Comprehensive Evaluation of Sentinel-2 Red Edge and Shortwave-Infrared Bands to Estimate Soil Moisture

Ying Liu, Jiaxin Qian, and Hui Yue

***Abstract*—This article aims to explore the applicability of SMMI (soil moisture monitoring index), MSMMI (modified soil moisture monitoring index), PDI (perpendicular drought index), and MPDI (modified perpendicular drought index) in estimating soil moisture (SM) in farmland. The random forest classifier was used to obtain two-stage land cover types maps. The sensitivity of Sentinel-2 spectral bands to the measured SM at a depth of 0–5 cm was optimized by random forest regression. According to the sensitive bands, SMMI and PDI from different feature spaces were constructed to explore their feasibility for monitoring SM under different land cover types. Second, fractional vegetation cover (FVC) in the study area was estimated by nine kinds of FVC estimation models and compared with the measured FVC. The effects of different FVC methods on estimating SM by MSMMI and MPDI were evaluated. The results show that red edge and short-wave infrared (SWIR) bands of Sentinel-2 had irreplaceable effects on the land cover classification. In terms of monitoring SM in bare soil areas, the SM indices with SWIR bands had high correlations with measured SM. For vegetation-covered areas, MSMMI from the FVCgr model (dimidiate pixel model with red edge bands) and the Short wave infrared1–Short wave infrared2 feature space had the highest correlation with the measured 0–5 cm depth SM. Whether vegetation-covered areas or bare soil areas, the combination of red edge and SWIR bands can effectively improve the estimation accuracy of SM. MSMMI can be used as the best SMMI in the study area. Sentinel-2 images, with great potential, can effectively estimate SM at a depth of 0–5 cm in farmland with complex environments.**

***Index Terms*—Red edge, Sentinel-2, short-wave infrared (SWIR), soil moisture (SM), Soil Moisture Active Passive Validation Experiment, 2016 (SMAPVEX16), spectral feature space.**

Abbreviations

SM	Soil moisture
SMMI	Soil moisture monitoring index
PDI	Perpendicular drought index
MSMMI	Modified soil moisture monitoring index
MPDI	Modified perpendicular drought index
FVC	Fractional vegetation cover
FVCg	$FVC_{\text{Gutman and Ignatov}}$
FVCgr	FVCg with red edge bands
FVCb	FVC_{Baret}
FVCbr	FVCb with red edge bands
FVCc	$FVC_{\text{Carlson and Ripley}}$
FVCcr	FVCc with red edge bands
FVCd	FVC with a scale-differential vegetation index
FVCdr	FVCd with red edge bands

Manuscript received January 13, 2021; revised May 21, 2021 and July 10, 2021; accepted July 15, 2021. Date of publication July 21, 2021; date of current version August 4, 2021. This work was supported in part by the Natural Science Basic Research Program of Shaanxi under Grant 2020JM-514, in part by the Research on Ecological Restoration and Protection of Coal Base in Arid Ecofragile Region under Grant GJNY2030XDXM-19-03.2, in part by the Project of Shaanxi Coal and Chemical Industry Group under Grant 2018SMHKJ-A-J-03, and in part by the Xi'an University of Science and Technology under Grant 2019YQ3-04. *(Corresponding author: Ying Liu.)*

The authors are with the College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China (e-mail: liuying712100@163.com; 467545659@qq.com; 13720559861@163.com).

Digital Object Identifier 10.1109/JSTARS.2021.3098513

I. Introduction

SOIL moisture (SM) plays an important role in climate change because it is a basic parameter in the process of formation, transformation, and consumption of land and water resources [1]. SM is also one of the basic conditions for crop growth and development [2]–[6]. Insufficient water supply will hinder the growth of crops, indirectly affecting yield. Excessive water supply will lead to soil hypoxia, further hinder the respiration of crop roots, resulting in reduced physiological functions and even death. Real-time and accurate monitoring of SM is one of the important links of agricultural water management and the realization of modern precision agriculture [6]. Remote sensing technology has become an efficient approach to monitoring SM [1]. Currently, remote sensing methods for monitoring SM are as follows: thermal inertia [2], vegetation water supply index [3], temperature vegetation dryness index [4], anomaly vegetation index [5], conditional vegetation index [6], vegetation temperature condition index [7], microwave [8]–[10], and spectral feature space method [11]–[15]. Methods derived from spectral feature spaces are widely applied to assess SM because of being simple and easy to operate.

Richardson and Wiegand [16] proposed the perpendicular vegetation index (PVI) based on the soil line in NIR (near-infrared)-Red space. PVI can eliminate the influence of soil background and is less sensitive to the atmosphere than other vegetation indices [16]. Fensholt and Proud [17] established a short-wave infrared (SWIR) water stress index with SWIR and NIR bands of MODIS, which provided a strong basis for monitoring SM in semi-arid regions [17]. Ghulam *et al.* [18] proposed the perpendicular drought index (PDI) in NIR-Red space, which was better in drought monitoring than other indices and suitable for drought monitoring in bare soil and low vegetation-covered areas [18]. Ghulam *et al.* [19] proposed

the modified perpendicular drought index (MPDI) considering fractional vegetation cover (FVC), which had roughly the same accuracy in monitoring SM in bare soil than PDI and was better in vegetation-covered areas [19]. Zhu *et al.* [20] indicated that PDI had a good negative correlation with SM which was also suitable for drought monitoring with low spatial resolution satellites in large areas [20]. Amani *et al.* [21] proposed the triangle SM index and the improved triangle SM index in NIR-Red space, which had higher accuracy than other SM models [21]. Chen *et al.* [22] proposed a decomposition SM index in NIR-Red space combining the normalized SM index and Soil Moisture Active Passive (SMAP) data, which was suitable for monitoring SM in a large area [22]. Rao *et al.* [23] noted that MPDI could effectively monitor vegetation pressure and forest decline on a macro scale [23]. Zormand *et al.* [24] indicated that MPDI could replace standardized precipitation index to monitor drought conditions in drought areas with little precipitation [24]. Liu *et al.* [25] expressed that soil moisture monitoring index (SMMI) was better than PDI in estimating SM, and both of them were suitable for monitoring SM at a depth of 0–5 cm [25]. Liu *et al.* [26] proposed the modified soil moisture monitoring index (MSMMI) based on SMMI and FVC, in which monitoring accuracy was the same as MPDI in semi-arid areas [26].

Previous studies of monitoring SM based on the spectral feature space mainly used NIR-Red or NIR-SWIR spaces, and rarely involving red-edge bands. However, these bands are unique to capture the vegetation spectral characteristics and more sensitive to the state of vegetative growth. Red edge cannot only increase the accuracy of crop analysis and improve the ability of crop identification but also improve the diagnostic accuracy of crop growth status [27]–[31]. Currently, many scholars have used red-edge bands to monitor vegetation health conditions [27], estimate crop leaf area index (LAI) [28], classify wetlands [29], and crops [30]–[31]. Few studies are using red-edge bands of Sentinel-2 to improve the ability to assess SM. Furthermore, FVC is a function of vegetation canopy morphology, vegetation spatial distribution, solar incidence angle, and observation angle, which can better represent the spatial pattern between vegetation and bare soil areas [19], [26]. Accurate estimation of FVC is the key point in revealing SM through MSMMI and MPDI. Few studies have made a qualitative assessment of the influence of FVC estimation on estimating SM.

This article mainly focused on the following questions:

1) Which bands and derived vegetation indices of Sentinel-2 are more important for land cover classification in agricultural areas? (Part A of Section IV)
2) For different land cover types, which bands of Sentinel-2 are more sensitive to the measured SM? Can red-edge bands of Sentinel-2 improve the estimation accuracy of SM? (Parts B and C of Section IV)
3) Which of the four SM indices (SMMI, PDI, MSMMI, and MPDI) can better accurately estimate SM at the regional scale? Which estimation model of FVC is more suitable for monitoring SM by MSMMI and MPDI? (Parts D and E of Section IV)

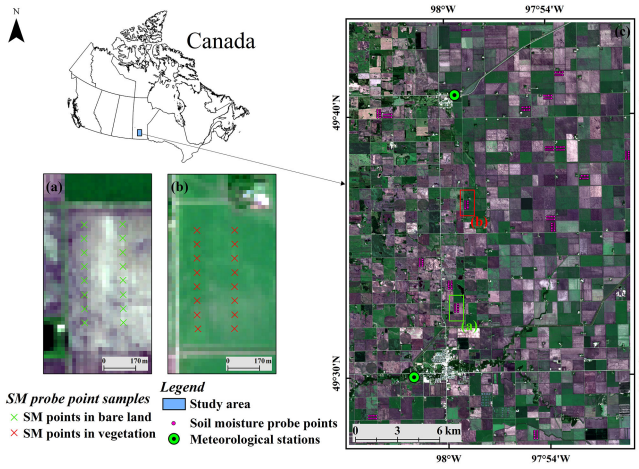


Fig. 1. Location of the study area and field measured SM probe distribution. (a) SM points in bare soil areas. (b) SM points in vegetation-covered areas. (c) Sentinel-2A imagery on June 10, 2016.

TABLE I
MEASURED DATA AND REMOTE SENSING DATA INFORMATION USED IN THE STUDY

Data	Date	Description
Sentinel-2A	2016/06/10&13&/20	Two images mosaicked
Measured SM	2016/06/14 9:00-11:00 2016/06/20 9:00-11:00	158 bare soil; 119 crops 60 bare soil; 46 crops
Measured FVC	2016/06/13&2016/06/20	70 samples; 60 samples
Measured soil texture	2016/06/01-2016/07/31	57 samples

II. STUDY AREA AND DATA

The study area locates in Elm Creek and Carman, Manitoba, and Canada (Fig. 1). It has a typical temperate continental climate, with hot and sunny summer, and long and cold winter but abundant sunshine. It is flat and open terrain, and mainly grows cash crops such as barley, wheat, canola, and soybean. The study area is about 27×45 km^2.

The field measured data were derived from the SMAPVEX16 dataset (Soil Moisture Active Passive Validation Experiment, 2016) (Table I), which included meteorological data, land cover type data, vegetation status data, field measured FVC data, field-measured SM data, and soil texture (ST) data [32]–[36]. The dataset was a part of an SM validation experiment in Canada. The experiment was designed to calibrate and improve the accuracy of NASA's SMAP product.

A. Sentinel-2A Data

Sentinel-2A L1C data were derived from the Copernicus Open Access Hub. The images were acquired on June 10, 13, and 20, 2016. The orbital number of the images is 55 and the identifications are 14UNA and 14UNV, respectively (Table I). L1C data is an apparent reflectance product after orthographic correction and geometric precision correction. Atmospheric correction is required to obtain the land surface reflectance of each band. Sentinel-2A data are mainly used for extracting land cover types, constructing SM indices, and estimating FVC in the study area.

B. Land Cover Type and Vegetation Status Data

The land cover type map was obtained from the Earth Observation Group of Agriculture and Agri-Food Canada based on optical and radar satellite images (Landsat-8, Sentinel-2, Gaofen-1, and Radarsat-2) using decision tree supervision classification. The map with 30-m resolution included crops (food crops, oil crops, and vegetable crops), woodland, water, bare soil land, urban, and so on. The overall accuracy of the dataset is more than 85%. The vegetation growth status data mainly included various plant characteristics of the study area, such as biomass, plant height, plant density, growth stage, etc.

C. Field Measured Fractional Vegetation Cover

The FVC data measured the ratio of the projected area of vegetation coverage to the total projected area, which was derived from digital photos of the sampling points and processed by ViewNX-2 and Can-Eye 5.1 software. Digital photos were taken by a Nikon camera with a fish-eye lens at least 50 cm below the vegetation canopy. There were three sampling points in each field, and at least ten photos were acquired from each sampling point. In this article, the measured FVC data on June 13 and 20, 2016 in the study area were selected to validate the estimated FVC by different models.

D. Field Measured Soil Moisture and Soil Texture

This article selected the SM probe data and the *in-situ* core data as the measured SM data. The data measured soil real dielectric constant (RDC) at a depth of 0–5 cm. Stevens Poke and Go (POGO) was used to measure the top, middle, and bottom of the 0–5 cm soil layer (perpendicular to the top, middle, and bottom of the furrow, respectively). The RDC value was converted into volume SM (cm^3/cm^3) using a field-specific calibration equation, and its expression is as follows:

$$\mathrm{SM} = A \times \mathrm{RDC}^{0.5} + B \tag{1}$$

where A and B are constants, which can be obtained by linear regression from *in-situ* core SM samples.

The soil samples were dried for at least 48 h in a 105 °C oven at the Soil Science Laboratory of the University of Manitoba, and the wet and dry weights were recorded. In this article, the mean results of three different depth measurements were used as the SM verification values. The POGO device collected SM data from 50 farmland and two meteorological stations. Each farmland has 16 collection points, each at a distance of at least 70 m [Fig. 1(a) and (b)]. All data had been quality-controlled and any erroneous records had been deleted. Garmin GPS devices with 3-m precision were used to locate the sampling points when FVC and SM were collected.

There was approximately 12 mm precipitation in the study area on June 12, 2016, and no rain around June 20, 2016. The air temperature in the study area on the 13th and 14th ranged from 15.5 °C to 27.5 °C, with a mean and standard deviation of 20.5 °C and 3.5 ° C, respectively. Since SM data were not collected on the 13th, the verification data on the 14th was used instead. There were 277 measured SM samples in the study area from 10:00 to 12:00 on June 14, 2016 (158 points were bare soil and 119 points were crops). There were 106 measured SM samples in the study area from 10:00 to 12:00 on June 20, 2016 (60 points were bare soil and 46 points were crops).

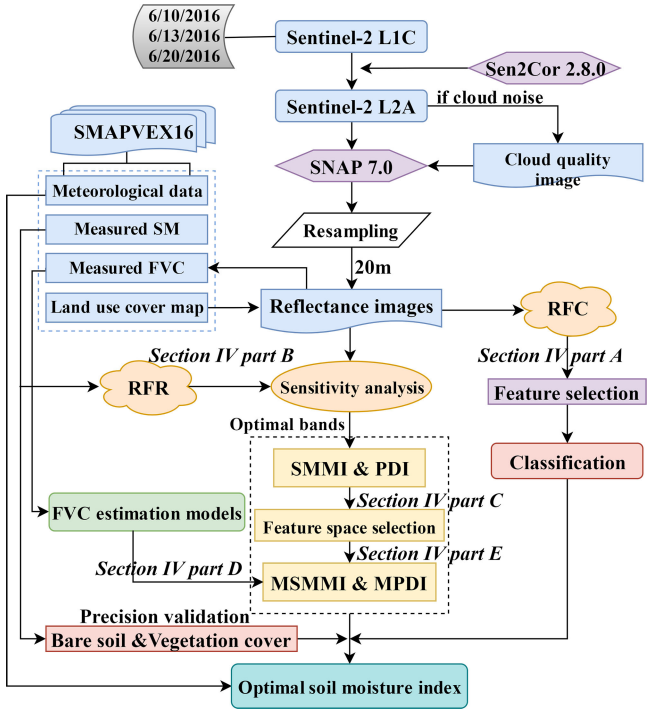


Fig. 2. Technical flow chart.

The soil content data contains ST and organic matter observations, including the percentage of soil contained in the sand fraction, silt fraction, and clay fraction (%). The ST was divided according to different content. The STs of measured SM sample points included heavy clay (HC, account for about 45.4%), clay (C, 9.1%), sandy loam (SL, 9.1%), loamy sand (LS, 18.2%), and sand (S, 18.2%). The percentage of organic matter in the soil was between 2.2% and 11.6%, and the mean and standard deviation were 7.0% and 3.1%, respectively.

III. Methods

The technical process of this article is as follows (Fig. 2):

First, Sentinel-2 L1C data were generated into L2A products using Sen2cor 2.8.0. The Sentinel-2 toolbox of SNAP 7.0 (Sentinel Application Platform) was invoked for resampling (10 to 20 m by mean aggregation method), mosaicking, and subset to generate the surface reflectance images with 20-m spatial resolution in three periods of the study area.

Second, Sentinel-2 images of June 10 and 20, 2016 were used for classification by random forest classifier (RFC) (Part A of Section IV).

Third, the sensitivity of each spectral band of Sentinel-2 to measured SM under different land cover types was analyzed by random forest regression (RFR) (Part B of Section IV). SMMI and PDI from different feature spaces were constructed by the bands with higher importance scores. The measured SM data

in bare soil and vegetation-covered areas were compared with SMMI and PDI to find out which feature space is more suitable for monitoring SM at different surfaces (Part C of Section IV).

Fourth, nine empirical models were used to estimate the FVC of the study area (Part D of Section IV). MSMMI and MPDI from the selected feature space (Part C of Section IV) were calculated under different FVC models to compare with measured SM in bare soil and vegetation-covered areas (Part E of Section IV).

Finally, the optimal SM index was selected according to land cover types, meteorological data, and measured SM data.

A. Classification of Land Cover Types

The land cover map of the SMAPVEX16 dataset reflected the year-round surface cover in 2016. However, many crops in the study period (in June 2016) had been harvested or just begun to sow. Compared with the Sentinel-2 images on June 10 and 20, canola and other crops developed rapidly in large areas. To better reveal the influence of different land cover types on estimating SM, land cover types needed to be re-evaluated.

In this article, spectral information, traditional vegetation indices, water indices, chlorophyll vegetation indices, red-edge vegetation indices, and texture features were selected to construct the classification feature dataset. Spectral information is relative to the Sentinel-2 bands of Blue, Green, Red, NIR, nNIR (narrow NIR), Red edge1, Red edge2, Red edge3, SWIR1 (Short wave infrared1), and SWIR2 (Short wave infrared2); Traditional vegetation indices contain DVI (difference vegetation index) [37], ratio vegetation index [38], NDVI (normalized difference vegetation index) [39], green NDVI [40], enhanced vegetation index [41], and modified soil adjust vegetation index [42]; water indices contain NDWI (normalized difference water index) [43], NDMI (normalized difference moisture index) [44], and MNDWI (modified normalized difference water index) [45]; Chlorophyll vegetation indices contain ARI1/2 (Anthocyanin reflectance index1/2) [46] and CRI1/2 (Carotenoid reflectance index1/2) [47]; Red-edge vegetation indices contain red edge normalized difference vegetation indices (NDre1/2/3, NDVIre1/2/3, and NDVIre21/31/32), MRENDVI (modified red edge normalized difference vegetation index) [48], MRESR (modified red edge simple ratio) [49], REPI (red-edge position index) [50], S2REP (Sentinel-2 REPI) [51], triangular vegetation index [52], novel inverted red-edge chlorophyll index [53], terrestrial chlorophyll index [54], MCARI (modified chlorophyll absorption ratio index) [55], MCARI2 [56], red-edge chlorophyll index [57], and TCARI (transformed chlorophyll absorption reflectance index) [58]. If the central wavelength used in the above indices could not be obtained in Sentinel-2, the nearest reflectance band was used instead.

The GLCM (gray level co-occurrence matrix) was employed as the texture feature to improve classification accuracy [59], [60]. In texture analysis, the mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, and correlation were used for quantitative analysis. If texture features are generated at each band of Sentinel-2, the classification efficiency will be reduced due to data redundancy. Therefore, the first principal component was selected through a principal component analysis of the Sentinel-2 imagery to calculate texture features. Generally, there are four directions for computing the GLCM. The mean eigenvalues of these four directions were taken as the final eigenvalue matrix.

The specific formulas for the vegetation indices and texture features used for classification are displayed as supplementary material of this article. There are altogether 51 spectral features.

B. Random Forest Algorithm

Random forest (RF) algorithm can predict the role of up to thousands of explanatory variables very well, which is praised as one of the best machine learning algorithms at present [61]. RF cannot only realize remote sensing image classification (RFC) but also plays an important role in feature selection and dimension reduction [62]. Out-of-bag (OOB) error of RF is an unbiased estimate of prediction error, which can assess the classification accuracy and calculate the importance of different feature variables for feature selection. The regression problem can also be handled by RFR. The prediction results of RFR are obtained by averaging the prediction results of all internal binary decision trees. In this article, the EnMAP-BOX tool developed by the Environment Mapping and Analysis Program of Germany was adopted to optimize features and extract land use cover maps. The importance of 51 feature variables was evaluated by RFC. The feature subset with the highest classification accuracy was selected to generate a land cover map in the study area. Ten independent experiments were conducted in the importance evaluation process. The mean of the ten importance scores was the final score result. There are two very important parameters in the construction of RF, namely the number of decision trees (N) and the number of features (m) extracted in the process of node splitting. Through a large number of experiments, the OOB errors gradually converged and tended to be stable with the number of decision trees ($N \geq 100$). Therefore, N was set to 100. The arithmetic square root was selected as the number of features (m). In this article, the parameters for regression and classification were similar.

C. Sample Selection and Accuracy Verification

Training samples and verification samples for classification were selected from Sentinel-2A on June 10 and 20, 2016. There were 14 types of land cover, namely water, urban, bare soil, woodland, grassland, pasture, barley, oats, rye, triticale, winter wheat, spring wheat, canola, and pea. At least 30 samples of each category were evenly selected and randomly sampled, with 70% as training samples and the remaining 30% as verification samples. Jeffries–Matusita (JM) distance was employed to quantitatively characterize the degree of separability among selected samples [63]. The selection of samples was adjusted to ensure the JM distance of every two samples greater than 1.90. The overall accuracy, Kappa coefficient, producer accuracy (PA), and user accuracy (UA) were adopted to evaluate the accuracy of classification results.

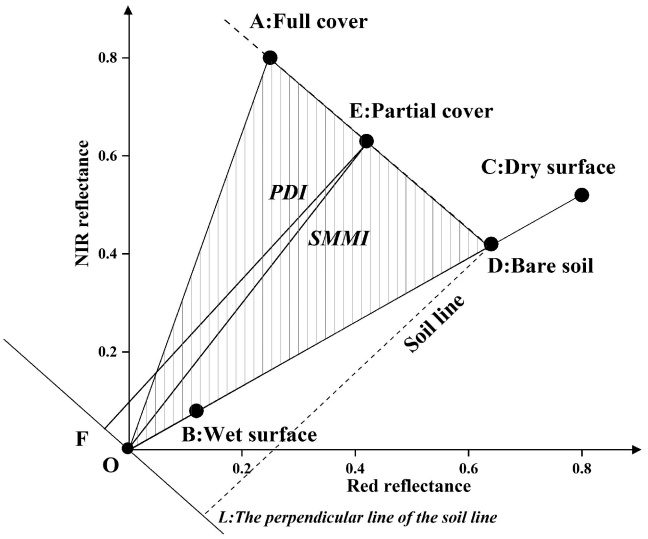


Fig. 3. Schematic diagram of SMMI and PDI.

D. Soil Moisture Indices Based on Spectral Feature Space

Ghulam *et al.* [18] put forward the PDI based on NIR-Red feature space, which is simple and effective in monitoring SM [18]. Fig. 3 shows that there is a fixed soil line in the triangular NIR-Red feature space, which is affected by ST, fertility, and other factors. To reduce the influence of the soil line, Liu *et al.* (2017) proposed the SMMI which does not involve the soil line [25]. SMMI works similarly to PDI, but with a slight difference. The distance from any point in the feature space to the origin O reflects the SM conditions. When the point locates at point B (Fig. 3), it is close to the origin O, indicating the SM in this area is wetter. When the point lies in point D (Fig. 3), indicating the SM in the region is drier. When the point locates at the origin O, the region is water or an extremely humid region. Therefore, the smaller the distance is from any point to the origin, the higher the SM is. The ratio of OB to OD is selected as the index of SM, where the distance of OD is a fixed value ($\sqrt{2}$).

PDI and SMMI are formulated as follows:

$$\mathrm{PDI} = \frac{R_{\mathrm{red}} + M \cdot R_{\mathrm{NIR}}}{\sqrt{M^2+1}} \tag{2}$$

$$\mathrm{SMMI} = \frac{OB}{OD} = \frac{\sqrt{{R_{\mathrm{red}}}^2 + {R_{\mathrm{NIR}}}^2}}{\sqrt{2}} \tag{3}$$

where R_{red} and R_{NIR} represent the reflectance of red and NIR bands, respectively. M is the slope of the soil line. The slopes of the soil line in different phases were calculated by interactive data language. The bands used in PDI and SMMI can be determined according to the feature space used.

E. Soil Moisture Indices Involving FVC

The accuracy of monitoring SM by PDI is low under vegetation-covered areas [18], [19]. Ghulam *et al.* [19] put forward the MPDI by introducing vegetation coverage [19]. The physical meaning of MPDI is that drought conditions at any point in the NIR-Red feature space are decided by SM and FVC. The farther the distance from the origin of coordinates is and the higher the MPDI value is, the more severe the drought is. The MSMMI was proposed through the introduction of vegetation coverage based on SMMI [26]. The principles of MSMMI and MPDI are also similar, but slightly different. The former does not require a soil line.

MPDI and MSMMI are calculated as follows:

$$\mathrm{MPDI} = \frac{R_{\mathrm{red}} + M \cdot R_{\mathrm{NIR}} - FVC \times (R_{v,\mathrm{red}} + M \cdot R_{v,\mathrm{NIR}})}{(1-FVC) \cdot \sqrt{M^2+1}} \tag{4}$$

$$\mathrm{MSMMI} = \frac{\sqrt{(R_{\mathrm{NIR}} - FVC \cdot R_{v,\mathrm{NIR}})^2 + (R_{\mathrm{red}} - FVC \cdot R_{v,\mathrm{red}})^2}}{\sqrt{2} \times (1 - \mathrm{FVC})} \tag{5}$$

where $R_{v,\ \mathrm{red}}$ and $R_{v,\ \mathrm{NIR}}$ are the reflectance of red and NIR bands of vegetation, respectively, usually taking 0.05 and 0.5 [19], [26]. These two bands can be adjusted according to the feature space used. If FVC approaches 1.0 (pure vegetation), the value of MSMMI and MPDI will be infinite (namely SM is extremely low). It is inconsistent with the actual situation. Therefore, the maximum FVC was limited to 0.95. Besides, the empirical reflectance values of pure vegetation also need to be changed according to crop types. In this article, the pixels of NDVI > 0.9 were considered as pure vegetation pixels. The mean reflectance used in the feature space (such as $R_{v,\ \mathrm{red}}$ and $R_{v,\ \mathrm{NIR}}$) of these pixels was taken as the empirical reflectance values of pure vegetation.

F. Methods for Estimating FVC

FVC can usually be determined by taking its relationship with the vegetation index. NDVI is the most widely used to retrieve the biophysical characteristics of vegetation canopy, such as LAI, FVC, and biomass. Over the past two decades, NDVI has been successfully used to monitor global vegetation coverage change [64]–[66]. There are three typical semi-empirical relationships between FVC and NDVI. Baret *et al.* [67] established the general semi-empirical relationship between NDVI and FVC. Carlson and Ripley [68] proposed the semi-empirical relationship between NDVI and FVC based on the simple radiation transmission model. Gutman and Ignatov [69] proposed the semi-empirical formulas between FVC and NDVI based on the dimidiate pixel model (DPM).

The expressions of these FVC models are as follows:

$$\mathrm{FVC_{Baret}}\ (\mathrm{FVC}b) = 1 - \left(\frac{\mathrm{NDVI}_v - \mathrm{NDVI}}{\mathrm{NDVI}_v - \mathrm{NDVI}_s}\right)^{0.6175} \tag{6}$$

$$\mathrm{FVC_{CarlsonandRipley}}\ (\mathrm{FVC}c) = \left(\frac{\mathrm{NDVI} - \mathrm{NDVI}_s}{\mathrm{NDVI}_v - \mathrm{NDVI}_s}\right)^2 \tag{7}$$

$$\mathrm{FVC_{GutmanandIgnatov}}\ (\mathrm{FVC}g) = \frac{\mathrm{NDVI} - \mathrm{NDVI}_s}{\mathrm{NDVI}_v - \mathrm{NDVI}_s} \tag{8}$$

where NDVI_v and NDVI_s are NDVI values of pure vegetation and pure bare soil pixels, respectively. They can be obtained

from field-measured data or NDVI images. Generally, the values corresponding to 1% and 99% of the cumulative frequency from NDVI images are selected as NDVI_s and NDVI_v.

NDVI can help highlight vegetation information and even distinguish healthy vegetation from unhealthy. However, NDVI has obvious limitations. On the surface of dense vegetation, NDVI mainly responds to the red band, but is relatively insensitive to NIR changes, so it cannot correctly describe the changes in canopy density [68], [70], [71]. There are many opinions and differences in the relation between NDVI and FVC. The principle of deriving FVC from NDVI is to correlate the NDVI of the mixed pixel with the reference NDVI value, such as the NDVI of dense vegetation and bare soil. It is assumed that the NDVI of each component of the mixed pixel can be represented by the reference NDVI. However, even if the mixed pixels can be separated correctly and estimated as the reference NDVI value, there is still uncertainty caused by the proportional effect of NDVI. Zhang *et al.* (2006) proposed a scale-differential vegetation index (SDVI) with DVI, which can be used for remote sensing images to invert vegetation-covered areas, especially surfaces under heterogeneous conditions [72]. DVI of shadowed soil is usually very small, which is not sensitive to vegetation coverage of shadowed soil. Its expression is as follows [72]:

$$\mathrm{FVC}_{\mathrm{SDVI}}\left(\mathrm{FVC}d\right) = \frac{\mathrm{DVI} - \mathrm{DVI}_s}{\mathrm{DVI}_v - \mathrm{DVI}_s} \tag{9}$$

$$\mathrm{DVI} = R_{\mathrm{NIR}} - R_{\mathrm{red}} \tag{10}$$

where DVI_s represent DVI of bare soil and DVI_v is DVI of dense vegetation.

Red-edge normalized difference vegetation index ($\mathrm{NDVI}_{\mathrm{red\ edge}}$) is a slight improvement over traditional NDVI. $\mathrm{NDVI}_{\mathrm{red\ edge}}$ uses the edge zone with chlorophyll absorption characteristics (such as 705 nm), which is more sensitive to the health status of vegetation [73]. Replacing the red and NIR bands of NDVI and DVI with the red-edge bands to obtain $\mathrm{NDVI}_{\mathrm{red\ edge}}$ and $\mathrm{DVI}_{\mathrm{red\ edge}}$ (red-edge difference vegetation index), whose expressions are as follows:

$$\mathrm{NDVI}_{\mathrm{rededge}} = \frac{R_{\mathrm{rededge3}} - R_{\mathrm{rededge1}}}{R_{\mathrm{rededge3}} + R_{\mathrm{rededge1}}} \tag{11}$$

$$\mathrm{DVI}_{\mathrm{rededge}} = R_{\mathrm{rededge3}} - R_{\mathrm{rededge1}} \tag{12}$$

where $R_{\mathrm{red\ edge1}}$ and $R_{\mathrm{red\ edge3}}$ represent the reflectance of the Band5 and Band7 of Sentinel-2, respectively.

Formulas 6–9 can be adapted for considering the red-edge bands as follows:

$$\mathrm{FVC}br = 1 - \left(\frac{\mathrm{NDVI}_{\mathrm{rededge},v} - \mathrm{NDVI}_{\mathrm{rededge}}}{\mathrm{NDVI}_{\mathrm{rededge},v} - \mathrm{NDVI}_{\mathrm{rededge,s}}}\right)^{0.6175} \tag{13}$$

$$\mathrm{FVC}cr = \left(\frac{\mathrm{NDVI}_{\mathrm{rededge}} - \mathrm{NDVI}_{\mathrm{rededge},s}}{\mathrm{NDVI}_{\mathrm{rededge},v} - \mathrm{NDVI}_{\mathrm{rededge},s}}\right)^2 \tag{14}$$

$$\mathrm{FVC}gr = \frac{\mathrm{NDVI}_{\mathrm{rededge}} - \mathrm{NDVI}_{\mathrm{rededge},s}}{\mathrm{NDVI}_{\mathrm{rededge},v} - \mathrm{NDVI}_{\mathrm{rededge},s}} \tag{15}$$

$$\mathrm{FVC}dr = \frac{\mathrm{DVI}_{\mathrm{rededge}} - \mathrm{DVI}_{\mathrm{rededge},s}}{\mathrm{DVI}_{\mathrm{rededge},v} - \mathrm{DVI}_{\mathrm{rededge},s}} \tag{16}$$

where the values corresponding to the $\mathrm{NDVI}_{\mathrm{red\ edge}}$ imagery cumulative frequency confidence of 1% and 99% were selected as the values of pure bare soil and pure vegetation (namely $\mathrm{NDVI}_{\mathrm{red\ edge},\ s}$ and $\mathrm{NDVI}_{\mathrm{red\ edge},\ v}$). $\mathrm{DVI}_{\mathrm{red\ edge},\ s}$ and $\mathrm{DVI}_{\mathrm{red\ edge},\ v}$ work in the same way.

Finally, four traditional semi-empirical FVC estimation models (formulas 6–9) and four semi-empirical FVC estimation models with red edge bands (formulas 13–16) were formed.

The L2A product of Sentinel-2 can generate LAI, fraction of absorbed photosynthetically active radiation, canopy chlorophyll content, CWC (canopy water content), and FVC in a large area by using the biophysical module in SNAP software. In this module, the PROSPECT+SAIL radiation transmission model and artificial neural network are used for estimation calculation. The construction of the neural network model is composed of 11 explicit input layer parameters including Green (the central wavelength is 560 nm), Red (665 nm), Red edge1 (705 nm), Red edge2 (740 nm), Red edge3 (783 nm), nNIR (842 nm), SWIR1 (1610 nm), SWIR2 (2190 nm), the cosine of zenith angle, the cosine of solar altitude angle, cosine of relative azimuth angle, and five implicit neuron parameters with tangent S-shape curve transfer function. This algorithm can calculate the relationship between vegetation pixels and canopy reflectance to the greatest extent, and ensure the optimal estimation simulation accuracy through the model self-verification function. Data products with 10 or 20 m resolution can be generated through resampling. Its accuracy is greatly improved compared with previous remote sensing data such as Landsat, providing a powerful application reference value for quantitative remote sensing analysis of regional vegetation ecological environment [74], [75]. In this article, the FVC estimated by the physical model of SNAP was named $\mathrm{FVC}_{\mathrm{S2}}$.

G. Accuracy Verification of Soil Moisture Estimation

The Pearson's correlation coefficient (R) was applied to evaluate the accuracy of each SM estimation model [76], which can be written as

$$R = \frac{\mathrm{Cov}\left(X, Y\right)}{\sqrt{\mathrm{Var}\left(X\right)\mathrm{Var}\left(Y\right)}} \tag{17}$$

where X is SM index, Y is field measured SM, Cov(X, Y) is the covariance of X with Y, Var(X) is the variance of X, Var(Y) is the variance of Y.

IV. Results and Discussions

A. Analysis of Land Cover Type Classification

To present the feature variables of high importance more intuitively and clearly, each feature was ranked according to the importance score. Fig. 4(a) demonstrates that the importance scores of feature variables were significantly different. MNDWI had the highest score (up to 15.132 ± 0.429, mean and standard deviation). GLCM_correlation had the lowest score which almost did not influence classification. On the whole, the contributions of the feature variables for extracting land cover were water indices > spectral information > chlorophyll

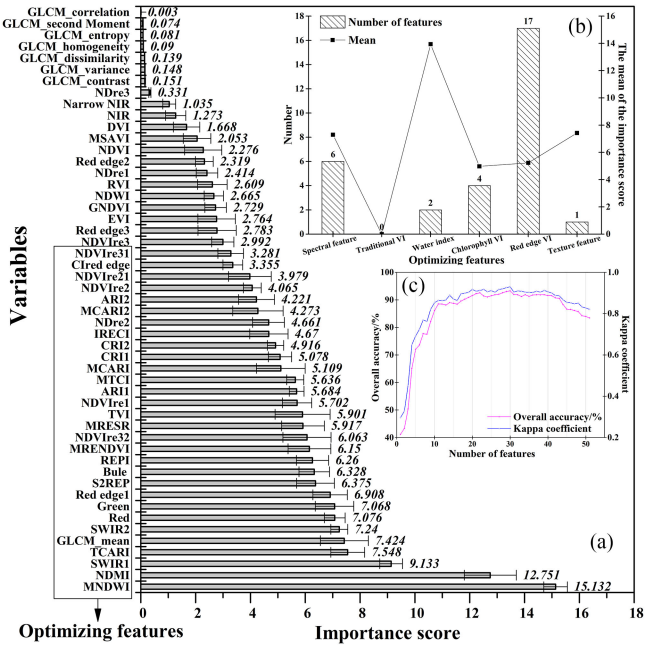


Fig. 4. (a) Distribution of feature importance scores. (b) Mean importance scores and the number of optimizing features. (c) Relationship between the number of features and classification accuracy.

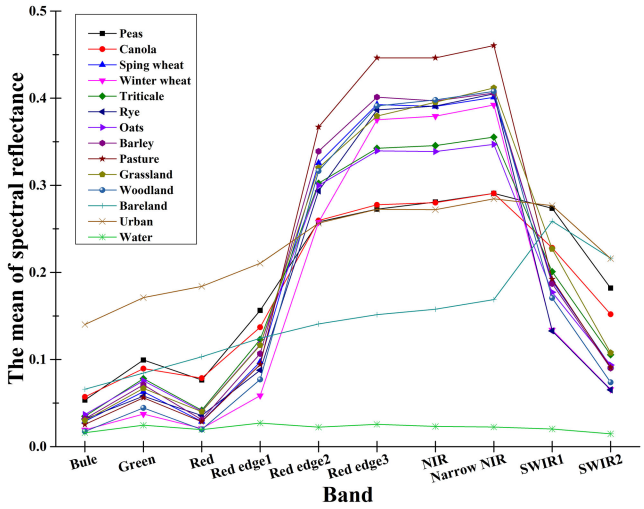


Fig. 5. Mean spectral reflectance of some typical feature samples.

vegetation indices > red edge vegetation indices > traditional vegetation indices > texture features. Fig. 4(c) demonstrates that the classification accuracy increases from 41.15% to 86.11% rapidly in the early stage with 10 feature variables. This was mainly due to the high importance score of feature variables that were included in the early stage. The classification accuracy was improved gradually from 86.11% to 91.51% in the middle stage with 11–41 features. The classification accuracy gradually decreased in the later period with 42–51 features. This was due to redundant features and irrelevant features that were involved. Most texture features with 20-m spatial resolution will not necessarily improve the classification accuracy.

Fig. 4(a) also indicates that the features with high important scores are MNDWI, NDMI, SWIR1, TCARI, GLCM_mean, SWIR2, Red, Green, Red edge1, and S2REP. When all features in the classification, information redundancy may reduce the overall accuracy. Only the feature variables that played a key role in classification were retained. The 30 optimization features were selected according to the highest classification accuracy [Fig. 4(c)], including six spectral features (the mean of importance score is 7.292), two water indices (13.942), four chlorophyll vegetation indices (4.975), seventeen red edge vegetation indices (5.232), and one texture feature (7.424), not including traditional vegetation indices.

Fig. 5 shows that the spectral reflectance of different land cover samples varies greatly. The reflectance of each band of water is extremely lower than other ground objects. The spectral curves of urban and bare soil are similar, which rise with the increase of wavelength. The reflectance of all kinds of crops has obvious spectral reflectance characteristic of vegetation, which is different at Green, Red, Red edge1, SWIR1, and SWIR2 bands. The NIR reflectance can effectively differentiate between crops like canola and pasture. The NIR reflectance of canola and peas is similar to each other which is difficult to distinguish them effectively. However, the reflectance difference between them in the red edge and SWIR bands is obvious. This also explains why the importance score of the traditional vegetation index and NIR band for land cover was low. Therefore, NDWI composed of green and NIR had a lower importance score, while MNDWI and NDMI participated by SWIR had higher scores. The importance score of chlorophyll vegetation indices was also higher, mainly because these indices used Blue, Green, and Red edge1 bands which have remarkable spectral differences. The contribution of red edge indices for land cover is significantly higher than that of traditional vegetation indices. It indicates that red edge bands can effectively identify complex land cover types. Only the GLCM_mean feature had a higher contribution among texture features, while other texture features contributed very little. It indicated that texture features extracted from Sentinel-2 images were only applicable to the land cover type with relatively simple texture information.

This is consistent with previous research results. Thanh and Kappas [77] evaluated the ability of Sentinel-2 data in August to identify crop types and tree species. The results showed that red-edge and SWIR bands had great value for vegetation mapping [77]. Zhang *et al.* [78] used the multitemporal Sentinel-2 data to classify land use types in the Yellow River Delta region of China. The results showed that RF classification results with optimal multitemporal features (up to 33) had the highest accuracy, with an overall accuracy of 90.93% and a Kappa coefficient of 0.90. Red edge, SWIR bands, and their respective indices were of great importance to the classification, but most texture features did not improve the final classification accuracy for Sentinel-2 data [78]. The growth status of crops in this study area was different. The red-edge bands are the sensitive characteristic spectral band of vegetation, and can effectively reflect the chlorophyll content, phenological period, and category of vegetation [73]. The SWIR bands are mainly affected by the content of water content, starch

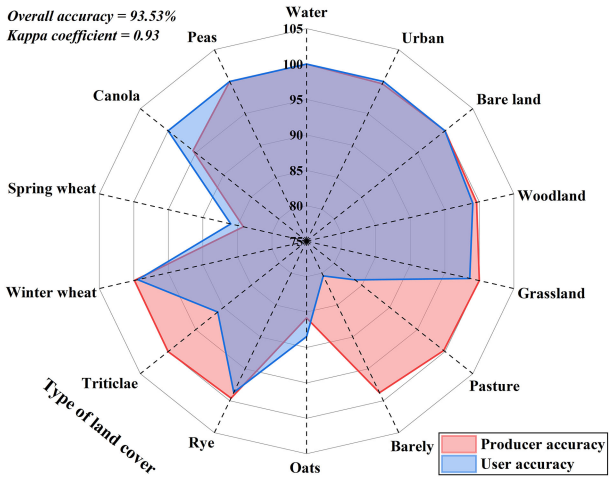


Fig. 6. Classification accuracy statistics of land cover types.

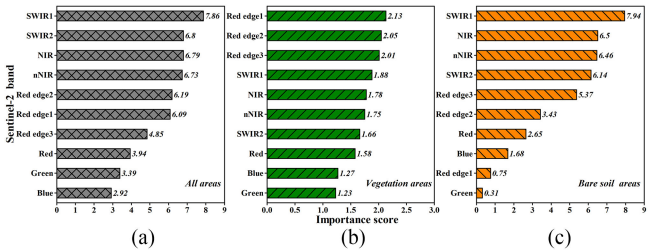


Fig. 7. Importance scores between each band of Sentinel-2A and field-measured soil moisture.

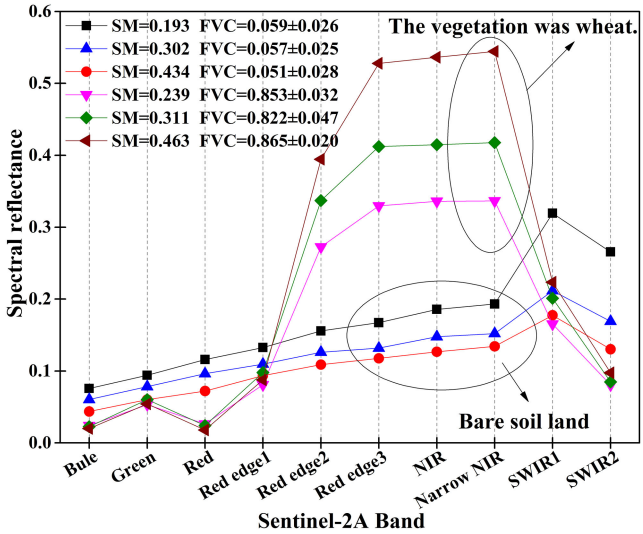


Fig. 8. Importance scores between each band of Sentinel-2A and field-measured soil moisture.

content, protein, oil, sugar, lignin, and cellulose in vegetation leaves, which are conducive to the recognition of vegetation [77], [78].

Fig. 4(c) shows that the overall accuracy and Kappa coefficient reach the maximum when the number of feature variables reaches 30. Therefore, we adopted the top 30 features in the feature importance score to classify [Fig. 4(a)]. Fig. 6 shows that the PA and UA of water, bare soil, and urban construction land are almost 100%, and the accuracy of woodland, grassland, winter wheat, rape, and peas is also relatively high (PA and UA $>$ 95%). The accuracy of spring wheat, barley, oat, rye, and other wheat crops ranges from 80% to 90%. The reason may be that winter wheat was at the mature stage while other wheat crops were at the growing stage. Therefore, the spectral curve of winter wheat was different from other crops. The phenomenon of the same spectrum with different objects in other wheat crops was more serious.

The bare soil areas accounted for 59.01% on June 10, 2016. Oats, canola, grassland, pasture, spring wheat, triticale, woodland, rye, barley, winter wheat, and peas accounted for 11.89%, 7.53%, 3.97%, 3.03%, 2.97%, 2.51%, 1.98%, 1.69%, 1.55%, 1.07%, and 0.62%, respectively. Urban was 1.98% and water was 0.2%. Land cover types classification on June 20, 2016 was obtained in the same way. Part of bare soil was changed into canola on June 20, 2016, accounting for 4.18% of the total area.

B. Sensitivity Analysis Between Measured Soil Moisture and Sentinel-2 Bands

RFR was used to assess the sensitivity between each band of Sentinel-2 and field measured SM under different land cover types. A total of 383 fields measured SM points were involved in regression (165 were vegetation and 218 were bare soil). Fig. 7 shows the importance scores of each band of Sentinel-2 to measured SM under different land cover types. Fig. 7(a)–(c) contains all measured sample points, vegetation sample points, and bare soil sample points, respectively. Fig. 7(a) indicates that SWIR1 of Sentinel-2, SWIR2, NIR, and nNIR had relatively high contributions to SM, of which SWIR1 had the highest score (7.86), followed by SWIR2 (6.8) and NIR (6.79). Fig. 7(b) expresses that Red edge1 had the highest contribution to SM in vegetation-covered areas (2.13), followed by red edge2 (2.05), red edge3 (2.01), and SWIR1 (1.88). Fig. 7(c) demonstrates that SWIR1 had the highest contribution to SM in bare soil areas (7.94), followed by NIR (6.5), nNIR (6.46), and SWIR2 (6.14).

We can see that the importance score of most bands in bare soil areas is significantly higher than that in vegetation-covered areas. This is because the reflectance spectrum of bare soil is a direct reflection of ST and SM. The reflectance spectrum in vegetation-covered areas reflects the characteristics of the vegetation canopy. The composition of the spectrum is relatively complex, which is affected by vegetation types, vegetation water content, vegetation water-holding capacity, and vegetation health. It can only indirectly reflect the SM on the underlying surface [44]. However, Fig. 7(b) also shows that red edge and SWIR bands are more sensitive to SM in vegetated areas compared with VIS (visible light) and NIR bands. This indicates that red edge bands could improve the accuracy of SM estimation in vegetated areas.

Fig. 8 shows the spectral reflectance curve of wheat crops and bare soil sample points from Sentinel-2A. The FVC of bare soil sample points was around 0.05, which was almost unaffected by

TABLE II
DIFFERENT COMBINATIONS OF SPECTRAL FEATURE SPACE

Group	Band	Spectral feature space
(1)	Red & NIR	B4-B8
(2)	Red & NIR & SWIR	B4-B11; B8-B11
(3)	SWIR	B11-B12
(4)	Red & Red edge & NIR	B4-B5; B4-B6; B4-B7; B5-B8; B6-B8; B7-B8
(5)	Red edge & SWIR	B5-B11; B6-B11; B7-B11
(6)	Red edge	B5-B6; B5-B7; B6-B7

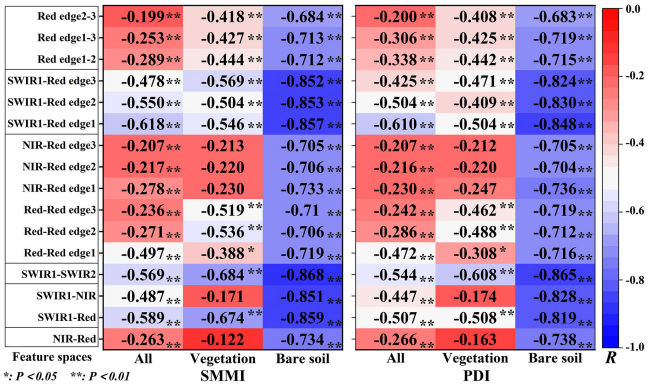


Fig. 9. Relationship was expressed by the correlation coefficient between soil moisture indices and field measured SM under different land cover types on June 13, 2016.

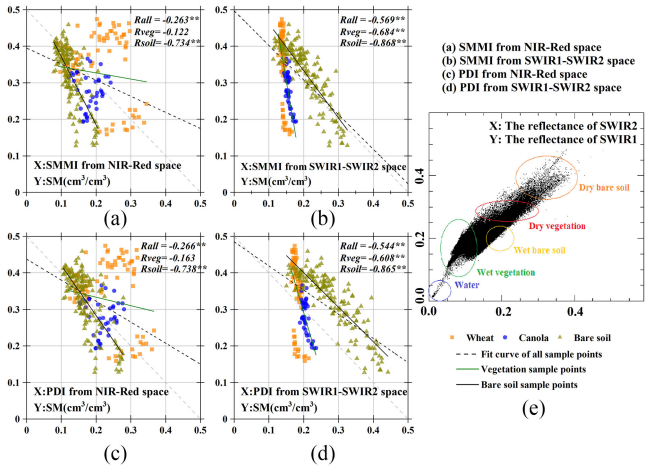


Fig. 10. Scatter plots of soil moisture indices and measured 0–5 cm depth SM. (a) SMMI from NIR-Red space. (b) SMMI from SWIR1–SWIR2 space. (c) PDI from NIR-Red space. (d) PDI from SWIR1–SWIR2 space. (e) Spectral feature space based on SWIR1–SWIR2.

vegetation. The changing trend of bare soil spectral curve with different SM status was almost the same, which continued to rise with increasing wavelength, reached a peak at SWIR1, and decreased at SWIR2. The difference in the reflectance of dry bare soil in VIS and NIR bands was small, while they were different at SWIR bands. Therefore, the SWIR band can reflect the SM distribution of bare soil better than other bands. The wheat crops were in the leaf development stage (three-leaf expansion stage) of the first major growth stage with a measured FVC around 0.8. The increase of SM led to vegetation spectra remarkably increase at red edge2, red edge3, and NIR, and a slight increase at red edge1, SWIR1, and SWIR2. There was no obvious impact on vegetation spectra at VIS bands. The geometric characteristics of vegetation leaves, canopy structure, and water demand all can affect the water content of vegetation. SM indirectly affects vegetation CWC. The influence of vegetation water content on vegetation canopy reflectance mainly includes red edge, NIR, and SWIR bands [17], [44]. Therefore, these bands can better reflect the distribution of SM in vegetated areas than other bands.

As shown above, Band4 (red), Band5 (red edge1), Band6 (red edge2), Band7 (red edge3), Band8 (NIR), Band11 (SWIR1), and Band12 (SWIR2) of Sentinel-2 is more sensitive to SM than others, which can better reflect the spectral differences of vegetation canopy and bare soil under different SM conditions. Different spectral feature spaces were constructed to obtain SMMI and PDI (Table II).

C. Analysis of SMMI and PDI From Different Feature Spaces

Fig. 9 indicates that the correlations between SM indices and measured SM in bare soil were significantly higher than that in vegetation-covered areas. For all measured SM sample points, there was a significant negative correlation between each index and measured SM ($P < 0.01$). The correlation between SMMI from red edge1–SWIR1 space and measured SM was the highest ($R = -0.618$), followed by PDI from red edge1–SWIR1 space ($R = -0.610$), SMMI from SWIR1–SWIR2 space ($R = -0.569$), and PDI from SWIR1–SWIR2 space ($R = -0.544$). For measured SM in vegetation-covered areas, the correlation between indices (from red–NIR, red edge1–NIR, red edge2–NIR, red edge3–NIR, and NIR–SWIR1 spaces) and measured SM did not pass the significance test, while the other indices passed the significance test ($P < 0.01$). The correlation between SMMI from SWIR1–SWIR2 space and measured SM in vegetation-covered areas was the highest ($R = -0.684$), followed by PDI from SWIR1–SWIR2 space ($R = -0.608$). For measured SM in bare soil areas, indices (from Red–SWIR1, NIR–SWIR1, SWIR1–SWIR2, Red edge1–SWIR1, Red edge2–SWIR1, and Red edge3-SWIR1 spaces) with SWIR band participation had high correlations with measured SM ($R > 0.85$). SMMI from SWIR1–SWIR2 space had the highest correlation ($R = -0.868$), followed by PDI from SWIR1–SWIR2 space ($R = -0.865$), which were significantly higher than that of SMMI from NIR-Red space ($R = -0.734$) and PDI from NIR-Red space ($R = -0.738$).

Although Red edge1–SWIR1 space had the highest correlation with all SM sample points, the correlations of vegetation areas or bare soil areas were both slightly lower than SWIR1–SWIR2 space. Therefore, in this article, only the SWIR1–SWIR2 feature space was presented and compared with the NIR-Red space. Fig. 10 shows that SM decreases with the increase of SMMI and PDI. Scatter plots of SMMI from NIR-Red space and SM in bare soil areas are evenly distributed on both sides of the fitting trend line, while in vegetation-covered areas are relatively scattered [Fig. 10(a)], so is the PDI from NIR-Red space [Fig. 10(c)]. Scatter plots of SMMI from SWIR1–SWIR2

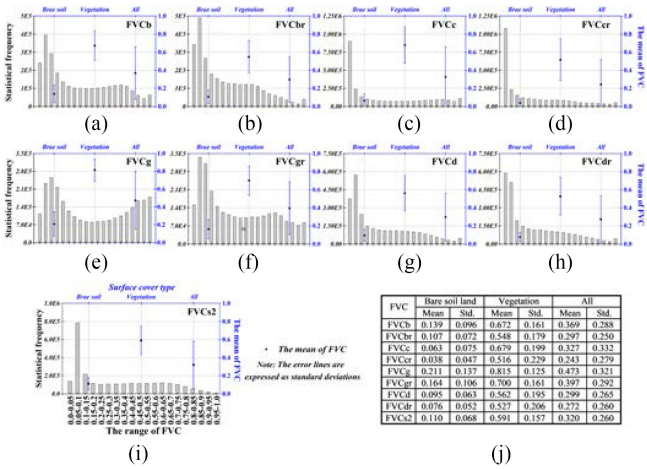
space and SM are evenly distributed on both sides of the fitting trend line [Fig. 10(b)], which are denser than PDI from SWIR1–SWIR2 space [Fig. 10(d)]. Fig. 10(b) demonstrates that although SMMI from SWIR1–SWIR2 space and measured SM scatter are relatively concentrated under a single land cover type, vegetation, and bare soil points are separated on the whole. This may be due to the inconsistent response of different land cover to SM [Fig. 10(e)]. Similarly, the SMMI and PDI from each feature space on June 20, 2016, were calculated. The correlation analysis was also conducted with field-measured SM. The results were consistent with the above.

The above results indicate that NIR-Red feature space can reflect the dry and wet conditions of bare soil, but cannot effectively monitor the dry and wet conditions of vegetated areas. This is consistent with the conclusion of Ghulam *et al.* [18] that PDI is more suitable for monitoring SM in bare soil or areas with sparse vegetation [18]. To address the weaknesses of PDI based on the NIR-Red feature space, Ghulam *et al.* [79] proposed a new index (shortwave infrared perpendicular water stress index) based on the NIR–SWIR feature space and showed that this index has potential in estimating vegetation water content and SM. Fensholt and Proud [17] used NIR and SWIR bands from MODIS data to establish short-wave infrared water stress index (SIWSI) and showed that there was a strong correlation between SIWSI and SM [16]. Ma *et al.* [80] used Sentinel-1 dual-polarized SAR data to invert SM in the semi-arid region of China and compared it with the SMMI based on Landsat images. The results show that SMMI from NIR-SWIR2 space could effectively reflect the surface dry and wet distribution in a semi-arid area, but for the land with sandy type, there was an overestimation [80]. Koley and Jeganathan [81] employed Sentinel-2 and Landsat8 data to monitor SM changes in northeast India and found that the accuracy of SM derived from the SWIR2 band was more accurate than the thermal infrared band. It can be found that although the band combinations of these SM indices are slightly different (which is greatly related to the actual surface cover of the study area), they all involve SWIR bands.

In this article, the correlations between the indices including red-edge bands and SM in vegetated areas were much higher than that from NIR-Red space but still lower than that from SWIR1–SWIR2 space. As a strong absorptive band of water vapor, SWIR bands can effectively reflect SM in bare soil and vegetated areas [16], [79], [81]. Besides, the overall accuracy of SM by SMMI was slightly higher than that of PDI, mainly because SMMI did not consider the soil line, which was significantly different under the influence of ST, soil fertility, vegetation cover types, and other factors in the study area [18], [19], [25], [26].

D. Analysis of FVC Estimation Algorithms

As can be seen in Part C of Section IV, since vegetation cover is not considered, SMMI and PDI have an obvious segmentation phenomenon of SM between vegetation areas and bare soil areas. FVC is an effective index to describe the difference between vegetation and bare soil areas. MSMMI and MPDI introduced FVC to improve the performance of monitoring SM. In this article, four semi-empirical models, four semi-empirical models with red-edge bands, and one biophysical model specifically designed for Sentinel-2 were employed to calculate the FVC. Fig. 10 shows that the frequency of FVCb ($\mathrm{FVC_{Baret}}$), FVCbr (FVCb with red-edge bands), FVCc ($\mathrm{FVC_{Carlson\ and\ Ripley}}$), FVCcr (FVCc with red-edge bands), FVCg ($\mathrm{FVC_{Gutman\ and\ Ignatov}}$), FVCgr (FVCg with red-edge bands), FVCd ($\mathrm{FVC_{SDVI}}$, FVC with an SDVI), FVCdr (FVCd with red-edge bands), and $\mathrm{FVC_{S2}}$ (FVC obtained by Sentinel-2 and SNAP) is consistent and present an obvious unimodal pattern. The field measured FVC $<$ 0.45 was regarded as sample points with no vegetation coverage and low vegetation coverage (a total of 40), and FVC $\geq$ 0.45 was regarded as sample points with medium and high vegetation coverage (a total of 30) [83].



FVC	Bare soil land		Vegetation		All	
	Mean	Std.	Mean	Std.	Mean	Std.
FVCb	0.139	0.096	0.672	0.161	0.369	0.288
FVCbr	0.107	0.072	0.548	0.179	0.297	0.250
FVCc	0.063	0.075	0.679	0.199	0.327	0.332
FVCcr	0.038	0.047	0.516	0.229	0.243	0.279
FVCg	0.211	0.137	0.815	0.125	0.473	0.321
FVCgr	0.164	0.106	0.700	0.161	0.397	0.292
FVCd	0.095	0.063	0.562	0.195	0.299	0.265
FVCdr	0.076	0.052	0.527	0.206	0.272	0.260
FVCs2	0.110	0.068	0.591	0.157	0.320	0.260

Fig. 11. Frequency histograms of estimated FVC. (a) FVCb. (b) FVCbr. (c) FVCc. (d) FVCcr. (e) FVCg. (f) FVCgr. (g) FVCd. (h) FVCdr. (i) $\mathrm{FVC_{S2}}$. (j) Mean and variance of FVC in bare soil, vegetation-covered areas, and the whole area.

Fig. 11(j) demonstrates that, under the same FVC model, the mean and standard deviation of FVC with red-edge bands in bare soil areas were lower than those without red-edge bands, while the standard deviation of vegetation areas was opposite. This suggests that the FVC estimation model with red-edge bands can effectively amplify vegetation canopy differences and better describe the heterogeneity of bare soil and vegetation.

Fig. 12 indicates that the determination coefficients (R^2) of FVC obtained from each model and measured FVC are greater than 0.9. The correlation between FVC with red-edge bands and measured FVC was higher than that without red-edge bands, which was especially obvious when FVC $<$ 0.45. Besides, due to the red-edge bands involved in $\mathrm{FVC_{S2}}$ extracted by SNAP, the accuracy is satisfactory but not the highest. The above results show that the FVC from the semi-empirical model with red-edge bands can more accurately reflect the vegetation coverage situation in the study area. This finding is consistent with previous research. Verrelst *et al.* [84] assessed the applicability of Sentinel-2 and Sentinel-3 bands in estimating biomass indices using four machine learning algorithms in Spain, which showed that the bands most relevant to FVC are concentrated red-edge bands of Sentinel-2 which the most important band is near 740 nm. Cai *et al.* [85] applied Sentinel-2 to estimate FVC by

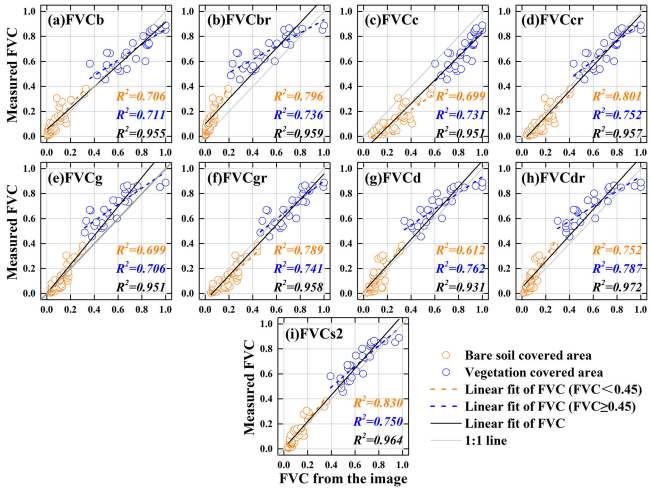


Fig. 12. Scatter diagram of FVC was obtained by different estimation models and measured FVC on June 13, 2016.

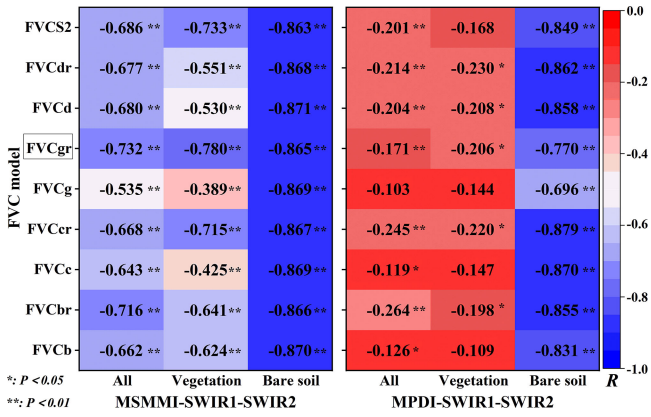


Fig. 13. Relationship was expressed by the correlation coefficient between soil moisture indices from SWIR1–SWIR2 space and measured SM under different land cover types on June 13, 2016.

object-based mixture analysis method in Changsha of China, which indicated that the estimation accuracy of FVC can be improved by red-edges bands, but not significantly. In our study, NDVI or DVI of the FVC model was replaced by $\mathrm{NDVI}_{\mathrm{red\ edge}}$ or $\mathrm{DVI}_{\mathrm{red\ edge}}$, which had higher estimation accuracy and simple operation.

E. Comparison of MSMMI and MPDI Under Different FVC Models

We found that after introducing vegetation coverage into each feature space (Part C of Section IV), the correlation between SWIR1–SWIR2 feature space and measured SM was significantly higher than that of other feature spaces. Therefore, the SWIR1–SWIR2 feature space was selected to calculate MSMMI and MPDI in different FVC models. The NIR-Red feature space was selected as a reference.

Fig. 13 expresses that there are distinct differences in the correlations between SM indices under different FVC models and measured SM. According to Part D of Section IV, although the

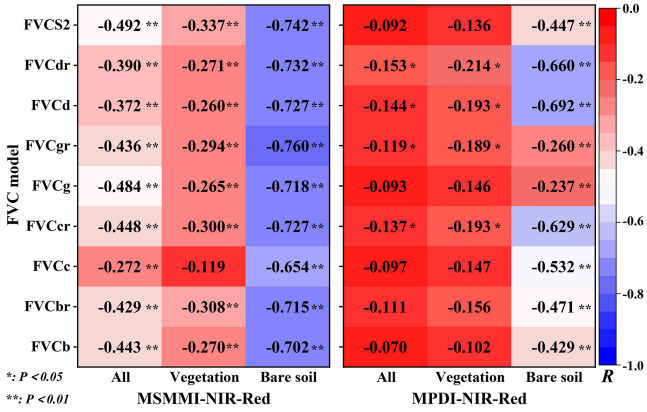


Fig. 14. Relationship was expressed by the correlation coefficient between soil moisture indices from NIR-Red space and measured SM under different land cover types on June 13, 2016.

correlation between the nine kinds of models and measured FVC is high, the distribution of different FVC histograms is greatly different. The inconsistency in the calculation method of FVC is one of the important reasons for the difference in SM monitoring by different models. The correlation between MSMMI from SWIR1–SWIR2 space and measured SM in vegetation-covered areas was significantly higher than that of MPDI from SWIR1–SWIR2 space, and was slightly higher than MPDI in bare soil areas. The correlation between SM indices and measured SM in bare soil was higher than that of SM in vegetation-covered areas. For all measured SM points, there was a significant negative correlation between MSMMI and measured SM ($P < 0.01$). MSMMI based on FVCgr model was the highest ($R = -0.732$), followed by FVCbr model ($R = -0.716$). For vegetated SM points, the correlation between MSMMI under the FVC model with red-edge bands and measured SM was significantly higher than that without red-edge bands. The highest correlation was MSMMI under FVCgr model ($R = -0.780$), followed by $\mathrm{FVC}_{\mathrm{S2}}$ ($R = -0.733$). The correlation between MPDI under each FVC model and measured SM in vegetation-covered areas was low ($R < 0.25$), indicating that MPDI from SWIR1–SWIR2 space was suitable for monitoring SM in bare soil and not good for vegetation-covered areas or mixed covered areas. In conclusion, for SM in bare soil areas, the monitoring accuracy by SM indices under FVC models with red-edge bands are equivalent to those without red-edge bands, while for SM in vegetation-covered areas, SM indices under FVC models with red-edge bands are more accurate.

Fig. 14 shows that the correlation between MSMMI from NIR-Red space and measured SM in bare soil was relatively high ($R > 0.65$), which was significantly lower than that from SWIR1–SWIR2 space (Fig. 13). There was a significant negative correlation between MSMMI from NIR-Red space and measured SM in vegetation-covered areas ($P < 0.01$), but the correlation was low ($R < 0.35$). This indicated that MSMMI from NIR-Red space can also monitor SM conditions in bare soil, but cannot accurately estimate SM in vegetated areas. The correlation between MPDI from NIR-Red space under each

FVC model and SM at all sample points was low ($R < 0.16$), which was also low ($R < 0.22$) in vegetation-covered areas. This showed that MPDI from NIR-Red space was only suitable for monitoring SM in bare soil. Due to the crops in the study area were in different growth stages, the NIR band's responses which characterize the vegetative state were varied. The accuracy of monitoring SM in vegetation areas by NIR-Red feature space introduced FVC is still insufficient [86].

MPDI was proposed by Ghulam *et al.* [19] to solve the shortage of monitoring SM in vegetation areas by PDI. The increase of SM and FVC will decrease the MPDI value, which means the land surface is moist. MSMMI works similarly to MPDI, which does not need a soil line. The correlation between MSMMI and SM in bare soil is the same as that of SMMI, and that in vegetation-covered areas is significantly higher than that of SMMI (Section 4.3). Liu *et al.* [26] compared MSMMI and MPDI using SPOT and Landsat8 images, which showed that MSMMI and MPDI from NIR-Red space had similar SM monitoring accuracy in the semi-arid region. Liu *et al.* [26] employed MODIS data to construct NIR-Red feature space and analyzed the effectiveness of SMMI, PDI, MSMMI, and MPDI in drought monitoring in the Guanzhong Plain of China. The results showed that there was a negative correlation between these indices and SM at a depth of 10 cm, and the correlation of SMMI was higher [87]. However, previous studies were only verified in the semi-arid areas with low vegetation coverage by medium–high resolution satellites, or the agricultural areas by low-resolution satellites [26], [87]. No studies employed Sentinel-2 to assess the accuracy of these indices in monitoring SM in farmland with crops and bare soil areas.

In this article, crop types are abundant and vegetation coverage includes low FVC to high FVC. Proved by practice, the correlation between MPDI and SM in bare soil is slightly lower than that of MSMMI, and that in vegetation-covered areas is much lower than that of MSMMI. MPDI is still not accurate enough for monitoring SM in vegetation-covered areas in the complex surface cover environment. It is more suitable to represent the drought condition of vegetation under a certain stress degree [23], [24]. This is mainly because MPDI requires a fixed parameter (soil line slope), which is susceptible to be changed in land cover types and leads to the decrease of monitoring effect [24], [26]. Moreover, there will be a large number of outliers in high vegetation-covered areas. MSMMI does not have the above limitations.

The estimation accuracy and the stability of SM under different land cover types monitored by MSMMI (from SWIR1–SWIR2 space) under FVCgr model were verified by using the LOOCV method. As shown in Fig. 15, the scattered points of MSMMI from SWIR1–SWIR2 space are evenly distributed on both sides of the diagonal of 1: −1 compared with SMMI from SWIR1–SWIR2 space [Fig. 10(b)]. The fitting line of sample points is closer to the diagonal and parallel to it. The nRMSE of all SM samples, vegetation samples and bare soil samples were 0.027, 0.038, and 0.021 cm^3/cm^3, respectively. In the same way, MSMMI and MPDI from SWIR1–SWIR2 spaces under different FVC estimation models on June 20, 2016 were calculated and analyzed with measured SM. The results showed

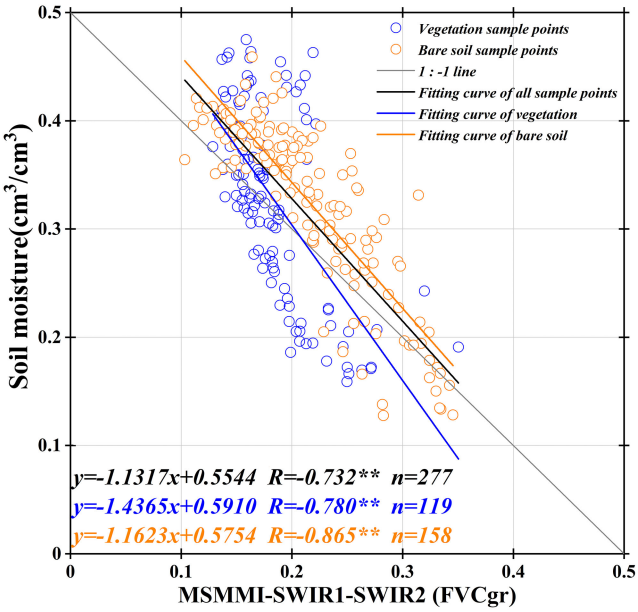


Fig. 15. Scatter plot of MSMMI from SWIR1–SWIR2 space and measured SM on June 13, 2016.

that the correlations between MSMMI from SWIR1–SWIR2 space based on FVCgr model and measured SM was still the highest ($R_{\text{all}} = -0.745$, $R_{\text{veg}} = -0.707$, and $R_{\text{soil}} = -0.897$).

As can be seen from Fig. 15, for vegetation-covered areas which measured SM ranging from 0.15 cm^3/cm^3 to 0.4 cm^3/cm^3, there is a significant negative linear correlation between MSMMI from SWIR1–SWIR2 space and measured SM. However, this index will significantly underestimate SM in vegetation-covered areas which measured SM greater than 0.4 cm^3/cm^3. This may be because when the SM on the underlying surface of the crops is saturated, the leaf water content will always remain within a certain range [86], [88]. Therefore, optical SM index may underestimate surface SM when crop growth is well developed (high FVC). In other words, SM on the underlying surface has a great influence on the SWIR reflectance of vegetation under low or medium FVC. This conclusion is consistent with the earlier findings of scholars [88], [89]. But MSMMI from SWIR1–SWIR2 space is still an effective index to monitor and quantify crop water shortages or extreme drought. It has great potential to be used to monitor timing irrigation in water-saving agriculture, completely independent of auxiliary data, and completely derived from optical remote sensing images.

As can be seen from Fig. 16, on the whole, there is no obvious relationship between FVC and measured SM for one phase. Although the FVC of most bare soil areas is less than 0.1, the measured SM is ranging from 0.1 to 0.5 cm^3/cm^3. In addition, the SM of most crop areas is greater than 0.3 cm^3/cm^3. Some crop areas have higher FVC, but the measured SM is between 0.1 and 0.25 cm^3/cm^3. This may be because the FVC of crops can remain at a high level for a period when the underlying surface is suffering from a short-term water shortage [90], [91]. Therefore, FVC obtained from a single temporal remote sensing image

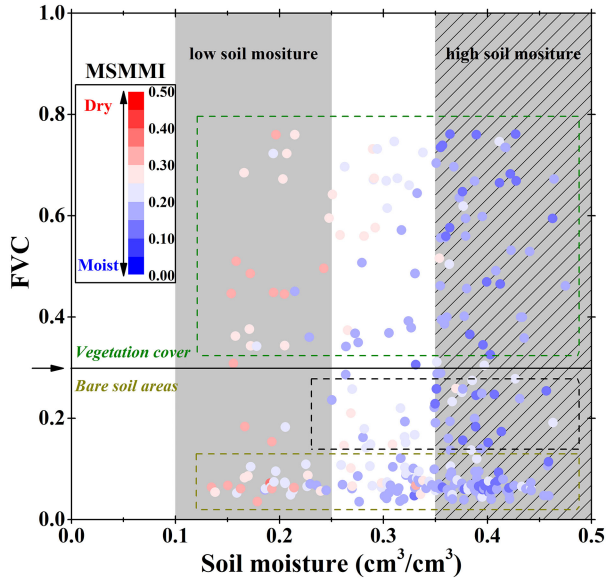


Fig. 16. Relationship between FVCgr measured SM and MSMMI.

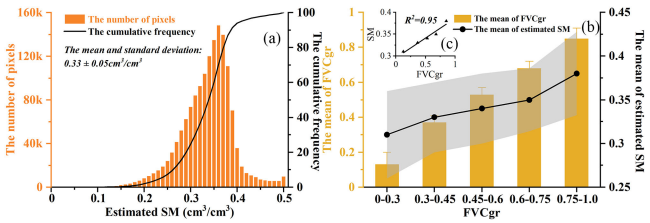


Fig. 17. (a) Frequency histogram of estimated soil moisture. (b) Mean of FVC and the mean of corresponding estimated soil moisture under different FVC conditions. (c) Correlation between the mean of FVC and estimated SM under different FVC conditions.

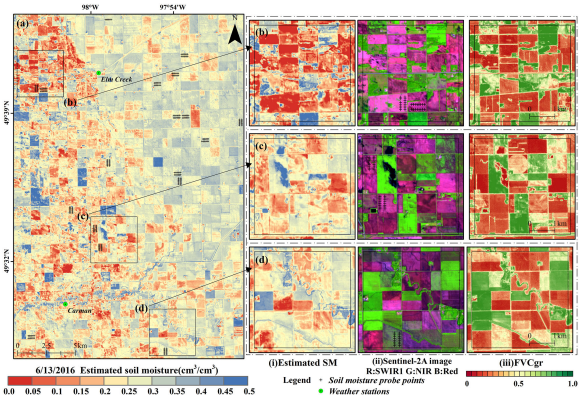


Fig. 18. (a) Estimated soil moisture map by MSMMI in the study area on June 13, 2016. (b)–(d) Magnified views of images. (i) Column: soil moisture map. (ii) Column: Sentinel-2A imagery. (iii) Column: FVCgr imagery.

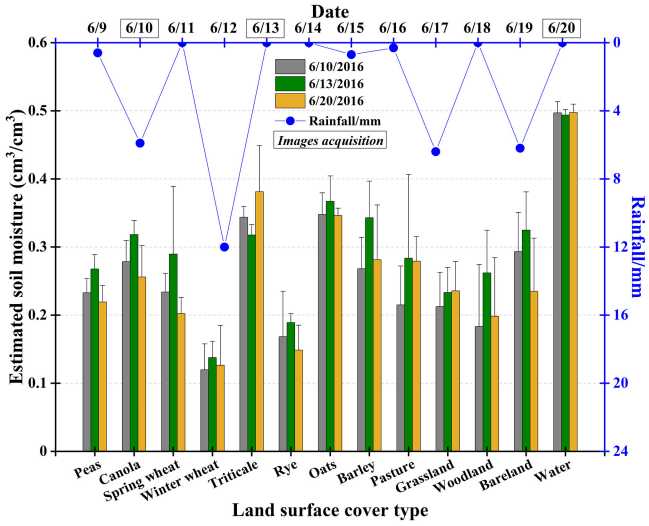


Fig. 19. (a) Estimated soil moisture map by MSMMI in the study area on June 13, 2016. (b)–(d) Magnified views of images. (i) Column: soil moisture map. (ii) Column: Sentinel-2A imagery. (iii) Column: FVCgr imagery.

cannot be used directly and timely to evaluate SM distribution in complex agricultural areas. MSMMI from SWIR1–SWIR2 space combined with FVC and SWIR bands which are sensitive to leaf moisture content or SM can effectively monitor the actual distribution of SM in both bare soil areas and vegetation-covered areas.

Combined with the scatters of the value of MSMMI and the corresponding measured SM (Fig. 15), SM levels of MSMMI from SWIR1–SWIR2 space were graded into six groups: 0.0 $<$ MSMMI $\leq$ 0.1 (extremely moist, $\geq$0.5cm^3/cm^3), 0.1 $<$ MSMMI $\leq$ 0.2 (moist, 0.4–0.5 cm^3/cm^3), 0.2 $<$ MSMMI $\leq$ 0.3 (normal, 0.3–0.4 cm^3/cm^3), 0.3 $<$ MSMMI $\leq$ 0.4 (mild drought, 0.2–0.3 cm^3/cm^3), 0.4 $<$ MSMMI $\leq$ 0.5 (drought, 0.1–0.2 cm^3/cm^3), and MSMMI $>$ 0.5 (extremely drought, 0–0.1 cm^3/cm^3). The SM in the study area on June 13, 2016 was estimated according to the linear fitting equation in Fig. 15. As shown in Fig. 17(a), the estimated SM by MSMMI was normally distributed on the whole, with the mean and standard deviation of 0.33 and 0.05 cm^3/cm^3, respectively. The SM of most pixels was between 0.3 and 0.4 cm^3/cm^3 and was in normal condition. The estimated FVC by FVCgr model was graded according to certain criteria [83], namely 0–0.3, 0.3–0.45, 0.45–0.6, 0.6–0.75, and 0.75–1.0. The mean and standard deviation of corresponding FVC were 0.13 $\pm$ 0.07, 0.37 $\pm$ 0.04, 0.53 $\pm$ 0.04, 0.68 $\pm$ 0.04, and 0.85 $\pm$ 0.06, respectively. Fig. 17(b) and (c) shows that, under different FVC conditions, the correlation between the mean FVC and the corresponding mean SM is extremely high ($R^2 > 0.95$). The greater the FVC, the greater the estimated SM. The mean and standard deviation of estimated SM under each level FVC were 0.31 $\pm$ 0.05, 0.33 $\pm$ 0.04, 0.34 $\pm$ 0.04, 0.35 $\pm$ 0.03, and 0.38 $\pm$ 0.05 cm^3/cm^3, respectively. Fig. 18(a) shows the estimated SM distribution by MSMMI from SWIR1–SWIR2 space. The extremely moist and moist areas were mainly located in the pond [Fig. 18(c)], rivers, woodland along the rivers [Fig. 18(d)], and part of crop areas. Drought and extremely drought areas were mainly in the northwest and central bare soil areas [Fig. 18(b)], which had been fallow for some time and had a reduced ability to retain SM. The estimated SM of bare soil areas in the northeast is relatively high. Fig. 19 shows that most crops were in mild drought conditions and tended to be normal.

TABLE III
EVALUATION OF SOIL MOISTURE MONITORING OF SOIL MOISTURE INDICES

Index \ Land cover	Vegetated areas	Bare soil	All	Evaluation
SMMI	Acceptable	Good	Acceptable	SWIR needed to be introduced
PDI	Acceptable	Good	Acceptable	
MSMMI	Good	Excellent	Good	SWIR needed to be introduced & Red edge bands needed to estimate FVC
MPDI	Acceptable	Good	Acceptable	

The estimated SM of water was always high and changed very little. Estimated SM of most crops showed a trend of rising first and then decreasing, which was mainly influenced by the precipitation of about 12 mm on June 12, 2016. Estimated SM of the triticale and oats plots was always relatively high and was in the normal region. SM of the winter wheat plots was always low, mainly because the crops were in the ripening stage and the water demand was relatively small. Therefore, combined with the estimated SM images and the actual crop growth period, it can accurately guide the irrigation of plots under soil water stress. The above results indicate that MSMMI (from SWIR1–SWIR2 space) based on the FVCgr model can effectively monitor the surface dry and wet conditions of different crops and can be used as the best SM index in the study area. Once again, we emphasize this SM index, whose calculation formulae are as follows: (18) shown at the bottom of this page,

$$\mathrm{FVCgr} = \frac{\mathrm{NDVI}_{\text{red edge}} - \mathrm{NDVI}_s}{\mathrm{NDVI}_v - \mathrm{NDVI}_s} \tag{19}$$

$$\mathrm{NDVI}_{\text{red edge}} = \frac{R_{\text{red edge3}} - R_{\text{red edge1}}}{R_{\text{red edge3}} + R_{\text{red edge1}}} \tag{20}$$

where R_{SWIR1} and R_{SWIR2} represent the SWIR1 and SWIR2 bands reflectance of Sentinel-2, respectively. $R_{v,\ \mathrm{SWIR1}}$ and $R_{v,\ \mathrm{SWIR2}}$ represent empirical values of pure vegetation in SWIR1 and SWIR2 images, respectively. FVCgr represents FVC based on the DPM with red-edge bands, $\mathrm{NDVI}_{\mathrm{red\ edge}}$ represents red-edge normalized difference vegetation index. NDVI_v and NVDI_s represent the value of pure vegetation and bare soil in $\mathrm{NDVI}_{\mathrm{red\ edge}}$ imagery, respectively. The values can be obtained from the cumulative confidence of the respective images at 1% and 99%.

V. OUTLOOK AND LIMITATIONS

It can be seen from this article that red-edge bands (Band5 and Band7 of Sentinel-2) and SWIR bands (Band11 and Band12 of Sentinel-2) can effectively improve the estimation accuracy of SM in farmland, which is especially obvious in vegetation-covered areas. Table III shows a qualitative evaluation of SM indices in monitoring SM according to the correlations with measured SM under different land cover and operability. SMMI, PDI, MSMMI, and MPDI can effectively monitor the dry and wet conditions in bare soil areas. Therefore, SM in bare soil areas can be retrieved by satellites carrying traditional four-band (Blue, Green, Red, and NIR) through these four indices. But the monitoring accuracy may not be as good as that of satellites carrying SWIR bands. Compared with visible and NIR bands, SWIR bands have longer wavelengths less affected by atmospheric scattering, which can significantly improve the ability to obtain surface SM and crop water content. MSMMI involving red-edge and SWIR bands can accurately estimate SM both in bare soil and vegetation-covered areas, which is suitable for satellites carrying red-edge and SWIR bands at the same time. However, at present, there are few satellites with these conditions. Of these, Sentinel-2 is the only open-access satellite, and Worldview-3 (the spatial resolution of multispectral and SWIR bands was 1.24 and 3.70 m, respectively) is the main commercial satellite [92]. China's GaoFen-6 has two red-edge bands with 16-m spatial resolution, but it lacks SWIR bands, which limits its application value to a certain extent [93]. Hyperspectral satellite sensors, such as China's ZiYuan-1 (ZY-1) 02D, GaoFen-5 (GF-5), and OuBiTe hyperspectral satellite-2/3 (OHS-2/3), have more detailed bands and higher spatial resolution, which are also effective data sources for monitoring SM in agricultural areas [94], [95].

The shortcoming of the spectral feature space method in monitoring SM is that many factors are affecting the surface reflection and emission spectra, such as atmospheric conditions, light conditions, soil structure, ST, soil surface roughness, organic matter content, etc., thus affecting the accuracy of the spectral eigenspace method in estimating SM [88], [96]. Besides, the response degree of different crops to underlying surface SM will also affect the accuracy and consistency of the spectral feature space method to estimate SM in vegetation-covered areas. In the future, the spectral feature space of long time series should be established according to measured SM data in different regions, different crop types, and deeper soil layers to further verify the reliability of the spectral feature space method in monitoring SM. The performance and sensibility of the spectral feature space method for monitoring crops or natural vegetation suffering extreme drought conditions will also need to be further validated. Moreover, SAR data has a certain penetration ability and is more sensitive to SM in vegetation-covered areas. It is also an interesting research direction to estimate SM in farmland by combining Sentinel-1 C-band SAR data and Sentinel-2's multispectral data [97]–[99].

$$\mathrm{MSMM}I = \frac{\sqrt{\left(R_{\mathrm{SWIR1}} - FVCgr \cdot R_{v,\mathrm{SWIR1}}\right)^2 + \left(R_{\mathrm{SWIR2}} - FVCgr \cdot R_{v,\mathrm{SWIR2}}\right)^2}}{\sqrt{2} \cdot (1 - \mathrm{FVCgr})} \tag{18}$$

Accurate monitoring of cultivated areas, species classification, nutrients, moisture content, diseases, and insect pests of crops are indispensable links for the modernization and intelligence of agriculture, which are also inseparable from remote sensing satellites. Compared with other fields, agricultural remote sensing monitoring has more complexity and particularity. Medium- and high-resolution remote sensing satellites with a red-edge and SWIR bands are more suitable for agriculture. These bands can effectively improve the diagnostic analysis accuracy of crop growth state and realize the fine monitoring of crop growth during the whole growth period. There will be an increasing number of agricultural applications of Sentinel-2 soon. In this article, we only employed the Pearson's correlation coefficient (R) to evaluate the performance of different indices, because R can directly reflect the linear relationship between the two variables, and can directly reflect the performance of different indices through the heat map (Figs. 9, 13, and 14). However, it cannot directly assess the accuracy of each SM estimation model in monitoring SM. We will adopt mean absolute error, mean relative error, root mean square error, and the leave-one-out cross-validation method to assess the stability of the estimation model in the next work.

VI. Conclusion

In this article, the potential of Sentinel-2 for crop extraction by RFC was explored. Second, this article assessed the performance of Sentinel-2 for estimating SM in bare soil areas and vegetation-covered areas by spectral feature space methods, namely SMMI, PDI, MSMMI, and MPDI.

The main conclusions are as follows:

1) For land cover classification, the red-edge and SWIR bands of Sentinel-2 are helpful for the accurate classification of complex land cover types. In the feature subset with the highest precision classification, red edge, SWIR bands, and their indices account for 87%. The spectral and vegetation indices important to the classification of land cover types mainly include the Sentinel-2 bands of SWIR1, SWIR2, Red, Green, Red edge1, Blue, and the derived vegetation indices such as MNDWI, NDMI, TCARI, S2REP (Sentinel-2 REPI), REPI, MRENDVI, red-edge vegetation index32, and MRESR. Only one texture feature (GLCM_mean) improves the classification accuracy, while the rest texture features at 20-m spatial resolution have little effect on the classification. Therefore, remote sensing satellites with a red-edge and SWIR bands are of greater application in the extraction of crop types in complex agricultural areas.
2) For SM estimation in bare soil areas, the importance score of the SWIR bands and measured SM is higher in RFR. For SM estimation in vegetation-covered areas, the red-edge bands scored higher. Both SMMI and PDI involving SWIR bands have universality for accurately monitoring SM in bare soil areas. The correlation between SMMI, PDI involving red-edge bands, and measured SM in vegetation-covered areas was significantly higher than that from the NIR-Red feature space but lower than that from the SWIR1–SWIR2 feature space. In general, SMMI is more accurate than PDI in monitoring SM in both bare soil and vegetated areas.
3) SMMI and PDI do not consider the impact of FVC, while FVC can better reduce the inconsistent response of different land cover to SM. Therefore, the accuracy in monitoring SM in vegetated areas is slightly lower. MSMMI and MPDI are the improvement of SMMI and PDI, respectively. The estimation of FVC is one of the important factors affecting the accuracy of SM estimation for the former two. The comparison results of different FVC models show that different FVC models can effectively estimate the vegetation cover of the study area. The correlation between FVC with red-edge bands and measured FVC was higher than that without red edge bands. FVC with red-edge bands can better enlarge the FVC difference of vegetation areas and weaken the FVC difference of bare soil areas. Also, the correlation between MSMMI based on FVC with red-edge bands and measured SM is higher. The improvement of the SM monitoring effect by red-edge bands is not directly caused by itself, but by introducing FVC with red-edge bands. The MSMMI based on the SWIR1–SWIR2 feature space and the FVCgr model (DPM involving red-edge bands) was the best SM index to monitoring regional SM in farmland. Whether vegetation-covered areas or bare soil areas or vegetation-bare soil mixed areas, the correlation between MSMMI with measured SM at a depth of 0–5 cm was the highest. MPDI has the worst monitoring performance. MSMMI has equivalent accuracy to SMMI and PDI in monitoring SM in bare soil areas, and significantly superior in vegetation-covered areas.

The upper soil is an area where crop growth and microbial activity are relatively active. MSMMI independent of the soil line can more accurately establish the general empirical model of regional SM in farmland. Sentinel-2 satellites with a red-edge and SWIR bands can map 0–5 cm depth SM at 10–20 m spatial resolution and provide the basis for agricultural water management with great potential.

Acknowledgment

The authors sincerely thank Copernicus Open Access Hub for providing Sentinel-2A images and NASA's Earth Observing System Data and Information System for providing the SMAPVEX16 field dataset. The authors are very grateful to every scientist who has contributed to the SMAPVEX16 dataset. The authors would also like to thank the anonymous reviewers whose comments have greatly contributed to the improvement of this study.

References

[1] N. Nicolai-Shaw *et al.*, "A drought event composite analysis using satellite remote-sensing based soil moisture," *Remote Sens. Environ.*, vol. 203, pp. 216–225, 2017.
[2] J. C. Price, "Thermal inertia mapping: A new view of the earth," *J. Geophys. Res.*, vol. 82, no. 18, pp. 2582–2590, 1977.
[3] J. C Price, "On the analysis of thermal infrared imagery: The limited utility of apparent thermal inertia," *Remote Sens. Environ.*, vol. 18, no. 1, pp. 59–73, 1985.

[4] I. Sandholt, K. Rasmussen, and J. Andersen, "A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status," *Remote Sens. Environ.*, vol. 79, no. 2–/3, pp. 213–224, 2002.

[5] R. B. Myneni, S. O. Los, and C. J. Tucker, "Satellite-based identification of linked vegetation index and sea surface temperature anomaly areas from 1982-1990 for Africa, Australia and South America," *Geophys. Res. Lett.*, vol. 23, no. 7, pp. 729–732, 2013.

[6] S. M. Quiring and S. Ganesh, "Evaluating the utility of the vegetation condition index (VCI) for monitoring meteorological drought in texas," *Agricultural Forest Meteorol.*, vol. 150, no. 3, pp. 330–339, 2010.

[7] P. Wang, J. Gong, and X. Li, "Vegetation-temperature condition index and its application for drought monitoring," *Geomatics Inf. Sci. Wuhan Univ.*, vol. 26, no. 5, pp. 141–143, 2001.

[8] C. Pathe, W. Wagner, D. Sabel, M. Doubkova, and J. B. Basara, "Using ENVISAT ASAR global mode data for surface soil moisture retrieval over oklahoma, USA," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 2, pp. 468–480, Feb. 2009.

[9] W. Dorigo *et al.*, "ESA CCI soil moisture for improved earth system understanding: State-of-the art and future directions," *Remote Sens. Environ.*, vol. 203, pp. 185–215, 2017.

[10] M. Hornacek *et al.*, "Potential for high resolution systematic global surface soil moisture retrieval via change detection using Sentinel-1," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1303–1311, Aug. 2012.

[11] J. Pan and T. Li, "Extracting desertification from landsat TM imagery based on spectral mixture analysis and albedo-vegetation feature space," *Natural Hazards*, vol. 68, no. 2, pp. 915–927, 2013.

[12] A. Ghulam *et al.*, "Sub-canopy soil moisture modeling in n-dimensional spectral feature space," *Photogramm. Eng. Remote Sens.*, vol. 77, no. 2, pp. 149–156, 2011.

[13] C. Wu *et al.*, "Soil moisture monitoring of vegetative area in farmland by remote sensing based on spectral feature space," *Trans. Chin. Soc. Agricultural Eng.*, vol. 30, no. 16, pp. 106–112, 2014.

[14] M. Sadeghi *et al.*, "The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to sentinel-2 and landsat-8 observations," *Remote Sens. Environ.*, vol. 198, pp. 52–68, 2017.

[15] E. Babaeian *et al.*, "Mapping soil moisture with the OPtical TRApezoid model (OPTRAM) based on long-term MODIS observations," *Remote Sens. Environ.*, vol. 211, pp. 425–440, 2018.

[16] A. J. Richardson and C. L. Wiegand, "Distinguishing vegetation from soil background information," *Photogramm. Eng. Remote Sens.*, vol. 43, no. 12, pp. 1541–1552, 1977.

[17] R. Fensholt and I. Sandholt, "Derivation of a shortwave infrared water stress index from MODIS near- and shortwave infrared data in a semiarid environment," *Remote Sens. Environ.*, vol. 87, no. 1, pp. 111–121, 2003.

[18] A. Ghulam, Q. Qin, and Z. Zhan, "Designing of the perpendicular drought index," *Environ. Geol.*, vol. 52, no. 6, pp. 1045–1052, 2007.

[19] A. Ghulam *et al.*, "Modified perpendicular drought index (MPDI): A real-time drought monitoring method," *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 2, pp. 150–164, 2007.

[20] L. Zhu *et al.*, "FY-3A-MERSI data in drought monitoring in northern china," *J. Remote Sens.*, vol. 14, no. 5, pp. 1004–1016, 2010.

[21] M. Amani *et al.*, "Two new soil moisture indices based on the NIR-red triangle space of landsat-8 data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 50, pp. 176–186, 2016.

[22] N. Chen, Y. He, and X. Zhang, "Nir-red spectra-based disaggregation of SMAP soil moisture to 250 m resolution based on Oznet in Southeastern Australia," *Remote Sens.*, vol. 9, no. 1, 2017, Art. no. 51.

[23] M. Rao *et al.*, "Mapping drought-impacted vegetation stress in California using remote sensing," *GISci. Remote Sens.*, vol. 54, no. 2, pp. 185–201, 2017.

[24] S. Zormand, R. Jafari, and S. S. Koupaei, "Assessment of PDI, MPDI and TVDI drought indices derived from MODIS aqua/terra level 1B data in natural lands," *Natural Hazards*, vol. 86, no. 2, pp. 757–777, 2017.

[25] Y. Liu *et al.*, "Comparison of SMMI, PDI and its applications in shendong mining area," *Earth Environ. Sci.*, vol. 57, no. 1, pp. 1–6, 2017.

[26] Y. Liu *et al.*, "Comparison and application of MPDI and MSMMI for drought monitoring in desert mining area," *Earth Environ. Sci.*, vol. 146, no. 1, 2018, Art. no. 012001.

[27] F. Canying, W. Lin, and X. U. Hanqiu, "A comparative study of different red edge indices for remote sensing detection of urban grassland health status," *J. Geo-Inf. Sci.*, vol. 19, no. 10, pp. 1382–1392, 2017.

[28] A. Kross *et al.*, "Assessment of rapideye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 34, no. 1, pp. 235–248, 2015.

[29] G. Kaplan and U. Avdan, "Evaluating the utilization of the red edge and radar bands from sentinel sensors for wetland classification," *Catena*, vol. 178, pp. 109–119, 2019.

[30] M. Immitzer, F. Vuolo, and C. Atzberger, "First experience with sentinel-2 data for crop and tree species classifications in central europe," *Remote Sens.*, vol. 8, no. 3, 2016, Art. no. 166.

[31] M. Belgiu and O. Csillik, "Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis," *Remote Sens. Environ.*, vol. 204, pp. 509–523, 2018.

[32] H. McNairn, K. Gottfried, and J. Powers, *SMAPVEX16 Manitoba Meteorological Data, Version 1*. Boulder, CO USA: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

[33] H. McNairn, K. Gottfried, and J. Powers, *SMAPVEX16 Manitoba Land Cover Classification Map, Version 1*. Boulder, CO USA: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

[34] H. McNairn, K. Gottfried, and P. J. Powers, *SMAPVEX16 Manitoba In Situ Vegetation Data, Version 1*. Boulder, CO USA: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

[35] H. McNairn, K. Gottfried, and J. Powers, *SMAPVEX16 Manitoba Leaf Area Index, Version 1*. Boulder, CO USA: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

[36] H. McNairn, K. Gottfried, and P. J. Powers, *SMAPVEX16 Manitoba Probe-Based In Situ Soil Moisture Data, Version 1*. Boulder, CO USA: NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

[37] C. J. Tucker, "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sens. Environ.*, vol. 8, no. 2, pp. 127–150, 1979.

[38] D. J. Major, F. Baret, and G. Guyot, "A ratio vegetation index adjusted for soil brightness," *Int. J. Remote Sens.*, vol. 11, no. 5, pp. 727–740, 1990.

[39] J. A. Schell and D. Deering, "Monitoring vegetation systems in the great plains with ERTS," *NASA Special Publication*, vol. 351, 1973, Art. no. 309.

[40] A. A. Gitelson and M. N Merzlyak, "Remote sensing of chlorophyll concentration in higher plant leaves," *Adv. Space Res.*, vol. 22, no. 5, pp. 689–692, 1998.

[41] A. Huete *et al.*, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1–/2, pp. 195–213, 2002.

[42] J. Qi *et al.*, "A modified soil adjusted vegetation index," *Remote Sens. Environ.*, vol. 48, no. 2, pp. 119–126, 1994.

[43] B. C Gao, "NDWI-A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.

[44] T. J. Jackson *et al.*, "Vegetation water content mapping using landsat data derived normalized difference water index for corn and soybeans," *Remote Sens. Environ.*, vol. 92, no. 4, pp. 475–482, 2004.

[45] H. Xu, "A study on information extraction of water body with the modified normalized difference water index (MNDWI)," *J. Remote Sens.*, vol. 9, no. 5, pp. 589–595, 2005.

[46] A. A. Gitelson, M. N. Merzlyak, and O. B. Chivkunova, "Optical properties and nondestructive estimation of anthocyanin content in plant leaves," *Photochem. Photobiol.*, vol. 74, no. 1, pp. 38–45, 2010.

[47] A. A. Gitelson *et al.*, "Assessing carotenoid content in plant leaves with reflectance spectroscopy," *Photochem. Photobiol.*, vol. 75, no. 3, pp. 272–281, 2010.

[48] B. Datt, "A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using eucalyptus leaves," *J. Plant Physiol.*, vol. 154, no. 1, pp. 30–36, 1999.

[49] D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote Sens. Environ.*, vol. 81, no. 2–/3, pp. 337–354, 2002.

[50] P. J. Curran, J. L. Dungan, and H. L. Gholz, "Exploring the relationship between reflectance red edge and chlorophyll content in slash pine," *Tree Physiol.*, vol. 7, no. 1–4, pp. 33–48, 1990.

[51] T. Majasalmi and M. Rautiainen, "The potential of sentinel-2 data for estimating biophysical variables in a boreal forest: A simulation study," *Remote Sens. Lett.*, vol. 7, no. 4–6, pp. 427–436, 2016.

[52] N. H. Broge and E. Leblanc, "Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density," *Remote Sens. Environ.*, vol. 76, no. 2, pp. 156–172, 2001.

[53] W. J. Frampton *et al.*, "Evaluating the capabilities of sentinel-2 for quantitative estimation of biophysical variables in vegetation," *ISPRS J. Photogramm. Remote Sens.*, vol. 82, pp. 83–92, 2013.

[54] J. Dash and P. J. Curran, "The MERIS terrestrial chlorophyll index," *Int. J. Remote Sens.*, vol. 25, no. 23, pp. 5403–5413, 2004.

[55] A. A. Gitelson, Y. Gritz, and M. N. Merzlyak, "Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves," *J. Plant Physiol.*, vol. 160, no. 3, pp. 271–282, 2003.

[56] A. A. Gitelson, G. P. Keydan, and M. N. Merzlyak, "Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves," *Geophys. Res. Lett.*, vol. 33, no. 11, pp. 431–433, 2006.

[57] C. S. T. Daughtry *et al.*, "Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance," *Remote Sens. Environ.*, vol. 74, no. 2, pp. 229–239, 2000.

[58] D. Haboudane *et al.*, "Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture," *Remote Sens. Environ.*, vol. 90, no. 3, pp. 337–352, 2004.

[59] B. Hua, F. L. Ma, and L. C. Jiao, "Research on computation of GLCM of image texture," *Acta Electronica Sinica*, vol. 1, no. 1, pp. 155–158, 2006.

[60] G. Beliakov, S. James, and L. Troiano, "Texture recognition by using GLCM and various aggregation functions," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, 2008, pp. 1–6.

[61] M. Belgiu and L. Drăguţ, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, 2016.

[62] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, "Random forest classification of multisource remote sensing and geographic data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2004, pp. 1049–1052.

[63] C. Wang, Z. Liu, and C. Yan, "An experimental study on imaging spectrometer data feature selection and wheat type identification," *J. Remote Sens.*, vol. 10, no. 2, 2006, Art. no. 249.

[64] R. Fensholt and S. R. Proud, "Evaluation of earth observation based global long-term vegetation trends—Comparing GIMMS and MODIS global NDVI time series," *Remote Sens. Environ.*, vol. 119, pp. 131–147, 2012.

[65] S. N. Goward *et al.*, "Normalized difference vegetation index measurements from the advanced very high-resolution radiometer," *Remote Sens. Environ.*, vol. 35, no. 2–3, pp. 257–277, 1991.

[66] C. Tucker *et al.*, "An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data," *Int. J. Remote Sens.*, vol. 26, no. 20, pp. 4485–4498, 2005.

[67] F. Baret, J. Clevers, and M. D Steven, "The robustness of canopy gap fraction estimates from red and near-infrared reflectances: A comparison of approaches," *Remote Sens. Environ.*, vol. 54, no. 2, pp. 141–151, 1995.

[68] T. N. Carlson and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol. 62, no. 3, pp. 241–252, 1997.

[69] G. Gutman and A. Ignatov, "The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models," *Int. J. Remote Sens.*, vol. 19, no. 8, pp. 1533–1543, 1998.

[70] A. R. Huete, R. D. Jackson, and D. F Post, "Spectral response of a plant canopy with different soil backgrounds," *Remote Sens. Environ.*, vol. 17, no. 1, pp. 37–53, 1985.

[71] B. M. Diaz and G. A. Blackburn, "Remote sensing of mangrove biophysical properties: Evidence from a laboratory simulation of the possible effects of background variation on spectral vegetation indices," *Int. J. Remote Sens.*, vol. 24, no. 1, pp. 53–73, 2003.

[72] Z. Jiang *et al.*, "Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction," *Remote Sens. Environ.*, vol. 101, no. 3, pp. 366–378, 2006.

[73] E. Christos and N Alexandre, "Red-edge normalized difference vegetation index (NDVI705) from sentinel-2 imagery to assess post-fire regeneration," *Remote Sens. Appl.: Soc. Environ.*, vol. 17, 2020, Art. no. 100283.

[74] F. Gascon and F. Ramoino, "Sentinel-2 data exploitation with ESA's sentinel-2 toolbox," EGUGA, 2017, Art. no. 19548.

[75] M. Amani and M. R. Mobasheri, "A parametric method for estimation of leaf area index using landsat ETM+ data," *GIScience Remote Sens.*, vol. 52, no. 4, pp. 478–497, 2015.

[76] A. Vehtari, A. Gelman, and J. Gabry, "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC," *Stats Comput.*, vol. 27, pp. 1–20, 2015.

[77] P. Thanh Noi and M Kappas, "Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery," *Sensors*, vol. 18, no. 1, 2018, Art. no. 18.

[78] L. Zhang *et al.*, "Wetland mapping of yellow river delta wetlands based on multi-feature optimization of sentinel-2 images," *J. Remote Sens.*, vol. 23, no. 2, pp. 313–326, 2019.

[79] A. Ghulam *et al.*, "A method for canopy water content estimation for highly vegetated surfaces-shortwave infrared perpendicular water stress index," *Sci. China Ser. D: Earth Sci.*, vol. 50, no. 9, pp. 1359–1368, 2007.

[80] W. Ma *et al.*, "Estimation of soil moisture in time-series dual-polarization SAR mining subsidence area," *J. Remote Sens.*, vol. 22, no. 03, pp. 521–534, 2018.

[81] S. Koley and C. Jeganathan, "Estimation and evaluation of high spatial resolution surface soil moisture using multi-sensor multi-resolution approach," *Geoderma*, vol. 378, 2020, Art. no. 114618.

[82] Y. Sun, Q. Qin, H. Ren, T. Zhang, and S. Chen, "Red-edge band vegetation indices for leaf area index estimation from Sentinel-2/MSI imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 826–840, Feb. 2019.

[83] L. Gao *et al.*, "Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 159, pp. 364–377, 2020.

[84] J. Verrelst *et al.*, "Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for sentinel-2 and-3," *Remote Sens. Environ.*, vol. 118, pp. 127–139, 2012.

[85] Y. Cai, M. Zhang, and H. Lin, "Estimating the urban fractional vegetation cover using an object-based mixture analysis method and sentinel-2 MSI imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 341–350, 2020.

[86] A. Ghulam *et al.*, "Estimating crop water stress with ETM+ NIR and SWIR data," *Agricultural Forest Meteorol.*, vol. 148, no. 11, pp. 1679–1695, 2008.

[87] Y. Liu *et al.*, "Comparison and application of drought remote sensing monitoring index in guanzhong plain," *Agricultural Res. Arid Areas*, vol. 36, no. 06, pp. 201–207, 2018.

[88] L. Wang *et al.*, "Sensitivity studies of the moisture effects on MODIS SWIR reflectance and vegetation water indices," *Int. J. Remote Sens.*, vol. 29, no. 24, pp. 7065–7075, 2008.

[89] H. Bach and W. Verhoef, "Sensitivity studies on the effect of surface soil moisture on canopy reflectance using the radiative transfer model GeoSAIL," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2003, pp. 1679–1681.

[90] J. O. Adegoke and A. M. Carleton, "Relations between soil moisture and satellite vegetation indices in the US corn belt," *J. hydrometeorol.*, vol. 3, no. 4, pp. 395–405, 2002.

[91] D. Zhang and G. Zhou, "Estimation of soil moisture from optical and thermal remote sensing: A review," *Sensors*, vol. 16, no. 8, 2016, Art. no. 1308.

[92] E. R. Hunt, Jr, C. S. T. Daughtry, and L Li, "Feasibility of estimating leaf water content using spectral indices from worldview-3's near-infrared and shortwave infrared bands," *Int. J. Remote Sens.*, vol. 37, no. 2, pp. 388–402, 2016.

[93] X. Jiang *et al.*, "Rice mapping and growth monitoring based on time series GF-6 images and red-edge bands," *Remote Sens.*, vol. 13, no. 4, 2021, Art. no. 579.

[94] H. Guo, D. Liang, and G. Liu, "Progress of earth observation in China," *Chin. J. Space Sci.*, vol. 40, no. 5, pp. 908–919, 2020.

[95] Y. Zhong *et al.*, "Advances in spaceborne hyperspectral remote sensing in China," *Geo-Spatial Inf. Sci.*, vol. 24, no. 1, pp. 95–120, 2021.

[96] J. Yue *et al.*, "Development of soil moisture indices from differences in water absorption between shortwave-infrared bands," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 216–230, 2019.

[97] Y. Liu, J. Qian, and H. Yue, "Combined Sentinel-1A with Sentinel-2A to estimate soil moisture in farmland," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1292–1310, 2021.

[98] M. El Hajj *et al.*, "Synergic use of sentinel-1 and sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas," *Remote Sens.*, vol. 9, no. 12, 2017, Art. no. 1292.

[99] Q. Gao *et al.*, "Synergetic use of sentinel-1 and sentinel-2 data for soil moisture mapping at 100 m resolution," *Sensors*, vol. 17, no. 9, 2017, Art. no. 1966.

Ying Liu received the Ph.D. degree in photogrammetry and remote sensing from China University of Mining and Technology, Beijing, China, in 2013.

She is currently an Associate Professor with Xi'an University of Science and Technology, Xi'an, China. Her research interests include remote sensing monitoring of the mine environment (multisource remote sensing and cooperative observation), remote sensing inversion of soil moisture (drought monitoring and drought prediction), and quantitative remote sensing (information extraction and spatial–temporal analysis of remote sensing data).

Jiaxin Qian received the B.S. degree in remote sensing science and technology from Nanjing University of Information Science and Technology, Nanjing, China, in 2018. He is currently working toward the M.S. degree in surveying engineering at the Xi'an University of Science and Technology, Xi'an, China.

His research interest includes the inversion of soil moisture and terrain deformation by optical and SAR images.

Hui Yue received the Ph.D. degree in environmental sciences from the China University of Mining and Technology, Beijing, China, in 2013.

He is currently a Lecturer with Xi'an University of Science and Technology, Xi'an, China. His research interests include remote sensing of the environment, extraction of remote sensing geoscience information, and monitoring and assessment of the ecological environment.