

# Mapping wetland using the object-based stacked generalization method based on multi-temporal optical and SAR data

Yaotong Cai<sup>a,b,c</sup>, Xinyu Li<sup>a,b,c,d</sup>, Meng Zhang<sup>a,b,c,\*</sup>, Hui Lin<sup>a,b,c,\*</sup>

<sup>a</sup> Research Center of Forestry Remote Sensing & Information Engineering Central South University of Forestry & Technology, Changsha 410004, Hunan province, PR China

<sup>b</sup> Key Laboratory of Forestry Remote Sensing Based Big Data & Ecological Security for Hunan Province, Changsha 410004, Hunan province, PR China

<sup>c</sup> Key Laboratory of State Forestry & Grassland Administration on Forest Resources Management and Monitoring in Southern Area, Changsha 410004, Hunan province, PR China

<sup>d</sup> School of Information Science and Engineering, Hunan First Normal University, Changsha 410205, PR China

## ARTICLE INFO

### Keywords:

Wetland  
Classification  
Sentinel-1/2  
Multi-Temporal  
Object-Based  
Stacked generalization

## ABSTRACT

Wetland ecosystems have experienced dramatic challenges in the past few decades due to natural and human factors. Wetland maps are essential for the conservation and management of terrestrial ecosystems. This study is to obtain an accurate wetland map using an object-based stacked generalization (Stacking) method on the basis of multi-temporal Sentinel-1 and Sentinel-2 data. Firstly, the Robust Adaptive Spatial Temporal Fusion Model (RASTFM) is used to get time series Sentinel-2 NDVI, from which the vegetation phenology variables are derived by the threshold method. Subsequently, both vertical transmit-vertical receive (VV) and vertical transmit-horizontal receive (VH) polarization backscatters ( $\sigma_0$  VV,  $\sigma_0$  VH) are obtained using the time series Sentinel-1 images. Speckle noise inherent in SAR data, resulting in over-segmentation or under-segmentation, can affect image segmentation and degrade the accuracies of wetland classification. Therefore, we segment Sentinel-2 multispectral images to delineate meaningful objects in this study. Then, in order to reduce data redundancy and computation time, we analyze the optimal feature combination using the Sentinel-2 multispectral images, Sentinel-2 NDVI time series, phenological variables and other vegetation index derived from Sentinel-2 multispectral images, as well as time series Sentinel-1 backscatters at the object level. Finally, the stacked generalization algorithm is utilized to extract the wetland information based on the optimal feature combination in the Dongting Lake wetland. The overall accuracy and Kappa coefficient of the object-based stacked generalization method are 92.46% and 0.92, which are 3.88% and 0.04 higher than that using the pixel-based method. Moreover, the object-based stacked generalization algorithm is superior to single classifiers in classifying vegetation of high heterogeneity areas.

## 1. Introduction

Wetland is an open, complex and interactive ecosystem, accounting for 6% of the earth surface area (Jahncke et al., 2018; Singh and Sinha, 2019). Wetland maintains biodiversity and regulates runoff, and plays an important role in improving water quality and regulating microclimate (Debanshi and Pal, 2020; Guan et al., 2020). In recent years, natural wetlands are dramatically shrinking, because of the impact of human activities and global climate change. Therefore, it is urgent to develop methods for monitoring and assessing the wetland changes and improving the scientific understanding of the wetland ecosystem.

Remote sensing techniques can investigate large wetland systems and monitor their changes over time. However, there is spectral overlap

between the vegetation types when mono-temporal multispectral images are used. Time series optical satellite images are able to periodically capture the phenological information of wetland vegetation and can generate wetland maps with high accuracy (Tong et al., 2019). Additionally, images of spring and autumn are able to optimize the separability between vegetation types (Deventer et al., 2019). MODIS, Landsat and SPOT images have been widely and successfully used for monitoring the wetland vegetation and detecting the presence and extent of flood (Evans and Costa, 2013; Hou et al., 2018; Zhou et al., 2016). However, their spectral and spatial resolutions may compromise detail wetland vegetation type identification. Moreover, traditional multispectral sensors, such as the Landsat TM/ETM + /OLI, often have saturation of reflected signals in the red and near-infrared (NIR)

\* Corresponding authors at: Research Center of Forestry Remote Sensing & Information Engineering, Central South University of Forestry & Technology, Changsha 410004, Hunan province, PR China.

E-mail addresses: [mengzhang@csuft.edu.cn](mailto:mengzhang@csuft.edu.cn) (M. Zhang), [Huolin@csuft.edu.cn](mailto:Huolin@csuft.edu.cn) (H. Lin).

spectral regions when map the vegetation community of subtropical regions (Deventer et al., 2019). With the development of the remote sensing technology, the availability of high spatial/spectral resolution imagery has been increased. Sentinel-2 Multispectral Instrument (MSI) launched by ESA recently, a new generation multispectral sensor, has been successfully used in land use/land cover research, because of its high spatio-temporal resolution, wide spatial coverage and broad spectrum. Furthermore, the three red-edge bands of Sentinel-2 images are particularly effective for vegetation monitoring (Kaplan and Avdan, 2019). However, optical satellite sensors are always vulnerable to the cloudy and rainy weathers, so it is difficult to acquire adequate and clear Sentinel-2 images in tropical and subtropical coastal areas. Spatio-temporal fusion algorithms were proposed to obtain time series remote sensing data with a fine resolution and have been successfully applied to land use/cover classification, burned area detection and time series biomass/GPP/NPP estimation research (Gevaert and García-Haro, 2015; Wang and Atkinson, 2018; Zhao et al., 2018).

Being immune to weather conditions, multi-temporal synthetic aperture radar (SAR) data (e.g. Sentinel-1, GF-3, RADARSAT, ALOS-2 and TanDEM-X) are frequently employed for flood detection, forest classification and crop identification, as well as biomass mapping (Li and Bijk, 2019; Martone et al., 2018; Zhang et al., 2018a). The application of SAR in scientific fields depends on SAR polarization and frequency (Mahdianpari et al., 2017). According to the scatter mechanisms of ground targets captured by Polarimetric Synthetic Aperture Radar (PolSAR), a range of studies have been conducted for discriminating various land cover types with different shapes, structures as well as roughness and permittivity. For instance, dual-polarization (vertical transmit-vertical receive (VV) / vertical transmit-horizontal receive (VH) data have been used for the identification of the water body and vegetation (Mahdianpari et al., 2017; Jahncke et al., 2018; Evans et al., 2013). Frequency (or wavelength) is another SAR parameter, which related to the penetration depths through ground targets, reflecting the land surface structure. Due to the higher penetration ability, longer wavelengths, such as P- and L-band, are able to detect the vegetation canopy and quantify soil moisture, whereas shorter wavelengths, like C- and X-band, perform better in classifying land cover in open area (e.g., open wetland). Consequently, multi-frequency and multi-source SAR data have been integrally exploited for wetland monitoring and achieved successful results (Mahdianpari et al., 2017; DeLancey et al., 2019). However, the high cost and acquisition difficulty limit the feasibility of multi-frequency and multi-source PolSAR data in many cases, especially in time series analysis. Sentinel-1 satellite with a C-band and dual-polarization (VV/VH) sensor provides a promising opportunity for time-series wetland land cover analysis at no cost. Besides, polarimetric features, such as backscatter coefficients, interferometry data (e.g., coherence), Stocks vector, the degree of polarization and linear polarization ratio, have been widely utilized to help identify land cover types (Mohammadianesh et al., 2018, 2019; Guo et al., 2019). Some studies reported that the backscatter coefficients in Sentinel-1 SAR time series have the greatest utility among SAR features in land cover classification and could obtain the highest classification accuracy (Li et al., 2020; Cai et al., 2019). Therefore, the Sentinel-1 backscatters ( $\sigma_0$  VV,  $\sigma_0$  VH) from all within-year observations are employed to discriminate wetland types in this study.

However, the land use/cover classification accuracy using SAR data is usually lower than that of the optical multispectral data with the same spatial resolution. Combining optical and SAR data for land cover and vegetation mapping brings accuracy higher than that of using only either optical or SAR images (Guo et al., 2019; Shuai et al., 2019; Zhang and Xu, 2018). Additionally, with the launch of many long term observation missions, the ever-increasing volume and accessibility of optical and SAR remote sensing images enable the accuracy improvement

in the quantitative research of remote sensing (Ienco et al., 2019; Wittke et al., 2019). However, very few studies on wetland monitoring have used dense time series optical and SAR images. Only recently, the time-series Sentinel-1 and Sentinel-2 images were combined to improve the land surface monitoring accuracy (Cai et al., 2019; Steinhausen et al., 2018). Therefore, combining Sentinel-1 and Sentinel-2 data is useful and has significant advantages in wetland mapping at regional or global scales. The object-based image analysis (OBIA) has some advantages over the pixel-based methods, such as adding object shape and context features, as well as reducing the “salt-and-pepper” phenomenon. Therefore, the OBIA has got wider application in classification and change detection in wetland ecosystem than pixel-based methods, and it is very useful when high/very high spatial resolution images are used for classification (Hossain and Chen, 2019; Phiri et al., 2018). Multi-resolution segmentation (MRS) is one of the most widely used and compatible segmentation algorithm. It is superior to other segmentation algorithms (e.g., chessboard segmentation, quadtree-based segmentation and spectral difference segmentation) as it allows to construct primitive objects with different sizes using heterogeneous thresholds and enhance the response of objects generation to landscape patch structure. Its feasibility in complex landscapes has been demonstrated by Mahdianpari et al. (2017) and Yang et al. (2019). To get high accuracy vegetation classification in the study area, we applied the MRS method to generate homogeneous image objects of ground targets for the subsequent classification.

Some machine learning algorithms (e.g., SVM and RF) based on OBIA have been extensively used to produce high accuracy land cover products (Mahdianpari et al., 2017; Mohammadianesh et al., 2018; Ghulam et al., 2014; Ma et al., 2010; Zhang et al., 2019). In addition, deep learning models have been extended for wetland mapping and drawn wide attention. Mahdianpari et al. (2018) and Rezaee et al. (2018) proposed a deep convolutional neural network approach for wetland classification. Mohammadianesh et al. (2019) used a fully convolutional neural network for semantic segmentation in a complex land cover ecosystem. The deep learning and XGBoost shallow learning for large-scale wetland classification were compared by DeLancey et al. (2019). However, all these classifiers have pros and cons, and none of them performs well for all datasets. Moreover, wetland classification is still challenging if the object-based method with a single classifier is used, because the wetland landscape is featured high heterogeneity and complicated compositions. Furthermore, the inherent characteristics of remote sensing images cannot be exactly extracted for further model learning, which partially explains why a single classifier is hard to handle the classification with multi-feature inputs. According to the “No Free Lunch” theorems (Healey et al., 2018; Wolpert and Macready, 1997), if a method performs well in one type of problems, it naturally decreases the accuracy on a set of remaining problems. Multiple classifier systems, such as bagging, boosting and stacking algorithms, which ensemble various algorithms can improve the global accuracy. Compared with the bagging- and boosting- based ensemble algorithm, the stacked generalization algorithm proposed by Wolpert (Wolpert, 1996), relies upon features from a learning dataset to facilitate the weighting of outputs on the basis of model performance, which is different from the simple combination rule like voting and averaging. It is conducive to study how to extract wetland information with high heterogeneity using the object-based stacked generalization method.

Sentinel-1 and Sentinel-2 have been applied independently in many studies, but seldom in the researches on wetland mapping (Kaplan and Avdan, 2019; Li and Bijk, 2019; Li et al., 2020). Furthermore, there are still some challenges in wetland classification when using the combination of dense Sentinel-1 and Sentinel-2 time series. On one hand, it is hard to obtain continuous Sentinel-2 optical time-series images over the wetland in tropical and subtropical areas. On the other

hand, the huge calculation load of Sentinel-1/2 time series cause difficulties in application. Additionally, many studies focused on seeking a single classifier for high accuracy wetland mapping (Dronova et al., 2011; Jahncke et al., 2018; Evans and Costa, 2013). However, they may ignore the realistic ecology situation of landscape units in heterogeneous area. Therefore, using an ensemble model based on OBIA may solve these problems.

In this research, we aim to study the wetland classification in a subtropical zone using time series Sentinel-1 / 2 data and the object-based stacked generalization method. The specific objectives of this study are: 1) to test the validity of dense Sentinel-2 MSI and Sentinel-1 data in classifying vegetation types of the subtropical wetland ecosystem using an object-based stacked generalization method; 2) to compare our classification results against that using the pixel-based stacked generalization method and that using the object-based method with single classifiers; 3) to discuss the stability of the proposed algorithm and search for better wetland classification solutions in tropical/subtropical areas. We expect that the combination of multi-temporal Sentinel-1 and Sentinel-2 images, NDVI and the derived phenology data will yield higher precise wetland classification of vegetation types than current available methods.

## 2. Study area and data

### 2.1. Study site

The study site, the Dongting lake wetland (latitude/longitude: 28°47'N - 29°35'N, 112°38'E -113°07'E), is located in the middle reaches of Yangtze River, China (Fig. 1) and covers an area of approximately  $1.33 \times 10^5$  ha (Cai et al., 2019). The Dongting Lake is a national wetland park located in the sub-tropical monsoon climate zone, with an annual average precipitation of 1650 mm and an annual average temperature of 16 °C. The area has a relatively flat terrain with the elevation ranging from 20 to 207 m (Yu et al., 2018). Poplar, sedge, reed, single season rice and double cropping rice are the most common wetland vegetation surrounding the Dongting Lake. These five

vegetation types have different growing cycles and distinct phenology in year from other vegetation (Table 1).

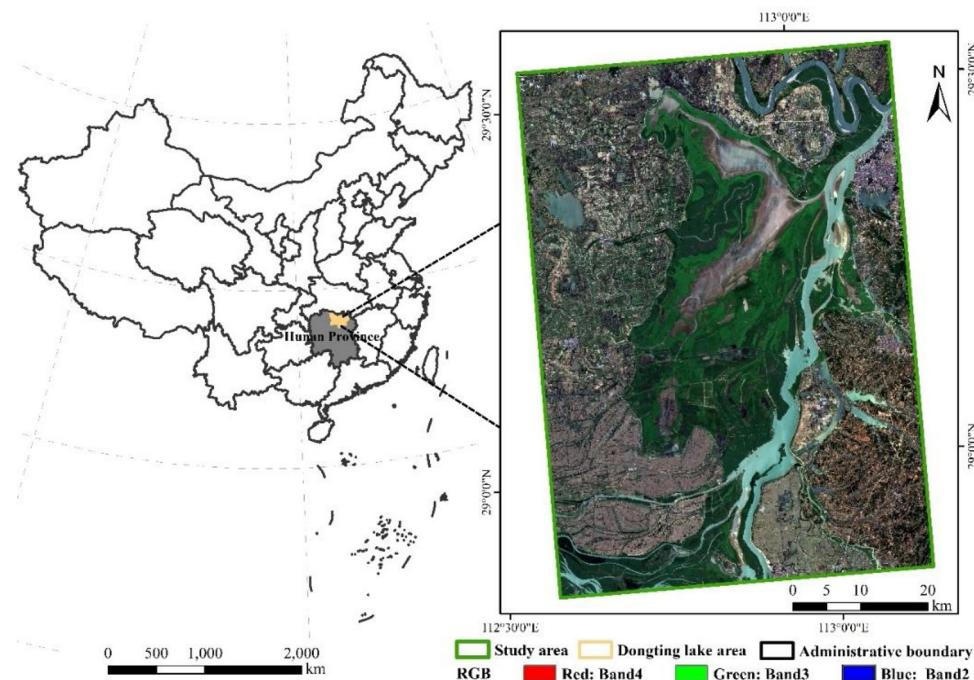
## 2.2. Data and preprocessing

### 2.2.1. Remote sensing data

A total of 30 Sentinel-1 GRD images in the Interferometric Wide swath (IW) mode (10 m resolution) with VV and VH polarizations were downloaded (Fig. 2). Sentinel-1 images have been preprocessed by the Sentinel Application Platform (SNAP, <http://step.esa.int/main/download/>) through the following steps: 1) orbit correction, 2) thermal noise removal, 3) border noise removal, 4) radiometric calibration to obtain sigma<sub>0</sub>( $\sigma^0$ ), 5) multi-looking with a  $2 \times 2$  window, 6) converting linear data to decibel (dB), 7) range Doppler terrain correction against the geometric distortions (layover, foreshortening and shadow) and 8) speckle filtering using the refined Lee filter. Finally, the backscattering coefficient in dB scale is got.

Sentinel-2 images have 13 bands with three different spatial resolutions (10 m, 20 m and 60 m) and a 5-day interval. In this research, only the 10 m (visible and near infrared bands) and 20 m (vegetation red edge and shortwave infrared bands) resolution bands were selected. All Sentinel-2 imagery (cloud percentage less than 5%) covering the whole study area in 2018 was downloaded from the European Space Agency (ESA, <https://scihub.copernicus.eu/>). The Level-1C Sentinel-2 images have been orthographically corrected and sub-pixel geometrically calibrated. We performed the radiometric calibration and atmospheric correction in the Sen2cor v2.8 plug-in (<http://step.esa.int/main/third-party-plugins-2/sen2cor/>). Both Sentinel-1 and Sentinel-2 images were resampled to 10 m resolution and re-projected onto the WGS-84 UTM coordinates.

23 MOD13Q1 time-series NDVIs (16-day interval and 250 m resolution) with sinusoidal projection obtained from the USGS website (<https://www.usgs.gov>) was utilized to generate the fused Sentinel-2 NDVI time series by the RASTFM algorithm. The MOD13Q1 NDVI images were firstly projected into the same projection of Sentinel image and resampled to 10 m resolution. Subsequently, the Sentinel-2 image

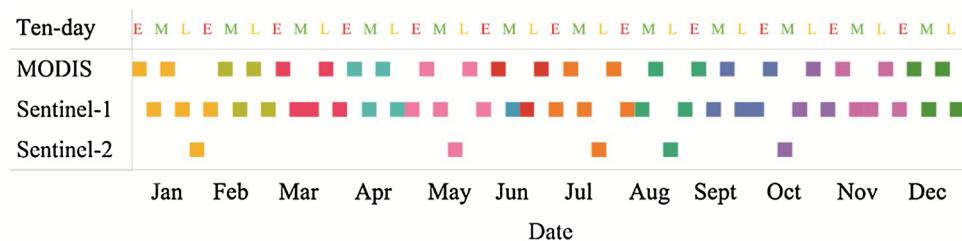


**Fig. 1.** Location of the study area (true color composite of Sentinel-2 multispectral image on 23 May 2018).

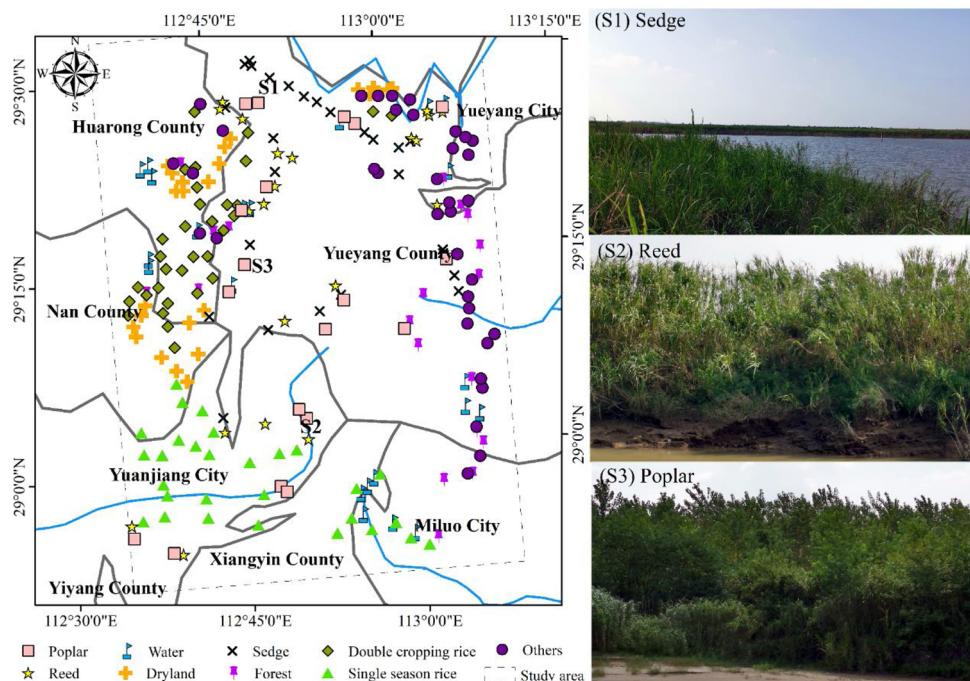
**Table 1**  
Phenology stages of the major plants in the Dongting lake.

Month	March			April			May			June			July			August			September			October			
	E	M	L	E	M	L	E	M	L	E	M	L	E	M	L	E	M	L	E	M	L	E	M	L	
Ten-day																									
Sedge	1			2			3			4			5			3			4			6			4
Reed				1																					
Poplar				1			2																		
Evergreen forest	1			2			1			2			3			4			5			6			4
Single season rice					1		2			3			4						7			8			10
Double cropping rice (early)			1				2						1						3						
Double cropping rice (late)																			2						3

\*“E”: the early ten-day period of a month; “M”: the middle ten-day period of a month; “L”: the last ten-day period of a month; Sedge: 1- Sprouting, 2- Tillering, “3- Flowering, 4- Ripening stage, 5- Growing, 6- Fade; Reed: 1-Germinating, 2-Leafing, 3-Flowering, 4-Fade; Poplar: 1-Sprouting, 2-Leafing, 3-Growing, 4-Defoliating; Evergreen forest: 1-Sprouting, 2-Leafing, 3-Growing, 4-Fade; Single season rice: 1-Sowing, 2-Seeding/flooding, 3-Transplanting/flooding, 4-Reviving, 5-Tillering, 6-Booting, 7-Heading, 8-Milky maturity, 9-Mature, 10-Harversting; Double cropping rice: 1-Sowing to transplantation, 2-Panicle initiation to flowering, 3-Ripening stage.



**Fig. 2.** Acquisition dates of the remote sensing data used in this study. “E” is the early ten-day period of a month, “M” is the middle ten-day period of a month and “L” is the last ten-day period of a month.



**Fig. 3.** Field survey sites. Left: the distribution of the field sites; s1-s3: photos of sedge, reed and poplar, respectively, in the field survey sites.

was taken as the base image to register the MOD13Q1 NDVI images. The Savizky-Golay (S-G) filter was applied to reconstruct the MODIS NDVI time series due to the disturbances of cloudy weather and speckles (Jönsson and Eklundh, 2004; Cai et al., 2019).

#### 2.2.2. Validation data

We did field survey following the principle of stratified sampling between April 1 st and July 31 st, 2018, when wetland vegetation types were in growing stage. The locations of field sites were recorded by

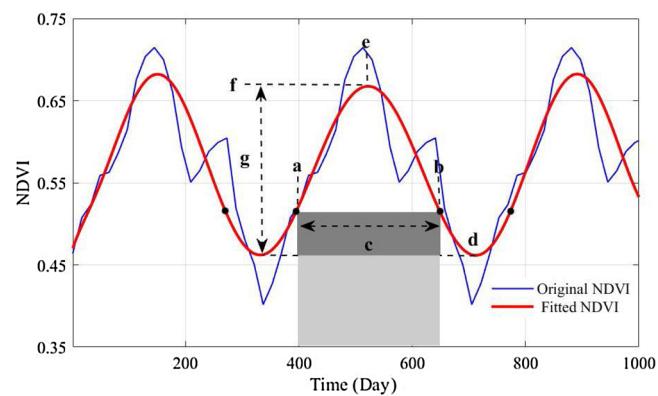
GPS. Each field site is larger than 400 m<sup>2</sup> (some larger than 1 ha) and has relatively “pure” cover types. There are 230 field sites with double cropping rice (32 sites), single season rice (30 sites), sedge (25 sites), reed (22 sites), poplar (20 sites), forest (19 sites), water (24 sites), dryland (22 sites) and others (e.g., built up, bare land) (36 sites) (Fig. 3). These sites were used to assess and validate the classification result.

**Table 2**  
Training sample information for wetland classification

Land cover types	Numbers of training samples (pixels)	Number of testing samples (pixels)	Total
Double cropping rice	8248	2693	10941
Single season rice	6725	1982	8707
Sedge	5573	1483	7056
Reed	7690	2507	10197
Poplar	3871	1025	4896
Forest	12752	4119	16871
Water	8936	2762	11698
Dryland	7364	2390	9754
Others	10374	3417	13791
Total	71533	22378	93911

### 2.2.3. Training and testing dataset

Google Earth images and the 1:10,000 land use/land cover (LULC) map of Hunan province in 2018 provided by the Hunan Provincial Department of Natural Resources were used to assist in identifying the cover types and collecting samples for training and testing models. The Hunan LULC map was generated from high-resolution aerial photography and satellite images (GF-1, GF-2 and ZY-3) by the man-machine interactive classification method. In the LULC map, the land cover types are categorized into 6 primary classes (cropland, forestland, grassland, water body, constructive land and unused land) and 25 sub-classes. The overall accuracy of the LULC product is higher than 95%, which has been validated by field survey (Cai et al., 2019; Zhang et al., 2018b). Each wetland cover type was delineated as polygons with ArcMap 10.2 software using multi-temporal Google Earth images, Hunan LULC map and the field data together. These polygons were divided into training

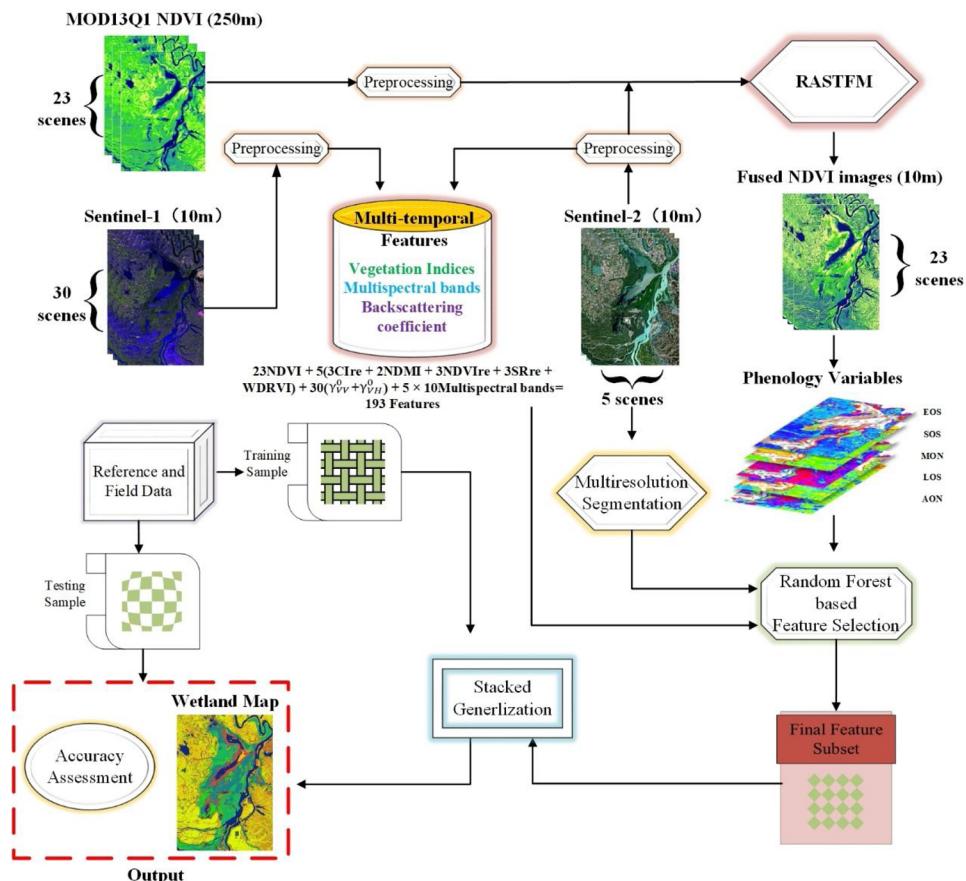


**Fig. 5.** Phenological parameters derived by the dynamic threshold method based on nonlinear least squares fits, which are the start of the growing season (point a), the end of the growing season (point b), the length of growing season (c), the base value of NDVI (d), the middle of the growing season (e), the maximum value of the NDVI (f) and the amplitude of the NDVI (g).

and testing samples at a ratio of 3:7. The information about the training and testing samples is shown in Table 2.

### 3. Method

In this study, we developed a method for wetland mapping using the object-based stacked generalization method based on multi-temporal and multi-source remote sensing images (Fig. 4). The proposed method mainly consists of three steps: 1) data collection, including the Sentinel-



**Fig. 4.** The flowchart of the proposed method for wetland classification.

**Table 3**  
Vegetation index used in this research.

Vegetation index/SAR backscattering coefficient	Formula	Source
Normalized Difference Moisture Index (NDMI)	$(\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR})$	(Gao, 1996)
Red-edge Chlorophyll Index ( $\text{Cl}_{re}$ )	$\text{NIR} / \text{Red-edge} - 1$	(Gitelson et al., 2005)
Normalized Difference Vegetation Index (NDVI)	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	(Rouse et al., 1974)
Red-edge Normalized Difference Vegetation Index ( $\text{NDVI}_{re}$ )	$(\text{NIR} - \text{Red-edge}) / (\text{NIR} + \text{Red-edge})$	(Gitelson and Merzlyak, 1997)
Modified red-edge Simple Ratio (MSR <sub>re</sub> )	$(\text{NIR} / \text{red-edge} - 1) / \text{sqrt}(\text{NIR} / \text{Red-edge} + 1)$	(Wu et al., 2008)
Wide Dynamic Range Vegetation Index (WDRVI)	$(\alpha^* \text{NIR} - \text{Red}) / (\alpha^* \text{NIR} + \text{Red}) (\alpha = 0.2)$	(Gitelson et al., 2003)
backscattering coefficient ( $\sigma^0$ )	$\sigma^0 = DN^2/A_\sigma^2$	(Periasamy, 2018)

\* Notes:  $DN$  represents Digital Number in the SAR imagery;  $A_\sigma$  is the calibration parameter.

#### Algorithm Stacked generalization:

```

Input: Training dataset  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ;
        Base learners  $\Phi_1, \Phi_2, \dots, \Phi_T$ ;
        Secondary learners  $\Phi$ ;
Process:
    for  $t = 1, 2, \dots, T$  do
         $h_t = \Phi_t(D)$ ;
    end for
     $D' = \emptyset$ ;
    for  $i = 1, 2, \dots, m$  do
        for  $t = 1, 2, \dots, T$  do
             $z_{it} = h_t(x_i)$ ;
        end for
         $D' = D' \cup ((z_{i1}, z_{i2}, \dots, z_{iT}), y_i)$ ;
    end for
     $h' = \Phi(D')$ ;
Output:  $H(x) = h'(h_1(x), h_2(x), \dots, h_m(x))$ 

```

Fig. 6. Stacked generalization algorithm.

2 NDVI time series predicted by RASTFM algorithm, the phenology data (e.g., the start, end and length of the growing season as well as the maximum and amplitude of NDVI that indicate the growth rhythm of vegetation) derived from the Sentinel-2 NDVI time series, the vegetation indices derived from Sentinel-2 multispectral images and the time series SAR backscatters ( $\sigma_{VV}^0, \sigma_{VH}^0$ ) data generated from Sentinel-1; 2) segmentation based on multi-temporal Sentinel-2 multispectral images; 3) separability analysis of land cover types and optimal feature combination (feature means the data used for wetland classification) identification at the object level, in which the Random Forest (RF) algorithm is utilized to select the optimal features for classification based on the abovementioned feature dataset and 4) mapping wetland using the stacked generalization method and optimal feature combination at the object level.

### 3.1. Sentinel-2 NDVI time series and vegetation phenology data generation

The RASTFM comprehensively considers the land cover and phenological changes in homogeneous and heterogeneous landscapes during the prediction period, so its results are more robust and precise than other spatio-temporal fusion models (Zhao et al., 2018). We employed the RASTFM to obtain Sentinel-2 like NDVI time series from the MOD13Q1 time-series NDVI and observed Sentinel-2 NDVI. The RASTFM contains 5 steps: 1) relative radiometric normalization and de-normalization; 2) Non-shape change prediction; 3) shape change detection; 4) shape change prediction; 5) regression based high-pass modulation. More details about RASTFM can be found in (Zhao et al., 2018).

The vegetation in Dongting Lake wetland ripens one or two times per year. Different vegetation shows distinct phenological phenomena in the same period (Table 1), which can be adopted to discriminate effectively vegetation types in wetland area (Zhang et al., 2018b; Kang et al., 2014). Particularly, some phenological parameters such as SOS (Start of the growing season), EOS (End of the growing season), LOS (Length of the growing season), MON (Maximum of NDVI) and AON (Amplitude of NDVI) have proved to be useful in vegetation identification (Cai et al., 2019; Zhang et al., 2018b; Kang et al., 2014). For considering the growth characteristics of vegetation types, these five vegetation phenological parameters were derived by the dynamic threshold method based on the nonlinear least squares fits of asymmetric Gaussian model functions (Fig. 5) (Jönsson and Eklundh, 2004).

### 3.2. Seasonal spectral characteristics, vegetation indices and SAR backscatters of different cover types

In the wetland ecosystem, there are different surface components with similar spectra and the same surface component with different spectra, which impacts the wetland mapping. Combining multi-

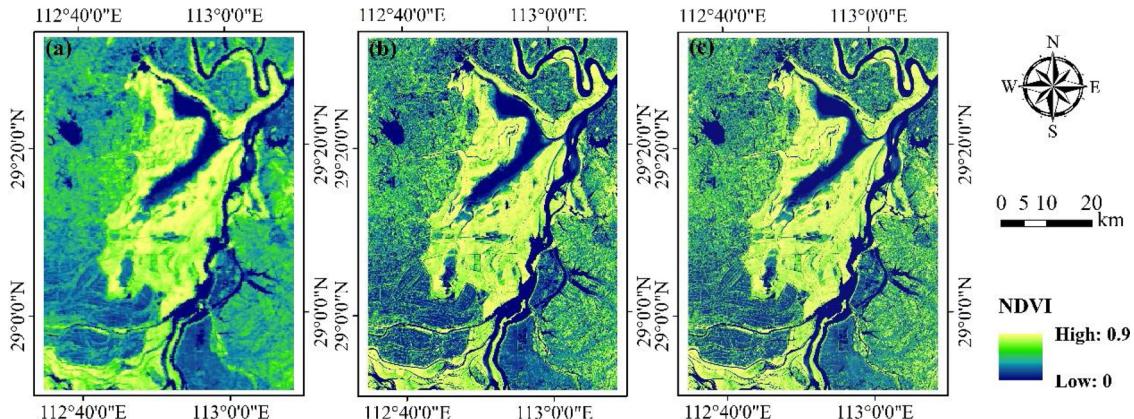
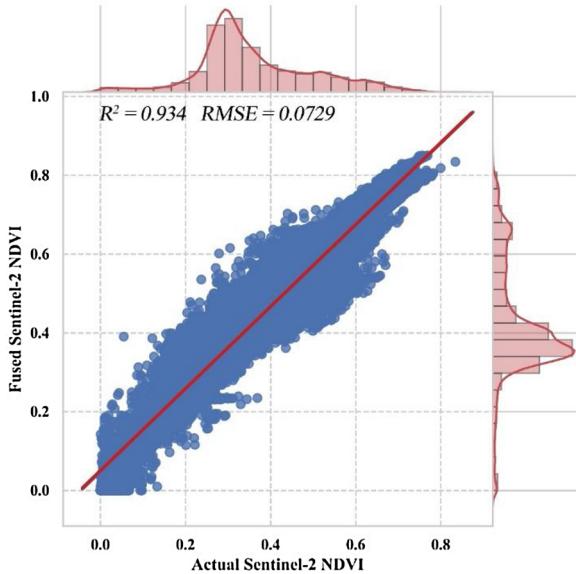


Fig. 7. (a) The MOD13Q1 NDVI, (b) the actual Sentinel-2 NDVI and (c) the fused Sentinel-2 NDVI on DOY 273, DOY 278 and DOY 273, respectively.

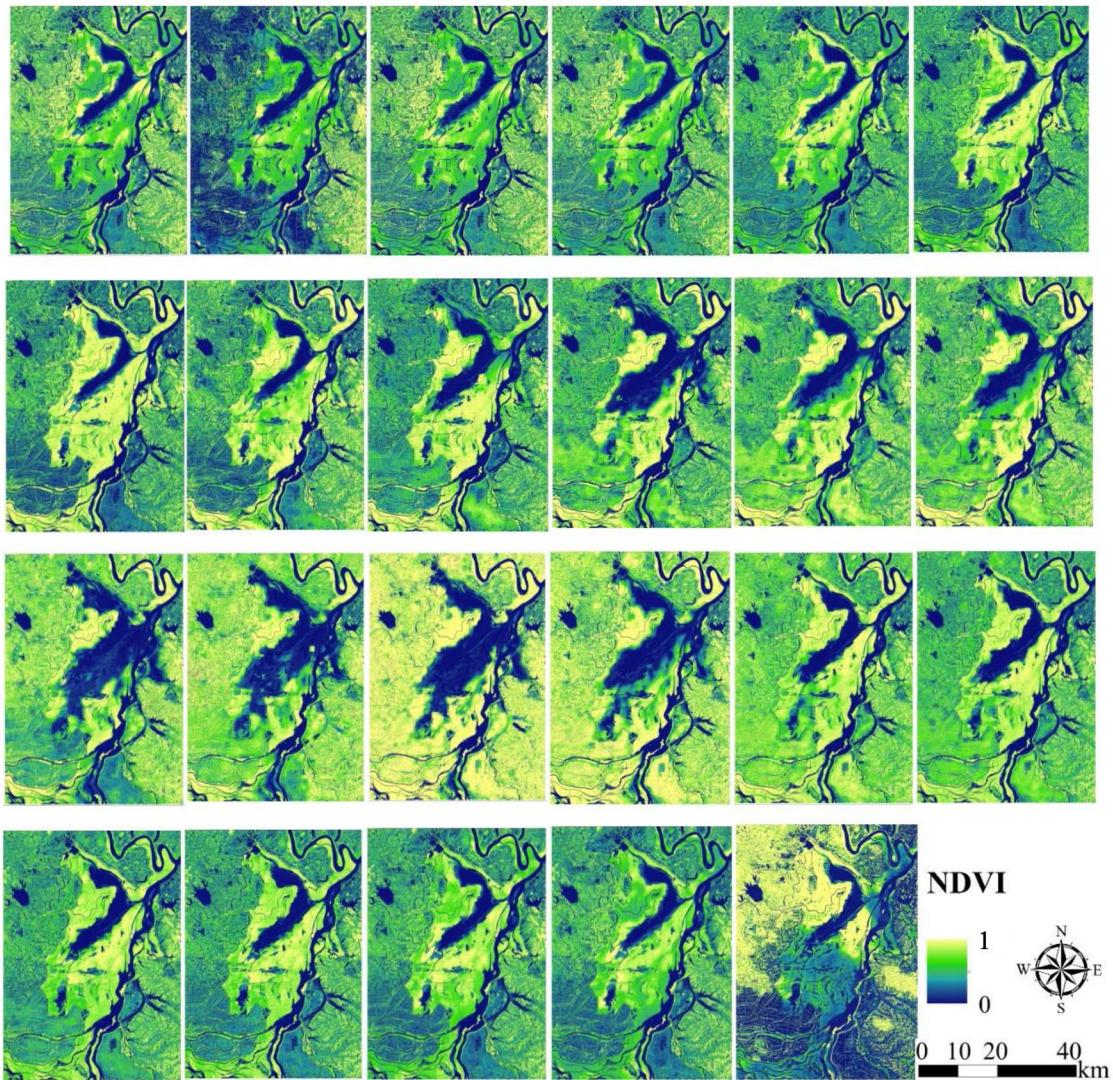


**Fig. 8.** The correlation between the actual Sentinel-2 NDVI and the fused Sentinel-2 NDVI on 2018-07-27.

temporal optical and SAR images can solve this problem, because this combination is sensitive to different seasonal characteristics, such as pigments, water, carbon, nitrogen, shape, structure and roughness. Therefore, the multi-temporal spectral characteristics, vegetation indices (VIs) and SAR backscattering coefficient were used for separability analysis. According to the width of the Sentinel-2 spectral band, we chose several VIs to compare wetland cover types. Some VIs are related to pigments, water and light use efficiency of vegetation. Additionally, due to the significant role of red-edge band in vegetation monitoring and the effectiveness of vegetation classification, we compared different red-edge vegetation indices in five dates, and finally selected 6 vegetation indices, whose definition and calculation are shown in Table 3 (Zhang et al., 2018b).

### 3.3. Image segmentation

According to Mahdavi et al. (2019), it is easier to segment the optical data than the SAR image and the combination of these two kinds of images, because of the speckle noise in SAR data. Therefore, we only applied segmentation on Sentinel-2 multispectral images in this study. The MSR algorithm implemented in eCognition Developer 9.0 software was employed to create the image objects using Sentinel-2 multispectral images. The MSR algorithm uses heterogeneity threshold to generate the primitive objects with different sizes and enhance the object



**Fig. 9.** The fused Sentinel-2 NDVI time series from DOY 001 to DOY 353.

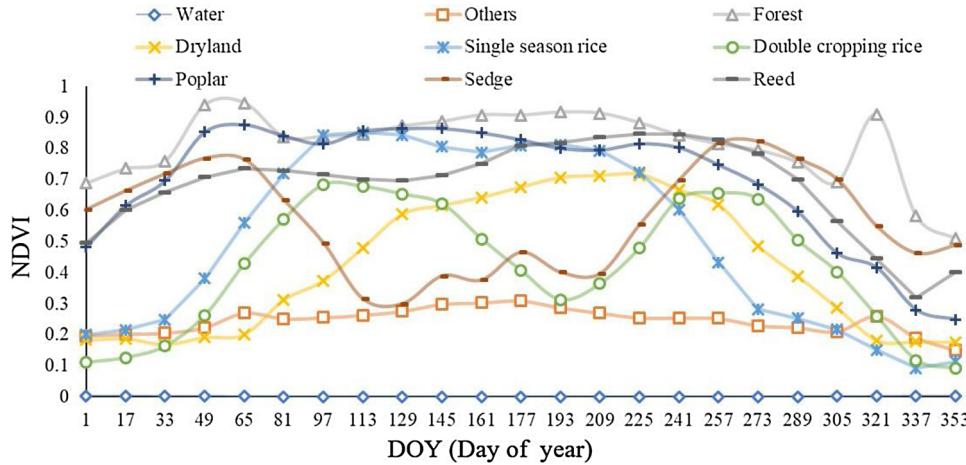


Fig. 10. The time-series Sentinel-2 NDVI of wetland cover types.

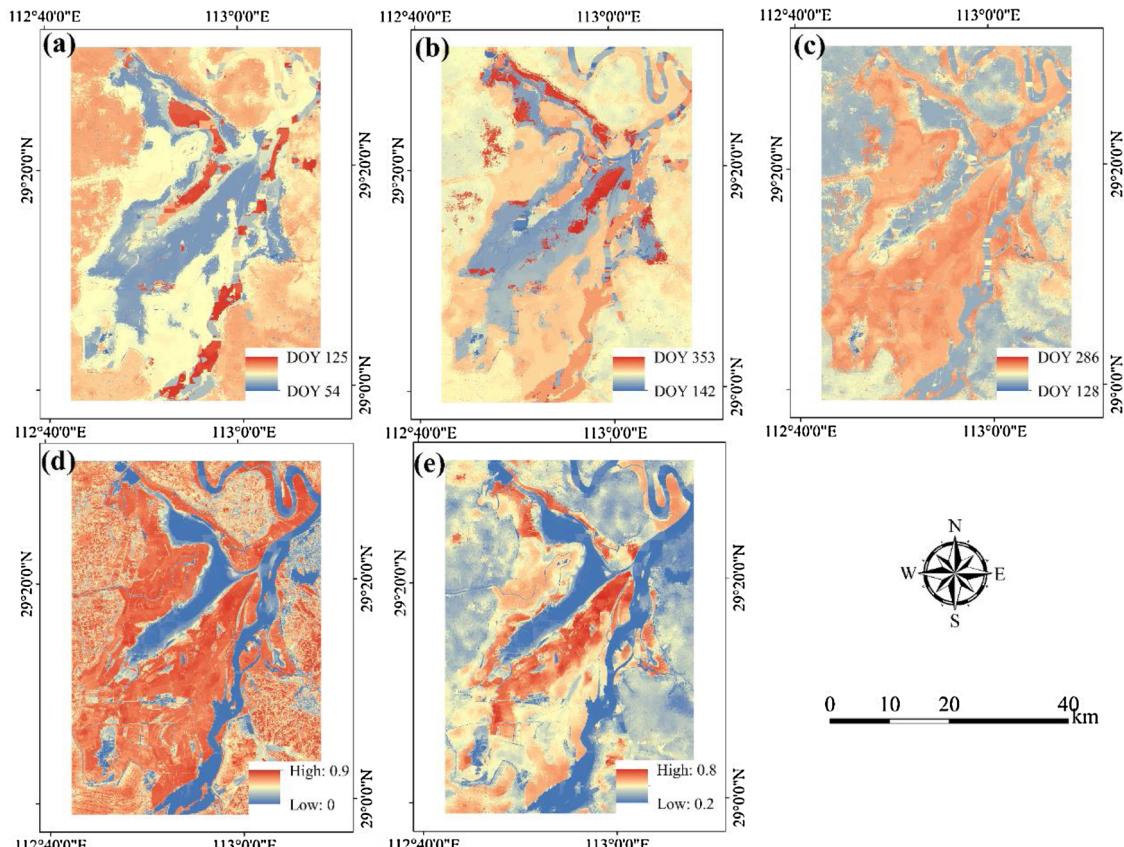


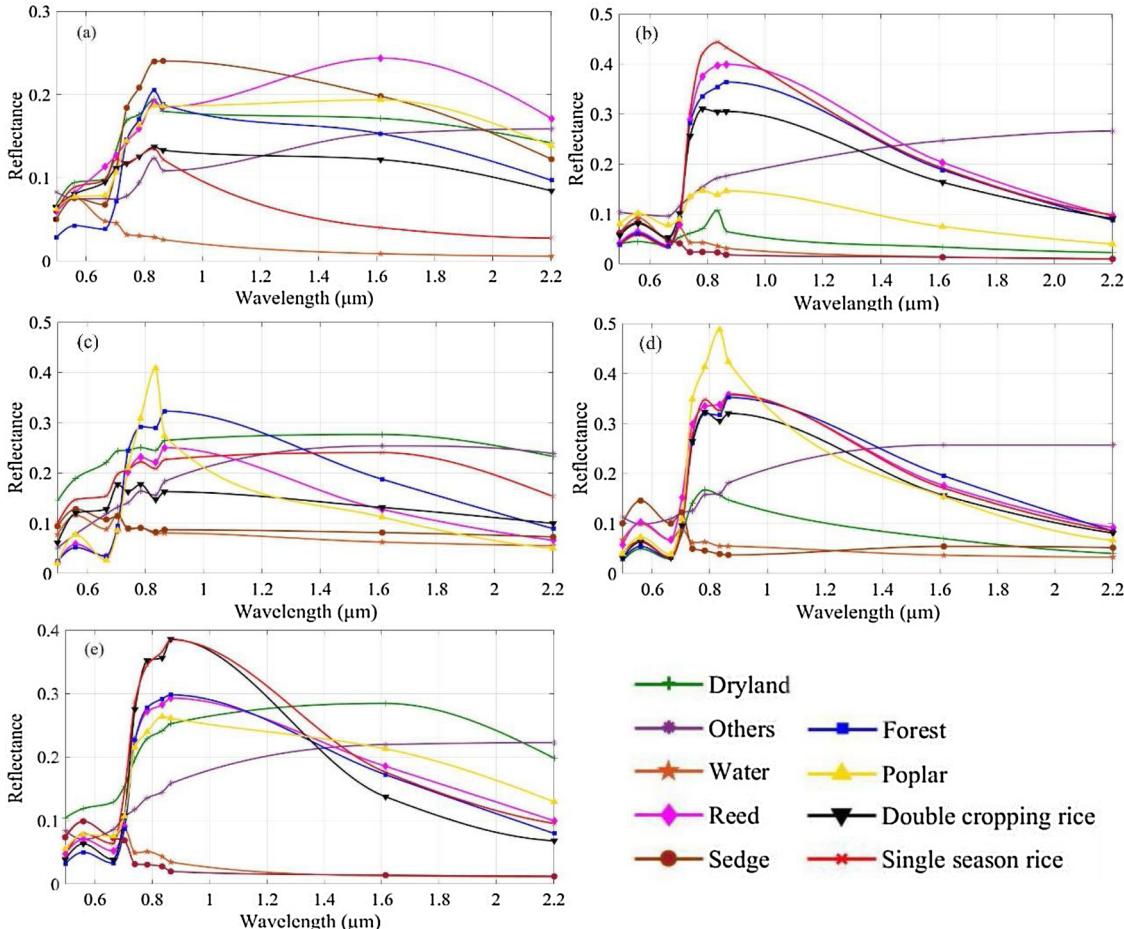
Fig. 11. Five phenological parameters generated from the fused Sentinel-2 NDVI time series. (a) SOS (Start of the growing season), (b) EOS (End of the growing season), (c) LOS (Length of the growing season), (d) MON (Maximum of NDVI) and (e) AON (Amplitude of NDVI). Color in (a)-(c) depicts the time (unit: day of year (DOY)) corresponding to phenology parameters, but represents the NDVI value in (d) and (e).

generation response to the landscape patch. Attributes of the generated image objects are determined by three parameters: segmentation scale, shape and compactness. Segmentation scale determines the size of objects, the compactness contributes to the fragmentation of landscape patches. In the segmentation, the dependence of image spectrum will decrease with the increase of shapes, resulting in more objects. In this study, the optimal segmentation scale (15 pixels), shape (0.2) and compactness (0.5) were determined by the ESP plug-in (Drăguț et al., 2010) of eCognition Developer 9.0. The ESP tool uses the rate of changes of LV (ROC-LV) to represent the standard deviation in the

segmentation result, and the local variance to represent the homogeneity within the segmentation object. When ROC reaches the peak, the corresponding segmentation scale is regarded as the optimal segmentation scale.

$$ROC - LV = \left[ \frac{L_i - L_{i-1}}{L_{i-1}} \right] \times 100\% \quad (1)$$

Where  $L_i$  and  $L_{i-1}$  represent the  $i$ -th and  $i-1$ -th segmented target layer, respectively.



**Fig. 12.** The spectral curves of Sentinel-2 on (a) 2018-01-23; (b) 2018-05-23; (c) 2018-07-27; (d) 2018-08-16; (e) 2018-10-05.

After the MSR segmentation, the image objects were yielded and output as a vector layer in eCognition Developer 9.0 software. Subsequently, the associated information of the MSR-generated image objects, including the spectral characteristics, phenology, vegetation indices and SAR backscatters, were extracted separately used the object vector after determining optimal feature combination based on the RF algorithm that was introduced in Section 3.4. Finally, the feature set of each object was taken as an input for stacked generalization (specifically, the base learners of stacked generalization).

#### 3.4. Optimal feature combination selection based on random forest

The features for classification in this research include the Sentinel-2 MSI multispectral images, VI, phenological parameters and Sentinel-1 backscattering coefficients. The mean spectral reflectance value of each object was first calculated during the segmentation. Then the segmentation result was overlaid on the SAR, VIs, phenological parameters to generate the average feature within object. Feature optimization transformed the selected features into a low-dimensional feature space to improve the classification accuracy and reduce data redundancy and computation load. Therefore, we used the RF algorithm with 500 decision trees based on a contribution metrics Gini (GI) to assign the variable importance measure (VIM) for each variable (Hariharan et al., 2018).

#### 3.5. Wetland classification by object-based stacked generalization algorithm

The stacked generalization fuses the base learning result utilizing “Meta Learner” (ML). In the stacked generalization, there are some individual learners called “base learner (BL)”, which are employed to extract the primary feature from the base training dataset and output secondary training dataset for the secondary learner (i.e., ML) (Peyret et al., 2018). In general, the BL is expected to be diverse and good enough to achieve a satisfactory prediction performance. By the stacked generalization method, the BLs are firstly trained using the training dataset. Then, the secondary training dataset is generated by BLs. Finally, the secondary learner is trained using the base prediction map that is then employed to produce the final result (Fig. 6). However, this algorithm usually leads to over-fitting. To overcome this shortcoming, we modified the fusion rule. Specifically, we separated the base training dataset into  $k$  parts, and used  $k$ -fold cross validation to train each individual BL. After training the BLs, we obtained a new feature by combining  $k$  prediction results. If there are  $T$  BLs in the stacked system, the training dataset of the secondary learner consists of  $T$  new features.

The more diverse the individual learners are, the higher classification accuracy will be achieved (Healey et al., 2018). Therefore, we chose both regression and classification models (e.g. RF, SVM, KNN, Logistic Regression and Naive Bayes model) as BLs for acquiring diverse secondary training datasets. Although the stacked generalization usually uses a regression model (e.g. logistic regression) as the secondary learner, using an ensemble learner like RF as the ML to combine BLs will improve the prediction performance (Healey et al., 2018). RF is

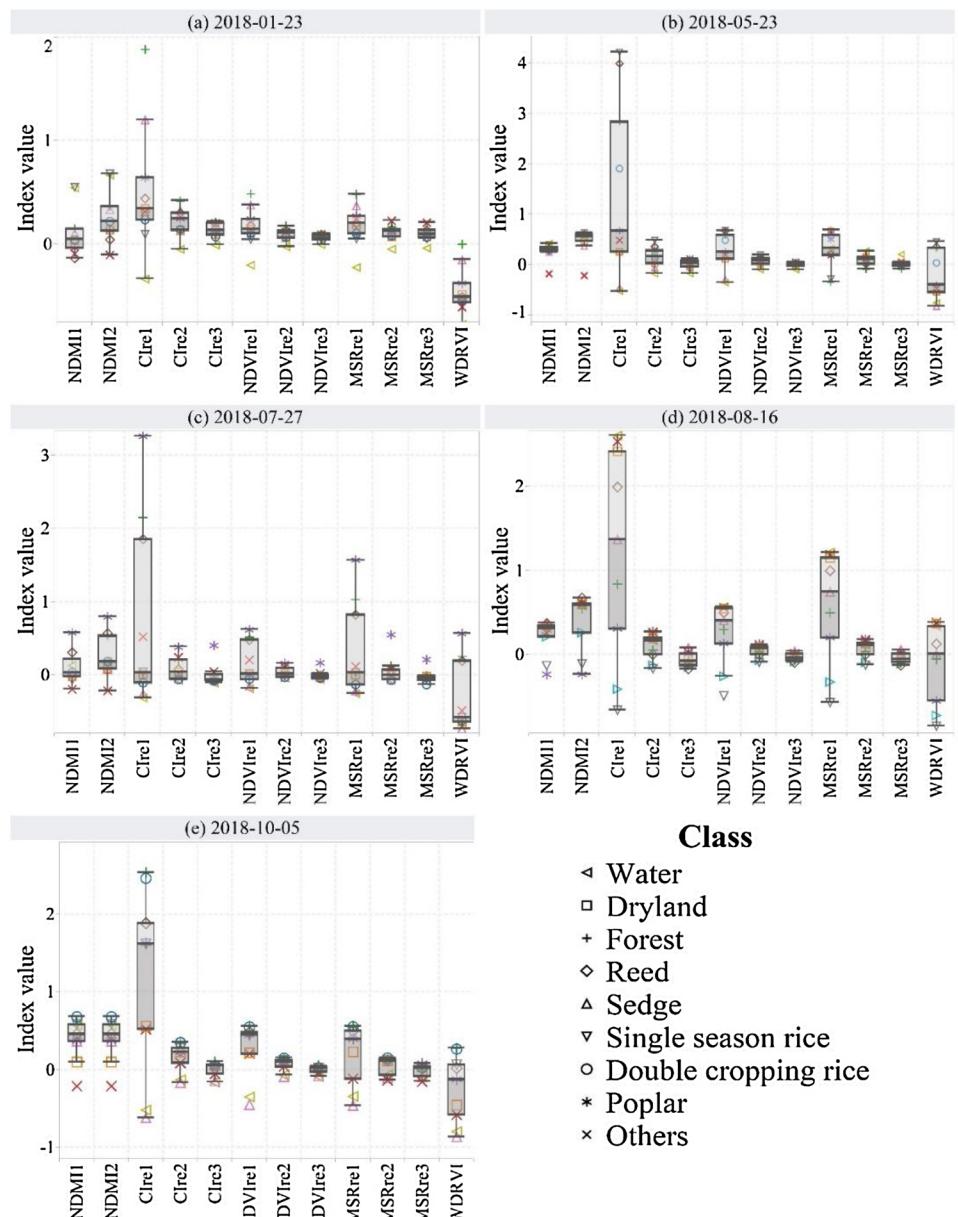


Fig. 13. Vegetation indices of each cover type at different growth stages: (a) 2018-01-23; (b) 2018-05-23; (c) 2018-07-27; (d) 2018-08-16 and (e) 2018-10-05.

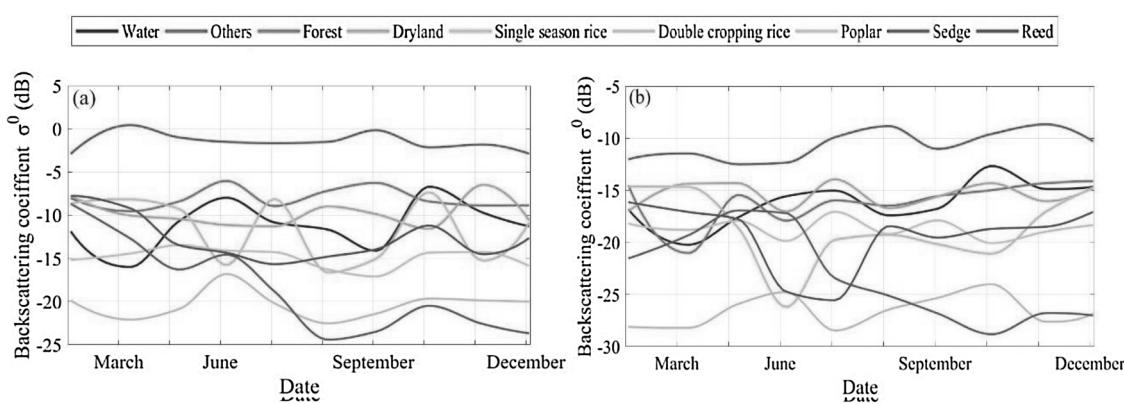


Fig. 14. The multitemporal SAR backscattering coefficients of all cover types with (a) VH polarization and (b) VV polarization.

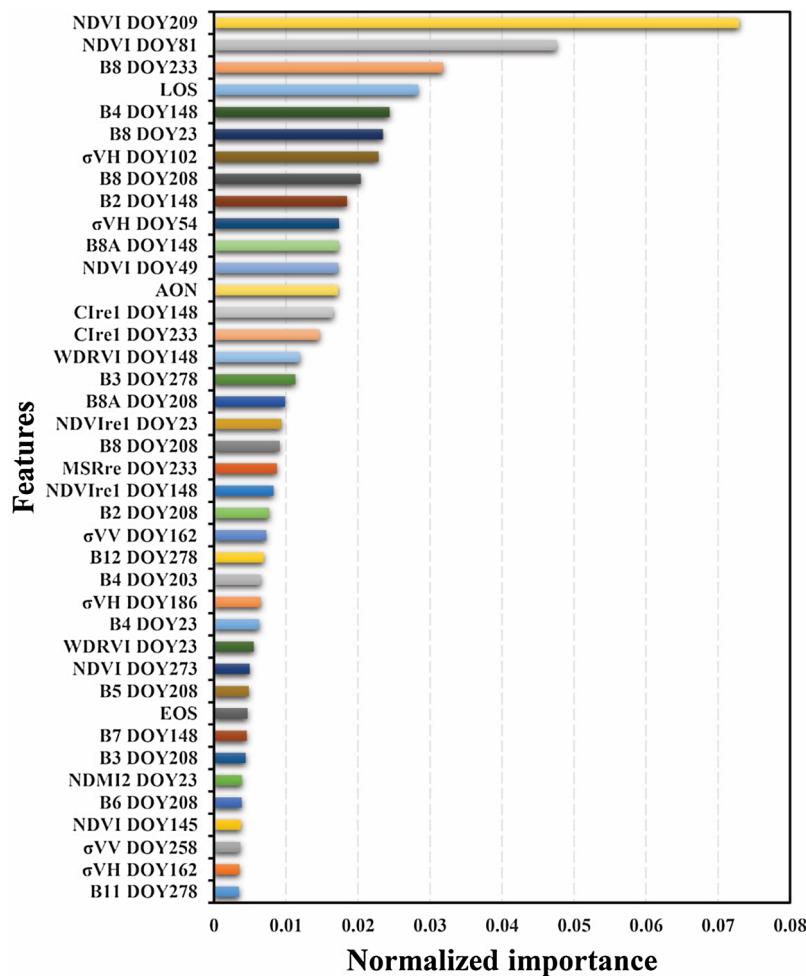


Fig. 15. Feature importance ranked by the RF algorithm.

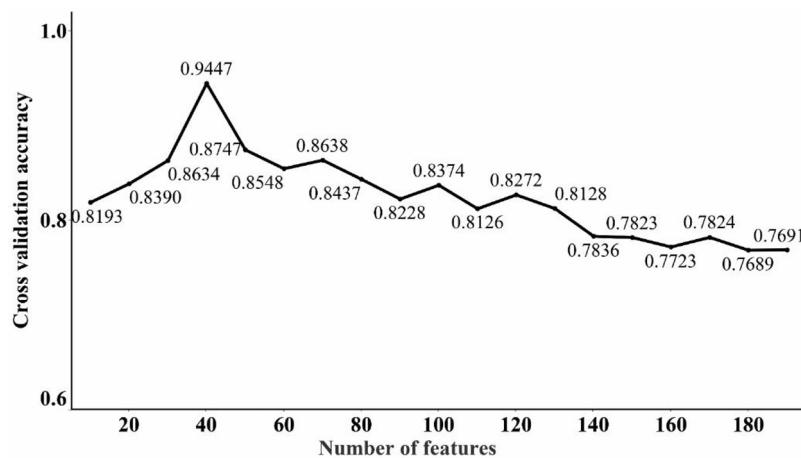


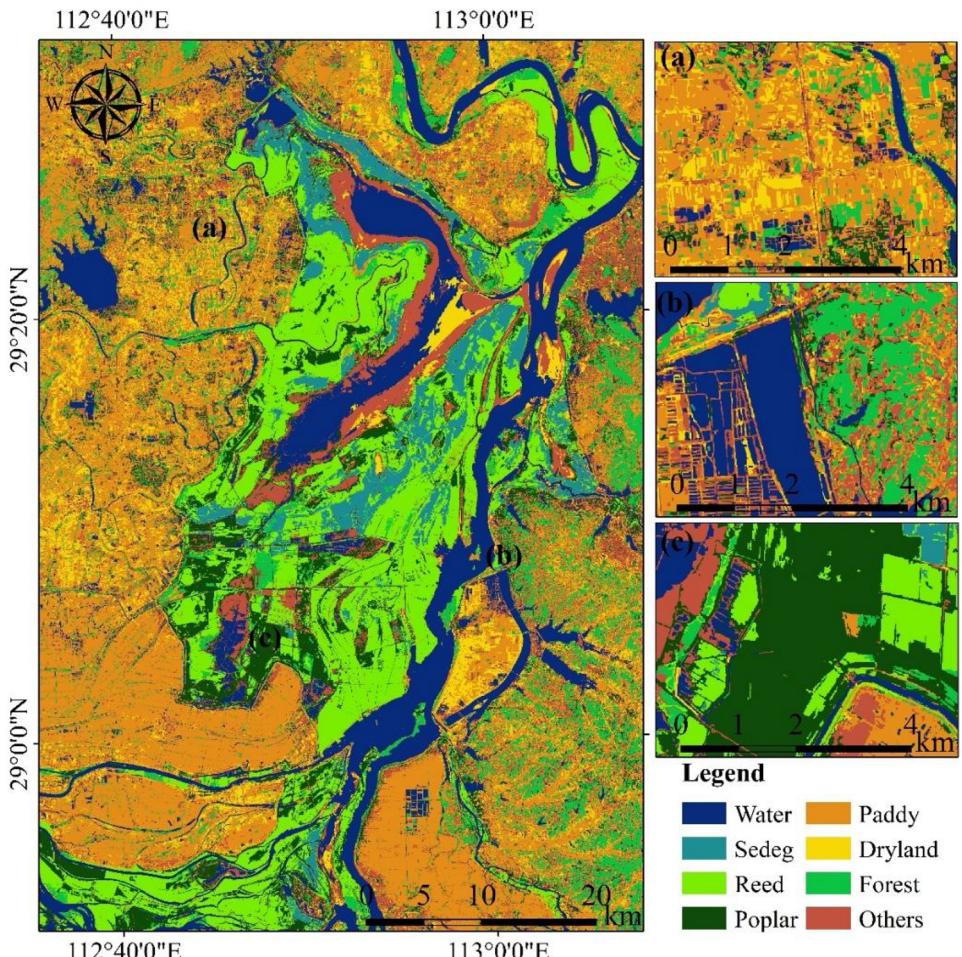
Fig. 16. Optimal feature numbers for combination.

an independent ensemble algorithm that uses the “bagging” method to integrate decision trees (Collins et al., 2018). The ensemble model applies the RF fusion rule, which has a three-layer structure consisting of the base learners, decision trees and bagging algorithm. Thus, the RF model was selected as the secondary learner to fuse BLs, which could improve the multiple classifier systems performance. The optimal hyperparameters of both BLs and secondary models were determined by

the random search approach (Jeżowski et al., 2005). The Scikit-Learn package in Python 3.7 was used for developing the object-based stacked generalization algorithm.

### 3.6. Classification accuracy assessments

According to the criteria of wetland classification in wetland



**Fig. 17.** Classification results of the object-based stacked generalization method.

convention and the current situation of LULC in Dongting Lake wetland, we classified the study area into nine land cover types: water, sedge, reed, poplar, single season rice, double cropping rice, dryland, forest and others (Cai et al., 2019; Dronova et al., 2011; Han et al., 2015). The testing samples of the object-based stacked generalization method were produced by combining the reference and field data based on multi-resolution segmentation. Firstly, the objects of the segmented image were considered to be “pure”, and the testing samples were generated by assigning the label of validation sample points (by GPS, visual interpretation and LULC map) to the corresponding segmented objects. The accuracies of both the pixel-based and object-based classification methods were validated using the testing samples by the 5-fold cross validation using a confusion matrix. The accuracies were assessed in terms of the overall accuracy (OA), producer accuracy (PA) and user accuracy (UA) as well as Kappa coefficient ( $\kappa$ ).

#### 4. Results and discussion

##### 4.1. Sentinel-2 NDVI time series and phenological parameters

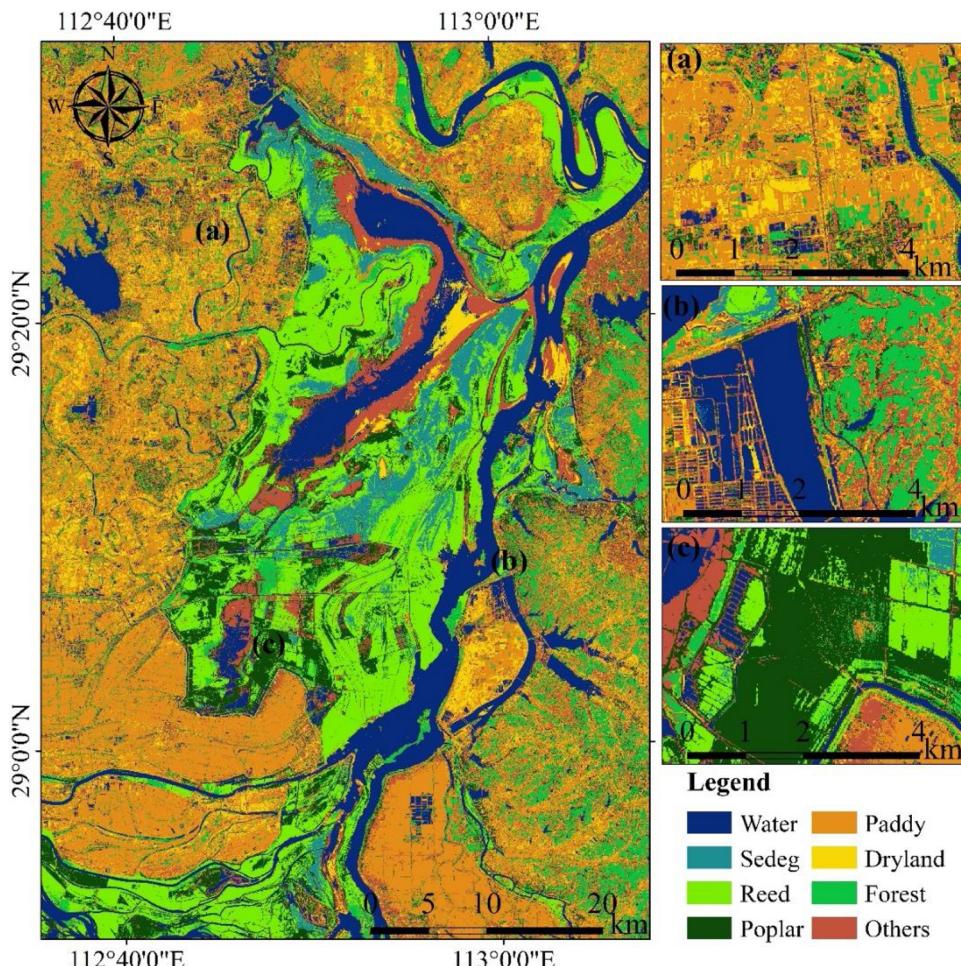
The MOD13Q1 NDVI, actual Sentinel-2 NDVI and the fused Sentinel-2 NDVI of the study area on DOY 273 are shown in Fig. 7. In general, the fused NDVI image has a high coherence in the spatial dimension with the actual image of the whole study area except the southwest part. Furthermore, the determination coefficient ( $R^2$ ) and mean square error (RMSE) between the fused Sentinel-2 NDVI

(DOY273) and actual Sentinel-2 NDVI (DOY278) were calculated for all pixels in the whole study region. As the scatterplots (Fig. 8) show, most points fit the 1:1 diagonal well, so the fused Sentinel-2 NDVI has achieved a high accuracy ( $R^2 = 0.934$  and RMSE = 0.0729). Therefore, the RASTFM can provide a credible and high spatial resolution image that considers the surface changes occurred between the acquisition date of input and output images. Generally, the fused Sentinel-2 NDVI time series are reliable to identify different wetland vegetation types (Fig. 9).

The temporal NDVI trends of nine wetland cover types in the Dongting lake area were extracted from the fused Sentinel-2 NDVI time series (Fig. 10). The phonological differences among these cover types help the vegetation classification. From the fused Setinel-2 NDVI time series, five phenological parameters including SOS, EOS, LOS, AON and MON were derived. As Fig. 11 shows, different vegetation types have different phenological parameter values. The phenological parameters were extracted according to the vegetation growth rythm, which probably benefits the identification of cover type.

##### 4.2. Multi-temporal spectrum, vegetation indices and time series backscattering coefficient

Five actual Sentinel-2 MSI images acquired in different growing stages were used to analyze the spectrum difference of nine vegetation types in the Dongting Lake wetland to determine the optimal features for classification (Fig. 12). The greatest difference of spectrum



**Fig. 18.** Classification results of the pixel-based stacked generalization method.

**Table 4**  
Classification accuracy of the pixel-based and object-based stacked generalization methods.

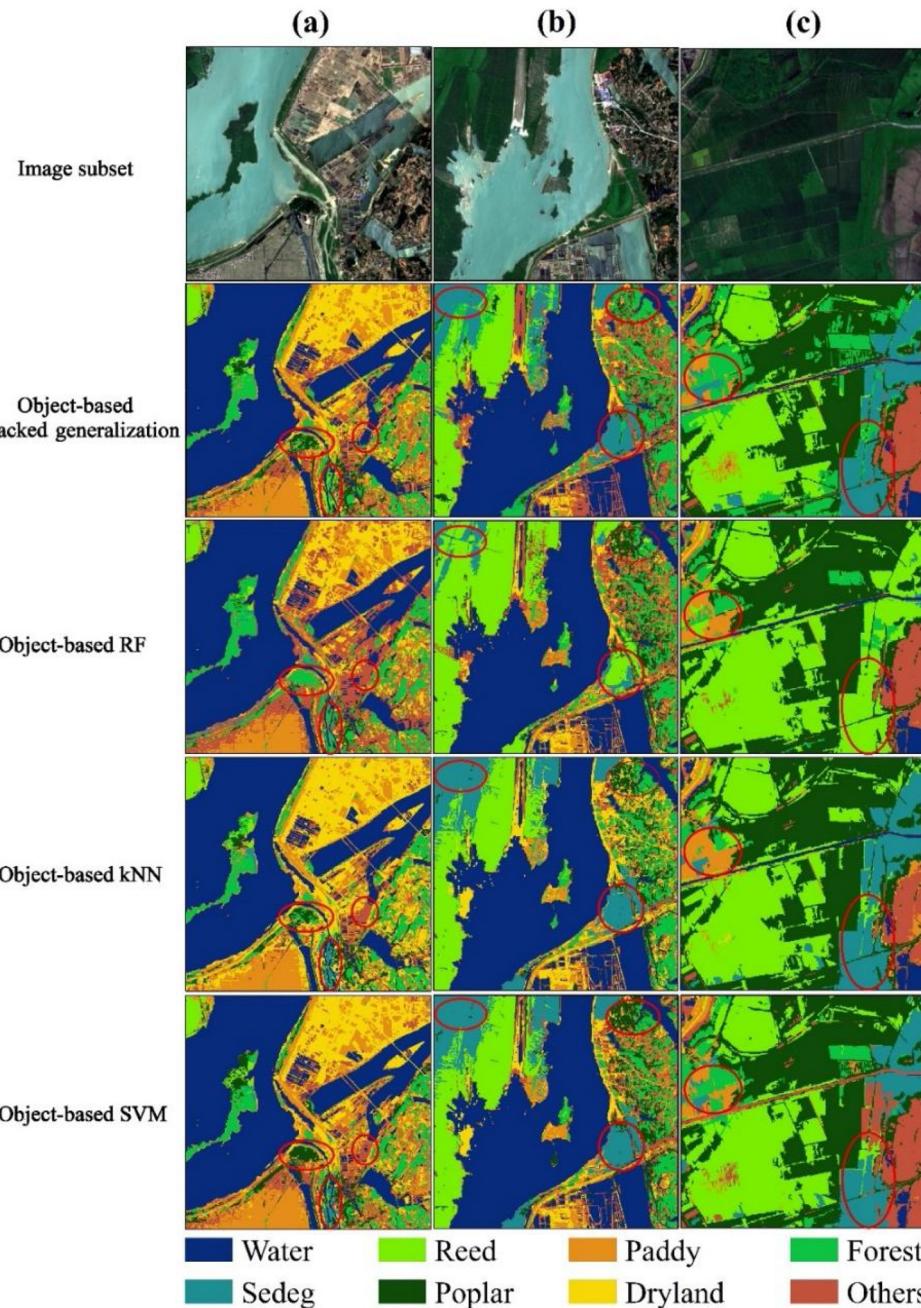
Land cover	Pixel-based		Object-based	
	PA (%)	UA (%)	PA (%)	UA (%)
Water	93.48	95.15	97.64	98.29
Sedge	84.06	84.63	86.26	87.66
Reed	84.78	85.72	88.10	90.37
Poplar	87.02	88.77	90.68	91.86
Paddy	89.37	89.96	92.16	94.43
Dryland	89.06	90.88	94.87	95.89
Forest	90.71	91.41	92.87	94.17
Others	89.42	90.13	90.04	90.76
Overall accuracy	88.58%		92.46%	
Kappa coefficient	0.88		0.92	

**Table 5**  
Classification accuracy of different feature combinations.

Feature combination	PA (%)	UA (%)	OA (%)	Kappa coefficient ( $\kappa$ )
Sentinel-2 MSI	81.37	82.12	81.69	0.81
Sentinel-2 MSI + Sentinel-1	86.94	89.13	88.54	0.86
Sentinel-2 MSI + Sentinel-1 + NDVI	89.87	91.64	91.02	0.90
Sentinel-2 MSI + Sentinel-1 + NDVI + phenology	91.58	92.93	92.46	0.92

characteristics appears on 23 May 2018 (Fig. 12b), when the spectral curves are also separated well. The separability is mainly contributed by the near infrared (NIR) and red-edge bands. In general, the peak reflectance values of the double cropping rice and single season rice appear in the tillage period, and the former is lower. These may help to identify them. Poplar has the reflectance in NIR and red-edge bands significantly different from other cover types in July and August, which may be the best time to distinguish them. Additionally, other cover types have relatively stable spectrum changes. In a word, the NIR and red-edge provide the potential and useful information for wetland classification, and the time-series spectral characteristics of the cover types are useful to distinguish the wetland vegetation.

Besides NDVI, we tested five VIs derived from the Sentinel-2 multispectral images, including NDMI, Clre, NDVIre, MSRre and WDRVI, for wetland classification. Subscripts were added to discriminate the VI origins. For example, the Sentinel-2 image has two Short-wave Infrared (SWIR) bands, band 11 (1610 nm) and band 12 (2190 nm), so NDMI<sub>1</sub> is the NDMI produced from band 11 and NDMI<sub>2</sub> is that calculated from band 12. Fig. 13 shows that the multi-temporal VIs of each vegetation type can distinguish different plants. Each VI has different separability for cover types at different growing stages. For example, MSRre<sub>1</sub> has the strongest separability on 2018-07-27, but the lowest separability on 2018-01-23. For the whole observation period, Clre<sub>1</sub> may be the most reliable index for classification, due to its wide value range. The VIs produced by the red-edge band show better separability than other VIs. NDMI and WDRVI without using red-edge spectrum also have the potential to identify different vegetation types, but their performances are influenced by the growing stage of wetland vegetation.



**Fig. 19.** The classification results of three study typical regions (a, b and c) in the study area using support vector machine (SVM), random forest (RF) and k-Nearest Neighbor (kNN) as well as stacked generalization method based on OBIA.

The curves of the multi-temporal SAR backscattering coefficients in the whole year are shown in Fig. 14. Compared with the curves of multispectral bands, the backscattering coefficient curves are irregular, because the backscattering is sensitive to the surface moisture and the structure of surface components. Furthermore, the backscattering coefficient curves of these vegetations have different and complicated changes throughout the year. Therefore, the time series backscattering coefficient can be applied to distinguish the wetland cover types. Most VH and VV curve peaks are found on May 23 and October 5, indicating that these dates may be the best time to distinguish wetland cover types. This finding is also consistent with by the time-series spectrum curves (Fig.12) and multi-temporal VIs dynamics (Fig.13).

#### 4.3. Optimal feature combination

The top 40 important features selected by the RF algorithm with 500 trees are listed in Fig.15. These features include VIs, visible bands, NIR bands, red-edge bands, backscattering coefficients and phenological parameters. The time-series NDVI, red-edge bands (includes red-edge VIs) and backscattering coefficients play significant roles in mapping wetlands. Therefore, it is feasible to use time-series and multi-source features for vegetation classification.

The optimal numbers of features selected by the RF method are shown in Fig. 16. With increasing features, the cross-validation accuracy is improved to 0.9447 and then quickly drops to the average level.

**Table 6**

Classification accuracy of three single classifiers: support vector machine (SVM), random forest (RF) and k-Nearest Neighbor (k-NN) based on OBIA.

Land cover	OBIA-SVM		OBIA-RF		OBIA-k-NN	
	PA (%)	UA (%)	PA (%)	UA (%)	PA (%)	UA (%)
Water	92.58	93.58	93.46	95.47	92.77	93.72
Sedge	83.65	84.62	86.87	86.72	83.42	84.28
Reed	84.28	85.77	86.34	87.37	87.08	87.67
Poplar	89.57	90.13	88.64	89.06	88.71	89.14
Paddy	89.24	89.98	90.73	91.52	88.16	88.51
Dryland	89.08	89.48	91.24	91.87	89.75	90.13
Forest	86.72	88.74	90.19	92.08	90.92	91.85
Others	85.58	86.97	88.07	88.97	86.73	88.29
Overall accuracy	87.18%		89.48%		87.78%	
Kappa coefficient	0.86		0.88		0.86	

So increasing the features cannot always increase the accuracy. We assessed the features' importance by the RF algorithm with 500 trees, and sorted features by the value from high to low. We selected the first 40 important features for classification. By doing this, the data redundancy and computation load can be effectively reduced and the mapping accuracy can be improved significantly.

#### 4.4. Classification result and accuracy

The classification results of the object-based stacked generalization method using the optimal feature combination are shown in Fig. 17. To simplify the analysis, single season rice and double cropping rice were taken as one class (paddy rice). For comparison, the pixel-based stacked generalization method was also used to classify the land cover using the test samples (Fig. 18). The results of the object-based stacked generalization method have meaningful geoinformation and much less "salt and pepper" phenomenon than the results of the pixel-based method. Additionally, the classification results based on the object-based method has clearer cover type borders and higher accuracy than the pixel-based one. Take region (a) as an example. The river is identified as a pure object by the object-based method (Fig. 17 a), but as a mixed object of vegetation and water by the pixel-based method (Fig. 18a). Also, the pond in Fig. 17b has a clearer border than that in Fig. 18b. The sedge, poplar as well as the reed are shown as meaningful landscape unit in Fig. 17c. However, in Fig. 18c, some pixels of poplar are misclassified as forest, reed or sedge.

We compared the classification accuracies of the pixel-based and object-based stacked generalization methods in Table 4. The overall accuracy and Kappa coefficient of the object-based method are 92.46% and 0.92, respectively, which are 3.88% and 0.04 higher than that of the pixel-based method. Both the PA and UA of each cover type get

improvements by the object-based method, especially the accuracy of dryland. Generally, the accuracies of all cover types except sedge and reed, are higher than 90%. Although the classification results of the pixel-based method have an acceptable OA, the PA and UA of some vegetation types are low. For example, sedge and reed have the PA of 84.06% and 84.78% and the UA of 84.63% and 85.72%. Therefore, the object-based stacked generalization method has a higher classification accuracy and better generalized performance for land cover classification, especially for the wetland ecosystem with high heterogeneity.

#### 4.5. The wetland classification ability of different feature combinations

In order to evaluate the feasibility and efficacy of different features in wetland mapping, we extracted the wetland information by the object-based stacked generalization method using four feature combinations (Table 5). For five actual Sentinel-2 multispectral images, the OA and Kappa coefficient are 81.69% and 0.81, respectively. When the Sentinel-1 backscattering coefficient and the fused Sentinel-2 NDVI time series are added, the OA and Kappa coefficient are improved to 91.02% and 0.90. The highest OA and Kappa coefficient (92.46%, 0.92) is obtained by the combination of Sentinel-2 MSI, Sentinel-1, NDVI and phenology. This result demonstrates the importance of combining Sentinel-1 SAR data and Sentinel-2 optical data, because the OA and Kappa coefficient is increased by approximate 7% and 0.05, respectively by adding the Sentinel-1 data. Multi-temporal optical images are likely to result in a precise classification for wetland vegetation communities, and adding SAR data can further improve the classification accuracy.

#### 4.6. Classification performance of single classifier and stacked generalization

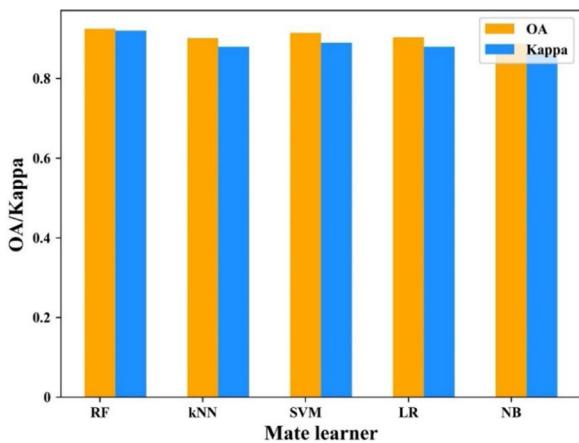
We used the three classifiers (SVM, RF and kNN) to map regions (a-c) in the study area separately, and compared their results with that using the proposed method (Fig. 19). The proposed method gets a better classification result. For example, the RF classifier misidentifies poplar as forest and the dryland and others are blended seriously in the results of the RF and kNN (Fig. 19a). In regions (b) and (c), single classifier cannot classify sedge and reed correctly, but the stacked generalization method can get an accurate classification. In some areas, the forest is misclassified as the paddy rice (e.g. region (c)) by RF and kNN. The stacked generalization can correctly differentiate between forest and paddy, because of its ensemble strategy and the favorable classification result of SVM. Even if there is a misclassification of BLs in stacking, as long as an algorithm only does one thing is a unique contribution, it may be a key progress (Healey et al., 2018).

Single classifiers get low classification accuracy (PA and UA) for

**Table 7**

Classification accuracy of object-based stacked generalization with different combinations of BLs.

Land cover	RF + SVM + kNN		RF + SVM + kNN + LR + NB + GLM							
	LR		LR + NB		NB		LR + NB + GBDT			
	PA (%)	UA (%)	PA (%)	UA (%)						
Water	95.23	95.84	97.64	98.29	94.47	94.57	97.32	97.54	98.37	98.58
Sedge	87.21	87.92	87.66	87.66	88.03	88.52	87.01	87.35	87.24	88.43
Reed	86.72	87.07	88.10	90.37	87.36	88.06	87.93	88.71	86.11	87.56
Poplar	89.39	90.03	90.68	91.86	87.72	89.21	90.23	90.54	90.72	91.54
Paddy	91.48	91.66	92.16	94.43	90.98	90.07	90.27	91.63	91.33	92.54
Dryland	92.25	92.86	94.87	95.89	92.84	93.20	93.23	92.47	93.03	93.42
Forest	91.77	92.53	92.87	94.17	92.39	94.07	90.94	91.20	92.54	93.05
Others	90.38	91.18	90.04	90.76	90.31	90.42	90.46	90.87	89.34	89.46
Overall accuracy	91.38%		92.46%		90.71%		91.86%		92.24%	
Kappa coefficient	0.90		0.92		0.91		0.92		0.92	



**Fig. 20.** The OA and Kappa coefficient of object-based stacked generalization with different MLs.

sedge and reed, while the stacked generalization method achieves higher classification accuracy (Table 6). Among the three single classifiers, RF performs better. The SVM algorithm has the lowest accuracy for water identification (92.58% for PA, 93.58% for UA), but the highest classification accuracy for poplar (89.57% for PA, 90.13% for UA). In high heterogeneity areas, single classifier usually suffers from the generalization error. The object-based stacked generalization algorithm is applicable for the vegetable classification in high heterogeneity areas and has better performance than single classifiers.

#### 4.7. Stability of the object-based stacked generalization approach

Fundamentally, the performance and generalization ability of stacking algorithms are determined by the model structure. For example, different combinations of BLs and MLs are variables in stacked generalization, and they have impacts on the classification results. Thus, three scenarios were developed in this study to analyze the influences of different BLs and MLs on the performance of the stacked generalization method. In the first scenario, in order to evaluate the influence of the number of BLs on the classification accuracy, the LR (Logistic Regression) and NB (Naive Bayesian) algorithms which achieve relative low accuracies when are used alone based on OBIA (LR: OA = 83.84%, Kappa = 0.79; NB: OA = 82.33%, Kappa = 0.77), were removed from the stacked generalization algorithm. In the second scenario, Gradient Boosting Decision Tree (GBDT) and Generalize Linear Model (GLM) were selected as the extra BLs of the stacked generalization algorithm to assess whether increasing the number of BLs affects the classification accuracy. Finally, all algorithms used in the BLs (RF, SVM, kNN, LR and NB) were used as ML to detect the influence of different MLs on classification accuracy. Their performances were assessed by the 5-fold cross validation.

The result of the first scenario (Table 7) shows that the OA and Kappa have reduced to 91.38%, 0.90 and 90.71%, 0.91, respectively, when the NB or LR algorithm are took out from the BLs, respectively. That is, the stability and generalization of the stacking algorithm is challenged by removing some BLs with a relative low classification accuracy. When the extra GBDT and GLM are added, the OA and Kappa have no obvious change. According to the results of the first and second scenarios, the classification accuracy changes caused by different configuration of BLs are uncertain. In the stacked generalization algorithm, the output result is produced by the ML base on the predictions of BLs. Therefore, the correlation of different classifiers in BLs may be the main factors impacting the classification accuracy (Shunmugapriya and Kanmani, 2013). The results using five different MLs indicate that non-

linear classifiers get higher accuracies than linear classifiers (LR: OA = 90.36%, Kappa = 0.88; NB: 88.72%, Kappa = 0.86). Note that, RF (OA = 92.46%, Kappa = 0.92) has the highest classification accuracy among three non-linear classifiers (kNN: OA = 90.14, Kappa = 0.88; and SVM: OA = 91.47%, Kappa = 0.88) (Fig. 20).

In general, the changes in classification accuracies caused by different configurations (different algorithms and combinations of BLs and different MLs) of the object-based stacked generalization are minimal. Thus, to achieve a stable and generalized performance, an adaptive object-based stacked generalization algorithm that can automatically search for the best combination of parameters is needed.

#### 4.8. Limitations and future works

Although this research has demonstrated the good performance of the object-based stacked generalization method by combining multi-temporal optical and SAR data, there are some limitations in the practical application. First, the 16-day fused Sentinel-2 time series NDVI was generated by RASTFM, which may have some prediction errors if the temporal land surface changes are too indiscernible in the low-resolution prior images (Cai et al., 2019; Zhao et al., 2018). The residual cloud contamination in the 16-day MODIS time series NDVI, which is quite common in tropical and sub-tropical areas, also has impacts on the prediction accuracy. Second, the optimal BLs and MLs in stacking algorithm should be further determined, and an adaptive object-based stacked generalization algorithm should be developed. Additionally, deep learning enables computer to automatically extract features from complex remote sensing images, and some remote sensing applications with deep learning as the core have achieved better recognition or classification performance than other algorithms in some cases (Mahdianpari et al., 2018; Rezaee et al., 2018; Mohammadimanesh et al., 2019; DeLancey et al., 2019). Therefore, integrating deep learning algorithms into stacked generalization will be an interesting topic. Third, the RF approach was used to select optimal features and reduce the data redundancy, but the pre-processing of the original images and the transformation of VI need massive computation, which lowers the processing speed. Cloud-based platform (e.g. Google Earth Engine) integrated with a large number of free remote sensing images and providing remote computation may be a feasible way to combine machine learning algorithms for ecosystem monitoring (Zurqani et al., 2018).

## 5. Conclusions

In this study, an object-based stacked generalization method was proposed for wetland classification using multi-temporal optical data and SAR data. The results demonstrate the superiority of the strategy that integrates multitemporal optical data and SAR data, and the potential of using object-based stacked generalization for wetland classification at regional or large scales. In spite of the practical barriers to broader application owing to the image acquisition, preprocessing costs and computation load, the object-based stacked generalization method gets higher classification accuracy than any single classifiers. The achieved OA and kappa coefficient are 92.46% and 0.92, respectively. The proposed method combines multiple features as the inputs of the training model, and it can provide higher accuracy wetland mapping for high heterogeneity area at a regional or large scale.

#### CRediT authorship contribution statement

**Yaotong Cai:** Data curation, Formal analysis, Writing - original draft. **Xinyu Li:** Resources, Validation. **Meng Zhang:** Conceptualization, Funding acquisition, Methodology, Project administration, Writing - review & editing. **Hui Lin:** Investigation, Resources, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

Our deepest gratitude goes to the editor and anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this manuscript. This study was funded by the National Natural Science Foundation of China (41901385), the Forestry Remote Sensing Application System based on GF satellites (Phase 2) (21-Y30B02-9001-19/22), and in part by the China Postdoctoral Science Foundation (2019M652815).

## References

- Cai, Y., Lin, H., Zhang, M., 2019. Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Adv. Space Res.* 64, 2233–2244.
- Collins, L., Griffioen, P., Newell, Mellor, A., 2018. The utility of Random Forests for wildfire severity mapping. *Remote Sens. Environ.* 216, 374–384.
- Debanshi, S., Pal, S., 2020. Wetland delineation simulation and prediction in deltaic landscape. *Ecol. Indic.* 108, 105757.
- DeLancey, E.R., Simms, J.F., Mahdianpari, M., Brisco, B., Mahoney, C., Kariyeva, J., 2019. Comparing Deep Learning and Shallow Learning for Large-Scale Wetland Classification in Alberta, Canada. *Remote Sens. (Basel)* 12 (1), 2.
- Deventer, H.V., Cho, M.A., Mutanga, O., 2019. Multi-season RapidEye imagery improves the classification of wetland and dryland communities in a subtropical coastal region. *Isprs J. Photogramm. Remote. Sens.* 157, 171–187.
- Drägut, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* 24, 859–871.
- Dronova, I., Gong, P., Wang, L., 2011. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at Poyang Lake, China. *Remote Sens. Environ.* 115, 0–3236.
- Evans, T.L., Costa, M., 2013. Landcover classification of the lower nhecolândia subregion of the brazilian pantanal wetlands using ALOS/PALSAR, RADARSAT-2 and ENVISAT/ASAR imagery. *Remote Sens. Environ.* 128, 118–137.
- Gao, B.-C., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* 58, 257–266.
- Gevaert, C.M., García-Haro, F.J., 2015. A comparison of STARFM and an unmixing-based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* 156, 34–44.
- Ghulam, A., Porton, I., Freeman, K., 2014. Detecting subcanopy invasive plant species in tropical rainforest by integrating optical and microwave (InSAR/PolInSAR) remote sensing data, and a decision tree algorithm. *Isprs J. Photogramm. Remote. Sens.* 88, 174–192.
- Gitelson, A.A., Gritz, Y., Merzlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282.
- Gitelson, A.A., Merzlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697.
- Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll in crops. *Geophys. Res. Lett.* 32, L08403.
- Guan, J., Qi, K., Wang, J., Zhuang, J., Yuan, X., Yan, B., Lu, N., Qu, J., 2020. Effects of conversion from boreal natural wetlands to rice paddy fields on the dynamics of total dissolved iron during extreme precipitation events. *Chemosphere* 242, 125153.
- Guo, Y., Jia, X., Paull, D., Benediktsson, J.A., 2019. Nomination-favoured opinion pool for optical-SAR-synergistic rice mapping in face of weakened flooding signals. *Isprs J. Photogramm. Remote. Sens.* 155, 187–205.
- Han, X., Chen, X., Feng, L., 2015. Four decades of winter wetland changes in Poyang Lake based on Landsat observations between 1973 and 2013. *Remote Sens. Environ.* 156, 426–437.
- Hariharan, S., Mandal, D., Tirolkar, S., Kumar, V., Bhattacharya, A., Lopez-Sanchez, J.M., 2018. A novel phenology based feature subset selection technique using random forest for multitemporal PolSAR crop classification. *Ieee J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 11, 4244–4258.
- Healey, S.P., Cohen, W.B., Yang, Z., Kenneth Brewer, C., Brooks, E.B., Gorelick, N., Hernandez, A.J., Huang, C., Joseph Hughes, M., Kennedy, R.E., Loveland, T.R., Moisen, G.G., Schroeder, T.A., Stehman, S.V., Vogelmann, J.E., Woodcock, C.E., Yang, L., Zhu, Z., 2018. Mapping forest change using stacked generalization: An ensemble approach. *Remote Sens. Environ.* 204, 717–728.
- Hou, X., Feng, L., Chen, X., Zhang, Y., 2018. Dynamics of the wetland vegetation in large lakes of the Yangtze Plain in response to both fertilizer consumption and climatic changes. *Isprs J. Photogramm. Remote. Sens.* 141, 148–160.
- Hossain, M.D., Chen, D., 2019. Segmentation for Object-Based Image Analysis (OBIA): a review of algorithms and challenges from remote sensing perspective. *Isprs J. Photogramm. Remote. Sens.* 150, 115–134.
- Ienco, D., Interdonato, R., Gaetano, R., Ho Tong Minh, D., 2019. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *Isprs J. Photogramm. Remote. Sens.* 158, 11–22.
- Jahncke, R., Leblon, B., Bush, P., LaRocque, A., 2018. Mapping wetlands in Nova Scotia with multi-beam RADARSAT-2 Polarimetric SAR, optical satellite imagery, and Lidar data. *Int. J. Appl. Earth Obs. Geoinf.* 68, 139–156.
- Jeżowski, J., Bochenek, R., Ziomek, G., 2005. Random search optimization approach for highly multi-modal nonlinear problems. *Adv. Eng. Softw.* 36, 504–517.
- Jönsson, P., Eklundh, L., 2004. TIMESAT—a program for analyzing time-series of satellite sensor data. *Comput. Geosci.* 30, 833–845.
- Kang, J., Hou, X., Niu, Z., Gao, S., Jia, K., 2014. Decision tree classification based on fitted phenology parameters from remotely sensed vegetation data. *Transactions Chin. Soc. Agricul. Eng.* 30 (9), 148–156.
- Kaplan, G., Avdan, U., 2019. Evaluating the utilization of the red edge and radar bands from sentinel sensors for wetland classification. *CATENA* 178, 109–119.
- Li, M., Bijkir, W., 2019. Vegetable classification in Indonesia using Dynamic Time Warping of Sentinel-1A dual polarization SAR time series. *Int. J. Appl. Earth Obs. Geoinf.* 78, 268–280.
- Li, Z., Chen, H., White, J.C., Wulder, M.A., Hermosilla, T., 2020. Discriminating treed and non-treed wetlands in boreal ecosystems using time series Sentinel-1 data. *Int. J. Appl. Earth Obs. Geoinf.* 85, 102007.
- Ma, L., Crawford, M.M., Tian, J., 2010. Local manifold learning-based -Nearest-Neighbor for hyperspectral image classification. *Ieee Trans. Geosci. Remote. Sens.* 48, 4099–4109.
- Mahdavi, S., Salehi, B., Amani, M., Granger, J., Brisco, B., Huang, W., 2019. A dynamic classification scheme for mapping spectrally similar classes: application to wetland classification. *Int. J. Appl. Earth Obs. Geoinf.* 83, 101914.
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Motagh, M., 2017. Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery. *Isprs J. Photogramm. Remote. Sens.* 130, 13–31.
- Mahdianpari, M., Rezaee, M., Zhang, Y., Salehi, B., 2018. Wetland classification using deep convolutional neural network. In: In IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE. pp. 9249–9252.
- Martone, M., Rizzoli, P., Wecklich, C., González, C., Bueso-Bello, J.-L., Valdo, P., Schulze, D., Zink, M., Krieger, G., Moreira, A., 2018. The global forest/non-forest map from TanDEM-X interferometric SAR data. *Remote Sens. Environ.* 205, 352–373.
- Mohammadimanesh, F., Salehi, B., Mahdianpari, M., Motagh, M., Brisco, B., 2018. An efficient feature optimization for wetland mapping by synergistic use of SAR intensity, interferometry, and polarimetry data. *Int. J. Appl. Earth Obs. Geoinf.* 73, 450–462.
- Mohammadimanesh, F., Salehi, B., Mahdianpari, M., Gill, E., Molinier, M., 2019. A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem. *Isprs J. Photogramm. Remote. Sens.* 151, 223–236.
- Periasamy, S., 2018. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: an attempt on Sentinel-1. *Remote Sens. Environ.* 217, 537–549.
- Peyret, R., Bouridane, A., Khelifi, F., Tahir, M.A., Al-Maadeed, S., 2018. Automatic classification of colorectal and prostatic histologic tumor images using multiscale multispectral local binary pattern texture features and stacked generalization. *Neurocomputing* 275, 83–93.
- Phiri, D., Morgenroth, J., Xu, C., Hermosilla, T., 2018. Effects of pre-processing methods on Landsat OLI-8 land cover classification using OBIA and random forests classifier. *Int. J. Appl. Earth Obs. Geoinf.* 73, 170–178.
- Rezaee, M., Mahdianpari, M., Zhang, Y., Salehi, B., 2018. Deep convolutional neural network for complex wetland classification using optical remote sensing imagery. *Ieee J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 11 (9), 3030–3039.
- Rouse Jr, J.W., Haas, R., Schell, J., Deering, D., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Special Publication 351–309.
- Shuai, G., Zhang, J., Basso, B., Pan, Y., Zhu, X., Zhu, S., Liu, H., 2019. Multi-temporal RADARSAT-2 polarimetric SAR for maize mapping supported by segmentations from high-resolution optical image. *Int. J. Appl. Earth Obs. Geoinf.* 74, 1–15.
- Shunmugapriya, P., Kanmani, S., 2013. Optimization of stacking ensemble configurations through artificial bee colony algorithm. *Swarm Evol. Comput.* 12 (Complete), 24–32.
- Singh, M., Sinha, R., 2019. Distribution, diversity, and geomorphic evolution of floodplain wetlands and wetland complexes in the Ganga plains of north Bihar, India. *Geomorphology* 106960.
- Steinhausen, M.J., Wagner, P.D., Narasimhan, B., Waske, B., 2018. Combining Sentinel-1 and Sentinel-2 data for improved land use and land cover mapping of monsoon regions. *Int. J. Appl. Earth Obs. Geoinf.* 73, 595–604.
- Tong, X., Tian, F., Brandt, M., Liu, Y., Zhang, W., Fensholt, R., 2019. Trends of land surface phenology derived from passive microwave and optical remote sensing systems and associated drivers across the dry tropics 1992–2012. *Remote Sens. Environ.* 232, 111037.
- Wang, Q., Atkinson, P.M., 2018. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sens. Environ.* 204, 31–42.
- Wittke, S., Yu, X., Karjalainen, M., Hyppä, J., Puttonen, E., 2019. Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *Int. J. Appl. Earth Obs. Geoinf.* 76, 167–178.
- Wolpert, D.H., 1996. Stacked generalization. *Neural Netw.* 5, 241–259.
- Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. *Ieee Trans. Evol. Comput.* 1, 67–82.

- Wu, C., Niu, Z., Tang, Q., Huang, W., 2008. Estimating chlorophyll content from hyperspectral vegetation indices: modeling and validation. *Agric. For. Meteorol.* 148, 1230–1241.
- Yang, L., Mansaray, L.R., Huang, J., Wang, L., 2019. Optimal Segmentation Scale Parameter, Feature Subset and Classification Algorithm for Geographic Object-Based Crop Recognition Using Multisource Satellite Imagery. *Remote Sens. (Basel)* 11, 514.
- Yu, Y., Mei, X., Dai, Z., Gao, J., Li, J., Wang, J., Lou, Y., 2018. Hydromorphological processes of Dongting Lake in China between 1951 and 2014. *J. Hydrol. (Amst)* 562, 254–266.
- Zhang, H., Lin, H., Wang, Y., 2018a. A new scheme for urban impervious surface classification from SAR images. *Isprs J. Photogramm. Remote. Sens.* 139, 103–118.
- Zhang, H., Xu, R., 2018. Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta. *Int. J. Appl. Earth Obs. Geoinf.* 64, 87–95.
- Zhang, M., Lin, H., Wang, G., Sun, H., Fu, J., 2018b. Mapping paddy rice using a convolutional neural network (CNN) with landsat 8 datasets in the Dongting Lake Area. *China. Remote Sensing* 10, 1840.
- Zhang, W., Brandt, M., Wang, Q., Prishchepov, A.V., Tucker, C.J., Li, Y., Lyu, H., Fensholt, R., 2019. From woody cover to woody canopies: how Sentinel-1 and Sentinel-2 data advance the mapping of woody plants in savannas. *Remote Sens. Environ.* 234, 111465.
- Zhao, Y., Huang, B., Song, H., 2018. A robust adaptive spatial and temporal image fusion model for complex land surface changes. *Remote Sens. Environ.* 208, 42–62.
- Zhou, Y., Xiao, X., Qin, Y., Dong, J., Zhang, G., Kou, W., Jin, C., Wang, J., Li, X., 2016. Mapping paddy rice planting area in rice-wetland coexistent areas through analysis of Landsat 8 OLI and MODIS images. *Int. J. Appl. Earth Obs. Geoinf.* 46, 1–12.
- Zurqani, H.A., Post, C.J., Mikhailova, E.A., Schlautman, M.A., Sharp, J.L., 2018. Geospatial analysis of land use change in the Savannah River Basin using Google Earth Engine. *Int. J. Appl. Earth Obs. Geoinf.* 69, 175–185.