Cropland Extraction Using Water Mask and Phenological Features based on Landsat-8 Imagery

Wenjuan Xiao¹, Jiasheng Wang¹, ², *
¹Faculty of Geography, Yunnan Normal University, Kunming, China
²The Engineering Research Center of GIS Technology in Western China, Ministry of Education, Kunming, China

Corresponding author: jersonwang@ynnu.edu.cn

Abstract. Accurate cropland information is crucial for food security assessment. At present, Landsat-8 remote sensing imagery and Google Earth Engine (GEE) cloud platform provide favorable conditions for mapping cropland. Despite the improved mapping conditions, the existence of mixed pixels is still one of the main difficulties in the cropland extraction. Especially in paddy fields, the influence of water bodies on farmland extraction should not be ignored. Also, the lack of time series data with temporal-spatial resolution impedes the accurate cultivated land mapping. Aiming at these problems, we proposed an effective method to extract cropland precisely based on the Landsat-8 imagery and GEE platform. First, we created cloud-free images and masked out water based on Landsat-8 images. Second, we constructed the normalized difference vegetation index time series data with 30-m resolution from Landsat-8 and MODIS images to obtain accurate phenological information. Third, other cropland features such as texture features, tasseled cap transformation features, terrain features and spectral features were acquired. Finally, we applied the Random Forest algorithm and obtained the cropland for 2015 in Qinzhou. The results show that terrain characteristics and phenological characteristics have the greatest influence on cropland extraction. Based on water mask, the overall accuracy of cropland improved by 6.9%. Overall, the product has high accuracy with the mean overall accuracy (OA) of 97.9%. Compared with MCD12Q1 and GFSAD30 cropland products, it is 6.2% and 24.3% higher in OA respectively. In brief, our method is helpful for the cropland extraction.

1. Introduction

Food security has been a significant issue on the global nowadays. We need to monitor cropland change rapidly and in real time. However, mixed pixels appear in the results of cropland extraction inevitably because of the cropland complexity. Although remote sensing technology is very mature, it is still difficult to obtain the cropland information accurately today. Some scholars proposed the method of eliminating other land-use types such as water bodies, wetlands, grasslands and impervious surfaces for cropland extraction. What these land-use types have in common is that they are easy to extract. Though this method is proved to be effective in improving the classification precision [1, 2], its application is not widespread.
Phenological characteristics reflects the seasonal variation information of cropland, which is important in the cropland extraction. At present, remote sensing time series is the main method to obtain phenological characteristics, especially for Landsat normalized difference vegetation index (NDVI) time series data [3-5]. The broad coverage and high spatial resolution make it was recognized widely. But limitations of temporal resolution needs to be improved to capture precise phenological information. Whereupon, many scholars make use of the advantage of MODIS images in the temporal resolution to make up for that [6-8]. This method is effective in phenological features extraction. Nevertheless, it is difficult to fuse the time series data with local computers because of finite storage and computational ability. With the development of cloud computing, some remote sensing image processing cloud platforms have emerged gradually. They perform well in remote sensing image storage and computation. For example, Google Earth Engine (GEE) platform is a representative.

Therefore, this study taked advantage of GEE cloud platform and explored the cultivated land extraction based on water mask and phenology characteristics, in order to enrich the cropland database and provide important datasets for ensuring food security.

2. Study area and data

2.1. Study area
Qinzhou is located on the Beibu Gulf coast, which is a prefecture-level city in Southern Guangxi (Figure 1). In total, the area of Qinzhou is 10, 895 km². Torridity, humidity and raininess are the major climatic characteristics of Qinzhou because of the maritime climate. The weather creates favourable conditions for farming, especially abundant water resources. The mainland coastline is 562.64 km long. Furthermore, the mean annual rainfall is just above 2, 170 mm. Qinzhou is one of the main grain bases in Guangxi province. The 40% of the cropland is paddy fields. The abundant water resources make field capacity to be 5, 470 m³ per acre. This research chose Qinzhou as the study area in order to discover the influence of water on cropland extraction better.

Figure 1. The location of the study area

2.2. Data sources and pre-processing

2.2.1. Satellite images. The Landsat-8 surface reflectance (SR) images, MODIS SR image collection, Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) V3.0 product, JRC Global Surface Water (GSW), MODIS land-use cover (LULC) types product and Global Food Security support Analysis Data (GFSAD) were used for this study. Among them, GFSAD product can be downloaded from NASA's (National Aeronautics and Space Administration) Earth Data website (https://search.earthdata.nasa.gov/search), while the others can be obtained from Google Earth Engine's
public data catalog (https://developers.google.com/earth-engine/datasets/catalog). These images were selected for the year 2015. A. Landsat-8 SR dataset. The Landsat-8 SR dataset for 2015 was a main data source in this study. It has a high spatial resolution at 30-m, but at long observation (every 16 days). Since geometric correction and radiance correction have been done before uploading GEE platform, so we skipped these tasks. First, we cleared cloud by CFMask algorithm [9] for cloud-free images with a total number of 60. Second, we filled the missing values by linear interpolation method. These processed images were used to obtain phenological characteristics. Finally, a single composite image was created by median composite method, which was used to classify. B. MODIS SR image collection. The MODIS SR image collection for 2015 has higher temporal resolution (8 days), but at a lower spatial resolution (250-m) than the Landsat-8 SR. They were data sources for the construction of NDVI time series with high spatial-temporal resolution for phenology information extraction. C. SRTM DEM V3.0 data. The SRTM DEM V3.0 data at a resolution of 30-m (1 arc-second) contains elevation data of 1200×1200 sample points. We used it for terrain features calculation. D. The JRC GSW data. The JRC GSW data was developed at 30-m spatial resolution by the Joint Research Centre (JRC) of the European commission. It contains maps of the location and temporal distribution of surface water from 1984 to 2018. So, it can be used for a water mask. E. MODIS LULC product. The MODIS LULC product at 500-m spatial resolution called MODIS Collection 6 (MCD12Q1) data provides data characterizing five global land cover classification systems, including cropland from 2001 to 2016. It was used for the cropland type assessment. F. GFSAD30. The GFSAD30 data for 2015 provides cropland and non-cropland data. This data not only helped us to evaluate our cropland products, but also helped in collecting samples.

2.2.2. Training and validation data. Training and validation data were obtained as follows: (1) We first collected the sample data for 7 LULC types: cropland, forest, grassland, bareland, wetland, impervious surface and water body according to the domestic classification system (GB/T21010-2017). First step, we obtained the sample points from GFSAD30 data by stratified sampling method based on GEE online platform. On this basis, the sample points were checked, corrected and expanded based on Google Earth offline platform. (2) Then, the sample data was redefined as cropland and non-cropland. (3) Finally, We generated training and validation data from the sample data randomly for classification. The ratio of training data to validation data is 3:1.

3. Methods

The whole workflow for the cropland extraction as shown below (Figure 2), including four major components: (1) Landsat-8 images preparation (done in section II already). (2) water mask for water-cleared of the study area. (3) cropland features calculation. (4) RF classification.

3.1. Water mask

Considering that water in the study region is confused with paddy fields, but it can be eliminated easily, we generated water mask for cropland classification. We believed that reducing the effects of water can be conducive to classification (this conjecture was verified in section 4). However, the classification based on a mask image is different from the direct. This method breaks the conventional way of remote sensing information extraction, but it relies on the pure pixels of the mask image. Ensuring the mask accuracy is the premise for us to obtain high precision cultivated land mapping. For this purpose, we chose the JRC GSW data with high resolution and precision for water mask. The JRC GSW data contains types of seasonal and permanent water, which overall accuracy reaches above 90%. Based on the JRC GSW data, we obtained water mask by the mask function on GEE platform. The water mask was the basis for classification.
3.2. Phenological features calculation

Phenological features is the key information separating cropland from other LULC types. At present, analyzing NDVI time series is the main way to capture phenological features. In this step, we constructed the NDVI time series from 30-m Landsat-8 SR and 500-m MODIS SR data sources, and obtained high spatial (30-m) and temporal resolution (8days) data by linear regression mode.

First, NDVI time series of Landsat-8 data and MODIS data were constructed respectively, as in equation (1). In the expression, NIR and R denotes the near infrared band and red band of the images separately.

\[
NDVI = \frac{\text{NIR}-\text{R}}{\text{NIR}+\text{R}}
\]  

(1)

Second, in order to make Landsat and MODIS data consistent in resolution, we resampled the time series data of MODIS NDVI to 30-m resolution by bilinear interpolation.

Third, we used image fusion for high spatial and temporal resolution data. A linear regression model [10] was used to establish the mathematical relationship between Landsat-8 NDVI time series and MODIS NDVI time series data, as in equation (2).

\[
k_{\text{MODIS}}^{\text{t}}(x, y, t) = \sum_{c=1}^{t} f_c(x, y, t_1) \times k_c^{\text{Landsat}}(x, y, t)
\]

(2)

In the equation, \(t\) denotes the time range from \(t_1\) to \(t_2\). \(k_{\text{MODIS}}^{\text{t}}(x, y, t)\) denotes the increment of MODIS pixel \((x, y)\) in range \(t\), while \(k_c^{\text{Landsat}}(x, y, t)\) denotes the increment of Landsat pixel with land cover \(c\). \(f_c(x, y, t_1)\) denotes the area proportion of land cover \(c\) in MODIS pixel.

Finally, we need to filter the NDVI time series data (Landsat-MODIS NDVI) because of noise. Based on Satizky-Golay algorithm, we reconstructed the NDVI time series data and drew the curve of Landsat-MODIS NDVI time series (Figure 3). The figure shows that cropland is most similar to grassland and forest in seasonal variation, but it changes more regularly. Cropland grows poorly in February but vigorously in autumn. During these two periods, cropland is different from other LULC types obviously.

Figure 2. Workflow for the cropland extraction
That is to say, growing and non-growing periods are the best time to distinguish cropland from non-cropland. In these two periods, the cropland has NDVI maximum values and the minimum respectively. Thus, we selected NDVI maximum and NDVI minimum as phenological parameters for classification.

![Figure 3](image-url). The curve of Landsat-MODIS NDVI time series

### 3.3. Other features calculation

#### 3.3.1. Texture features. In addition to phenological features, texture features is one of the most important features of cropland, which reflects the spatial correlation features of image gray scale. Because texture is not as complex as phenology, we can obtain the texture features in a traditional but very efficient way called Gray-lever Co-occurrence Matrix (GLCM). GLCM is used for texture metrics commonly [8-9], which is a tabulation of how often different combinations of pixel brightness values (grey levels) occur in an image. In this step, we created the GLCM by the `glcmTexture` function on GEE platform and selected the common metrics of Angular Second Moment (ASM), Contrast (CON) and Entropy (ENT) for classification.

#### 3.3.2. Terrain features. In order to improve classification accuracy, terrain features was used in many LULC types classification [11]. In the study area, hills comprise the main features. But most of the cropland is in low-lying area. The topography factors such as elevation and slope helps to separate cropland from non-cropland to a great extent. So, we calculated the terrain parameters of slope and elevation from the SRTM DEM V3 product by the `ee.Algorithms.Terrain` function on GEE platform.

#### 3.3.3. Tasseled cap transformation features. Tasseled cap transformation (TCT) called K-T transformation is a special principal component analysis (PCA) method in essence. The first three principal components denotes brightness, greenness and wetness, which reflects the information of soil rock, vegetation and moisture respectively. The TCT method is also very important for farmland extraction because it can make the image enhance. But for different remote sensing data, the coefficient matrix of TCT is different. Considering that the data we used is Landsat-8, we chose the coefficient matrix calculated by Baig et al [12]. The process of TCT as in equation (3).

\[ Y = cX \]  

(3)
In the expression, $Y$ denotes the image bands after TCT. $C$ represents the coefficient matrix of TCT, while $X$ denotes the raw image bands.

3.4. Feature selection

The original feature space was constructed from the features above and the original bands (blue, green, red, NIR, SWIR1 and SWIR2). The four kinds of features include two phenology metrics, three texture metrics, two terrain metrics and three TCT metrics. Considering that these 16 features may be correlated with each other and have data redundancy [13]. Because of the advantages of parametric evaluation, Jeffries-Matusita (J-M) distance is used generally in feature selection. The J-M distance is an important indicator for the separability of statistical categories based on Bhattacharyya Distance. In this way, the different features can be distinguished from the training samples. Based on training data, we calculated the J-M distance of different feature combinations, as in equation (4) and equation (5).

\[
J_M = \sqrt{2(1 - e^{-B_{ij}})}
\]

\[
B_{ij} = \frac{(M_i - M_j)^T (V_i + V_j)^{-1}}{8} + \frac{1}{2} \ln \left[ \frac{|V_i + V_j|}{|V_i| |V_j|} \right]
\]

Where $B_{ij}$ denotes the Bhattacharyya Distance, $V_i$ and $V_j$ denotes the covariance matrix of cropland and non-cropland respectively, $M_i$ and $M_j$ denotes the corresponding mean vectors, $i$ and $j$ denotes the cropland and non-cropland training data respectively. From the equation (4), we can see that the maximum J-M distance is $2^{1/2}$. The closer it is to $2^{1/2}$, the better features. The results are shown below (Table 1). From the table, we can find that the fifth group has the largest J-M distance, which is closer to $2^{1/2}$. In this group, the feature combination contains all the feature categories, but only has 9 features: NDVI maximum, NDVI minimum, Angular Second Moment, Contrast, Entropy, elevation, slope, wetness and SWIR 1. These features are used for classification.

| Group | feature combination | J-M distance |
|-------|---------------------|--------------|
| 1     | M+T                 | 1.41110      |
| 2     | M+D                 | 1.40514      |
| 3     | M+D+T               | 1.41393      |
| 4     | M+D+T+W             | 1.41410      |
| 5     | M+D+T+W+S           | 1.41417      |
| 6     | T+D                 | 1.35914      |
| 7     | T+D+W               | 1.36990      |
| 8     | T+D+W+S             | 1.37598      |

M, T, D, T, W and S denotes phenological features, texture features, terrain features, TCT features and spectral features respectively.

3.5. Random forest classification

Random forest (RF) is an integrated machine learning algorithm based on Boosting. Different from the traditional classification, the method combines multiple classifiers to make up for a single classifier in traditional classification. This integration way optimizes the generalization capacity. For this reason, RF classification has strong robustness and was applied in land-use classification widely. However, its result are influenced by the number of decision trees easily. Hence, setting the number of decision trees is a key to RF classification. When the number of decision trees increases to a certain level, RF gets a stable
classification. For this purpose, we need to carry out many experiments. At last, we built 300 decision trees and extracted cropland by the `ee.Classifier.randomForest` function on GEE platform. However, Due to noise, the result of RF classification also has the salt and pepper effect (Figure 4a). Some of the tiny patches are not cultivated land actually, but are classified as cultivated land because of noise. These small patches not only affect the visual quality, but also affect the classification accuracy to some extent. So we need to optimize the results. Morphology processing was chose in this study. First, we expanded the boundary of patches by the dilation operation. Then, the small patches were eliminated based on erosion operation. As a result, we inhibited the salt and pepper effect (Figure 4b).

![Figure 4](image-url)

**Figure 4.** The example of (a) RF classification;(b) the merged results with RF classification and morphology processing

4. Results

4.1. Classification accuracy

The quantitative evaluation was applied to analyze the classification accuracy such as overall accuracy (OA), user accuracy (UA), producer accuracy (PA) and Kappa coefficient. On the whole, based on RF method and morphology processing, the cropland we mapped has high classification accuracy with OA to 0.979, UA to 0.923, PA to 0.925 and kappa coefficient to 0.933.

4.2. Influence of water on classification

In order to evaluate the effects of water on classification, we compared the results of cropland classification before and after water mask. The evaluation indicators we used include OA and root mean squared error (RMSE). We first compared OA of the both and found that cropland accuracy was improved after water mask. The OA was 97.9%, while the original was 91.8%. Considering that RMSE performs better on accuracy assessments, we calculated the RMSE and drew a scatter plot for facilitate analysis (Figure 5). In the figure, the Y-axis denotes the cropland area of reference data, while the X-axis denotes the metrical based on 254 cropland sample polygons. The oblique solid line is a fitted line between reference data and the metrical. The lower RMSE, the higher R2, the higher classification accuracy. The figure shows that the classification accuracy is higher after water removed with RMSE of 0.0167, while RMSE improved by 4% approximately. The results reflect that water hinders the cropland extraction of the study area, while water mask facilitates classification.
4.3. Influence of different features on classification

In order to analyze the influence of different characteristics on classification accuracy, we compared the mean overall accuracy before and after the features removed from the cropland feature space. The experimental results are shown as in Fig. 6.

The figure shows although phenological features is the most important cropland feature, its contribution to precision is not the largest, which average loss accuracy is between 4.9% and 5.5%. By contrast, terrain features has dominant influence on classification accuracy, especially elevation, with a mean loss accuracy of 7.6%. The original spectral band of SWIR 1 has the least influence on the classification accuracy, with a mean loss accuracy of 1% only. Due to the large loss of precision in terrain features and phenological features, we removed the both at the same time to observe the change in precision. As a result, the accuracy dropped from 97.9% to 87.5%, with a loss of 10.4%. When there is no feature in the cropland feature space, the accuracy of cropland classification is 70.6%, downing by 27.3%. In summary, the rank of these features is terrain features, phenological features, texture features, TCT features and spectral features according to the degree of contribution to classification accuracy. Among them, terrain features and phenological features are the main factors affecting of cropland extraction. In the cropland feature space, the more classification features are reduced, the more accuracy loss.

4.4. Accuracy comparison between cropland products

The cropland product of this study compared with other cropland products such as MCD12Q1 and GFSAD30 (Table 2). In terms of the cropland area, it is inevitable that the cropland area are different among the three because of different classification systems. So we need to compare their overall accuracy. In terms of accuracy, this study product has the highest accuracy, which was 6.2% higher than...
GFSAD30 and 24.3% higher than MCD12Q1. The results show the cropland mapping of this study has a high accuracy.

| Product      | Area(km²) | Proportion | UA (%) | PA (%) | Kappa | OA(%) |
|--------------|-----------|------------|--------|--------|-------|-------|
| this study   | 2366      | 22.1       | 92.3   | 92.5   | 90.3% | 97.9  |
| MCD12Q1      | 2811      | 25.9       | -      | -      | -     | 73.6  |
| GFSAD30      | 3906      | 36         | 78.3   | 83.4   | -     | 91.7  |

5. Conclusion

The problems such as cropland mixed pixels, imprecise phenological information and heavy workloads are the main barriers in the cropland extraction. Aiming at the first two problems, we used a water mask and phenological features based on Landsat 8 images to extract the cropland in GEE platform, we mapped the cropland extent and verified the classification accuracy. The results showed that the cropland mapping has high accuracy. The mean OA was 97.9%, which was .2% higher than GFSAD30 and 24.3% higher than MCD12Q1. The mean UA is 92.3%, while the mean PA is 92.5% and the kappa coefficient is 90.3%. In terms of the image data, the multi-spectral Landsat-8 image collection at 30-m resolution provides enough spatial information for us to acquire the cropland characteristics. On the other hand, the NDVI time series with high spatial and temporal resolution from Landsat-8 and MODIS images is helpful to obtain more accurate phenological features. More importantly, we used a water mask to reduce the obstruction of water in farmland extraction. The water mask relies on high precision and high resolution JRC GSW water data. Because of the water data with high accuracy, the water mask has a positive effect on cropland classification. The water mask makes the overall accuracy of cropland improved by 6.9%.

Although our method is helpful for the cultivated land identification, but also still have some limitations. In particular, the mask method relies on high-precision mask products. Otherwise, it can have cascading effects. In the future, we can focus on using high-resolution images and better mapping techniques to identify cropland.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (No. 41961056).

References

[1] X. Chao, J. R. Zhang, Y. Z. Pan, and W. Q. Zhu, Extraction of cropland information based on multi-temporal TM images, Remote Sens. Land Resour. 25(2013) 166-173.
[2] J. Chen, T. Q. Chen, H. M. Liu, X. M. Mei, Q. B. Zhao and M. Deng, Hierarchical extraction of farmland from high-resolution remote sensing imagery, Trans. Chin. Soc. Agric. Eng. 31(2015) 190-198.
[3] Y. Y. Bai, J. L. Gao and B. L. Zhang, Extraction of crop planting structure based on time-series NDVI of Landsat8 images, Arid Land Geogr. 42(2019) 893-901.
[4] W. X. Wang, Detection Method of Land Use Change Based on Landsat Image Time Series, Nanjing University, 2018.
[5] Y. d. Xu, Y. Le, F. R. Zhao, et al, Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change detection and post classification approach Experiments from three sites in Africa, Remote Sens. Environ. 218(2018) 13-31.
[6] Y. Q. Ge, Y. R. Li, D. C. Li, et al, Two-way fusion experiment of Landsat and MODIS satellite data, Sci. Surv. Mapp. 44(2019) 107-114.
[7] J. G. Peng, W. J. Luo, X. B. Ning, et al, Remote Sensing Image Infusion Based on STARFM, Cent. South Forest Inventory Plann. 37(2018) 32-37.

[8] X. Cao, X. H. Chen, W. W. Zhang, et al, Global cultivated land mapping at 30 m spatial resolution, Sci. China Earth Sci. 59(2016) 2275-2284.

[9] Z. Zhu, S. X. Wang, C. E. Woodcock, Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images, Remote Sens. Environ. 15(2015) 269-277.

[10] Y. Rao, X. Zhu, J. Chen et al, An Improved Method for Producing High Spatial-Resolution NDVI Time Series Datasets with Multi-Temporal MODIS NDVI Data and LANDSAT TM/ETM+ Images, Remote Sens. 7(2015) 7865-7891.

[11] Z. X. He, M. Zhang, B. F. Wu, et al, Extraction of summer crop in Jiangsu based on Google Earth Engine, J. Geo-inf. Sci. 5(2019)752-766.

[12] M. H. A. Baig, L. F. Zhang, T. Shuai, et al, Derivation of a tasseled cap transformation based on LANDSAT 8 at-satellite reflectance, Remote Sens. Lett. 5(2014) 423-431.

[13] M. Y. Zhu, Research on remote sensing image classification based on multi-feature fusion, Yanshan university, 2016.