapeseed were developed to automatically generate samples in cloud-free pixels during the predicted flowering period throughout China. Second, all available features composited based on the rapeseed phenological calendar were used for classification to eliminate the phenology differences in different regions. Third, a specific sample augmentation that removes the observation in the flowering period was employed to improve the generalization to the pixels without cloud-free observation in the flowering period. Finally, to avoid the need for diverse samples of nonrapeseed classes, a typical one-class classifier, positive unlabeled learning implemented by random forest (PUL-RF) trained by the generated samples, was applied to map rapeseed. With the proposed method, China rapeseed was mapped at 20 m resolution during 2017–2021 based on the Google Earth Engine (GEE). Validation on six typical rapeseed planting areas demonstrates that RSG-OC achieves an average accuracy of 94.90%. In comparison, the average accuracy of the other methods ranged from 83.33% to 88.25%, all of which were poorer than the proposed method. Additional experiments show that the performance of RSG-OC was not sensitive to cloud contamination, inaccurate predicted flowering time and the threshold of sample selection rule. These results indicate that the rapeseed maps produced in China are overall reliable and that the proposed method is an effective and robust method for annual rapeseed mapping across China.

ARTICLE HISTORY

Received 9 June 2022
Accepted 24 December 2022

KEYWORDS


Rapeseed mapping; rule-based sample generation; sample augmentation; positive and unlabeled classification; random forest

1. Introduction

Rapeseed (*Brassica napus*, *Brassica campestris* L) is one of the most widely planted oil crops, serving as an important source of edible oil, animal feed, and green biofuel (Raboanatahiry et al. 2021; Firrisa, van Duren, and Voinov 2013; van Duren et al. 2015; Zhang and He 2013). Global rapeseed production has undergone

sustained growth over the past 20 years, becoming the second-most produced oilseed behind soybeans (FAO 2021, 2022; Carré and Pouzet 2014). And China is one of the largest producers and consumers of rapeseed in the world. (FAO 2022; USDA 2022; Hu et al. 2017; Tian et al. 2018). With the development of China's livestock industry and renewable energy

CONTACT Xuehong Chen  chenxuehong@bnu.edu.cn

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/15481603.2022.2163576>.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

sector, the demand for rapeseed has increased in recent years (Tian et al. 2021; Yin and Wang 2012). However, rapeseed production in China is constrained by limited cultivated land resources (Bonjean, Dequidt, and Sang 2016; Fu et al. 2016). Therefore, monitoring the extent of rapeseed in China is important for macro policy regulation of land use and national food security (Tao et al. 2019; Zang et al. 2020).

Remote sensing is an effective tool for large-extent crop mapping (Dong et al. 2016; Weiss, Jacob, and Duveiller 2020; Gumma et al. 2018; Huang et al. 2022; Mansaray et al. 2019; Zhou et al. 2019). Several methods have been proposed to identify rapeseed using remote sensing, which can be grouped into empirical index-based and machine learning-based methods (Ashourloo et al. 2019; Han et al. 2021; Tao et al. 2019; Zang et al. 2020; Sulik and Long 2020; Zhang, Liu, and Zhang 2022). Empirical index-based methods commonly utilize the spectral features of flowering, i.e. the distinctive bright yellow color of rapeseed flowers (Fang et al. 2016; Sulik and Long 2015). Thus, several yellowness indices, which enhance the spectral feature of increasing reflectance of red and green bands for the rapeseed flower, were developed to identify rapeseed (Ashourloo et al. 2019; Sulik and Long 2016, 2020; Zang et al. 2020). For example, Ashourloo utilized a Canola Index time-series, computed as $(\text{Green} + \text{Red}) \times \text{Nir}$, to classify rapeseed in 4 Iran and America test areas (Ashourloo et al. 2019). Han et al. developed an empirical index-based method to map rapeseed in 33 countries by combining the high normalized difference yellow index (NDYI) values on Sentinel-2 images during the flowering period and the high VH values on Sentinel-1 images during the podding period (Han et al. 2021). Sulik & Long combined NDYI and visible atmospherically resistant index (VARI) to identify flowering transition date of rapeseed (Sulik and Long 2020). Zang et al. developed an enhanced area yellow index (EAYI) by combining the features of the different yellow index (DYI) and NDVI during flowering period and mapped rapeseed flower coverage in five typical rapeseed planting areas in China (Zang et al. 2020). However, these methods heavily rely on cloud-free images acquired during the flowering period. Unfortunately, cloud-free observations during the flowering period in China are often lacking because rapeseed often flowers in the rainy season (Zang et al.

2020). This issue would lead to serious omission errors for empirical index-based methods. Furthermore, the flowering period is needed to obtain flowering images, which is difficult to be predicted accurately because rapeseed phenology varies spatially and yearly. Machine learning-based methods can address this issue because both flowering and nonflowering features can be learned from the training samples (Meng et al. 2020; Tao et al. 2019). For example, Tao et al. utilized an artificial neuron network to classify rapeseed on the Jiangnan Plain in China (Tao et al. 2019). Meng et al. used random forest to classify winter wheat and rapeseed, which showed that post-flowering features are also important for classification (Meng et al. 2020). However, collecting sufficient training samples required in machine learning-based methods is laborious and expensive (Ashourloo et al. 2019; Zhong, Gong, and Biging 2014). To overcome the above shortcomings, a hybrid method combining the advantages of empirical index-based and machine learning-based methods named seamless automated rapeseed mapping (SARM) was proposed very recently (Zhang, Liu, and Zhang 2022). First, a temporal difference in the winter rapeseed index (WRI) during the flowering season and before-flowering season, which empirically captures the unique phenological feature of rapeseed, was employed to automatically select rapeseed samples from cloud-free pixels. Then, based on these samples, multiple random forest classifiers were trained to classify each temporal image of the time series, and the classified results were integrated by voting on the classification probabilities of all single temporal images. This method works effectively in mapping rapeseed in the Yangtze River Basin. However, it still cannot meet the needs of large-scale and long-term rapeseed mapping in China for the following reasons. First, rapeseed is widely planted with a large phenological and spectral variation in China (Hu et al. 2017; Zang et al. 2020), and the non-rapeseed classes are more diverse than in the Yangtze River Basin. The WRI difference was designed for distinguishing winter rapeseed and winter wheat, which is inapplicable for other regions with more complicated rapeseed and non-rapeseed land cover types. Second, the SARM classified each temporal image independently to avoid the classifier learning the WRI difference itself as determining features; however, it would neglect the relationship among multitemporal features.

To address the above issue, a method integrating rule-based automatic sample generation and one-class classification (RSG-OC) was proposed to map rapeseed in China. A set of rules was developed based on empirical indices of flowering features to automatically select rapeseed samples from cloud-free pixels during the predicted flowering period. The samples were then augmented to improve the generalization to the pixels without cloud-free flowering features. Finally, a typical one-class classifier, positive unlabeled learning (PUL) implemented by random forest (PUL-RF), which only requires rapeseed samples and unlabeled samples randomly generated from all pixels, was applied to avoid the need for samples of diverse non-rapeseed classes (Chen et al. 2016; Lu and Wang 2021).

With the proposed method, rapeseed was mapped in China at 20 m resolution for 2017–2021 based on the Google Earth Engine (GEE) (Gorelick et al. 2017). Quantitative accuracy evaluation was conducted in 6 typical rapeseed planting areas, and 3 methods of rapeseed mapping were compared with the proposed method.

The rest of the paper is organized as follows. The materials and methodology are described in Sections 2 and 3, respectively. The results and discussion are presented in Section 4 and Section 5. Section 6 concludes the study.

2. Data and study area

2.1. Study area

The study aims to map rapeseed across China (73.5°E–135°E, 4°N–53.5°N) (Figure 1(a)). The main planting area of rapeseed is distributed widely in China, spanning from most southern in Yunnan (24.5°N, 101.5°E) to most northern in Hulun Buir (119.77°E, 49.21°N) and from most eastern in Jiangsu (119.8°N, 32.9°E) to most western in Yili (84.99°E, 44.45°N) (Wang, Guan, and Zhang 2007; Zang et al. 2020), with various climate types, including subtropical monsoon, temperate monsoon, and temperate continental climates. As a result, rapeseed shows a large difference in phenology and cultivar (Qian et al. 2006). Based on the field phenological observation records (Zhang, Liu, and Zhang 2022; Zhang et al. 2022; Li et al. 2021), the rapeseed phenological calendar in four typical rapeseed planting sites, Qujing (site 1), Jingzhou

(site 2), Haibei (site 3), and Hulun Buir (site 4), is presented in Figure 1(b). The phenological stages of rapeseed include seeding, emergence, budding, flowering, podding, mature, and harvest. All of these phenological stages are delayed from warm areas to cold areas. For example, the flowering dates ranged from February in the southernmost region (site 1) to August in the northmost region (site 4). Moreover, the length of phenological stages could differ among different cultivars. Especially, the winter rapeseed sowed from October to November in the southern region has a much longer emergence stage than spring rapeseed sowed from April to July in the northern region.

2.2. Data

2.2.1. Satellite data

Based on the growth calendar of rapeseed, Sentinel-1 and Sentinel-2 imagery during October 2016–October 2021 accessed on the GEE platform were used in this study.

Sentinel-1 is a C-band Synthetic Aperture Radar (SAR) satellite with dual polarization (VV and VH). To match Sentinel-2, we used the ground-ranging detection (GRD) product in IW mode, of which the spatial resolution is 10 m and the temporal resolution is 12 days (Torres et al. 2012). Previous studies have shown that the Sentinel-1 data has specific responses to the canopy structure of rapeseed, which helps to distinguish rapeseed from other crops (D'Andrimont et al. 2020; Veloso et al. 2017; Han, Zhang, and Cao 2021). Specifically, the VV scattered signal drops in the flowering season, and the VH scattered signal increases in the podding stage. Therefore, Sentinel-1 data plays an important role in identifying rapeseed in the case of cloudy weather.

Sentinel-2 is a multispectral optical imaging satellite with 10–60 m spatial resolution, 5–10 days temporal resolution, and 12 bands (Drusch et al. 2012). Both the top-of-atmosphere reflectance products at the L1C level (hereafter referred to as L1C data) and ground reflectance products at the L2A level (hereafter referred to as L2A data) were used in this study. The L2A data are used for sample generation with empirically designed rules because the thresholds in the rules could be more stable if the atmospheric effects are corrected (Zang et al. 2020). However, the availability of L2A data on GEE for most areas of China

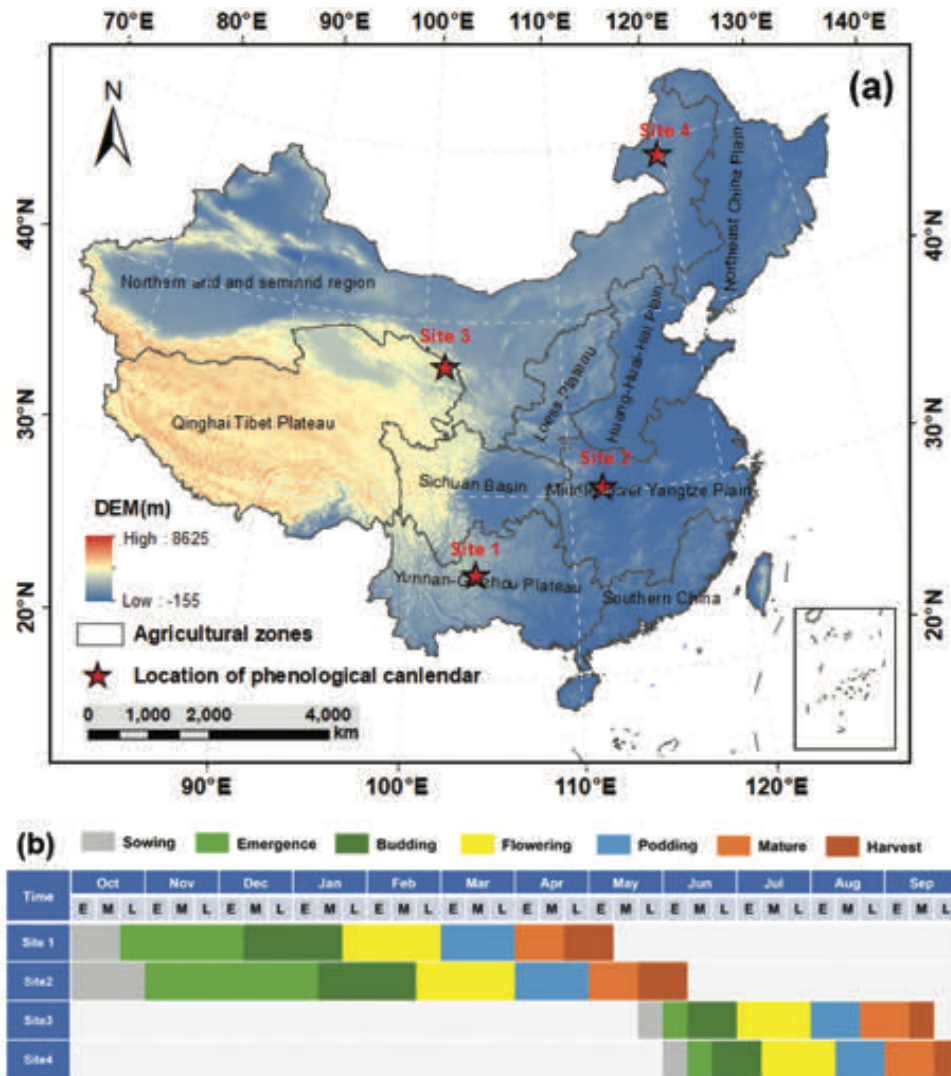


Figure 1. (a) Study area and distribution of rapeseed phenological calendar. (b) Phenological calendar of rapeseed in four typical sites. "E," "M," and "L" represent the early, middle, and late periods of the month, respectively.

(December 2018-present) was shorter than the availability of L1C data (August 2015-present). Therefore, L1C data-derived features were used for the classification of rapeseed considering that machine-learning techniques are not that sensitive to atmospheric effects (Jin et al. 2019; Pott et al. 2021). Moreover, the Sentinel-2 Cloud Probability product was used to remove clouds (ESA 2020).

2.2.2. Land cover data

We focused our study on cultivated land considering that most rapeseed is planted in cultivated land except for a negligibly small amount of rapeseed planted in urban green spaces. Such exclusion of non-cultivated land helps to save computational resources and avoid the interference of the non-cultivated land

in rapeseed mapping (Konduri et al. 2020; Zhang et al. 2017). To avoid the omission error of available cultivated land cover datasets, we integrated three widely used global land cover products, GlobeLand30-2020 (Chen et al. 2015), ESA World Cover V100 (Zanaga et al. 2021), and ESRI 2020 Global Land Use Land Cover (Karra et al. 2021), to take the union of their cropland layers to derive the regions to map.

2.2.3. Reference data

Ground truth data for evaluating the mapping results were obtained in six typical rapeseed planting regions by field survey or interpretation from very high-resolution images. As shown in Figures 2, 3 field surveys were conducted in Qujing city (Figure 2(d)), Sichuan Province

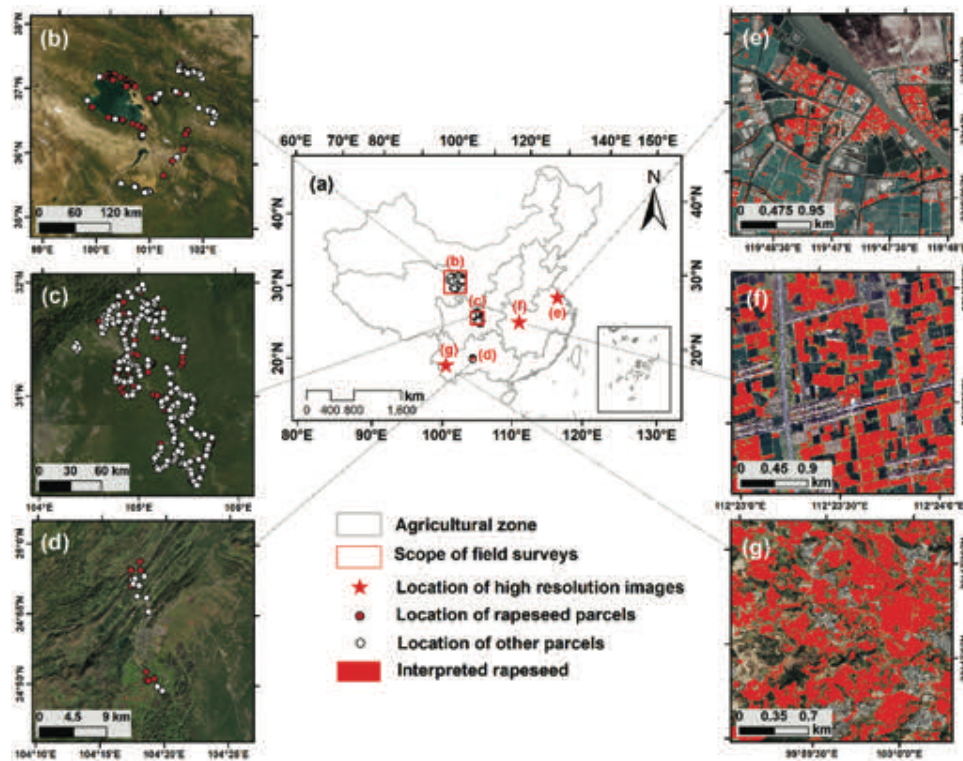


Figure 2. The locations of ground truth data. (b) to (d) are the distributions of validated parcels obtained by field campaigns in Qinghai, Sichuan, and Qujing. (e) to (g) are rapeseed maps interpreted from very high-resolution images acquired during the flowering period in Taizhou, Jingzhou, and Lincang.

(Figure 2(c)), and Qinghai Province (Figure 2(b)) during the flowering season (i.e. February, March, and July, respectively) of 2021. The GPS coordinates of crop parcels were recorded using mobile GPS devices. A total of 576 rapeseed and 494 non-rapeseed parcels were collected. These parcels were converted to binary images with 20 m resolution, and 105,179 rapeseed samples and 28,663 non-rapeseed samples were obtained. Two high-resolution Google Earth images in Taizhou (Figure 2(e)) and Lincang (Figure 2(g)) at a 0.5 m resolution and an unmanned aerial vehicle (UAV) image in Jingzhou (Figure 2(f)) at a 0.07 m resolution were collected during the rapeseed flowering period in 2020. All the high-resolution images were manually interpreted as rapeseed and non-rapeseed, resampled to 20 m resolution, and randomly selected as validation samples. A total of 13,744 rapeseed samples and 14,979 non-rapeseed samples were selected from the images as ground truth data for accuracy evaluation.

Census data of rapeseed area of 32 provinces in 2017–2020 were acquired from the China statistical

yearbooks released by the National Bureau of Statistics of China (<http://www.stats.gov.cn/tjsj/ndsj/>) to compare with the mapped rapeseed area at the provincial level.

3. Methodology

The proposed method includes five main stages: data preprocessing, sample generation, feature composition, sample augmentation, and PUL-RF classification (Figure 3).

3.1. Data preprocessing

As mentioned above, the input data for classification consist of Sentinel-1, Sentinel-2, and cropland data. The preprocessing of the Sentinel-1 data includes thermal noise removal, radiometric calibration, and topographic correction, implemented by a GEE toolbox (Thorpe and Drajat 2021). The preprocessing of the Sentinel-2 data includes cloud filtering, band selection, index calculation, compositing, and interpolation. First, the Sentinel-2 images with cloud cover below 95% were selected, and the pixels with cloud

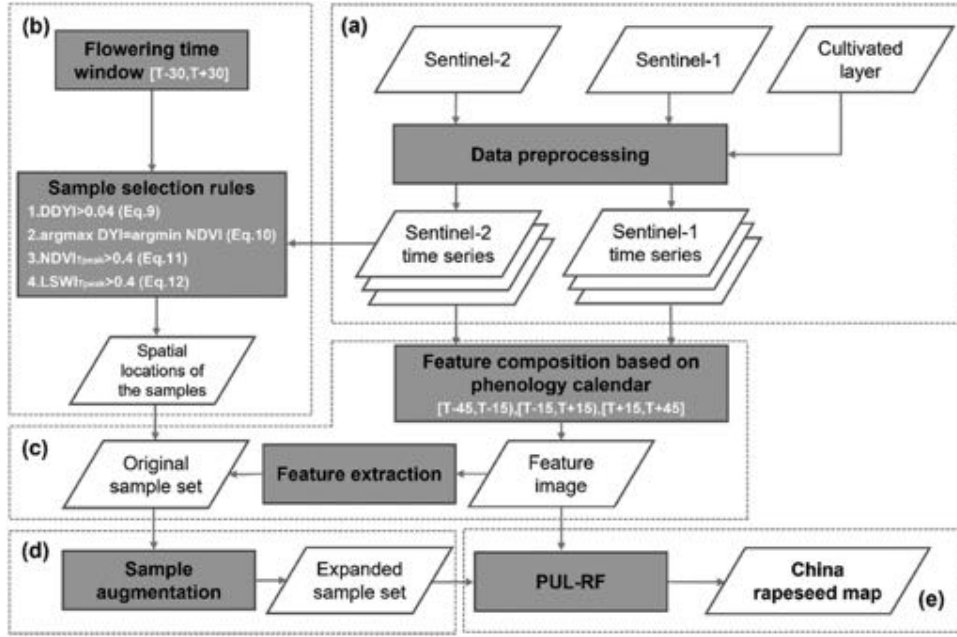


Figure 3. Flowchart of the proposed method. (a) Data preprocessing, (b) sample generation, (c) feature composition, (d) sample augmentation, and (e) PUL-RF classification.

probability above 50% were masked out using the ESA Sentinel-2 cloud probability product (ESA 2020). Second, the blue (B2), green (B3), red (B4), and near-infrared (B8) bands with a 10-m spatial resolution and the four red edge (B5 to B7, and B8A) and two short-wave infrared (B11 and B12) bands with a 20-m resolution were used. Third, the commonly used NDVI, DYI (Zang et al. 2020), NDYI (Sulik and Long 2016), and LSWI (Chandrasekar et al. 2010) were calculated for each pixel according to Eq. (1) to (4),

$$\text{NDVI} = \frac{B8 - B4}{B8 + B4} \quad (1)$$

$$\text{DYI} = B3 - B2 \quad (2)$$

$$\text{NDYI} = \frac{B3 - B2}{B3 + B2} \quad (3)$$

$$\text{LSWI} = \frac{B8 - B11}{B8 + B11} \quad (4)$$

The Sentinel-2 time series of each band and index were then composited by the 10-day median synthesis. Finally, the null values in the time series are linearly interpolated using the valid values from the adjacent times.

3.2. Generation of rapeseed samples

Sufficient samples are the basis for machine learning-based methods to perform large-area mapping (Foody and Mathur 2004; Maxwell, Warner, and Fang 2018). However, sample acquisition usually requires costly manual selection or field investigation (Friedl et al. 2002; Tian et al. 2020; Xiong et al. 2017). Considering that the yellow rapeseed flower is a distinctive feature compared to other crops (Fang et al. 2016; Sulik and Long 2020), a series of rules can be developed to select rapeseed samples automatically from cloud-free pixels, which only requires at least one clear observation during the flowering period (Zhang, Liu, and Zhang 2022).

The flowering period of each area was first determined by geolocation using a multilinear regression model according to (Zang et al. 2020). The model (Equation 5) was established between the multiyear average flowering date (T) recorded by agrometeorological stations and longitude (Lon), latitude (Lat), and altitude (Alt) (R^2 close to 0.9).

$$T = 7.07 \times \text{Lat} + 1.508 \times \text{Lon} + 0.03 \times \text{Alt} - 318.11 \quad (5)$$

Previous studies have examined several temporal features of rapeseed (Ashourloo et al. 2019; Han et al. 2021; Zang et al. 2020). Rapeseed flowers can increase

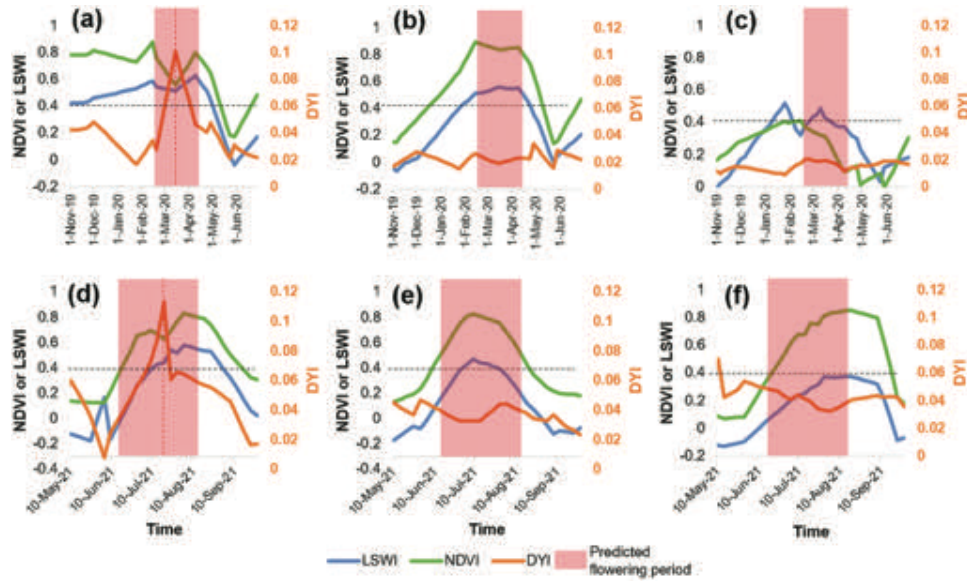


Figure 4. Temporal features of rapeseed and other crops. One parcel of each crop was selected for analysis. (a) Winter rapeseed, (b) winter wheat, and (c) water field in Jingzhou; (d) spring rapeseed, (e) barley, and (f) corn in Qinghai. The red dashed line indicates that the time when the DYI reaches a maximum is equal to the time when the NDVI reaches a minimum; the black dashed line is the threshold of the NDVI and LSWI.

the DYI and decrease the NDVI (Shen et al. 2014, 2010; Zang et al. 2020). Thus, we analyzed the DYI and NDVI of rapeseed on the Sentinel-2 time series. In addition, we also analyzed the LSWI, which can reflect the water content of rapeseed due to the high water content of rapeseed (Chandrasekar et al. 2010; Mandal et al. 2020). Comparing the time series of DYI, NDVI, and LSWI of rapeseed and other crops in two typical rapeseed growing areas, Jingzhou and Qinghai, we found several temporal features that can be used to develop the rules for sample selection (Figure 4). During the predicted flowering period, the DYI of rapeseed peaked, while the NDVI decreased to a local minimum because of the yellow flowers (Zang et al. 2020). In contrast, the DYI values of other crops remained low, and there was no such NDVI valley for them during the period. The LSWI of rapeseed is high during the flowering period, which will be used to remove interference pixels.

The DYI peak in the Moderate Resolution Imaging Spectroradiometer (MODIS) time series has been demonstrated to be a consistent feature of rapeseed in different regions of China by subtracting the base DYI at nonflowering periods from the peak DYI (Zang et al. 2020). We adapted the idea for the Sentinel-2 time series and derived the differential-DYI (DDYI) with a modified base DYI (Eq. (7)) due to the poorer data availability of Sentinel-2 compared with MODIS.

DDYI, calculated by Eqs. 6 to 8, is a global adaptable flowering index across China and is used as a main feature for selecting rapeseed samples.

$$DDYI = DYI_{[T-30, T+30]}^{\max} - DYI_{\text{base}}^{\text{mean}} \quad (6)$$

$$DYI_{\text{base}}^{\text{mean}} = \frac{DYI_{[T_{\text{peak}}-60, T_{\text{peak}}-15]}^{\text{mean}} + DYI_{[T_{\text{peak}}+15, T_{\text{peak}}+60]}^{\text{mean}}}{2} \quad (7)$$

$$T_{\text{peak}} = \text{argmax}DYI_{[T-30, T+30]} \quad (8)$$

where T_{peak} is the time when DYI achieves the maximum during the period $[T - 30, T + 30]$, and $DYI_{\text{base}}^{\text{mean}}$ is the basic DYI during the nonflowering period $[T_{\text{peak}} - 60, T_{\text{peak}} - 15]$ and $[T_{\text{peak}} + 15, T_{\text{peak}} + 60]$. Note that a time window of $[T-30, T+30]$, which was longer than the flowering period (20–40 days in usual), was applied here to ensure that the flowering image can be found within the period. A pixel with a higher DDYI indicates that the pixel is more likely to be flowering rapeseed. To select rapeseed samples as accurately as possible, an empirically strict threshold is applied to DDYI:

$$DDYI > 0.04 \quad (9)$$

Some non-rapeseed crops could also have a high DDYI if their greening-up stage overlaps with the flowering period of rapeseed. In this case, the reflectance in the green band increases dramatically,

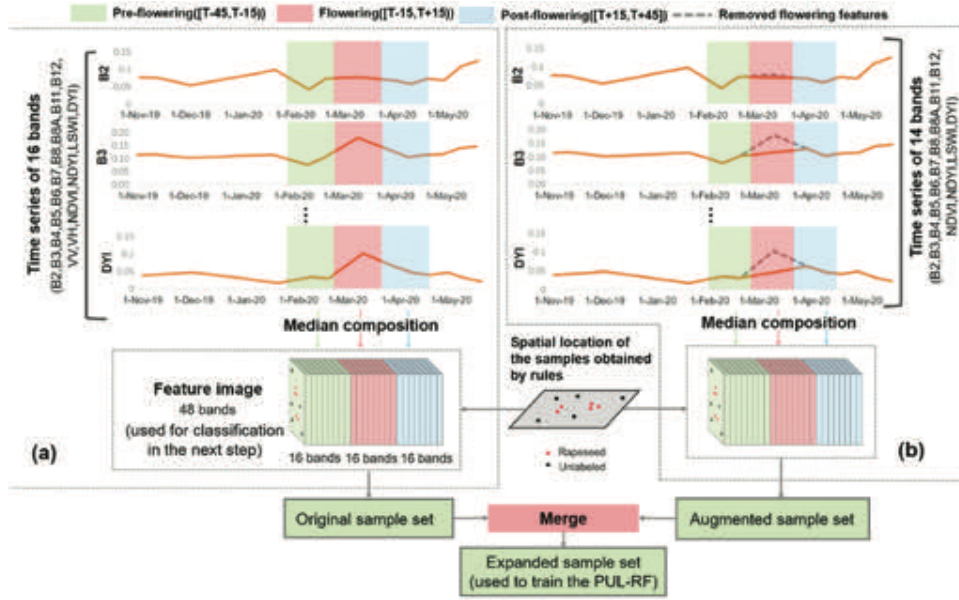


Figure 5. Schematic diagram of (a) feature composition and (b) sample augmentation.

leading to a high DYI value. Fortunately, the NDVI valley induced by yellow flowers could be used for distinguishing rapeseed from such crop types (Zang et al. 2020). Therefore, these nonrapeseed crop pixels can be excluded by equation (10):

$$\operatorname{argmax}DYI_{[T-30, T+30]} = \operatorname{argmin}NDVI_{[T-30, T+30]} \quad (10)$$

Which means that the observed DYI peak should correspond to an NDVI valley for the rapeseed pixels (Zang et al. 2020). Equation 9 and are the main rules for selecting rapeseed samples across all of China.

However, some occasional non-rapeseed pixels with undesired high DDYI values remain, including rainbowlike clouds (Frantz et al. 2018; Han et al. 2021) and blue impervious surfaces (Figure S1). Therefore, two additional criteria Equation (11) and Equation (12) were used to remove such pixels from the samples selected by the main rules of Eq. (9) and (10):

$$LSWI_{T_{\text{peak}}} > 0.4 \quad (11)$$

$$NDVI_{T_{\text{peak}}} > 0.4 \quad (12)$$

As such types of nonrapeseed pixels have much lower LSWI or NDVI values than rapeseed pixels, the thresholds are relatively easy to determine. A more detailed description of these two additional criteria is included in the supplementary materials.

Finally, with the proposed rules (Equation 9-Equation 12), sufficient rapeseed samples from 2019 to 2021 were generated across China based on the Sentinel-2 L2A data (Figure 6). Meanwhile, the same number of unlabeled samples were also randomly generated from all cultivated land pixels, which is also an essential input of the one-class PUL classifier that will be introduced in Section 3.5. Then, the spatial locations of these generated samples across different years were merged to generate a training sample set based on Sentinel-2 L1C data.

3.3. Feature composition based on phenology calendar

Variation in rapeseed phenology in different regions of China and availability of cloud-free satellite data challenges classification. To alleviate their negative effects, the phenology of rapeseed was first divided into three periods according to the flowering date (T): preflowering ($[T-45, T-15]$), flowering ($[T-15, T+15]$) and postflowering ($[T+15, T+45]$). The three periods could roughly correspond to the budding, flowering, and podding stages, which all last approximately 30 days in different geographic locations (Figure 1). Then, the median composition was employed for each time series at each stage (Phan, Kuch, and Lehnert 2020), and a total of 48 composited feature

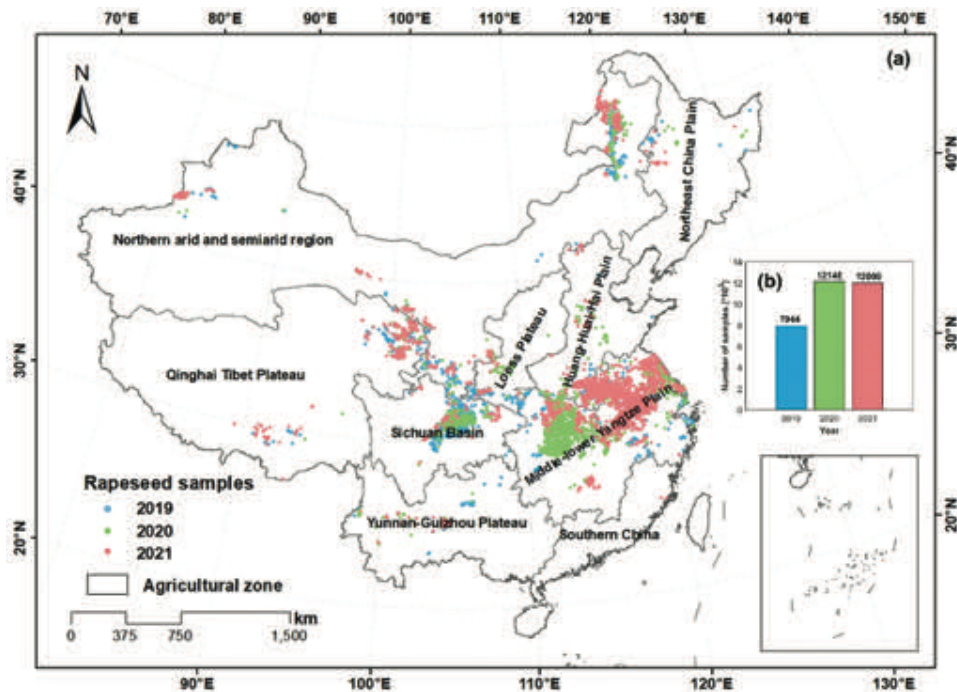


Figure 6. (A) Spatial distribution of generated rapeseed samples; (b) Number of generated rapeseed samples for 2019–2021.

images were generated (three stacks of 16 features from Sentinel-1 and –2) (Figure 5(a)) for the consequent classifier training and classification.

3.4. Sample augmentation

Note that the above generation of samples and the corresponding composited features were mainly based on the flowering signal. These samples thus would fail to represent the rapeseed pixels without adequately clear observation of flowering due to cloud blocking, because the interpolation step in Section 3.1 could produce distorted values in case of continuous cloudy weather. To improve the representativeness of the rapeseed samples, the samples were augmented by purposefully replacing the original cloud-free observations in the flowering period ($[T-15, T+15]$) with the linearly interpolated values from the valid observations in the pre and postflowering periods (i.e. $[T-45, T-15]$ and $[T+15, T+45]$) (Figure 5), thus simulating the poor observation condition. Note that the simulation was employed for all the features except VV and VH because Sentinel-1 is also effective under cloudy weather (D’Andrimont et al. 2020; Mercier et al. 2020). Sample augmentation was applied to both the rapeseed samples and the unlabeled samples. The expanded sample set merged the original and

augmented sample sets in equal proportions. With the expanded sample set including adequate clear and unclear samples, machine-learning method is expected to learn and adaptively use both flowering and non-flowering features for rapeseed classification.

3.5. Application of PUL-RF

Considering the diversity of the nonrapeseed classes, the one-class random forest classifier implemented by the PUL strategy (PUL-RF), which only needs the positive samples of the target land cover type and the unlabeled samples randomly generated from all pixels in cultivated land layers (Chen et al. 2016; Elkan and Noto 2008; Li, Guo, and Elkan 2010), was employed to map rapeseed. Here, the unlabeled samples include both the positive and negative samples, thus, were generated automatically without laborious labeling efforts.

PUL is designed to transform the traditional binary classifier into a one-class classifier based on the Bayesian rule and has been proven effective for mapping other land cover types (Lei et al. 2021; Lu and Wang 2021). It includes two steps. First, a binary random forest classifier is trained to calculate the probability that a pixel is labeled. Then, the probability that a pixel belongs to rapeseed is adjusted by a factor (c)

that represents the labeling probability of positive pixels

$$f(y = 1|x) = \frac{p(s=1|x)}{c} \quad (13)$$

where $f(y = 1|x)$ represents the probability that a pixel is rapeseed, x is the feature of the pixel and $y \in \{-1, 1\}$ represents the land cover types (non-rapeseed and rapeseed). $p(s = 1|x)$ is the probability that a pixel is labeled, and $s \in \{0, 1\}$ represents whether it is labeled or not. The adjusting factor (c) was estimated through an independent validation set (V).

$$c = \frac{1}{n} \sum_{x \in V} p(s = 1|x) \quad (14)$$

Here, the independent validation set (V) with n pixels was divided from the expanded rapeseed samples. This validation set (V) is excluded in the training phase but used for estimating the adjusting factor (c) and determining the hyperparameter of the random forest. Finally, a single PUL-RF classifier trained with the expanded sample set was used to map rapeseed during 2017–2021.

3.6. Accuracy assessment

To evaluate the mapping accuracy, the overall accuracy (OA), F1-score, producer accuracy (PA), and user accuracy (UA) were calculated using the ground truth data in the six experimental regions (Section 2.2.3). The mapping accuracy and spatial details of RSG-OC were also compared with those of the NDVI-based method (Han et al. 2021), the CI-based method (Ashourloo et al. 2019), and SARM (Zhang, Liu, and Zhang 2022).

4. Results

4.1. Spatial distribution of rule-based generated samples

Based on the sample generation rules and sample augmentation, a total of 61,464 samples were generated, including 32,090 rapeseed samples and 29,374 unlabeled samples. The number of rapeseed samples for 2019, 2020, and 2021 was 7944, 12164, and 12,000, respectively. Generally, the spatial distribution of the generated rapeseed samples is consistent with the main rapeseed production regions, and it varies slightly from year to year due to the cloud cover variation during the flowering period (Figure 6).

4.2. Map of rapeseed in China

The generated expanded sample set is divided into a training set and a validation set (V) by a ratio of 7:3 to train the PUL-RF. The tree number of the RF was set as 225 based on the highest OA on the validation set. The parameter c in PUL was derived from Eq.14, equaling to 0.87 in this study. The other parameters of the PUL-RF, such as the number of variables per split, were set to the default values in the GEE function of *smileRandomForest*. Using the trained PUL-RF, 20 m rapeseed maps of China for 2017–2021 were produced. The data link is provided in the data availability section. Consistent with prior knowledge (Wang, Guan, and Zhang 2007; Bonjean, Dequidt, and Sang 2016), rapeseed was widely cultivated in China in several regions, including the middle-lower Yangtze Plain, Sichuan Basin, Yunnan-Guizhou Plateau, north-eastern Qinghai Tibet Plateau, and northeastern and northwestern North arid and semiarid regions (Figure 7(a)). The mapped area of rapeseed in China during 2017–2021 is 2.46, 2.34, 2.35, 2.81, and 2.75 million hectares respectively (Figure 7(b)). Although the mapped area seems to have a similar temporal fluctuation with the census data, there is a large underestimation of the mapped area compared to the census data, indicating a nonnegligible difference between the two data sources.

To further illustrate the reasonability of rapeseed change in the produced annual maps, Figure 8 compared the temporal change in the rapeseed map over 2017–2021 in detail in four rapeseed planting sites. The Sentinel-2 images at the flowering period are also presented for visual comparison if they are cloud-free. Otherwise, the most cloud-free and temporally close images were used, and fortunately, rapeseed can still be identified on most of these images because the green color of rapeseed is subtly different from that of other vegetation. Visually, the maps match the rapeseed parcels every year and present clear details.

4.3. Visual comparison with other methods

Figure 9 compares the Sentinel-2 images during the flowering period, the ground truth (ground truth rapeseed maps interpreted from very high-resolution images in Jingzhou, Taizhou, and Lincang and crop parcels acquired from field surveys in Sichuan, Qujing, and Qinghai from Section 2.2.1),

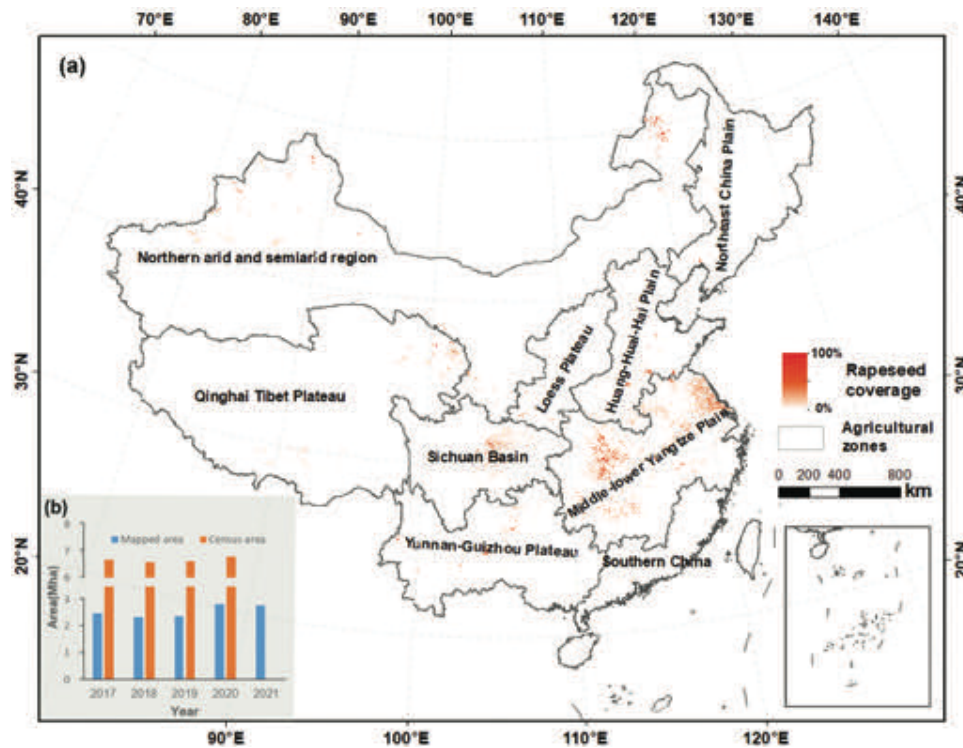


Figure 7. (a) Chinese rapeseed coverage map at 500 m resolution in 2020 (the original 20 m resolution was resampled to 500 m resolution for a better illustration in the thumbnail figure); (b) China rapeseed mapping area and census area, 2017–2021 (Note: census area in 2021 is unavailable).

and the mapping results of the NDYI-based method, the CI-based method, SARM, and the proposed RSG-OC at six validation areas. It should be noted that owing to constant cloud cover, the compared Sentinel-2 images in Sichuan and Qujing are not appropriate (and the one is at the nonflowering period and the other is partly covered by clouds) but are already the best we could find. In general, the mapping results of RSG-OC were in good agreement with the spatial distribution at all six validation areas, whereas the other methods produced large classification errors in some areas. Specifically, the NDYI-based and CI-based methods have a large omission error in the Qujing and Sichuan because they rely heavily on cloud-free observations during the flowering period. The results of SARM are poorly consistent with the actual rapeseed distribution at the Qinghai and Sichuan areas because their sample selection rule is inapplicable at these two areas.

4.4. Quantitative evaluation

The quantitative assessment is summarized here using the ground truth dataset (Table 1) and the agricultural

census area data (Figure 10). Among all of the compared methods, the proposed RSG-OC has the best performance with an average OA of 94.90% and an average F1-score of 93.50%. RSG-OC outperformed the other three methods in Taizhou, Sichuan, and Qinghai and achieved the second-highest OA in Jingzhou, Lincang, and Qujing. Although RSG-OC achieved stable performance in different areas, the PA in Sichuan was relatively low, indicating a large omission in this region. It is probably induced by very high cloud coverage (>54%) in Sichuan areas. SARM performed more poorly in Sichuan and Qinghai, with OAs of 81.58% and 56.59% and F1-scores of 61.75% and 67.70%, respectively, because the sample generation rules of SARM do not apply to some land cover in these two regions. The CI-based method has large omission errors in Sichuan and Qujing, with PAs of 42.69% and 43.69%, respectively, in these two regions. Similarly, the NDYI-based method has a poorer performance, with PAs of 0.35% in Sichuan and 22.13% in Qujing.

The RSG-OC mapped rapeseed area in each province was further compared with the official census area for 2017–2020. As shown in Figure 10, the mapped area is significantly correlated with the

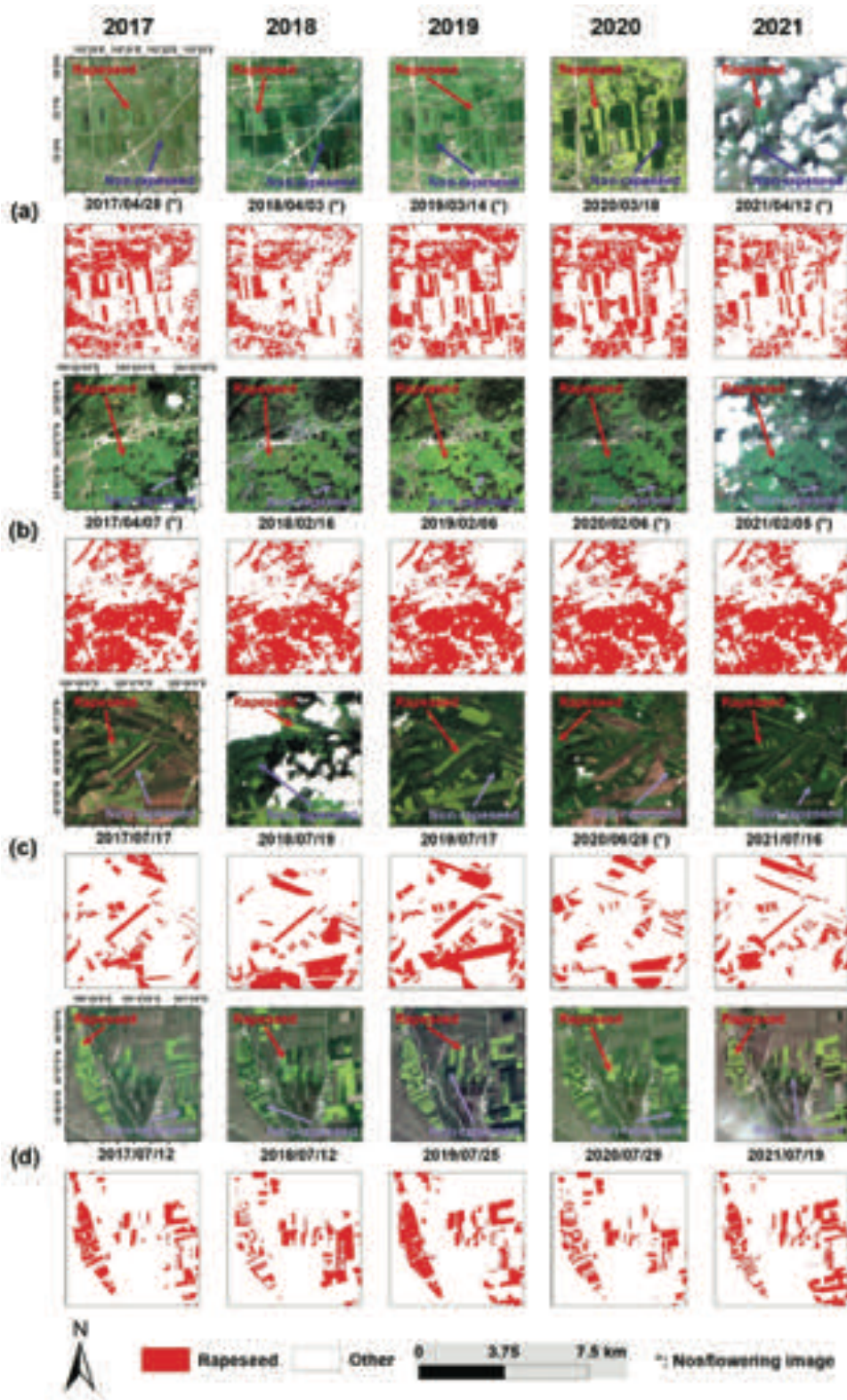


Figure 8. Detailed comparison of multiyear mapping results in 4 typical rapeseed planting areas. (a) Jingzhou, (b) Qujing, (c) Hulun Buir, (d) Qinghai. The first line in each region is the Sentinel-2 true color images, and the second line of each area is the mapped rapeseed. Images with '*' after the annotated date were acquired in the nonflowering period.

census area. However, the mapped areas are much lower than the census areas for most provinces; and the coefficients of determination (R^2) are generally

low, with some outliers like Jiangsu Province. The low consistency between the mapped area and census data might be induced by the uncertainties in

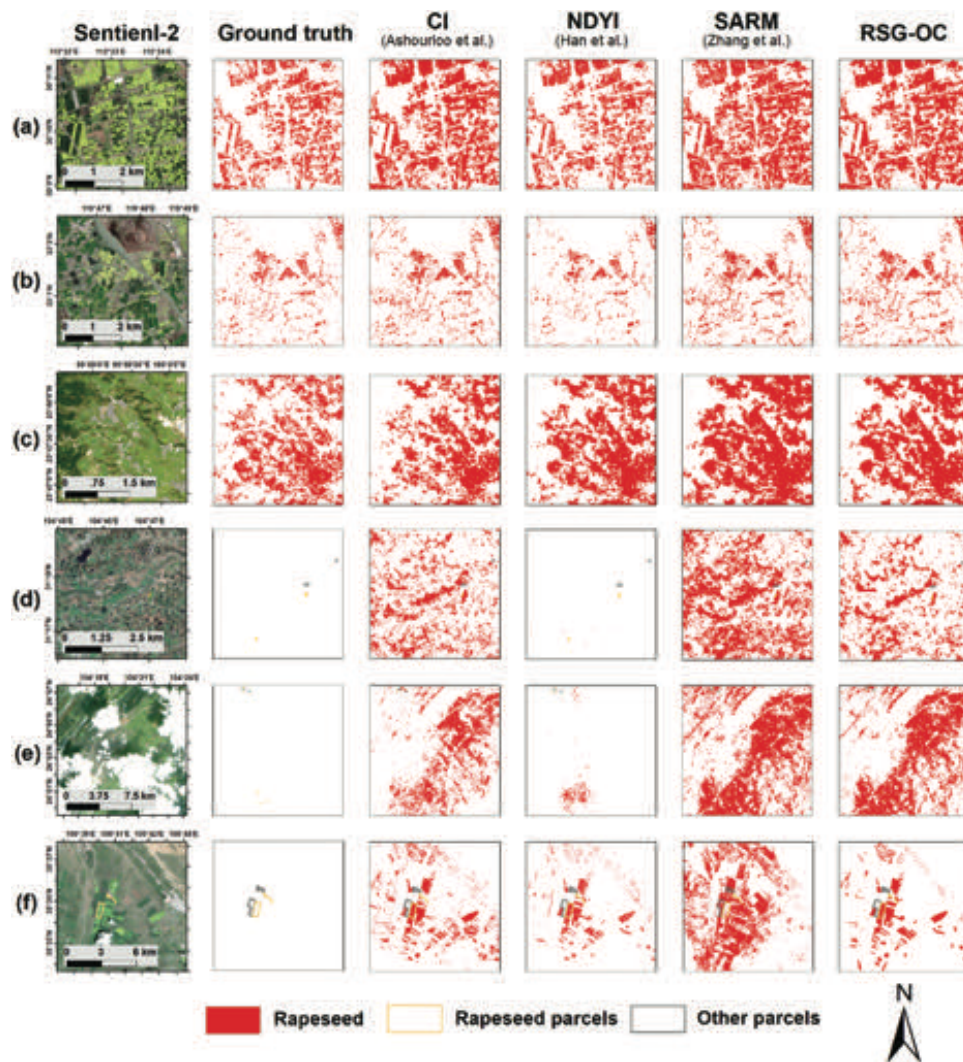


Figure 9. Comparison of the mapping details of the four methods in the validation areas (a) Jingzhou, (b) Taizhou, (c) Lincang, (d) Sichuan, (e) Qujing, and (f) Qinghai. The first column shows the RGB images of Sentinel-2 during the flowering periods in the validation areas except for Sichuan area, where the flowering images were totally cloud contaminated and the image acquired after flowering was shown instead. The second column shows the ground truth maps, which were interpreted from very high-resolution images for (a-c) and obtained from field-survey for (d-f). Columns 3 to 6 show the mapping details of CI, NDYI, SARM, and proposed RSG-OC respectively.

both the remotely sensed maps and the census data (Liu et al. 2020), which should be further explored in the future.

5. Discussion

5.1. Dependence on clear observation in the flowering period

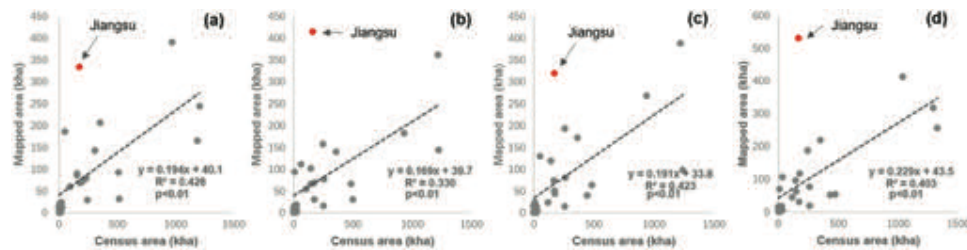
As aforementioned, both the index-based (CI and NDYI) and the machine learning-based methods (SARM and RSG-OC) require clear flowering observations but to different extents. The key difference here

is the method's sensitivity to the availability of clear observations in flowering season and should be investigated. Considering that RSG-OC employed sample augmentation and Sentinel-1 SAR data to address the issue of cloud contamination, two incomplete RSG-OC methods without sample augmentation and Sentinel-1 data were also compared to illustrate the importance of these two steps. To simulate the time-series data with weakening flowering signals induced by cloud contamination, we sequentially removed the observations during the flowering period with the top 33%, 66%, and 100% (namely, all removed) DYI values and checked how

Table 1. Comparison of the accuracy of the four methods.

Region	Metrics	NDYI	CI	SARM	RSG-OC
Jingzhou	OA	91.52%	95.39%	88.76%	92.64%
	UA	96.58%	<u>91.97%</u>	87.19%	87.58%
	PA	86.05%	<u>99.46%</u>	90.85%	99.53%
	F1-score	91.01%	95.57%	88.98%	93.17%
Taizhou	OA	91.17%	<u>93.72%</u>	93.89%	94.43%
	UA	98.51%	<u>94.30%</u>	92.62%	93.37%
	PA	80.69%	<u>90.88%</u>	93.21%	93.71%
	F1-score	88.71%	92.56%	<u>92.91%</u>	93.54%
Lincang	OA	90.87%	86.44%	94.41%	93.47%
	UA	98.09%	97.77%	96.33%	<u>97.86%</u>
	PA	83.36%	74.58%	92.34%	88.89%
	F1-score	90.13%	84.62%	94.29%	<u>93.16%</u>
Sichuan	OA	83.24%	86.47%	81.58%	95.49%
	UA	57.14%	<u>64.64%</u>	47.38%	90.66%
	PA	0.35%	42.69%	88.63%	81.53%
	F1-score	0.70%	51.42%	<u>61.75%</u>	85.85%
Qujing	OA	64.79%	74.53%	98.91%	98.78%
	UA	100.00%	100.00%	100.00%	100.00%
	PA	22.13%	43.69%	97.59%	97.30%
	F1-score	36.24%	60.81%	98.78%	<u>98.63%</u>
Qinghai	OA	78.52%	92.93%	56.59%	94.58%
	UA	97.66%	<u>99.17%</u>	87.44%	99.69%
	PA	75.74%	<u>92.19%</u>	55.23%	93.72%
	F1-score	85.31%	<u>95.55%</u>	67.70%	96.61%
Average OA*		83.35%	<u>88.25%</u>	85.69%	94.90%
Average F1-Score*		65.35%	<u>80.09%</u>	84.07%	93.50%

*Average OA and average F1-score denote the averaged values of OA and F1-score of the six areas.

**Figure 10.** Comparison mapped rapeseed area with census area at the provincial level for (a)2017, (b)2018, (c)2019, (d)2020.

the performances of the compared methods worsened. As shown in Figure 11, both RSG-OC and SARM were stable against the decreasing flowering observations, whereas the average OA and F1-score of the CI-based and NDYI-based methods dropped largely as the clear flowering observations decreased, and they completely failed when there were no clear flowering observations. The performance of two incomplete RSG-OC methods without samples augmentation and Sentinel-1 data also decreased with decreasing clear observation in the flowering season. Especially, when none of the clear observations was available during the flowering season, two incomplete RSG-OC methods performed much poorer than standard RSG-OC. These results confirmed the contribution of both the Sentinel-1 SAR data and sample augmentation on resisting the issue of cloud contamination.

The role of sample augmentation could be further explained by the differences between the feature importance of RFs in standard RSG-OC and the incomplete RSG-OC without sample augmentation. As Figure 12 shows, the optical flowering features, e.g. $LSWI_{flower}$, $NDYI_{flower}$, and $B11_{flower}$, exhibited great importance for RSG-OC without sample augmentation. These optical features unfortunately would be largely weakened by cloud contamination, and thus is highly dependent on adequate clear observation. For standard RSG-OC, the importance of optical flowering features decreased, whereas SAR features and postflowering optical features became more important. Specifically, the VH_{post} , VH_{flower} , and VV_{pre} ranked 1st, 3rd, and 5th respectively, and $B11_{post}$, $B5_{post}$, and $LSWI_{post}$ ranked 2nd, 4th and 6th respectively. Therefore, the RSG-OC method is not that dependent on cloud-free observations during the flowering period.

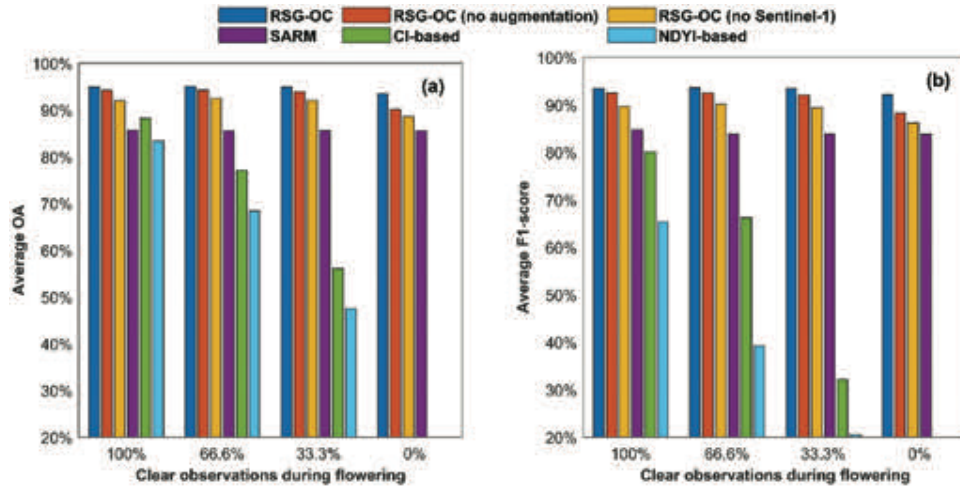


Figure 11. The sensitivity of the different methods to clear observation in the flowering period.

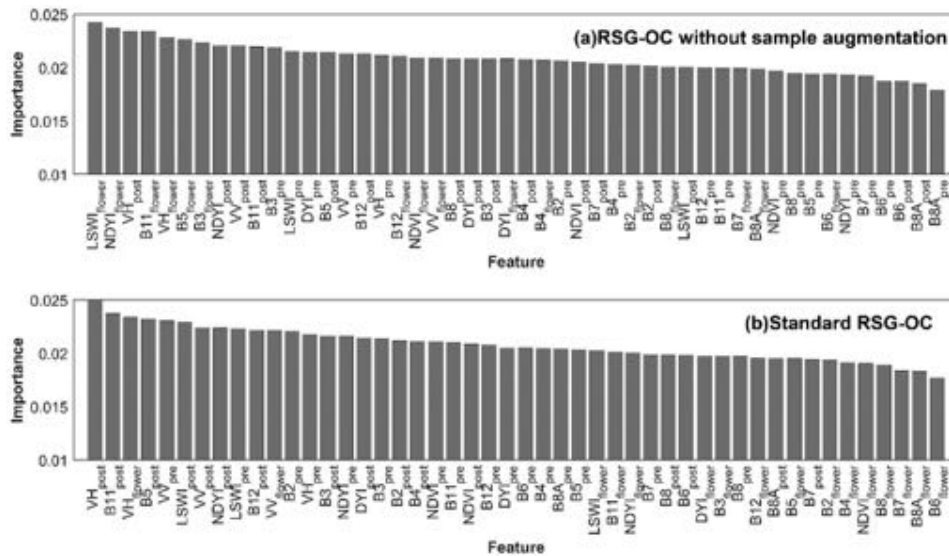


Figure 12. Ranking of the feature importance in (a) RSG-OC without sample augmentation and (b) standard RSG-OC. The subscripts after the features indicate the period of feature composition. “pre” represents the “preflowering” period, “flower” represents the “flowering” period, and “post” represents the “postflowering” period.

5.2. Comparison with SARM

RSG-OC adapted the strategy of SARM that combines empirical index-based sample generation and machine-learning methods to map rapeseed in China. Compared with SARM, there are three main improvements of RSG-OC. First, an empirical rule set with a stricter “yellowness” feature compared with the WRI index was employed to automatically select rapeseed samples from cloud-free pixels during the flowering period. The stricter “yellowness” rule helps to better select correct rapeseed samples across the whole of China; however, the representativeness of the generated samples could be biased. Therefore,

the second improvement compared to SARM is a specially designed sample augmentation that removes the flowering features of the generated rapeseed samples. Such sample augmentation helps to avoid the classifier learning the strict yellowness rule as a classification feature and improves the generalization to the pixels without cloud-free flowering features. Finally, a typical one-class classifier, positive unlabeled learning (PUL) implemented by random forest (PUL-RF), which only requires rapeseed samples, was applied to avoid the need for samples of diverse nonrapeseed classes (Chen et al. 2016; Lu and Wang 2021).

Especially, the rules employed in SARM method was designed for rapeseed in Yangtze River Basin (Zhang, Liu, and Zhang 2022), which was not applicable for other areas and lead to poor performance in Qinghai area (Table 1). The SARM selects rapeseed samples with following rules,

$$\text{NDVI}_{\max} > 0.5 \quad (15)$$

$$\text{WRI}_{\text{bf}} - \text{WRI}_{\text{f}} > 0 \quad (16)$$

$$\text{WRI} = \frac{\text{B8} - \text{B3}}{\text{B8} + \text{B3}} \times \frac{\text{B2}}{\text{B3} + \text{B4}} \quad (17)$$

where NDVI_{\max} is the maximum NDVI during the growing season, WRI_{bf} is the WRI before flowering and WRI_{f} is the WRI at flowering. In the Yangtze River Basin, the above rules are adequate to distinguish the winter rapeseed from other winter crops. However, such rules failed to accurately select rapeseed samples when they were applied to a larger extent. Some other usual vegetation types outside the Yangtze River Basin were likely selected by the SARM rules. In Qinghai, for example, the barley and natural grasslands also satisfy the rules of Equation 16 (Figure 13(a,b)). In Sichuan, the evergreen forest also satisfies the rules of Equation 16 (Figure 13(c)). In summary, the sample selection rule of SARM is too relaxed to be extended to a larger extent. In contrast, the strict sample selection rules of RSG-OC are more universal to generating accurate rapeseed samples, and these types of vegetation will not be selected by the rules of RSG-OC due to the low DDYI. However, a strict rule of flowering features could generate a biased sample set that cannot represent the rapeseed pixels without clear flowering features. Therefore, another improvement of RSG-OC compared to SARM is the employment of a specifically

designed sample augmentation that removes the observations during the flowering period. Such augmentation avoids the classifier from learning the biased selecting rule itself, thus enhancing the generalization performance of the trained classifier.

5.3. Sensitivity of sample selection rule

In this study, the sample generation rules can generate accurate rapeseed samples with easy-determining thresholds. To illustrate the robustness to the varying threshold, the number of pixels satisfying the sample selection rules, the total area of mapped rapeseed (denoted as A_{rapeseed}), and the classification accuracy at different thresholds of the main rules (Eq. (9)). Considering that the OA and F1-score do not reflect the error in the mapped area due to the inconsistency between the proportion of rapeseed in the ground truth data and the proportion of the actual area, the proportion correct (PC) was also used as an indicator of accuracy (Pontius and Millones 2011), which was calculated using Equation 18 to Equation 19:

$$\text{PC} = \sum_{i=1}^J p_{ii} \quad (18)$$

$$p_{ij} = \frac{n_{ij}}{\sum_{j=1}^J n_{ij}} \times \frac{N_j}{\sum_{i=1}^J N_i} \quad (19)$$

where J is the number of classes, which is two here; n_{ij} is the number of samples attributed to class i that are classified as class j ; and N_i is the total number of pixels classified as category i .

As shown in Figure 14, as the DDYI threshold becomes strict, the number of pixels satisfying the sample selection rules always decreases steadily. However, A_{rapeseed} decreases at different rates at different thresholds. When the threshold is relaxed, the

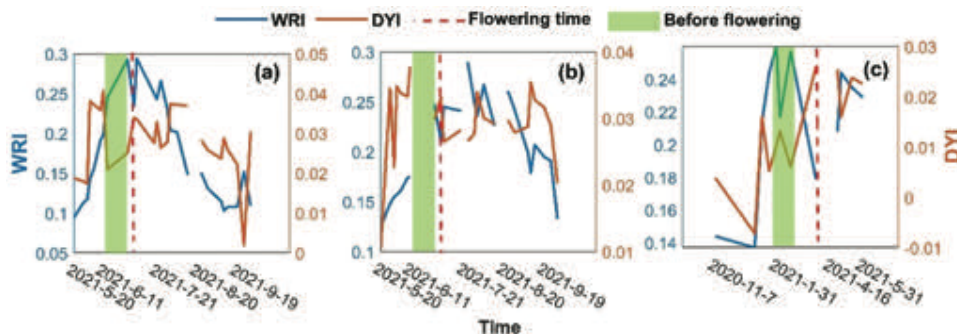


Figure 13. The WRI and DYI time series of (a) barley, (b) grass in Qinghai and (c) evergreen forest in Sichuan. The time series was selected from one typical parcel for each crop type.

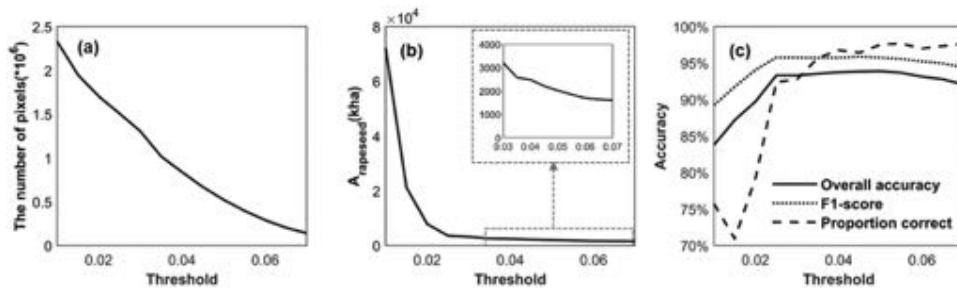


Figure 14. The variation of (a) number of pixels satisfying the sample selection rules, (b) $A_{rapeseed}$, and (c) accuracy with DDYI threshold.

$A_{rapeseed}$ drops largely. As the threshold tends to be strict, the $A_{rapeseed}$ gradually stabilizes. Similarly, as the threshold becomes stricter, the OA, F1-score, and PC generally become stable, although the PC is much lower when the threshold is relaxed. The cause can be explained as follows. When the threshold is relaxed, the selected samples are mixed with many inaccurate samples. Thus, more nonrapeseed was misclassified as rapeseed, leading to large classification error. As the threshold becomes stricter, the flowering features of selecting samples could be more biased for representing ordinary rapeseed pixels. Fortunately, the sample augmentation that removes the flowering features allows the classifier to focus more on nonflowering features. Therefore, bias in flowering features does not lead to large changes in $A_{rapeseed}$ and accuracy. Therefore, it is easier to select a strict global threshold for large-extent rapeseed mapping.

5.4. Cross-year generalization ability of RSG-OC

In this study, the PUL-RF trained by the samples generated from 2019–2021 was also used for mapping

rapeseed in 2017–2018 (Figure 8). Unfortunately, there was a lack of ground truth samples in 2017–2018 for evaluation. Thus, to confirm the reliability of such a cross-year mapping strategy, we regenerate two subsample sets by excluding the generated samples in 2020 and 2021 and retrain two PUL-RF classifiers by these two regenerated sample sets (abbreviated as the non2020-model and non2021-model, hereafter). Then, the ground truth samples in 2020 and 2021 were used to evaluate the non2020-model and non2021-model, respectively. The performance was also compared with the standard PUL-RF classifier trained by all samples (abbreviated as all-model hereafter).

As shown in Figure 15, the average OA of the non2020-model and non2021-model only decreased by 0.14% and 0.32% compared to the all-model, respectively. Similarly, the average F1-score of the two models only decreased 0.08% and 0.28% compared to the all-model. Thus, classifiers trained by the generated samples from other years are able to map rapeseed in the years without generated samples, suggesting that the 2017–2018 rapeseed delineated by the classifier for 2019–2021 is also reliable.

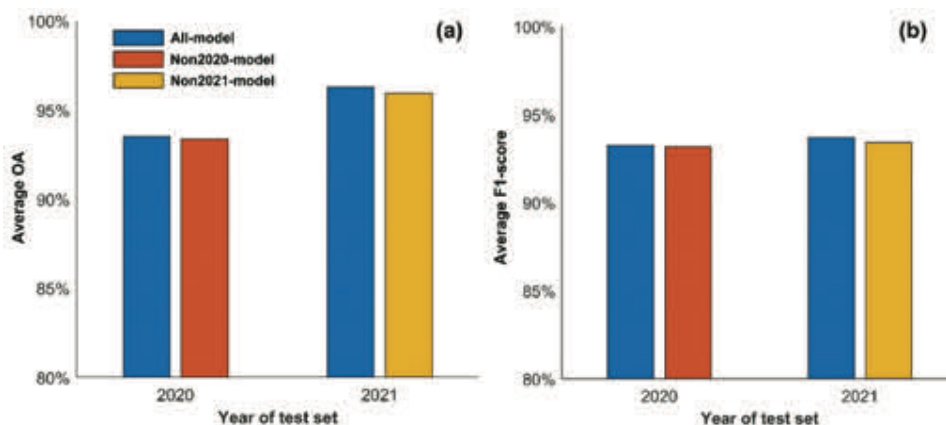


Figure 15. Comparison of (a) average OA and (b) average F1-score before and after excluding the 2020 and 2021 training samples.

5.5. Influence of the predicting error of flowering date

In this study, the flowering features used for PUL-RF were extracted from the flowering period of $[T-15, T+15]$, where the flowering date T is predicted at each location by our previously proposed model (Zang et al. 2020). However, the predicted flowering date may be biased because the actual flowering date varies from year to year due to climate interannual variability. Therefore, we examined how the predicting error of flowering date could influence the classification accuracy. Firstly, we compared the 734 flowering dates recorded at the agrometeorological stations from 2003–2017 and their corresponding predicted flowering dates. The predicted RMSE of flowering date was 13.09 days (Figure 16(a)), and the predicting error of flowering date obeyed a normal distribution with a mean of 0.23 and a standard deviation of 13.09 (Figure 16(b)). It means that the predicting error is less than 13 days with a probability of 68.2% and is less than 26 days with a probability of 95.5%. Thus, we manually added -30 to 30 -day offsets (one offset every 5 days) to the predicted flowering dates to simulate the possible errors of the predicted flowering date; and then evaluated the classification performance under each offset (Figure 16(c)). It is found that average OA and average F1-score stabilized above 94% and 93% respectively if the predicting error of flowering date is less than 15 days. If the predicted flowering date deviated 25 days, the average OA and average F1-score could slightly drop to 89% and 87% respectively. Thus, the proposed RSG-OC could tolerate the predicting error of flowering date to a large degree, which helps to relieve the influence of interannual fluctuations in the flowering time of rapeseed. However, if more accurate flowering dates were predicted by some advanced phenological models (Chuine and Régnière

2017) in the future, the mapping accuracy could be further stabilized.

5.6. Implications for future research

This study firstly produced the annual rapeseed maps at 20-m resolution for the whole of China, which exhibited detailed spatial patterns and temporal variation of rapeseed cultivation in China. This data could provide an important basis for agricultural subsidies, land use planning, and other policy formulation or evaluation (Zhang, Liu, and Zhang 2022; Zang et al. 2020). For example, a jump in rapeseed areas from 2019 to 2020 could reflect the effect of the policy reform in agriculture. In 2015, the Chinese government abolished the national rapeseed storage policy (Han 2015), resulting in a relatively low rapeseed price and less motivation for rapeseed planting for farmers in 2017–2019. Since 2019, several provinces with leading rapeseed productions started introducing subsidizing policies to promote rapeseed cultivation (e.g. General Office of Zhejiang Provincial Government 2020; Jiangsu Rural Statistics Division 2019; Anhui Provincial Department of Agriculture 2019), stimulating a rapid increase of rapeseed area from 2019 to 2020. The produced rapeseed maps provide evidence independent from census data for evaluating the policy effect. With much more spatial details than census data, the annual rapeseed maps derived from remote sensing could further support fine-grain analysis of the spatial pattern and the driving forces of the rapeseed area change. Moreover, the rapeseed maps are also the basic data layer for rapeseed growth monitoring with remote sensing. It thus provides opportunities for revealing more detailed spatial patterns of the rapeseed phenology, yield, and other agricultural information derived from remote sensing, compared

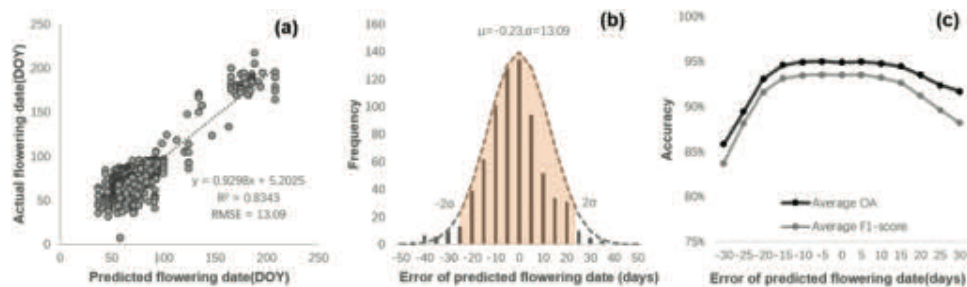


Figure 16. (a) Relationship between predicted flowering date and actual flowering date. (b) Distribution of errors in predicting flowering date. (c) Effect of bias in predicted flowering date on classification accuracy.

to traditional plot-scale studies (D'Andrimont et al. 2020; Sulik and Long 2020; Mercier et al. 2020).

The proposed RSG-OC employed a promising strategy for effectively combining expert (or domain) knowledge and machine-learning techniques. Recently, deep learning has shown great superiority in improving crop classification accuracy (Xu et al. 2021; Yang et al. 2022; Ge et al. 2021). However, deep learning techniques often rely on a large number of training samples, which is challenging for large-scale crop mapping. On the other hand, expert knowledge for crop classification has been intensively explored for a long time in remote sensing community, especially for the widely-cultivated crop types like wheat, corn, peanut and rice (Dong et al. 2016; Qiu et al. 2017, 2018, 2021; Zhan, Zhu, and Li 2021). Effectively integrating the domain knowledge might help to reduce the relying on the large training samples for deep (Dash et al. 2022; Xie et al. 2021). RSG-OC converts the expert knowledge into automatically generated samples, which provides a straightforward strategy to integrate expert knowledge into machine learning techniques. Such a strategy does not need to modify the available machine learning techniques, and thus should not be difficult to be applied for deep learning and also for the other crop classification.

5.7. Remaining issues

There remain several issues in this study. Firstly, the ground truth data for accuracy evaluation were only obtained in the major rapeseed planting regions, thus is not adequate to represent the whole of China. Especially, the commission error in non-

major rapeseed planting regions might be underestimated. For example, we found that some classified rapeseed areas (35.04N, 116.37E) in Shandong province might be actually garlic (Figure 17) based on our recent communication with local farmers. Therefore, more validations in more areas should be further conducted to better evaluate the accuracy of the produced maps. Secondly, the rapeseed area derived from the produced map is much less than the census data. Such inconsistency could be induced by the asymmetric commission/omission classification error, the effect of mixing pixels, and the uncertainty of census data (Boschetti, Flasse, and Brivio 2004; Czaplewski and Catts 1992; Gallego 2004; Waldner and Defourny 2017; Ozdogan and Woodcock 2006; Liu et al. 2020). Even with accurate land cover maps, there is still a possible area for the data with mixed pixels, especially in the case of fragmented landscape (Ozdogan and Woodcock 2006). The exact reasons for the inconsistency between mapped area and census data would be further explored in future research. Finally, the RSG-OC generated samples from the whole of China and trained only one national-wide classifier for rapeseed mapping, which might lead to underfitting for the regions where the generated samples are inadequate. Introducing a zoning strategy on sample generation and classifier training might further improve the performance of RSG-OC.

6. Conclusion

Chinese rapeseed for 2017–2021 was mapped using the RSG-OC method in this paper. With specifically

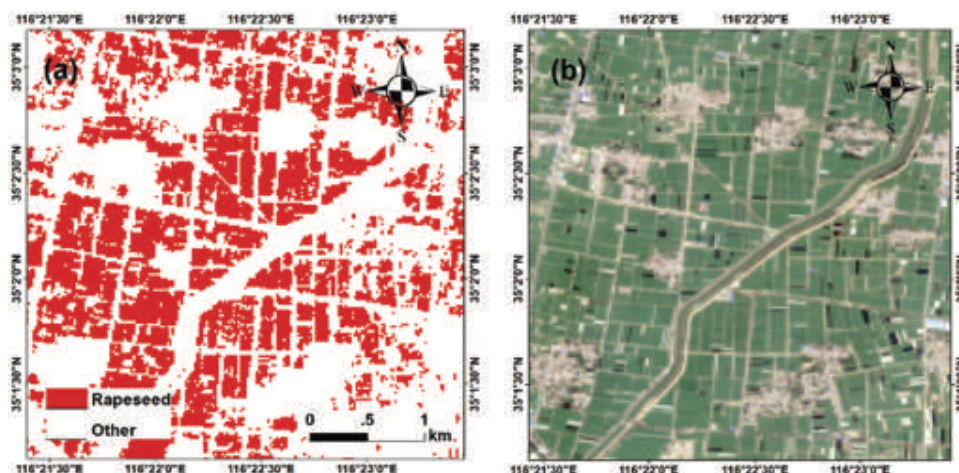


Figure 17. Example of misclassified rapeseed fields (35.04N, 116.37E) without yellowness in flowering season. (a) classification map, (b) true color image in flowering season.

designed selection rules and augmentation methods for generating rapeseed samples, RSG-OC trained a classifier with less dependence on optical flowering features and is thus generalizable to pixels with serious cloud contamination during the flowering period. With the implementation of the PUL one-class classification strategy, the RSG-OC method avoids the difficulty of collecting diverse nonrapeseed samples across all of China. Evaluating the ground truth samples confirms that RSG-OC performs better in rapeseed mapping than the other methods. The rapeseed map produced for 2017–2021 filled the gap in long-term-span and large-extent rapeseed maps in China, which is of importance to national food security and land use regulation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study is supported by the National Key Research and Development Program of China (No. 2022YFD2001101) and National Natural Science Foundation of China (No. 41871224).

Data availability statement

The Sentinel data (Sentinel-1, Sentinel-2 L2A, Sentinel-2 L1C) used in this paper are available in GEE; The produced rapeseed maps are also shared in GEE (<https://code.earthengine.google.com/b457540737690943e0f0ac1d7dd41ed6>) and Zenodo (<https://doi.org/10.5281/zenodo.7047270>). The processing code is available at (<https://code.earthengine.google.com/7e7484b3ca4d9d1bd5896956c9c45ddc>).

The validation data in this study are available from the corresponding author, X.H.Chen, upon reasonable request.

References

- Anhui Provincial Department of Agriculture. 2019. "Office of the Anhui Provincial Department of Agriculture and Rural Affairs on the Issuance of Technical Guidance for Oilseed Rape Production in 2019." Anhui Provincial Department of Agriculture. Accessed 3 September 2022. <http://nync.ah.gov.cn/public/7021/11243741.html>.
- Ashourloo, D., H. S. Shahrabi, M. Azadbakht, H. Aghighi, H. Nematollahi, A. Alimohammadi, and A. A. Matkan. 2019. "Automatic Canola Mapping Using Time Series of Sentinel 2 Images." *Isprs Journal of Photogrammetry and Remote Sensing* 156: 63–76. doi:10.1016/j.isprsjprs.2019.08.007.
- Bonjean, A. P., C. Dequidt, and T. Sang. 2016. "Rapeseed in China." *Ocl* 23 (6): D605. doi:10.1051/ocl/2016045.
- Boschetti, L., S. P. Flasse, and P. A. Brivio. 2004. "Analysis of the Conflict Between Omission and Commission in Low Spatial Resolution Dichotomic Thematic Products: The Pareto Boundary." *Remote Sensing of Environment* 91 (3–4): 280–292. doi:10.1016/j.rse.2004.02.015.
- Carré, P., and A. Pouzet. 2014. "Rapeseed Market, Worldwide and in Europe." *Ocl* 21 (1): D102. doi:10.1051/ocl/2013054.
- Chandrasekar, K., M. V. R. S. Sai, P. S. Roy, and R. S. Dwevedi. 2010. "Land Surface Water Index (LSWI) Response to Rainfall and NDVI Using the MODIS Vegetation Index Product." *International Journal of Remote Sensing* 31 (15): 3987–4005. doi:10.1080/01431160802575653.
- Chen, J., J. Chen, A. P. Liao, X. Cao, L. J. Chen, X. H. Chen, C. Y. He, et al. 2015. "Global Land Cover Mapping at 30 M Resolution: A POK-Based Operational Approach." *Isprs Journal of Photogrammetry and Remote Sensing* 103: 7–27. doi:10.1016/j.isprsjprs.2014.09.002.
- Chen, X. H., D. M. Yin, J. Chen, and X. Cao. 2016. "Effect of Training Strategy for Positive and Unlabelled Learning Classification: Test on Landsat Imagery." *Remote Sensing Letters* 7 (11): 1063–1072. doi:10.1080/2150704x.2016.1217437.
- Chuine, I., and J. Régnière. 2017. "Process-Based Models of Phenology for Plants and Animals." *Annual Review of Ecology, Evolution, and Systematics* 48 (1): 159–182. doi:10.1146/annurev-ecolsys-110316-022706.
- Czaplewski, R. L., and G. P. Catts. 1992. "Calibration of Remotely Sensed Proportion or are Estimates for Misclassification." *Remote Sensing of Environment* 39 (1): 29–43. doi:10.1016/0034-4257(92)90138-A.
- D'Andrimont, R., M. Taymans, G. Lemoine, A. Ceglar, M. Jordanov, and M. van der Velde. 2020. "Detecting Flowering Phenology in Oil Seed Rape Parcels with Sentinel-1 and -2 Time Series." *Remote Sensing of Environment* 239: 111660. doi:10.1016/j.rse.2020.111660.
- Dash, T., S. Chitlangia, A. Ahuja, and A. Srinivasan. 2022. "A Review of Some Techniques for Inclusion of Domain-Knowledge into Deep Neural Networks." *Scientific Reports* 12 (1): 1. doi:10.1038/s41598-021-04590-0.
- Dong, J., X. Xiao, M. A. Menarguez, G. Zhang, Y. Qin, D. Thau, C. Biradar, and B. Moore. 2016. "Mapping Paddy Rice Planting Area in Northeastern Asia with Landsat 8 Images, Phenology-Based Algorithm and Google Earth Engine." *Remote Sensing of Environment* 185: 142–154. doi:10.1016/j.rse.2016.02.016.
- Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, et al. 2012. "Sentinel-2: Esa's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120: 25–36. doi:10.1016/j.rse.2011.11.026.
- Elkan, C., and K. Noto. 2008. "Learning Classifiers from Only Positive and Unlabeled Data." *Proceedings of the 14th ACM*

- SIGKDD international conference on Knowledge discovery and data mining 213–220. doi: [10.1145/1401890.1401920](https://doi.org/10.1145/1401890.1401920)
- ESA. 2020. "Sentinel-2: Cloud Probability." Accessed 3 September 2022. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_CLOUD_PROBABILITY.
- Fang, S. H., W. C. Tang, Y. Peng, Y. Gong, C. Dai, R. H. Chai, and K. Liu. 2016. "Remote Estimation of Vegetation Fraction and Flower Fraction in Oilseed Rape with Unmanned Aerial Vehicle Data." *Remote Sensing* 8 (5): 416. doi:[10.3390/rs8050416](https://doi.org/10.3390/rs8050416).
- Firrisa, M. T., I. van Duren, and A. Voinov. 2013. "Energy Efficiency for Rapeseed Biodiesel Production in Different Farming Systems." *Energy Efficiency* 7 (1): 79–95. doi:[10.1007/s12053-013-9201-2](https://doi.org/10.1007/s12053-013-9201-2).
- Food and Agriculture Organization of the United Nations (FAO). 2021. "Crops and Livestock Products." Food and Agriculture Organization of the United Nations. Accessed 3 September 2022. https://www.fao.org/3/cb4477en/online/cb4477en.html#chapter-2_1.
- Food and Agriculture Organization of the United Nations (FAO). 2022. "FAOSTAT." Accessed 3 September 2022. <https://www.fao.org/faostat/en/#data/QL>.
- Foody, G. M., and A. Mathur. 2004. "A Relative Evaluation of Multiclass Image Classification by Support Vector Machines." *IEEE Transactions on Geoscience and Remote Sensing* 42 (6): 1335–1343. doi:[10.1109/Tgrs.2004.827257](https://doi.org/10.1109/Tgrs.2004.827257).
- Frantz, D., E. Hass, A. Uhl, J. Stoffels, and J. Hill. 2018. "Improvement of the Fmask Algorithm for Sentinel-2 Images: Separating Clouds from Bright Surfaces Based on Parallax Effects." *Remote sensing of environment* 215: 471–481. doi:[10.1016/j.rse.2018.04.046](https://doi.org/10.1016/j.rse.2018.04.046).
- Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, et al. 2002. "Global Land Cover Mapping from MODIS: Algorithms and Early Results." *Remote Sensing of Environment* 83 (1–2): 287–302. doi:[10.1016/S0034-4257\(02\)00078-0](https://doi.org/10.1016/S0034-4257(02)00078-0).
- Fu, D. H., L. Y. Jiang, A. S. Masons, M. L. Xiao, L. R. Zhu, L. Z. Li, Q. H. Zhou, C. J. Shen, and C. H. Huang. 2016. "Research Progress and Strategies for Multifunctional Rapeseed: A Case Study of China." *Journal of Integrative Agriculture* 15 (8): 1673–1684. doi:[10.1016/S2095-3119\(16\)61384-9](https://doi.org/10.1016/S2095-3119(16)61384-9).
- Gallego, F. J. 2004. "Remote Sensing and Land Cover Area Estimation." *International Journal of Remote Sensing* 25 (15): 3019–3047. doi:[10.1080/01431160310001619607](https://doi.org/10.1080/01431160310001619607).
- General Office of Zhejiang Provincial Government. 2020. "The General Office of the Zhejiang Provincial People's Government on the Enhancement of Oil Supply Security Capacity to Promote the Implementation of High-Quality Industrial Development Views." Accessed 3 September 3. https://www.zj.gov.cn/art/2020/10/28/art_1229017139_2037686.html.
- Ge, S., J. S. Zhang, Y. Z. Pan, Z. Yang, and S. Zhu. 2021. "Transferable Deep Learning Model Based on the Phenological Matching Principle for Mapping Crop Extent." *International Journal of Applied Earth Observation and Geoinformation* 102: 102451. doi:[10.1016/j.jag.2021.102451](https://doi.org/10.1016/j.jag.2021.102451).
- Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote sensing of environment* 202: 18–27. doi:[10.1016/j.rse.2017.06.031](https://doi.org/10.1016/j.rse.2017.06.031).
- Gumma, M. K., P. S. Thenkabail, K. Charyulu Deevi, I. A. Mohammed, P. Teluguntla, A. Oliphant, J. Xiong, T. Aye, and A. M. Whitbread. 2018. "Mapping Cropland Fallow Areas in Myanmar to Scale Up Sustainable Intensification of Pulse Crops in the Farming System." *GIScience & Remote Sensing* 55 (6): 926–949. Taylor & Francis. doi:[10.1080/15481603.2018.1482855](https://doi.org/10.1080/15481603.2018.1482855).
- Han, Y. 2015. "National Storage Cancellation of the Main Production Areas Prices Fell Chongqing Rapeseed Supply and Demand Stable Prices Firm." Accessed 3 September 2022. https://www.moa.gov.cn/xw/qg/201905/t20190523_6314442.htm.
- Han, J. C., Z. Zhang, and J. Cao. 2021. "Developing a New Method to Identify Flowering Dynamics of Rapeseed Using Landsat 8 and Sentinel-1/2." *Remote Sensing* 13 (1): 1. doi:[10.3390/rs13010105](https://doi.org/10.3390/rs13010105).
- Han, J. C., Z. Zhang, Y. C. Luo, J. Cao, L. L. Zhang, J. Zhang, and Z. Y. Li. 2021. "The RapeseedMap10 Database: Annual Maps of Rapeseed at a Spatial Resolution of 10m Based on Multi-Source Data." *Earth System Science Data* 13 (6): 2857–2874. doi:[10.5194/essd-13-2857-2021](https://doi.org/10.5194/essd-13-2857-2021).
- Huang, X., J. Huang, L. Xuecao, Q. Shen, and Z. Chen. 2022. "Early Mapping of Winter Wheat in Henan Province of China Using Time Series of Sentinel-2 Data." *GIScience & Remote Sensing* 59 (1): 1534–1549. doi:[10.1080/15481603.2022.2104999](https://doi.org/10.1080/15481603.2022.2104999).
- Hu, Q., W. Hua, Y. Yin, X. K. Zhang, L. J. Liu, J. Q. Shi, Y. G. Zhao, L. Qin, C. Chen, and H. Z. Wang. 2017. "Rapeseed Research and Production in China." *Crop Journal* 5 (2): 127–135. doi:[10.1016/j.cj.2016.06.005](https://doi.org/10.1016/j.cj.2016.06.005).
- Jiangsu Rural Statistics Division. 2019. "The Province's Rapeseed Production This Year to Stop the Decline to Increase." Accessed 3 September 2022. http://tj.jiangsu.gov.cn/art/2019/12/4/art_4027_8834637.html.
- Jin, Z. N., G. Azzari, C. You, S. Di Tommaso, S. Aston, M. Burke, and D. B. Lobell. 2019. "Smallholder Maize Area and Yield Mapping at National Scales with Google Earth Engine." *Remote Sensing of Environment* 228: 115–128. doi:[10.1016/j.rse.2019.04.016](https://doi.org/10.1016/j.rse.2019.04.016).
- Karra, K., C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby. 2021. "Global Land Use/Land Cover with Sentinel 2 and Deep Learning." In 2021 IEEE international geoscience and remote sensing symposium IGARSS 4704–4707. doi: [10.1109/IGARSS47720.2021.9553499](https://doi.org/10.1109/IGARSS47720.2021.9553499).
- Konduri, V. S., J. Kumar, W. W. Hargrove, F. M. Hoffman, and A. R. Ganguly. 2020. "Mapping Crops Within the Growing Season Across the United States." *Remote Sensing of Environment* 251: 112048. doi:[10.1016/j.rse.2020.112048](https://doi.org/10.1016/j.rse.2020.112048).
- Lei, L., X. Y. Wang, Y. F. Zhong, H. W. Zhao, X. Hu, and C. Luo. 2021. "DOCC: Deep One-Class Crop Classification via Positive and Unlabeled Learning for Multi-Modal Satellite Imagery."

- International Journal of Applied Earth Observation and Geoinformation* 105: 102598. doi:10.1016/j.jag.2021.102598.
- Li, W., Q. Guo, and C. Elkan. 2010. "A Positive and Unlabeled Learning Algorithm for One-Class Classification of Remote-Sensing Data." *IEEE Transactions on Geoscience and Remote Sensing* 49 (2): 717–725. doi:10.1109/TGRS.2010.2058578.
- Liu, G., X. Wang, G. Baiocchi, M. Casazza, F. Meng, Y. Cai, Y. Hao, W. Feng, and Z. Yang. 2020. "On the Accuracy of Official Chinese Crop Production Data: Evidence from Biophysical Indexes of Net Primary Production." *Proceedings of the National Academy of Sciences* 117 (41): 25434–25444. doi:10.1073/pnas.1919850117.
- Li, S., W. Zhang, L. Zhao, and X. Wang. 2021. "Phenological Period Identification of Oilseed Rape Based on Time-Series PolSAR Image and Decision Tree Model." *Acta Agriculturae Zhejiangensis* 11 (33): 2116–2127. doi:10.3969/j.issn.1004-1524.2021.11.14.
- Lu, Y., and L. Wang. 2021. "How to Automate Timely Large-Scale Mangrove Mapping with Remote Sensing." *Remote Sensing of Environment* 264: 112584. doi:10.1016/j.rse.2021.112584.
- Mandal, D., V. Kumar, D. Ratha, S. Dey, A. Bhattacharya, J. M. Lopez-Sanchez, H. McNairn, and Y. S. Rao. 2020. "Dual Polarimetric Radar Vegetation Index for Crop Growth Monitoring Using Sentinel-1 SAR Data." *Remote Sensing of Environment* 247: 111954. doi:10.1016/j.rse.2020.111954.
- Mansaray, L. R., L. Yang, V. T. S. Kabba, A. S. Kanu, J. Huang, and F. Wang. 2019. "Optimising Rice Mapping in Cloud-Prone Environments by Combining Quad-Source Optical with Sentinel-1A Microwave Satellite Imagery." *GIScience & Remote Sensing* 56 (8): 1333–1354. Taylor & Francis. doi:10.1080/15481603.2019.1646978.
- Maxwell, A. E., T. A. Warner, and F. Fang. 2018. "Implementation of Machine-Learning Classification in Remote Sensing: An Applied Review." *International Journal of Remote Sensing* 39 (9): 2784–2817. doi:10.1080/01431161.2018.1433343.
- Meng, S. Y., Y. F. Zhong, C. Luo, X. Hu, X. Y. Wang, and S. X. Huang. 2020. "Optimal Temporal Window Selection for Winter Wheat and Rapeseed Mapping with Sentinel-2 Images: A Case Study of Zhongxiang in China." *Remote Sensing* 12 (2): 226. doi:10.3390/rs12020226.
- Mercier, A., J. Betbeder, S. Rapinel, N. Jegou, J. Baudry, and L. Hubert-Moy. 2020. "Evaluation of Sentinel-1 And-2 Time Series for Estimating LAI and Biomass of Wheat and Rapeseed Crop Types." *Journal of Applied Remote Sensing* 14 (2): 1. doi:10.1117/1.JRS.14.024512.
- Ozdogan, M., and C. E. Woodcock. 2006. "Resolution Dependent Errors in Remote Sensing of Cultivated Areas." *Remote Sensing of Environment* 103 (2): 203–217. doi:10.1016/j.rse.2006.04.004.
- Phan, T. N., V. Kuch, and L. W. Lehnert. 2020. "Land Cover Classification Using Google Earth Engine and Random Forest Classifier—the Role of Image Composition." *Remote Sensing* 12 (15): 2411. doi:10.3390/rs12152411.
- Pontius, R. G., and M. Millones. 2011. "Death to Kappa: Birth of Quantity Disagreement and Allocation Disagreement for Accuracy Assessment." *International Journal of Remote Sensing* 32 (15): 4407–4429. doi:10.1080/01431161.2011.552923.
- Pott, L. P., T. J. C. Amado, R. A. Schwalbert, G. M. Corassa, and I. A. Ciampitti. 2021. "Satellite-Based Data Fusion Crop Type Classification and Mapping in Rio Grande Do Sul, Brazil." *Isprs Journal of Photogrammetry and Remote Sensing* 176: 196–210. doi:10.1016/j.isprsjprs.2021.04.015.
- Qian, W., J. Meng, M. Li, M. Frauen, O. Sass, J. Noack, and C. Jung. 2006. "Introgression of Genomic Components from Chinese Brassica Rapa Contributes to Widening the Genetic Diversity in Rapeseed (*B. Napus* L.), with Emphasis on the Evolution of Chinese Rapeseed." *Theoretical and Applied Genetics* 113 (1): 49–54. doi:10.1007/s00122-006-0269-3.
- Qiu, B., Y. Huang, C. Chen, Z. Tang, and F. Zou. 2018. "Mapping Spatiotemporal Dynamics of Maize in China from 2005 to 2017 Through Designing Leaf Moisture Based Indicator from Normalized Multi-Band Drought Index." *Computers and Electronics in Agriculture* 153: 82–93. doi:10.1016/j.compag.2018.07.039.
- Qiu, B., F. Jiang, C. Chen, Z. Tang, W. Wenbin, and J. Berry. 2021. "Phenology-Pigment Based Automated Peanut Mapping Using Sentinel-2 Images." *GIScience & Remote Sensing* 58 (8): 1335–1351. Taylor & Francis. doi:10.1080/15481603.2021.1987005.
- Qiu, B. W., Y. H. Luo, Z. H. Tang, C. C. Chen, D. F. Lu, H. Y. Huang, Y. Z. Chen, N. Chen, and W. M. Xu. 2017. "Winter Wheat Mapping Combining Variations Before and After Estimated Heading Dates." *Isprs Journal of Photogrammetry and Remote Sensing* 123: 35–46. doi:10.1016/j.isprsjprs.2016.09.016.
- Rabonatahiry, N., L. Huaixin, Y. Longjiang, and L. Maoteng. 2021. "Rapeseed (*Brassica Napus*): Processing, Utilization, and Genetic Improvement." *Agronomy* 11 (9): 1776. doi:10.3390/agronomy11091776.
- Shen, M., J. Chen, X. Zhu, and Y. Tang. 2014. "Yellow Flowers Can Decrease NDVI and EVI Values: Evidence from a Field Experiment in an Alpine Meadow." *Canadian Journal of Remote Sensing* 35 (2): 99–106. doi:10.5589/m09-003.
- Shen, M., J. Chen, X. Zhu, Y. Tang, and X. Chen. 2010. "Do Flowers Affect Biomass Estimate Accuracy from NDVI and EVI?" *International Journal of Remote Sensing* 31 (8): 2139–2149. doi:10.1080/01431160903578812.
- Sulik, J. J., and D. S. Long. 2015. "Spectral Indices for Yellow Canola Flowers." *International Journal of Remote Sensing* 36 (10): 2751–2765. doi:10.1080/01431161.2015.1047994.
- Sulik, J. J., and D. S. Long. 2016. "Spectral Considerations for Modeling Yield of Canola." *Remote Sensing of Environment* 184: 161–174. doi:10.1016/j.rse.2016.06.016.
- Sulik, J. J., and D. S. Long. 2020. "Automated Detection of Phenological Transitions for Yellow Flowering Plants Such as Brassica Oilseeds." *Agrosystems, Geosciences & Environment* 3 (1): e20125. doi:10.1002/agg2.20125.
- Tao, J. B., W. B. Liu, W. X. Tan, X. B. Kong, and M. Xu. 2019. "Fusing Multi-Source Data to Map Spatio-Temporal Dynamics of Winter Rape on the Jiangnan Plain and Dongting Lake Plain, China." *Journal of Integrative*

- Agriculture* 18 (10): 2393–2407. doi:10.1016/S2095-3119(19)62577-3.
- Thorp, K. R., and D. Drajat. 2021. “Deep Machine Learning with Sentinel Satellite Data to Map Paddy Rice Production Stages Across West Java, Indonesia.” *Remote Sensing of Environment* 265: 112679. doi:10.1016/j.rse.2021.112679.
- Tian, Z., Y. H. Ji, L. X. Sun, X. L. Xu, D. L. Fan, H. L. Zhong, Z. R. Liang, and G. Ficsher. 2018. “Changes in Production Potentials of Rapeseed in the Yangtze River Basin of China Under Climate Change: A Multi-Model Ensemble Approach.” *Journal of Geographical Sciences* 28 (11): 1700–1714. doi:10.1007/s11442-018-1538-1.
- Tian, Z., Y. H. Ji, H. Q. Xu, H. G. Qiu, L. X. Sun, H. L. Zhong, and J. G. Liu. 2021. “The Potential Contribution of Growing Rapeseed in Winter Fallow Fields Across Yangtze River Basin to Energy and Food Security in China.” *Resources Conservation and Recycling* 164: 105159. doi:10.1016/j.resconrec.2020.105159.
- Tian, J. Y., L. Wang, D. M. Yin, X. J. Li, C. Y. Diao, H. L. Gong, C. Shi, et al. 2020. “Development of Spectral-Phenological Features for Deep Learning to Understand *Spartina Alterniflora* Invasion.” *Remote sensing of environment* 242: 111745. doi:10.1016/j.rse.2021.112679.
- Torres, R., P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, et al. 2012. “GMES Sentinel-1 Mission.” *Remote sensing of environment* 120 (May): 9–24. doi:10.1016/j.rse.2011.05.028.
- USDA. 2022. “Oilseeds: World Markets and Trade.” *USDA Foreign Agricultural Service*. Accessed 3 September 2022. <https://www.fas.usda.gov/data/oilseeds-world-markets-and-trade>.
- van Duren, Iris, A. Voinov, O. Arodudu, and M. Tesfaye Firrisa. 2015. “Where to Produce Rapeseed Biodiesel and Why? Mapping European Rapeseed Energy Efficiency.” *Renewable Energy* 74: 49–59. doi:10.1016/j.renene.2014.07.016.
- Veloso, A., S. Mermoz, A. Bouvet, T. L. Toan, M. Planells, J. F. Dejoux, and E. Ceschia. 2017. “Understanding the Temporal Behavior of Crops Using Sentinel-1 and Sentinel-2-Like Data for Agricultural Applications.” *Remote Sensing of Environment* 199: 415–426. doi:10.1016/j.rse.2017.07.015.
- Waldner, F., and P. Defourny. 2017. “Where Can Pixel Counting Area Estimates Meet User-Defined Accuracy Requirements?” *International Journal of Applied Earth Observation and Geoinformation* 60: 1–10. doi:10.1016/j.jag.2017.03.014.
- Wang, H. Z., C. Y. Guan, and C. L. Zhang. 2007. Studies on rapeseed production and cultivation science and technology in China. In *The 12th International Rapeseed Congress Proceeding Science*, edited by Fu, T.D., and Guan, C.Y., 2–7. Wuhan, China: Press USA Inc.
- Weiss, M., F. Jacob, and G. Duveiller. 2020. “Remote Sensing for Agricultural Applications: A Meta-Review.” *Remote Sensing of Environment* 236: 111402. doi:10.1016/j.rse.2019.111402.
- Xie, X., J. Niu, X. Liu, Z. Chen, S. Tang, and S. Yu. 2021. “A Survey on Incorporating Domain Knowledge into Deep Learning for Medical Image Analysis.” *Medical Image Analysis* 69: 69. doi:10.1016/j.media.2021.101985.
- Xiong, J., P. S. Thenkabail, M. K. Gumma, P. Teluguntla, J. Poehnelt, R. G. Congalton, K. Yadav, and D. Thau. 2017. “Automated Cropland Mapping of Continental Africa Using Google Earth Engine Cloud Computing.” *Isprs Journal of Photogrammetry and Remote Sensing* 126: 225–244. doi:10.1016/j.isprs.2017.01.019.
- Xu, J., J. Yang, X. Xiong, L. Haifeng, J. Huang, K. C. Ting, Y. Ying, and T. Lin. 2021. “Towards Interpreting Multi-Temporal Deep Learning Models in Crop Mapping.” *Remote Sensing of Environment* 264: 264. doi:10.1016/j.rse.2021.112599.
- Yang, L., R. Huang, J. Huang, T. Lin, L. Wang, R. Mijiti, P. Wei, et al. 2022. “Semantic Segmentation Based on Temporal Features: Learning of Temporal-Spatial Information from Time-Series SAR Images for Paddy Rice Mapping.” *IEEE Transactions on Geoscience and Remote Sensing* 60. doi:10.1109/TGRS.2021.3099522.
- Yin, Y., and H. Wang. 2012. “Achievement, Problem and Scientific Policy of Rapeseed Industry Development in China.” *Journal of Agricultural Science and Technology* 14 (4): 1–7. doi:10.3969/j.issn.1008-0864.2012.04.01.
- Zanaga, D., R. van de Kerchove, W. de Keersmaecker, N. Souverijns, C. Brockmann, R. Quast, A. G. Jan Wevers, A. Paccini, and S. Vergnaud. 2021. “ESA WorldCover 10 M 2020 V100.” doi:10.1038/s41467-021-24227-0.
- Zang, Y. Z., X. H. Chen, J. Chen, Y. G. Tian, Y. S. Shi, X. Cao, and X. H. Cui. 2020. “Remote Sensing Index for Mapping Canola Flowers Using MODIS Data.” *Remote Sensing* 12 (23): 3912. doi:10.3390/rs12233912.
- Zhang, X. L., and Y. He. 2013. “Rapid Estimation of Seed Yield Using Hyperspectral Images of Oilseed Rape Leaves.” *Industrial Crops and Products* 42: 416–420. doi:10.1016/j.indcrop.2012.06.021.
- Zhang, H. Y., W. B. Liu, and L. P. Zhang. 2022. “Seamless and Automated Rapeseed Mapping for Large Cloudy Regions Using Time-Series Optical Satellite Imagery.” *Isprs Journal of Photogrammetry and Remote Sensing* 184: 45–62. doi:10.1016/j.isprs.2021.12.001.
- Zhang, G., X. Xiao, C. M. Biradar, J. Dong, Y. Qin, M. A. Menarguez, Y. Zhou, et al. 2017. “Spatiotemporal Patterns of Paddy Rice Croplands in China and India from 2000 to 2015.” *The Science of the Total Environment* 579: 82–92. doi:10.1016/j.scitotenv.2016.10.223.
- Zhang, Y. X., W. Zhang, K. Xu, and J. Li. 2022. “Phenological Phase Identification of Oilseed Rape (*Brassica Napus* L.) Using Typical Stokes Parameters.” *Geomatics and Information Science of Wuhan University*. doi:10.13203/j.whugis.20210394.
- Zhan, P., W. Q. Zhu, and N. Li. 2021. “An Automated Rice Mapping Method Based on Flooding Signals in Synthetic

- Aperture Radar Time Series." *Remote Sensing of Environment* 252: 112112. doi:[10.1016/j.rse.2020.112112](https://doi.org/10.1016/j.rse.2020.112112).
- Zhong, L. H., P. Gong, and G. S. Biging. 2014. "Efficient Corn and Soybean Mapping with Temporal Extendability: A Multi-Year Experiment Using Landsat Imagery." *Remote Sensing of Environment* 140: 1–13. doi:[10.1016/j.rse.2013.08.023](https://doi.org/10.1016/j.rse.2013.08.023).
- Zhou, Y., J. Luo, L. Feng, Y. Yang, Y. Chen, and W. Wei. 2019. "Long-Short-Term-Memory-Based Crop Classification Using High-Resolution Optical Images and Multi-Temporal SAR Data." *GIScience & Remote Sensing* 56 (8): 1170–1191. Taylor & Francis. doi:[10.1080/15481603.2019.1628412](https://doi.org/10.1080/15481603.2019.1628412).