Monitoring of Rice Irrigation Area Based on GF-1 Data and Spectral Matching Technology - A Case Study of Chuanhang Irrigation Area in Suqian City

Tong Daobin¹, Duan Yuanyuan², Chen Xiuhua², Lu Yizhu², Liu Hongjie², Zhao Guanliang³, Song Wenlong²*

¹Water Resources Bureau of Sucheng District, Suqian City, Jiangsu Province, 223800, P.R. China,
²China Institute of Water Resources and Hydropower Research, Beijing, 100000, P.R. China,
³Management Office of Chuanhang Irrigation District, Sucheng District, Suqian City, Jiangsu Province, 223800, P.R. China
*Corresponding author’s e-mail: songwl@iwhr.com

Abstract. China is the world's largest rice producer and consumer. Accurate identification of rice planting and irrigation area is of great significance to China's food security. Based on GF-1 multispectral data, spectral matching and OTSU algorithm were used to monitor the rice irrigation area in Chuanhang Irrigation Area from 2016 to 2020. The accuracy was verified by confusion matrix. The overall accuracy was 98.1691%, 91.8706%, 95.2451%, 93.0604%, 91.9697%, respectively. The Kappa coefficients were 0.9541, 0.7898, 0.8777, 0.8330 and 0.7824, respectively, which proved that this method was suitable for rice irrigation area extraction based on high spatial resolution remote sensing data, and could provide technical support for rice irrigation area monitoring in irrigated areas.

1. Introduction
Rice is one of the main food crops in China, with its planting area accounting for 30% of China's grain crop planting area and its yield accounting for about half of the total grain output in China [1]. Therefore, it is of great significance to accurately monitor the variation of spatial and temporal distribution of rice for food security and water resources utilization [2]. Irrigation areas rely on traditional methods based on survey statistics to obtain rice planting area, which consumes manpower and material resources, has poor timeliness, and lacks spatial information [3-5]. Remote sensing method has the advantages of economy, accuracy and efficiency, and can quickly carry out large-scale monitoring and identification, which is widely used in the research of rice planting area extraction. Domestic and foreign scholars have carried out a large number of studies on this [6-8]. Low resolution satellite data has a short return period and high temporal resolution, which can obtain the complete time series data of rice growth cycle. Rice planting area can be extracted through multi-phase images, which is the most commonly used data source for remote sensing monitoring of rice planting area [9-11]. However, southern China is the main region for rice cultivation, with complex terrain, diverse planting structures and serious fragmentation of land plots. Therefore, the spatial resolution of low-resolution satellite data cannot meet the precision requirements of rice planting area extraction in...
southern China, while high-resolution satellite data can adapt to the actual situation in southern China [12-13]. Due to the limitation of time resolution, high-resolution satellite data mainly uses spectral features, texture features, terrain features and other information of single-period images to extract rice planting area through supervised classification methods such as decision tree, support vector machine and random forest [14]. High resolution satellite data are susceptible to weather, so there are relatively few studies on rice planting area extraction based on high resolution image time series, and the supervised classification method requires a large number of ground sample points to ensure the accuracy of the results. In this paper, Chuanhang Irrigation District in Suqian, Jiangsu Province in southern China was selected as the research area, GF-1 multispectral image was used as the data source to extract the rice planting area of Suqian Irrigation District from 2016 to 2020 by spectral matching method based on pixel scale using the time series of high-resolution remote sensing images, evaluated the extraction accuracy of the method, and verified the recognition effect of rice planting area extraction method based on high resolution image time series.

2. Research area and data

2.1. Study area

Chuanhang Irrigation Area is located in the southwest of Sucheng District, Suqian City, Jiangsu Province, between latitude 34°02' ~ 34°13' N, longitude 118°100' ~ 118°40' E, bordering Suqian City and Zaohe Irrigation Area in the north, adjacent to the Beijing-Hangzhou Grand Canal in the east, Xishahe River and Xuhong River in the west, and facing Yunnan Irrigation Area in the south with Yaohe River and Dongfangdagou as the boundary. It belongs to the Huaihe River basin and is one of the large irrigation areas utilizing the water source of the Beijing-Hangzhou Grand Canal. The irrigation area covers 7 townships (towns and street offices), including Sankeshu, Buzi, Chenji, Luowei, Nancai, Yangbei, and Longhe, with a total controlled area of 325.1km², cultivated area of 320,000 mu, effective irrigation area of 314,000 mu, and designed irrigation area of 320,000 mu (see Figure 1). Chuanhang Irrigation Area belongs to the Yellow Flood Alluvial Plain. The terrain in the area is relatively flat and slopes from northwest to southeast, with the highest elevation of 24.7m and the lowest elevation of 16.6m. The main crops planted in the irrigated area are rice, wheat, cotton, corn, vegetables and other crops. Rice - wheat rotation and wheat - jade rotation are the main local planting patterns. The growth cycle of rice is shown in Table 1.

![Figure 1. Schematic diagram of the location of the Chuanhang irrigation District](image-url)
Table 1. Periodic table of rice growth

| Months | Ten days | Growth stage          |
|--------|----------|-----------------------|
| 4      | 1        | Raising seedlings     |
|        | 2        |                       |
|        | 3        |                       |
| 5      | 2        | Trefoil stage         |
|        | 3        |                       |
| 6      | 2        | Transplant            |
|        | 3        |                       |
| 7      | 1        | Rejuvenation period   |
|        | 2        |                       |
|        | 3        |                       |
| 8      | 1        | Booting stage         |
|        | 2        |                       |
|        | 3        |                       |
| 9      | 2        | Milk maturity         |
|        | 3        |                       |
| 10     | 2        | Maturity              |
|        | 3        |                       |

2.2. The data source

2.2.1. High resolution satellite remote sensing data. The data source is the 2016-2020 GF-1 16m multispectral data, the source is the China Resources Satellite Application Center (http://www.cresda.com/CN/), and the detailed parameters are shown in Table 2. Use the images with cloud coverage of less than 10% in the study area from 2016 to 2020, and a total of 60 scene data. This data is used to accurately obtain the time-series spectral characteristic information of rice.

Table 2. GF-1 multispectral sensor information

| Spectrum range/(μm) | Spatial resolution/(m) | Width/(Km) | Side swing ability | Revisit cycle/(d) |
|---------------------|------------------------|------------|--------------------|-------------------|
| Multispectral camera| 0.45-0.52              | 16         | 800                | ±35°              | 2                 |
|                     | 0.52-0.59              |            |                    |                   |                   |
|                     | 0.63-0.69              |            |                    |                   |                   |
|                     | 0.77-0.89              |            |                    |                   |                   |

The remote sensing data were preprocessed by orthographic correction, radiometric calibration, geometric correction, atmospheric correction, etc., and NDVI was calculated to obtain the NDVI time series curves of each pixel in the study area. NDVI can better reflect the vegetation growth state, and the normalized processing can better limit the data range, which is conducive to the subsequent spectral matching calculation and can be used to extract the irrigation area.

2.2.2. Measured sample point data. The spectral matching algorithm at pixel scale requires the NDVI time series curve of rice in the irrigated area as the standard spectrum, so field investigation and sampling work are carried out in the study area to obtain the real ground sample points of the main features in the study area, which can be used to extract the irrigated area and verify the accuracy.
Combined with the history of irrigation with high-resolution satellite remote sensing image, the selection of typical significance of rice irrigation areas and the rice irrigation area, the selection of the irrigation area should pay attention to avoid the interference of nearby buildings and roads, to avoid farmland boundaries, in case when take sample point object that contains both farmland and other interference. Finally, a total of 101 real sample points were selected in the study area (see Figure 2) to record the location of sample points, crop types, irrigation conditions and other information.

In order to ensure sufficient sample points for the confusion matrix, a number of sample areas, a total of 3400 pixels, were selected from the measured sample point data of rice irrigation area and non-rice irrigation area on the basis of field investigation and visual interpretation of the image.

3. Methods

In this paper, the spectral matching method is applied to pixel scale for high-resolution remote sensing images, so as to ensure that the time-series spectral curves of all pixels participate in the matching calculation and reduce the errors caused by clustering process. In addition, the OTSU adaptive threshold algorithm is introduced to automatically determine the extraction threshold of irrigation area. The introduction of OTSU algorithm can ensure that each pixel can automatically determine the
reasonable variation range of similarity while calculating similarity information. It has good adaptability to change in different years and different data conditions, and the irrigation area can be extracted without human intervention. The remote sensing monitoring method of irrigation area based on pixel scale spectral matching calculation is suitable for high-resolution satellite remote sensing data, meets the demand of extracting irrigation area of small plots, and can improve the accuracy of monitoring results of irrigation area.

3.1. Spectral matching method
Spectral matching is a method to quantify the similarity between the endmember spectrum (sample spectrum) and the target spectrum (spectrum to be measured). In this paper, statistical algorithm and spectral waveform feature algorithm were used to calculate spectral similarity. Through spectral similarity, the matching degree between the endmember spectra of main crops in the study area and the target spectrum was quantitatively analyzed.

Spectral Similarity Value (SSV) combines the characteristics of SCS (shape quantification) and EDS (distance quantification), and measures the physical similarity between endmember spectrum and target spectrum from the curve shape of NDVI time series and the spatial distance of spectral features. End-member spectrum $C = (S_1, S_2, ..., S_n)$, and the target spectrum $X_{i,j} = (b_1, b_2, ..., b_n)$, the calculation formula of SSV is as follows:

$$
SCS = \frac{1}{n+1} \left[ \frac{\sum_{i=1,j=1}^{n} (X_{i,j} - \mu_{i,j})(C - \mu_S)}{\sigma_X \sigma_C} \right]
$$

$$
EDS = \left( \sum_{i=1,j=1}^{n} (X_{i,j} - C)^2 \right)^{1/2}
$$

$$
EDS_{normal} = (EDS - m)(M - m)^{-1}
$$

$$
SSV = \sqrt{EDS_{normal}^2 + (1 - SCS)^2}
$$

Where, SCS is spectral correlation similarity; EDS is Euclidean distance; $X_{i,j}$ is the NDVI value of the target spectrum; $C$ is the NDVI value of end-member spectrum; $i$ and $j$ are respectively the ith row and the j column; $N$ is the number of layers of data set; $\mu_{i,j}$ are the mean NDVI of the target spectrum; $\mu_s$ is the mean NDVI of end-member spectrum; $\sigma_x$ is the NDVI standard deviation of the target spectrum, and $\sigma_c$ is the NDVI standard deviation of the end-member spectrum. $M$ and $M$ represent the maximum and minimum Euclidean distance respectively. $EDS_{normal}$ is the result of Euclidean distance normalization.

The smaller the SSV value is, the more similar the spectra are, and the range is usually between 0 and 1.414. The difference of NDVI time series curves among different crops was great, and the SSV value was high. For the same crop, the NDVI time series curve in irrigated area was more consistent than that in non-irrigated area, and the SSV value was lower.

3.2. OTSU algorithm
The NDVI time series curves of rice in the irrigation area obtained by field sampling were taken as endmember spectra respectively, and the SSV values of each pixel in the research area were calculated by using the above index to measure spectral similarity. The higher the spectral similarity is, the lower the SSV value is. The lower the spectral similarity of different classes, the higher the SSV value. It is necessary to determine a reasonable segmentation threshold, and when the SSV value is less than the threshold value, it is identified as the same category as the endmember spectrum. Therefore, the OTSU adaptive threshold algorithm was introduced to calculate the SSV segmentation threshold to determine whether it was the irrigated area of rice. If the threshold value was less than this, it was the irrigated area, so as to identify the spatial distribution of the irrigated area in the study area and greatly reduce the impact caused by human intervention.
OTSU algorithm is an adaptive threshold selection method without parameters [15]. This method can set a temporary threshold and then calculate the variance value of the pixel SSV in the range on both sides of the threshold. When the variance reaches the maximum, the corresponding pixel SSV is the optimal threshold value, that is, the difference between the target area and the background area is maximized. Then, those whose similarity is greater than or equal to the threshold are merged into one class, and those whose similarity is less than the threshold are merged into another class to obtain the corresponding binarization image. The specific principle and algorithm are as follows:

When the image data magnitude is \( L \) (\( G = 0, 1, ..., L-1 \)), the initial threshold is \( h \), and the image is divided into target region \( T \) and background region \( B \). \( P_i \) represents the probability of the occurrence of pixels whose SSV is \( I \), \( n_i \) represents the number of pixels whose SSV is \( I \), \( N \) represents the total number of pixels, and the probabilities of parts \( T \) and \( B \) are:

\[
P_T(h) = \sum_{i=0}^{h} \frac{n_i}{N}, \quad P_B(h) = \sum_{i=h+1}^{L-1} \frac{n_i}{N}
\]

(2)

The mean value of corresponding parts \( T \) and \( B \) is:

\[
\mu_T(h) = \frac{\sum_{i=0}^{h} iP_i}{P_T(h)}, \quad \mu_B(h) = \frac{\sum_{h+1}^{L-1} iP_i}{P_B(h)}
\]

(3)

Then there is a threshold \( H_0 \), which is the optimal threshold:

\[
h_0 = \arg \max (P_T P_B (\mu_T - \mu_B)^2)
\]

(4)

4. Results and analysis

4.1. Standard spectral extraction of irrigated rice

The spectral curve of rice growth period from 2016 to 2020 extracted from GF-1 NDVI time series image data is shown in Fig. 3. The time series spectrum curve of rice is shown in the figure, which is obtained by calculating the mean spectrum based on the spectral value extracted from the time series NDVI image of sampling points. The mean spectrum is used for spectral matching to obtain the irrigation area of rice. As can be seen from the spectral curve, the NDVI value of rice began to increase after the transplanting stage, reached the highest value at the heading stage (from late August to early September), and then gradually decreased to the harvest stage (late October). The spectral curves of rice were basically consistent with the actual growth period.

Figure 3. Spectral curve of rice growth period
4.2. Analysis on the characteristics of irrigated area and spatial distribution of rice

The statistical table (see Table 3) and spatial distribution (see Figure 4) of the irrigated area of rice were obtained by spectral matching method. The spatial distribution of rice irrigated area from 2016 to 2020 showed that the irrigated area of rice reached the maximum value of 237,000 mu in 2020, accounting for about 39% of the total area of the study area and about 68.7% of the cultivated area of the study area. In 2016, the irrigated area of rice was the smallest, accounting for 197,400 mu, 32.5% of the area of the study area, and 57.2% of the cultivated area. In terms of spatial distribution, rice was mainly distributed in the west, middle and east of the study area in 2016 and 2018, and the rice area in the south of the study area was broken. In 2017 and 2019, there was also a large area of concentrated distribution of rice in the south of the study area. In 2020, the distribution of rice in the southern part of the study area was also relatively concentrated, but the overall area was smaller than that in 2017 and 2019.

Table 3. Statistical table of rice irrigation area in Chuanhang Irrigation Area from 2016 to 2020

| Year | 2016 | 2017 | 2018 | 2019 | 2020 |
|------|------|------|------|------|------|
| Irrigated area of rice (km²) | 142.2 | 190  | 154.27| 199.87| 167.87|

(a) Rice irrigated area in 2016  (b) Rice irrigated area in 2017
Figure 4. 2016-2020 distribution map of rice irrigation area in Chuanhang Irrigation Area

(c) Rice irrigated area in 2018

(d) Rice irrigated area in 2019

(e) Rice irrigated area in 2019
4.3. Verify the accuracy

The accuracy of the confusion matrix was verified for the extraction results of the irrigated area of rice from 2016 to 2020, and the results were shown in Table 4. The results showed that the overall accuracy of rice irrigation area reached more than 91% from 2016 to 2020, among which, the overall accuracy reached 98.1691% in 2016. In 2016, 2018 and 2019, the Kappa coefficient was above 0.8, with the highest value of 0.9541 in 2016, indicating that the monitoring results of rice irrigation area in these three years were in high consistency with the actual situation. The Kappa coefficient in 2017 and 2020 is about 0.78, indicating that the monitoring results of rice irrigation area in these two years are basically consistent with the actual situation. In general, spectral matching based on GF-1 data can accurately extract the irrigated area of rice. In 2019, the commission error was large, which was 14.94%. This may be because the area of corn planted in Chuanhang Irrigation Area in 2019 was small and fragmented, and the harvest time of corn in 2019 was close to that of rice, which resulted in the similarity of spectral curves of the two crops, leading to the misclassification of the irrigated area of corn into the irrigated area of rice. In 2020, the omission error is large. Through NDVI time series images, it was found that the rice in some areas of Chuanhang irrigation area was harvested earlier in 2020 than in previous years, leading to some rice areas not being identified. However, when magnifying the extraction effect, it can be seen that this method can effectively distinguish farmland, road and residential land, and identify the broken farmland, which is suitable for China's cultivated land with complex planting structure and scattered distribution, and reduces the classification error caused by mixed pixels.

Table 4. Result evaluation based on pixels

| Year | Overall Accuracy | Kappa Coefficient | Commission Error | User’s Accuracy | Omission Error | Producer’s Accuracy |
|------|-----------------|-------------------|------------------|----------------|----------------|--------------------|
| 2016 | 98.1691%        | 0.9541            | 0.11%            | 99.89%         | 6.34%          | 93.66%             |
| 2017 | 91.8706%        | 0.7898            | 9.01%            | 90.99%         | 21.23%         | 78.77%             |
| 2018 | 95.2451%        | 0.8777            | 0.61%            | 99.39%         | 16.15%         | 83.85%             |
| 2019 | 93.0604%        | 0.8330            | 14.94%           | 85.06%         | 8.4%           | 91.6%              |
| 2020 | 91.9697%        | 0.7824            | 0                | 100%           | 28.57%         | 71.43%             |

5. Conclusion

In this study, spectral matching and OTSU algorithm were adopted at pixel scale to obtain the irrigation area and distribution of rice in Chuanhang Irrigation Area from 2016 to 2020 by utilizing high-resolution NDVI time series images from 2016 to 2020 in cooperation with field survey and sampling, which verified the monitoring ability of spectral matching method on pixel scale. Based on the spectral matching method and OTSU algorithm, the irrigated area of rice from 2016 to 2020 is 213300, 285000, 231400, 299800 and 251,800 mu respectively, which is basically consistent with the field survey results. The monitoring results were verified by confusion matrix: in 2016-2020, the overall accuracy of rice irrigated area were 98.1691%, 91.8706%, 95.2451%, 93.0604%, 91.9697%, and the Kappa coefficients were 0.9541, 0.7898, 0.8777, 0.8330, 0.7824, respectively. It shows that the monitoring results have high accuracy. The pixel-scale spectral matching calculation method adopted in this paper is suitable for the extraction of irrigation area based on high spatial resolution remote sensing data, which ensures the high accuracy of extraction of irrigation area.

Acknowledgments

This work was supported by the National Key R&D Program of China (2018YFC1508702, 2016YFC0400106-2), by the Research Program of China Institute of Water Resources and
Hydropower Research (JZ0145B472016, JZ0145B862017), by the Technical Service Project (JZ0203A072021), and by the Natural Science Foundation of China (51609259).

References

[1] Wu Mingquan, Wang Changyao, Niu Zheng. Transactions of the CSAE, 2010(07):240-244. (in Chinese with English abstract)

[2] Wang Y, Zang S, Tian Y. Mapping paddy rice with the random forest algorithm using MODIS and SMAP time series [J]. Chaos Solitons & Fractals, 2020, 140:110116.

[3] Xiao W, Xu S, He T. Mapping Paddy Rice with Sentinel-1/2 and Phenology-, Object-Based Algorithm Implementation in Hangjiahu Plain in China Using GEE Platform. 2021.

[4] Sun Shujuan, Li Minlu, Wang Ping, et al. Extraction of rice planting area based on GF-1/WFV EVI time series data [J]. Fujian Agricultural Journal, 2018, v.33; No.175(06):28-33

[5] Biggs, T. W., Thenkabail, P. S., Gumma, M. K., Scott, C. A., Parthasaradhi, G. R., Turrall, H. N., 2006. Irrigated areamapping in heterogeneous landscapes with MODIS time series, ground truth and census data, krishna basin, India. International Journal of Remote Sensing, 27(19), 4245-4266. https://doi.org/10.1080/01431160600851801.

[6] Lu Y, Song W, Su Z, et al. MAPPING IRRIGATED AREAS USING RANDOM FOREST BASED ON GF-1 MULTI-SPECTRAL DATA [J]. 2020.

[7] Song Wenlong, Li Meng, Lu Jingxuan, et al. Study on Monitoring Irrigation Area Based on GF-1 Satellite Data: A Case Study of Donglei Phase II Yellow Irrigation Area [J]. Journal of Hydraulic Engineering, 2019, 50(7):854-863.

[8] Breiman, L., 2001. Random Forests. Machine Learning, 45,5-32. https://doi.org/10.1023/A:1010933404324.

[9] Song Panpan, Du Xin, Wu Liangcai, et al. Remote Sensing of Rice in Southern China Based on Spectral Time Series Fitting [J]. Journal of Geo-Information Science, 2017, 19(001):117-124.

[10] Wang Jianxun, Hua Li, Deng Shichao, et al. Study on rice extraction in hilly areas of South China based on GF-1 and MODIS [J]. Agricultural Resources and Zoning of China, 2019(5):37-46.

[11] Yang Xiaohuan, Zhang xingping, Jiang dong. method of extracting sown area of multiple crops based on NDVI feature value of MODIS time series [J]. Journal of resource science, 2004, 26(006):17-22.

[12] Yang Hequn, Ma Bingxin, Gu Wei. Extraction of rice planting distribution in Chongming County based on time-series HJ-CCD data [C]// The 33rd annual meeting of the Chinese Meteorological Society S14 to improve the ability of meteorological science and technology innovation and ensure high yield and efficiency of agriculture. 2016.

[13] Xin Chuan, Xu Xingang, Wang Jihua, et al. Crop classification and identification based on time series environmental satellite image [J]. Transactions of the Chinese Society of Agricultural Engineering, 2013(02):169-176. (in Chinese)

[14] Teluguntla P, Thenkabail P S, Xiong J, et al. Mapping cropland extent and areas of Australia at 30-m resolution using multi-year time-series Landsat data and Random Forest machine learning algorithm through Google Earth Engine (GEE) Cloud Computing. 2017.

[15] OHTSU N. A threshold selection method from Gray-Level histograms [J]. IEEE Transactions on Systems Man & Cybernetics, 2007, 9(1): 62-66.