



Research papers

Applicability of three remote sensing based soil moisture variables for mapping soil organic matter in areas with different vegetation densities

Chenconghai Yang^a, Lin Yang^{a,b,*}, Lei Zhang^{a,d}, Feixue Shen^a, Di Fu^a, Shengfeng Li^a, Zhiqiang Chen^{e,f}, Chenghu Zhou^{a,c}

^a School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China

^b Frontiers Science Center for Critical Earth Material Cycling, Nanjing University, Nanjing 210023, China

^c State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

^d Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

^e Key Laboratory for Humid Subtropical Eco-geographical Processes of the Ministry of Education, Fujian Normal University, Fuzhou 350117, China

^f School of Geographical Sciences, Fujian Normal University, Fuzhou 350117, China

ARTICLE INFO

This manuscript was handled by Dr Y Huang, Editor-in-Chief, with the assistance of Jan Vanderborght, Associate Editor

Keywords:

Soil organic matter
Soil moisture
NSDSIs
OPTRAM
OPTRAM-NSDSI
Digital soil mapping

ABSTRACT

Obtaining accurate spatial information on soil organic matter (SOM) is crucial for understanding global carbon cycle. Digital soil mapping (DSM) has become an effective method for mapping SOM, in which selection of influential environmental covariates plays an important role. Soil moisture (SM) can serve as a potential covariate, especially it can be estimated at large spatial scales thanks to remote sensing. The normalized shortwave-infrared difference bare soil moisture indices (NSDSIs) based on Landsat SWIR bands generated at bare soil period has been employed in SOM mapping previously. However, soil is usually covered by vegetation, it is thus necessary to develop new SM indices applicable to areas covered with vegetation, and examine how SM indices perform in areas with different vegetation densities. In this paper, we developed a new SM index by introducing NSDSIs to the Optical TRapezoid Model (OPTRAM-NSDSI), and compared it with the original OPTRAM with the shortwave infrared transformed reflectance (OPTRAM-STR), as well as NSDSIs. SM indices were generated across two study areas, i.e. Zhuxi, Fujian (104 samples and 43.93 km² with forestland and farmland as main land uses) and Heshan, Heilongjiang (106 samples and 60 km² with primarily farmland) in China. The Integrated Nested Laplace Approximation with the Stochastic Partial Differential Equation approach was utilized as the SOM prediction model. The results suggest that adding SM variables into the commonly-used environmental covariates improves the prediction accuracies. The highest accuracy improvement of 26.8% in terms of Lin's concordance correlation coefficient in Zhuxi is obtained by NSDSIs, and the highest improvement of 56.7% in Heshan is obtained by OPTRAM-NSDSI. This may indicate that OPTRAM-NSDSI is more effective in areas with higher vegetation densities while NSDSIs in areas with lower densities. Furthermore, the optimal image dates for SM estimation are probably at the vegetation "green-up" stage. This study provides a reference for using SM information to improve SOM mapping in areas covered with vegetation.

1. Introduction

Soil organic matter (SOM) provides nutrients for plant growth (Yan et al., 2023), contributes to the maintenance of soil security (Wang et al., 2022), and is of significant importance in the global carbon and water cycle. Accurately mapping the spatial distribution of SOM is crucial for soil carbon accounting and land resource management (Lal, 2004; Lal, 2020; Sanderman and Baldock, 2010). With the development of

geographic information systems, land surface data acquisition technology, and machine learning models, digital soil mapping (DSM) has recently become an efficient and cost-effective method for predicting the spatial distribution of SOM (Arrouays et al., 2021; Huang et al., 2022; McBratney et al., 2003; Minasny and McBratney, 2016). When employing DSM, the relationships between the target soil properties and its environmental covariates established based on soil sample points are utilized to predict the spatial distribution of the soil properties

* Corresponding author at: School of Geography and Ocean Science, Nanjing University, Nanjing 210023, China.

E-mail address: yanglin@nju.edu.cn (L. Yang).

(McBratney et al., 2003). Due to the high spatial heterogeneity of soil, generating accurate and reasonable soil maps is a difficult task, therefore, ongoing research has been directed toward improving the DSM performance. One direction is to develop effective environmental covariates capable of indicating the spatial variability of the target soil property because environmental covariates determine the quality of the established soil-environment relationships and then map results and accuracies to a large extent (McBratney et al., 2003).

Soil moisture (SM) is a crucial variable in the water cycle and energy exchange between the Earth's surface and atmosphere (Li et al., 2021; Robinson et al., 2008), regulates essential hydrological processes such as evaporation, infiltration, and runoff (Babaeian et al., 2018), thereby influencing crucial physiological processes in soils (Ambrosone et al., 2020). Several experiments have shown that SM has a significant impact on SOM by influencing its dissolution, mineralization, synthesis, material migration, and physical protection (Krull et al., 2003; Yoshida et al., 2018). SM variables that are related to SOM could serve as predictors for mapping SOM. Yet, the traditional way to obtain SM conditions of an area by sampling is usually costly, especially for a large area, and thus not suitable for application in DSM (Yang et al., 2023a).

Compared with labor-intensive, range-limited and time-consuming ground-based point measurements, remote sensing is effective for characterizing and monitoring SM information at a large spatial scale (Chen et al., 2020; Li et al., 2021). However, the utilization of remote sensing-based SM covariates in DSM remains very limited. Various sensors, such as passive microwave, thermal, and optical remote sensing sensors, have been utilized to capture SM information on a large scale (Li et al., 2021; Susha Lekshmi et al., 2014). Passive microwave remote sensing provides accurate observations of SM, but its spatial resolution, usually 25 km (Li et al., 2021), is generally coarse for soil mapping. Thermal remote sensing relies on the sensitivity of soil to apparent thermal inertia to obtain information about the surface water status (Li et al., 2021; Sadeghi et al., 2015; Sandholt et al., 2002), but the generated SM information is heavily influenced by atmospheric factors, such as near-surface air temperature (Sandholt et al., 2002). Optical remote sensing has also been used to estimate SM information, as demonstrated by Sadeghi et al. (2017) and Yue et al. (2019). Sadeghi et al. (2015) introduced the concept of shortwave infrared transformed reflectance (STR) and reported a linear relationship between STR and SM. Yue et al. (2019) introduced the normalized shortwave-infrared difference bare soil moisture indices (NSDSIs), which makes use of the short wavelength infrared (SWIR) band with a spatial resolution of 10 m. The optical remote sensing approach provides a means of acquiring SM information with relatively high spatial resolution, such as using Landsat-8 (30 m) or Sentinel-2 images (10 m). Due to those advantages of the optical remote sensing generated SM indices, Yang et al. (2023a) utilized NSDSI indices for SOM mapping and proved their effectiveness in improving SOM mapping accuracies compared with commonly-used climate and topography variables. However, their study only validated the indices in one study area during winter (mid-February), which represented a bare soil condition (Yang et al., 2023a). It is unknown how the NSDSI indices perform under different levels of vegetation coverage since the soil is often covered by vegetation.

In the case of soil covered by vegetation, researchers have developed trapezoidal (triangular) models to estimate SM with vegetation cover (Babaeian et al., 2018; Ma et al., 2022; Sadeghi et al., 2017). Thermal Inertia derived from land surface temperature (LST) is highly correlated with SM (Zhang and Zhou, 2016). By investigating the correlation between LST and vegetation indices (VI), researchers found that scatterplots of LST and VI typically form triangular shapes (Carlson et al., 1994; Price, 1990). When the scatterplot represents the full range of fractional vegetation cover and SM contents, it tends to be trapezoidal (Moran et al., 1994a; Sandholt et al., 2002). Based on this observation, the Thermal Optical Trapezoid Model (TOTRAM) has been developed to estimate SM based on the distance of each pixel to the bottom (wet edge) and top (dry edge) of the pixel distribution within the LST-VI space

(Sadeghi et al., 2017; Sandholt et al., 2002). However, TOTRAM requires images obtained from optical and thermal remote sensing bands. Additionally, individual parameterization and calibration are required for each remote sensing image at observation dates, which may limit its practical application (Ma et al., 2022; Sadeghi et al., 2017). To address these limitations, Sadeghi et al. (2017) proposed the Optical TRApezoID Model (OPTRAM) for SM estimation using STR instead of LST (Longo-Minnolo et al., 2022; Sadeghi et al., 2017). OPTRAM is suitable for images with only optical bands, such as Sentinel-2. Previous studies have applied OPTRAM to estimate SM in areas with varying vegetation densities (Ma et al., 2022), mixed vegetation types including forests and grasslands (Chen et al., 2020), and as an effective indicator for agricultural drought (Ambrosone et al., 2020; Hassanpour et al., 2020). It suggests that OPTRAM may have broader seasonal applicability and is effective in estimating SM information in the presence of vegetation coverage on soil.

In the original OPTRAM, a single spectral reflectance index, i.e. STR, is adopted to construct relationships for SM estimation. However, one single soil spectral reflectance method can be affected by various factors, such as particle size distribution, mineral composition, sun angle, viewing angle, terrain, and instrumental errors (Galvão and Vitorello, 1998; Jacquemoud, 1993; Yue et al., 2019). Differently, the NSDSIs represent the normalized transformation of two spectral reflectances, which can mitigate the influence of various factors on single soil reflectance (Yue et al., 2019). It has also demonstrated that the correlations between NSDSIs and SM were higher than those of STR. Therefore, using NSDSIs instead of STR within the trapezoid model, called OPTRAM-NSDSI, may provide a more reliable approach for SM estimation in areas with vegetation covering.

In this study, we developed a new model, OPTRAM-NSDSI, for regions with vegetation coverage, and compared it with the original OPTRAM with STR and the NSDSI indices for SOM mapping. The experiments were conducted in two study areas, Zhuxi in Fujian Province and Heshan in Heilongjiang Province, China, characterized by varying vegetation densities in different climate zones and land use. The objectives of this study are: (1) to examine whether adding the SM indices acquired from the proposed OPTRAM-NSDSI into the commonly-used environmental covariates could improve the SOM mapping accuracy; (2) to compare the improvement in SOM mapping accuracy using OPTRAM-NSDSI estimated SM versus the original OPTRAM estimated SM (hereafter called OPTRAM-STR), and NSDSIs; (3) to determine the optimal image dates for the three types of SM variables in the two case study areas with different vegetation densities. This study aims to establish a framework for the generation of SM variables to improve SOM mapping, especially in areas with vegetation coverage.

2. Study area and data

2.1. Study areas and soil sampling

The two study areas are the Zhuxi watershed (located in Changting County of Fujian Province, hereafter referred to as Zhuxi) and Heshan farm (located in Nenjiang County of Heilongjiang Province, hereafter referred to as Heshan). Zhuxi has a subtropical monsoon climate and diverse land use with forestland and farmland as the main types. Conversely, Heshan is predominantly characterized by a cold temperate monsoon climate and generally homogeneous farmland. Although the farmland in Heshan would be bare in winter, the vegetation density of Heshan during the lush vegetation period (with the highest average fractional vegetation cover (FVC) value of 0.76 on Sept. 16th) is higher than that of Zhuxi (with the highest average FVC value of 0.56 on Aug. 25th). The specific details of the two study areas are as follows.

2.1.1. The Zhuxi watershed study area

The Zhuxi watershed, with an area of 43.93 km², is located at 25°38'15"-25°42'55" N and 116°23'30"-116°30'30" W. The main river

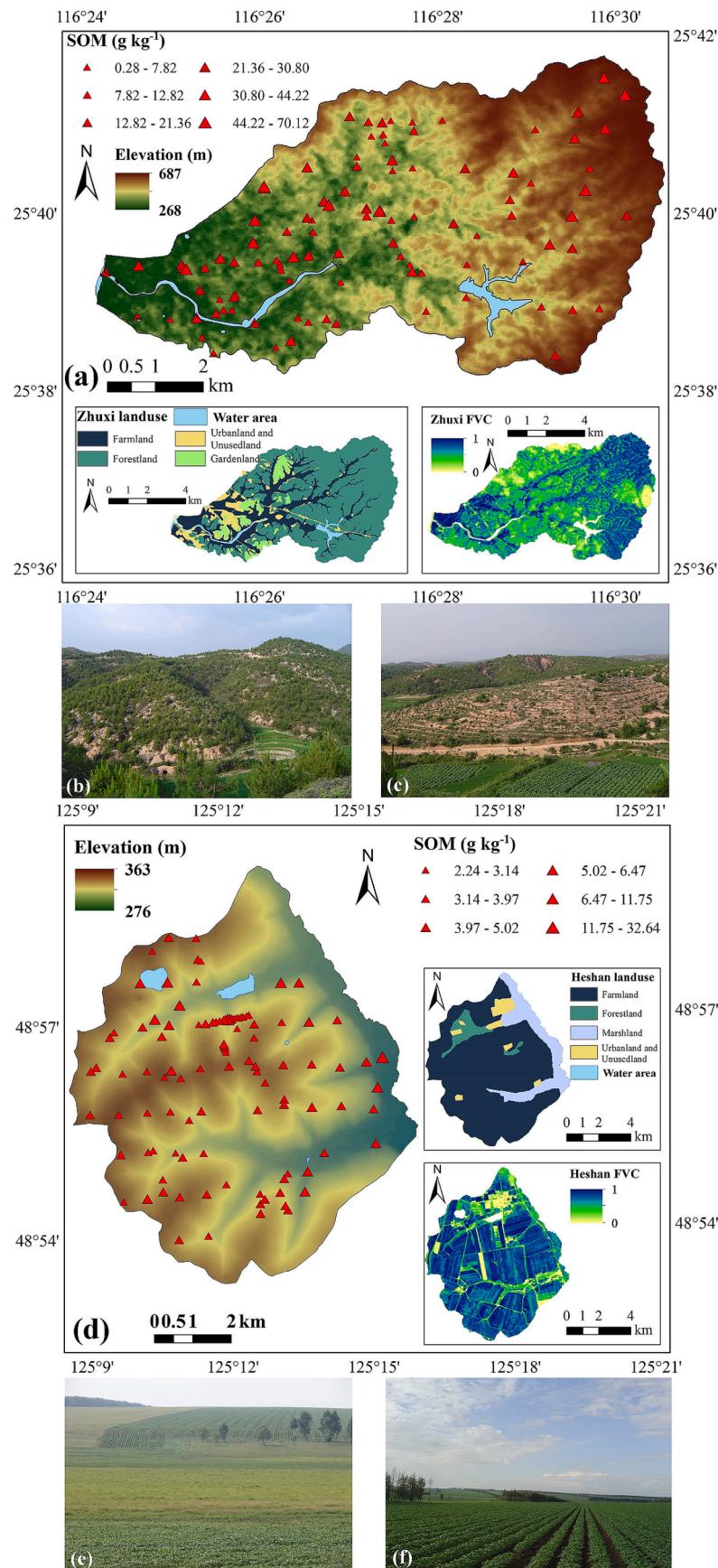


Fig. 1. The DEM, land use, and FVC from Landsat-5 images (on Sept. 16th of Zhuxi and Aug. 25th of Heshan, respectively, the growth peak stage), and soil samples of Zhuxi (a) and Heshan (c). FVC: fractional vegetation cover. (b) and (c), (e) and (f) demonstrate the photographs of field investigations of Zhuxi (b, c) and Heshan (e, f).

within the watershed, Zhuxi, extends over 13.93 km. The region exhibits a subtropical monsoon humid climate, featuring warm and humid springs and summers, as well as cold and dry winters and autumns. It experiences a prolonged summer season and a brief winter season. The mean annual temperature is 18.4 °C, and annual precipitation varies from 1400 to 2450 mm, with the majority occurring between March and July. The topography of the watershed is characterized by hills in the northeast and more gentle areas in the west. The elevation varies between 268 and 687 m, with an average slope of 11.8°. Zhuxi exhibits a variety of land use types, with forestland representing the largest proportion at 63.8% of the area, followed by farmland at 22.9%, urbanland and unusedland at 6.8%, and gardenland at 6.5% (Fig. 1a). This area exhibited serious soil erosion due to logging and steep slopes before the 1990s. Soil and water conservation and vegetation restoration projects have been conducted since 1995, turning the zonal evergreen broad-leaved forest into secondary and plantation forests, with *Pinus massoniana* as the primary species in forestland (Chen et al., 2019). The main local crop on the farmland is rice. Among all the dates, the highest average value of FVC of the whole study area on Sept. 16th is 0.56, and the range of NDVI at this date is -0.44 to 0.70 with an average of 0.31. Skeletal red soil and Percogenic paddy soil in Chinese Soil Taxonomy system are the dominant soil types in this watershed, with Skeletal red soil encompassing 61.4% of the total area.

2.1.2. The Heshan farm study area

Heshan Farm covers an area of approximately 60 km², which is located at 48°53'9"-48°59'20" N and 125°8'22"-125°16'21" W. The local climate is characterized by a cold temperate monsoon climate, with an average annual temperature ranging from -1.4 to 0.8 °C and an average annual precipitation ranging from 500 to 600 mm. This study, with elevations ranging from 276 to 363 m and an average slope gradient of less than 4°, has a very gentle terrain (hilly area) with long slopes and a very wide flat valley in the east. The parent materials are mainly silt loam loess over the whole area except in the valley bottom, which is mainly occupied by fluvial deposits. Different from the Zhuxi watershed with diverse land uses, the main land use of the Heshan study area is farmland, accounting for 76.8% of the total study area, with soya bean and maize as main crops, followed by marshland (14.5% of the total area) (Fig. 1b). Among all the dates, the highest average value of FVC of the whole study area on Aug. 25th is 0.76, and the range of NDVI at this date is -0.37 to 0.74 with an average value of 0.57. In the Chinese soil taxonomy system, the main soil types in this study area include Cambosols, Isohumosols, and Gleyosols (Zhang et al., 2021).

2.2. Soil sample data

A soil census sampling was carried out in 2002 in the Zhuxi watershed, and 106 soil samples in 0 ~ 20 cm depths were collected with a sampling strategy based on the Soil and Terrain Digital Database (SOTER) model (Oldeman and van Engelen, 1993). Finally, 49.0% of the samples were located in forestland, 28.3% in gardenland, 19.8% in farmland. SOM (g kg⁻¹) of the samples was measured using the potassium dichromate-concentrated sulphuric acid oxidation method.

At Heshan, 104 soil samples were collected from 2004 to 2005 at A horizon with three sampling strategies: systematic sampling, integrative hierarchical stepwise sampling and transect sampling (Yang et al., 2013; Zeng et al., 2016; Zhu et al., 2010), and SOM (g kg⁻¹) was measured using the dichromate oxidation method (external heat applied). 86.5% of the samples were located in farmland, while the remaining 5.8% and 7.7% samples were obtained from forestland and marshland, respectively.

The original left-skewed SOM data were converted to a normal distribution through logarithmic transformation for model building and transformed back to their original values for accuracy assessment.

2.3. Environmental covariates for SOM mapping

Commonly-used environmental covariates are those currently prevalent in the DSM, including climate, topography, parent material and land use. According to previous studies (Yang et al., 2023a), the soil-forming factors in Zhuxi study area are complex, thus topography, soil parent material, climate and land use were chosen as the commonly-used environmental covariates. The following environmental covariates were generated, namely elevation (ELEV), slope (SLP), topographic wetness index (TWI), soil type (SType) (representing soil parent material), annual mean temperature of the 2000–2010 year (TEM), annual mean precipitation of the 2000–2010 year (PRE), land use (LU). Topographic covariates, such as elevation, slope, and TWI, were obtained from the Digital Elevation Model (DEM). The DEM was created by vectorizing the contour map derived from the 1:10,000 topographic map to construct the Triangulated Irregular Network and subsequently generate the DEM for the watershed. The terrain analysis toolbox in ArcGIS was employed for the generation of the three topographic variables. The soil type data was obtained from the 1:50,000 Soil Map of Changting County. The soil type data can represent soil parent materials as well. Climate covariates were acquired from the Resource and Environment Science and Data Center (RESDC, <https://www.resdc.cn/>). The data from RESDC were produced by ANUSPLIN using interpolation of daily meteorological data from more than 2400 stations nationwide (Hutchinson, 1998). The original 1 km data of the two climate variables were resampled to a spatial resolution of 30 m, consistent with DEM, using the Cubic resampling method in ArcGIS. The land use is produced through image interpretation based on SPOT5 satellite images taken during the same period as soil sampling, and ground truthing and patch-by-patch verification verified its accuracy (Chen, 2011).

In Heshan, the macroclimate, soil parent lithology, and vegetation conditions are generally uniform, with topography exerting a dominant influence on soil formation (Zeng et al., 2016; Zhang and Zhu, 2019; Zhang et al., 2021). Consequently, we employed five topographic variables, ELEV, SLP, TWI, plan curvature (plan), profile curvature (prof), and land use (LU). We employed a 1:10,000 topographic map issued by the Chinese Bureau of Surveying and Mapping (1987) to produce a digital elevation model with a 10 m resolution. The DEM was generated using the TOPOGRID and TINLATTICE functions in Arc/Info, as described by Zhang et al. (2021). The topographic environmental covariates were derived from this DEM using the terrain analysis toolbox in ArcGIS. The land use of Heshan in 2005 was obtained from the Chinese land use remote sensing monitoring data set of RESDC (<https://www.resdc.cn/DOI/doi.aspx?DOIid=54>) with an original resolution of 30 m.

In addition to the above-mentioned environmental covariates, we also adopted NDVI in SOM mapping. NDVIs generated from remote sensing images have been increasingly used in soil mapping and considered as an effective predictor for mapping SOM (Wang et al., 2023a; Wu et al., 2021). Thus, we also evaluate the effectiveness of the proposed SM variables by comparing them with NDVIs. The NDVIs were generated from Landsat-5 images at the date with the strongest correlation with the measured SOM selected for each of the two areas.

2.4. Remote sensing data for calculating SM indices

The Landsat-5 satellite images in the year of soil sampling were utilized to extract NSDSIs and estimate SM with OPTRAM-NSDSI and OPTRAM-STR. The images were acquired from the USGS Earth Explorer platform (<https://earthexplorer.usgs.gov/>). For Zhuxi, all Landsat imagery from the entire year of 2002 was acquired, whereas imagery from the end of March to the end of October was obtained for Heshan. Heshan was covered with snow from November to early March, making optical remote sensing unable to extract SM information during this period. For those images affected by cloud contamination, a replacement image from the same date or a nearby date in the neighboring year was utilized

to ensure a comprehensive time series for the entire growing seasons (Zhang and Zhou, 2016). Table 1 displays the dates of the images utilized for the two study areas.

3. Methodology

3.1. The overall framework of the method

The methodology employed in our study is depicted in the flow chart Fig. 2. We utilized five types of environmental covariates: commonly-used environmental covariates (C), NDVI, normalized shortwave-infrared difference bare soil moisture indices (NSDSIs), normalized shortwave-infrared difference bare soil moisture indices with Optical TRApezo Model estimated SM (OPTRAM-NSDSI) and shortwave infrared transformed reflectance with Optical TRApezo Model estimated SM (OPTRAM-STR). The covariates were combined into four combinations to assess their ability in SOM prediction: "C", "C + NDVI", "C + NSDSIs", "C + OPTRAM-NSDSI" and "C + OPTRAM-STR". The mapping accuracies obtained from different covariate combinations were compared for each date to determine the most effective image date.

3.2. Normalized shortwave-infrared difference bare soil moisture indices (NSDSIs)

The NSDSI indices for estimating SM, proposed by Yue et al. (2019), is a normalized transformation of two spectral reflectances based on the difference in water absorption in the short wavelength infrared (SWIR) band. As the SM increases, the reflectance of the soil decreases (Twomey et al., 1986; Yue et al., 2019), leading to a darker appearance of the wet soil (Sadeghi et al., 2015). The reflectance in the SWIR band exhibits a strong sensitivity to variations in SM, (Sadeghi et al., 2015; Tian et al., 2021). Due to water absorption, there are significant differences in soil reflectance across various SWIR bands, such as SWIR1 (1.55 ~ 1.75 mm) and SWIR2 (2.08 ~ 2.35 mm). Since the observed difference exhibits a linear correlation with the SM, Yue et al. (2019) introduced the NSDSIs as a means to effectively represent SM information. The calculation formula for NSDSIs is presented in Table 2. In this study, the SM indices based on NSDSIs include the NSDSI1, NSDSI2 and NSDSI3 (Table 2).

3.3. Original Optical TRApezo model (OPTRAM-STR) for SM estimation

The trapezoidal (triangular) model is commonly used to estimate SM with vegetation cover (Babaeian et al., 2018; Ma et al., 2022; Sadeghi et al., 2017). The Optical TRApezo Model (OPTRAM) was initially proposed by Sadeghi et al. (2017) and depicts a physically based trapezoidal space representing the pixel distribution within the STR-NDVI space (Chen et al., 2020). OPTRAM assumes that SM status affects vegetation growth and that changes in SM content lead to changes in the spectral characteristics of vegetation, such that STR and VI are linearly related and STR-VI feature space forms a trapezoid shape (Ambrosone et al., 2020; Chen et al., 2014). When the study area encompasses a wide range of SM and fractional vegetation cover, the aggregation of all pixels will form a geometric shape, such as a triangle or trapezoid, each with distinct physical implications (Carlson et al., 1995; Moran et al., 1994b; Wang et al., 2023b). Considering a linear relationship between soil saturation degree, W (0 for completely dry and 1 for saturated soil) and STR can be expressed as (Sadeghi et al., 2017; Sadeghi et al., 2015):

$$STR = \frac{(1 - B_{SWIR})^2}{2B_{SWIR}} \quad (1)$$

$$W(OPTRAM - STR) = \frac{\theta - \theta_d}{\theta_w - \theta_d} = \frac{STR - STR_d}{STR_w - STR_d} \quad (2)$$

where B_{SWIR} represent the spectral reflectance of SWIR1 (1.55 ~ 1.75

μm) or SWIR2 (2.08 ~ 2.35 μm) wave bands in Landsat-5 imagery, STR is the shortwave infrared transformed reflectance of SWIR1 or SWIR2, W is the local normalized SM content by the local minimum dry SM content θ_d and local maximum wet SM content θ_w , STR_d and STR_w are the STR at θ_d and θ_w , representing the STR in dry and wet soil. OPTRAM-STR1 or OPTRAM-STR2 is generated using SWIR1 or SWIR2, respectively. STR_d and STR_w are obtained from the specific STR-VI trapezoidal space and determined by the following equation (Mokhtari et al., 2023; Sadeghi et al., 2015):

$$STR_d = i_{d1} + s_{d1} VI \quad (3)$$

$$STR_w = i_{w1} + s_{w1} VI \quad (4)$$

where i_{d1} and s_{d1} are the intercept and the slope of the dry edge, i_{w1} and s_{w1} are the intercept and the slope of the wet edge, in OPTRAM-STR. Based on the equation (3) and equation (4), the equation (2) can be transformed into the following form:

$$W(OPTRAM - STR) = \frac{i_{d1} + s_{d1} VI - STR}{i_{d1} - i_{w1} + (s_{d1} - s_{w1}) VI} \quad (5)$$

Based on OPTRAM-STR, we generated two OPTRAM-STR covariates, OPTRAM-STR1 and OPTRAM-STR2. Fig. 3 illustrates the schematic diagram for OPTRAM-STR and OPTRAM-NSDSI. The four vertices of the trapezoid correspond to the following conditions: dry bare soil, saturated bare soil, water-stressed vegetation, and well-watered vegetation.

For OPTRAM-STR and OPTRAM-NSDSI, previous studies have commonly utilized NDVI as the VI on the x-axis of trapezoidal models (Mokhtari et al., 2023; Tang et al., 2010). Nevertheless, NDVI values frequently fail to reach 0 in regions with bare soil or 1 in densely vegetated areas. To confine the x-axis between 0 and 1, FVC instead of NDVI in trapezoidal space in our study. The pixel dichotomy model was employed for the estimation of FVC (Gao et al., 2020):

$$NDVI = \frac{B_{NIR} - B_{Red}}{B_{NIR} + B_{Red}} \quad (6)$$

$$FVC = \frac{NDVI - NDVI_{soil}}{NDVI_{veg} - NDVI_{soil}} \quad (7)$$

where B_{NIR} and B_{Red} are the spectral reflectance of NIR (0.63 ~ 0.69 μm) and Red (0.76 ~ 0.90 μm) wave bands in Landsat-5, the $NDVI_{veg}$ and $NDVI_{soil}$ are the NDVI of complete vegetation cover areas (99.5 percent quartile of NDVI) and bare soil areas (0.5 percent quartile of NDVI), respectively.

In constructing a trapezoidal model, it is essential to ensure alignment between the wet edge and dry edge (Babaeian et al., 2018; Ma et al., 2022). In the majority of studies (Babaeian et al., 2018; Chen et al., 2020; Ma et al., 2022; Sadeghi et al., 2017), the differentiation between dry and wet edges relied on visual interpretation, a method that may introduce subjectivity. In addressing this issue, our study employed a more automated and objective method to extract the wet and dry edges, as outlined by Tang et al. (2010). Initially, the STR-VI and NSDSI-VI space was partitioned based on the distribution of image elements. The STR and NSDSI were specifically divided into twenty intervals according to the FVC values. Each interval was subsequently divided into five smaller intervals. The maximum and minimum STR or NSDSI values were calculated for each interval. The wet and dry edges were then generated through linear regression between STR or NSDSI and NDVI, then the modeled maximum STR or NSDSI value based on the regression corresponds to the wet edge and the minimum STR or NSDSI value corresponds to the dry edge, respectively. This approach not only reduces human subjectivity in wet and dry edge determination, but also decreases labour costs. The SM value within the trapezoidal space can be calculated by using the intercepts and slopes of the wet and dry edges in the trapezoidal space fit.

Table 1

Date of the Landsat-5 images used for the two study areas.

Study area	Image date
Zhuxi	Jan. 03rd, Jan. 08th, Feb. 05th, Feb. 15th, Apr. 20th, May. 13rd, May. 28th, Jun. 04th, Jun. 27th, Jul. 09th, Jul. 30th, Aug. 10th, Sept. 16th, Nov. 22nd, Dec. 13rd, Dec. 29th
Heshan	Mar. 23rd, Apr. 03rd, Apr. 12nd, May. 05th, May. 14th, May. 20th, May. 29th, Jun. 06th, Jun. 15th, Jul. 17th, Aug. 11st, Aug. 25th, Sept. 01st, Nov. 04th, Nov. 14th, Nov. 20th

3.4. Normalized shortwave-infrared difference bare soil moisture indices with Optical TRApeZoid model (OPTRAM-NSDSI) for SM estimation

Similar to the OPTRAM-STR framework, we introduce a novel method termed OPTRAM-NSDSI, which employs NSDSIs as the y-axis instead of STR in OPTRAM. NSDSIs calculated based on normalizing the transformation of two spectral reflectances have the potential to mitigate the influence of diverse factors on single soil reflectance. SM based on OPTRAM-NSDSI is calculated using the equation (8):

$$W(OPTRAM - NSDSI) = \frac{\theta - \theta_d}{\theta_w - \theta_d} = \frac{NSDSI - NSDSI_d}{NSDSI_w - NSDSI_d} \quad (8)$$

where W is the SM content for each pixel in trapezoidal space, NSDSI are the normalized shortwave-infrared difference bare soil moisture indices proposed by Yue et al. (2019), $NSDSI_d$ and $NSDSI_w$ are the NSDSI at θ_d and θ_w , representing the NSDSI at dry and wet soil. Based on the assumption made by STR-VI that SM status influences vegetation growth and that variations in SM content result in alterations in the spectral properties of vegetation (Ambrosone et al., 2020; Chen et al., 2014), the $NSDSI_w$ and $NSDSI_d$ can be determined by the following VI related functions:

$$NSDSI_w = i_{w2} + s_{w2} VI \quad (9)$$

$$NSDSI_d = i_{d2} + s_{d2} VI \quad (10)$$

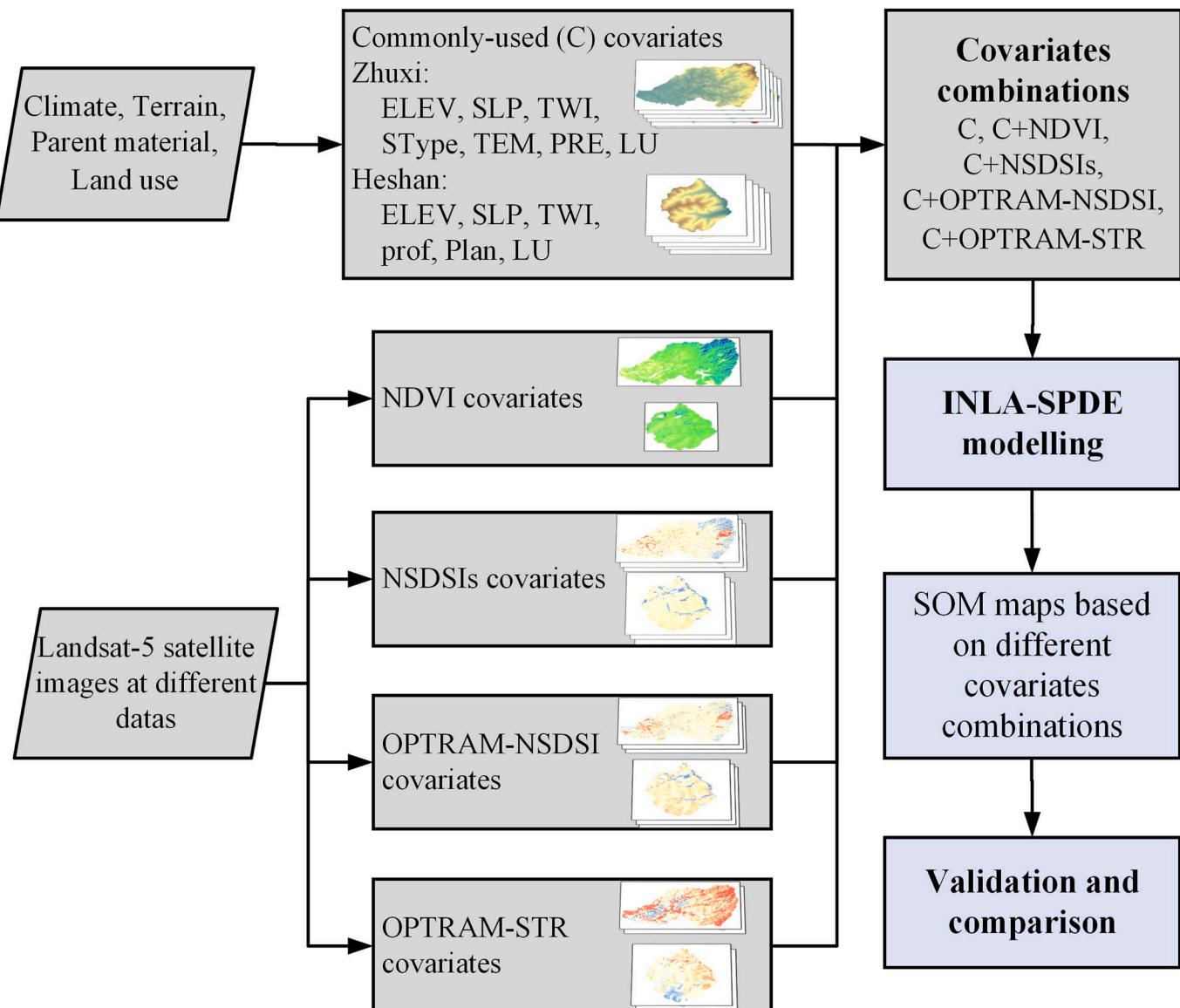


Fig. 2. The flow chart using C, NDVI, NSDSIs, OPTRAM-NSDSI estimated SM and OPTRAM-STR estimated SM covariates for SOM mapping. C: commonly-used; NDVI: Normalized Difference Vegetation Index; NSDSIs: normalized shortwave-infrared difference bare soil moisture indices; OPTRAM-NSDSI: normalized shortwave-infrared difference bare soil moisture indices with Optical TRApeZoid Model; OPTRAM-STR: shortwave infrared transformed reflectance with Optical TRApeZoid Model.

Table 2

The descriptions of Normalized Shortwave-infrared (SWIR) Difference Bare Soil moisture Indices (NSDSIs, 1 ~ 3).

Indices	Formula	Reference
NSDSI1	$\frac{B_{SWIR1} - B_{SWIR2}}{B_{SWIR1}}$	Yue et al. (2019)
NSDSI2	$\frac{B_{SWIR1} - B_{SWIR2}}{B_{SWIR2}}$	
NSDSI3	$\frac{B_{SWIR1} - B_{SWIR2}}{B_{SWIR1} + B_{SWIR2}}$	

Notes: the B_{SWIR1} and B_{SWIR2} represent the spectral reflectance of SWIR1 (1.55 ~ 1.75 μm) and SWIR2 (2.08 ~ 2.35 μm) wave bands in Landsat-5 imagery, respectively.

where, i_{d2} and s_{d2} , i_{w2} and s_{w2} are the intercept and the slope of the dry and wet edge in OPTRAM-NSDSI. Based on equation (9) and equation (10), the W for each pixel can be defined as a function of VI and NSDSI:

$$w(\text{OPTRAM} - \text{NSDSI}) = \frac{i_{d2} + s_{d2} \text{VI} - \text{NSDSI}}{i_{d2} - i_{w2} + (s_{d2} - s_{w2}) \text{VI}} \quad (11)$$

where the NSDSI is the NSDSI1, NSDSI2 and NSDSI3 (Table 1). Three OPTRAM-NSDSI models (OPTRAM-NSDSI1, OPTRAM-NSDSI2, and

OPTRAM-NSDSI3) were developed to estimate the SM content.

3.5. Covariates combinations and screening

Subsequently, four distinct combinations of covariates were generated for each observation date: “C”, “C + NDVI”, “C + NSDSIs”, “C + OPTRAM-NSDSI” and “C + OPTRAM-STR”. The combination denoted as “C” exclusively consisted of commonly-used environmental covariates and “C + NDVI” consisted of commonly-used environmental covariates and NDVI, while the combinations “C + NSDSIs” and “C + OPTRAM-NSDSI” included the commonly-used environmental covariates with NSDSIs (NSDSI1, NSDSI2, and NSDSI3) and SM covariates estimated with OPTRAM-NSDSI (OPTRAM-NSDSI1, OPTRAM-NSDSI2, and OPTRAM-NSDSI3), respectively. In comparison to OPTRAM-NSDSI estimated SM effectiveness, we also employed the original OPTRAM-STR (OPTRAM-STR1 and OPTRAM-STR2) estimated SM to form the covariate combination “C + OPTRAM-STR”. In covariates screening of different covariates combinations, the Deviance Information Criterion (DIC) was utilized to determine the most suitable combination of environmental covariates for each environmental covariates of Zhuxi and Heshan (Arshad et al., 2020; Huang et al., 2017; Li et al., 2018; Rue et al., 2009). A smaller Deviance Information Criterion (DIC) suggests a better model fit. Environmental covariates were systematically removed from the model, and their DIC values were

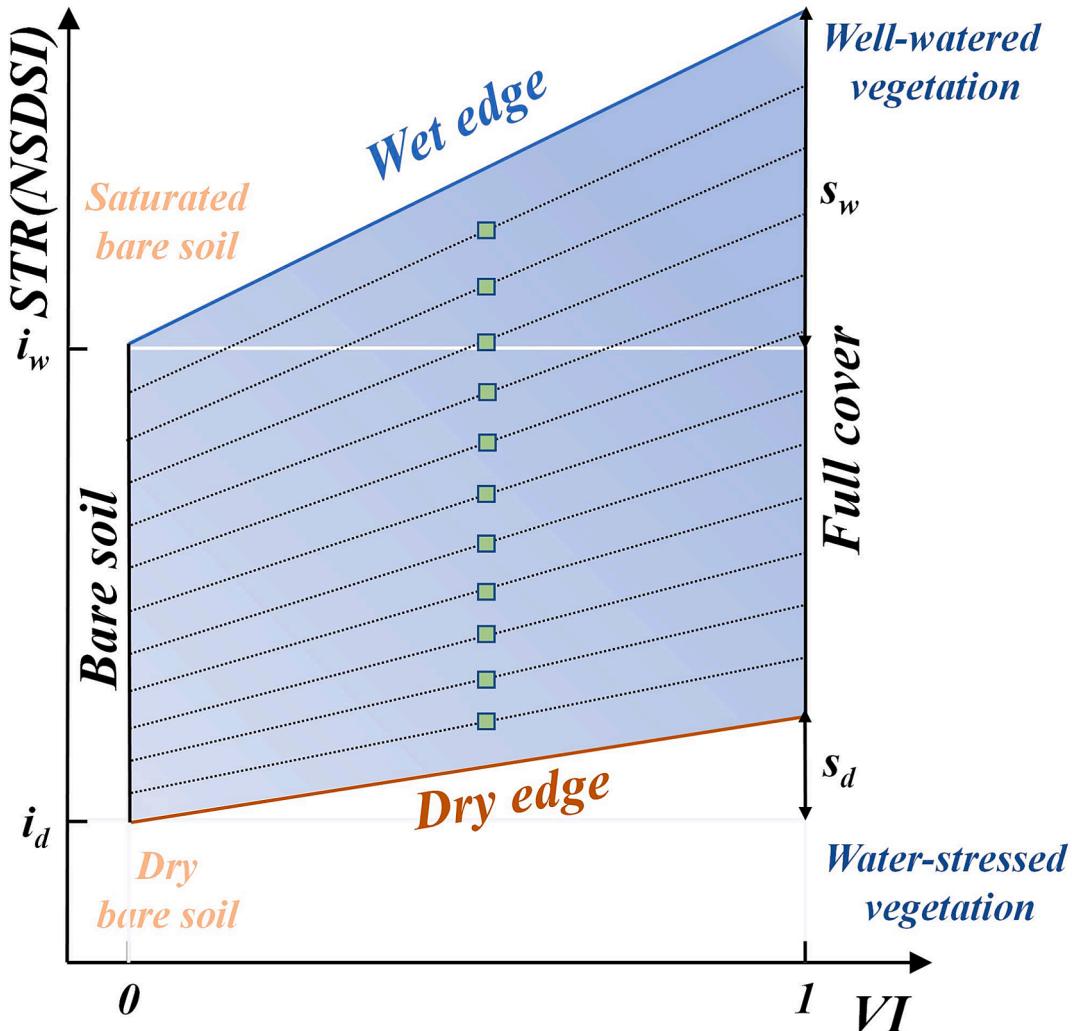


Fig. 3. The schematic diagram of STR-VI and NSDSI-VI trapezoid space of OPTRAM-STR and OPTRAM-NSDSI for SM estimation. The green dot represents an arbitrary pixel within an image. NSDSI: normalized shortwave-infrared difference bare soil moisture indices; STR: shortwave infrared transformed reflectance; OPTRAM: Optical TRApezioid Model. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

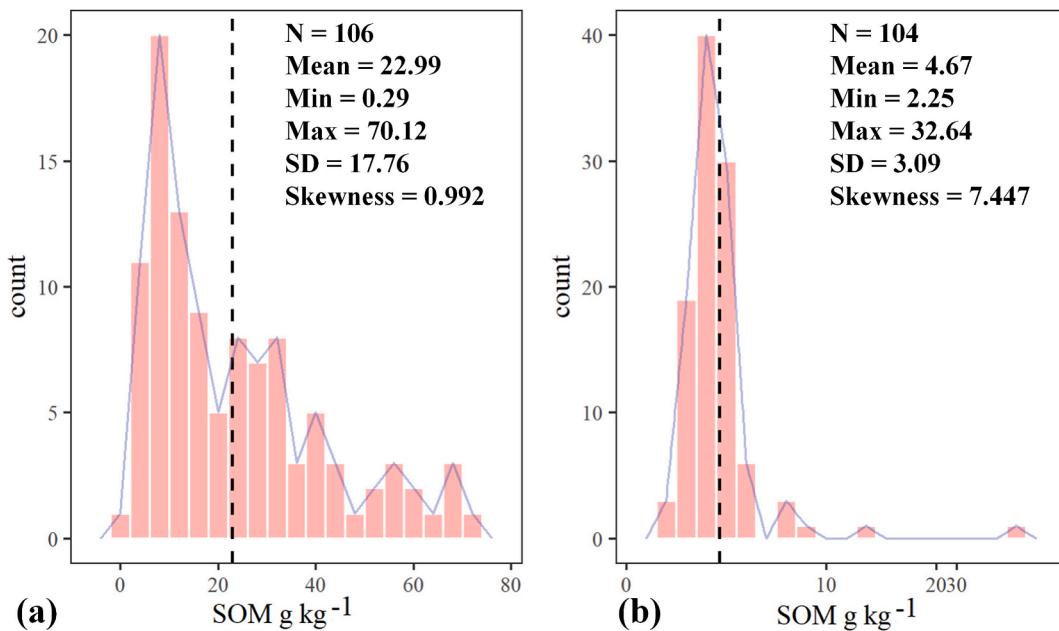


Fig. 4. The histograms of SOM content based on samples in the Zhuxi (a) and Heshan (b) study area.

compared to identify the combination of covariates with the lowest DIC values.

3.6. INLA-SPDE for SOM prediction

The Integrated Nested Laplace Approximation with the Stochastic Partial Differential Equation (INLA-SPDE) approach was utilized as the SOM prediction model. By taking into account the spatial correlation of soil sample points and environmental covariates, this model can provide higher modeling accuracies compared with random forest, residual maximum likelihood, or ordinary kriging (Huang et al., 2017; Yang et al., 2023a). The hierarchical model implemented in INLA-SPDE was utilized in this study to predict SOM (Huang et al., 2017; Wu, 2021). This model comprises three components: the intercept, the spatial fixed effects encompassing covariates and their coefficient matrices, and the spatial random effect associated with spatial locations. The three components can be articulated as demonstrated in previous studies (Arshad et al., 2020; Huang et al., 2017; Wang and Zuo, 2021; Wu, 2021; Yang et al., 2023a):

$$\eta_k = \beta_0 + \sum_{b=1}^B \beta_b x_{bk} + \xi(s_i, s_j) \quad (12)$$

Where η_k is an additive linear estimation representing the spatially predicted soil properties; the first term β_0 is the intercept; x_{bk} is the environmental covariate of k , β_b is the coefficient of the i th environmental covariate, $\sum_{b=1}^B \beta_b x_{bk}$ represents the linear fixed effects consisting of environmental covariates, and $\xi(s_i, s_j)$ is used to represent spatial random effect and can be expressed as the Matérn covariance function (Li et al., 2018; Wang and Zuo, 2021; Yang et al., 2023a), which is a function that reflects spatial autocorrelation (Huang et al., 2017; Li et al., 2018; Wu, 2021).

In the mesh setting of the INLA-SPDE model, the critical parameters including max.edge, cutoff, and offset were configured as follows: (0.03, 0.1) for max.edge, 0.02 for cutoff, and (-0.5, 0.1) for offset. The INLA-SPDE model was implemented utilizing the “INLA” package in the R language (<https://www.r-inla.org/>).

3.7. Evaluation of the prediction

Five-fold cross-validation was utilized to assess the prediction ac-

curacy using different covariate combinations in each study area, and the average of cross-validation was used as an accuracy result. The sampling points in each study area were divided into five subsets of equal size. Four segments in turn were employed for training the models, with the remaining subset allocated for validation (Abriha et al., 2023). We utilized Lin's concordance correlation coefficient (LCCC) (Carrasco et al., 2013; Lawrence, 1989) and coefficient of determination (R^2) to assess the mapping accuracy. The computation of these two indices is outlined as follows (Yang et al., 2023a):

$$LCCC = \frac{2r\sigma_y\sigma_{\hat{y}}}{\sigma_y^2 + \sigma_{\hat{y}}^2 + (\bar{y} - \bar{\hat{y}})^2} \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

where n is the sample size of validation points, y_i and \hat{y}_i is observed and predicted SOM value at the corresponding validation points, r is the correlation coefficient between y_i and \hat{y}_i , \bar{y} and $\bar{\hat{y}}$, σ_y and $\sigma_{\hat{y}}$ are the average and variance of y_i (observed SOM) and \hat{y}_i (predicted SOM).

4. Results

4.1. Statistical description of the observed SOM

Fig. 4 illustrates the SOM histograms for Zhuxi (a) and Heshan (b) study areas. Zhuxi demonstrated a higher average SOM content and lower skewness in comparison to Heshan. Both study areas exhibited substantial variation in SOM content, with a coefficient of variation (CV) of 77.3% in Zhuxi and 66.2% in Heshan.

4.2. The generated NSDSIs and SM variables estimated by OPTRAM

Fig. 5 depicts scatter plots illustrating the dry and wet edges and the fitting equation of the NSDSI-FVC trapezoid space using OPTRAM at the optimal date with the highest prediction accuracy as an example. The spatial distribution of STR and STR-FVC trapezoid space is shown in **Fig. S1** and **Fig. S2**. The spatial distributions of NSDSI-FVC features in both study areas exhibit trapezoid-like shapes, suggesting a positive correlation between NSDSI and FVC. Furthermore, the trapezoidal

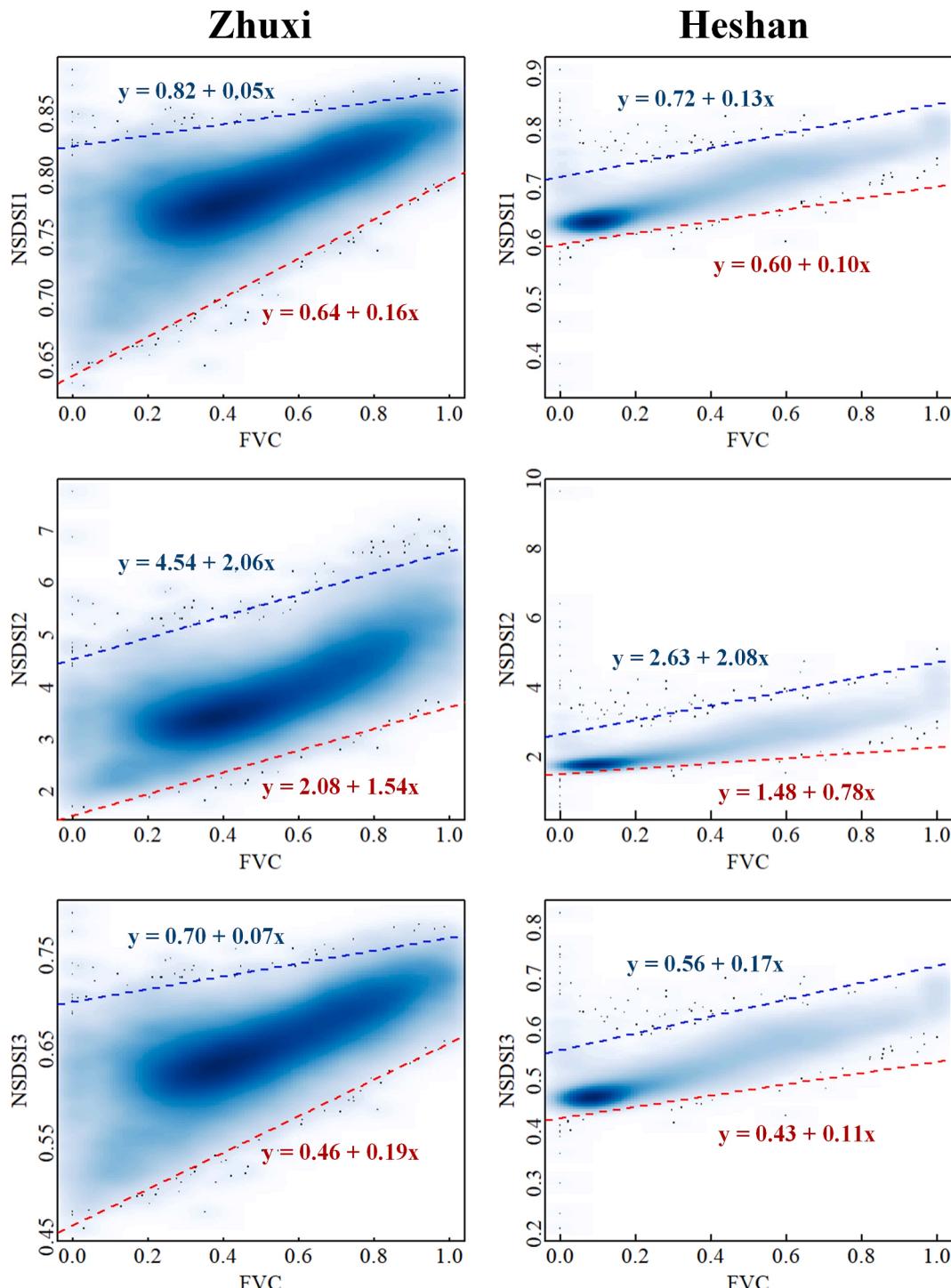


Fig. 5. The scatter plots of dry and wet edges of NSDSI-FVC trapezoid space and their fitting equations of dry and wet edges (image on the optimal date with the highest prediction accuracy was taken as an example). The red and blue dotted line represents the dry edge and the wet edge, respectively. NSDSI: normalized shortwave-infrared difference bare soil moisture indices; FVC: fractional vegetation cover. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

distribution of NSDSI-FVC space indicates a significant correlation between SM and NSDSI, even in areas with dense vegetation. Moreover, Zhuxi displays a larger trapezoidal area in comparison to Heshan, which may indicate a higher level of variability in SM.

Fig. 6 illustrates the spatial distribution of three covariates: NSDSIs, OPTRAM-NSDSI estimated SM, and OPTRAM-STR estimated SM, of the two study areas. In Zhuxi, higher NSDSIs and OPTRAM-NSDSI estimated SM values appear in the western hills and some valley

areas. In Heshan, higher values of NSDSIs and OPTRAM-NSDSI estimated SM are observed in the eastern wide floodplains and valleys with lower altitudes, and lower values usually are shown in gentle slopes with farmland. The spatial distribution of NSDSIs is close to that of the SM variables generated using OPTRAM-NSDSI. However, the spatial distribution of OPTRAM-STR estimated SM, which is similar to STR (Fig. S1), differs from either NSDSIs or OPTRAM-NSDSI estimated SM. In Zhuxi, OPTRAM-STR estimated SM is lower in high-altitude hilly areas, while

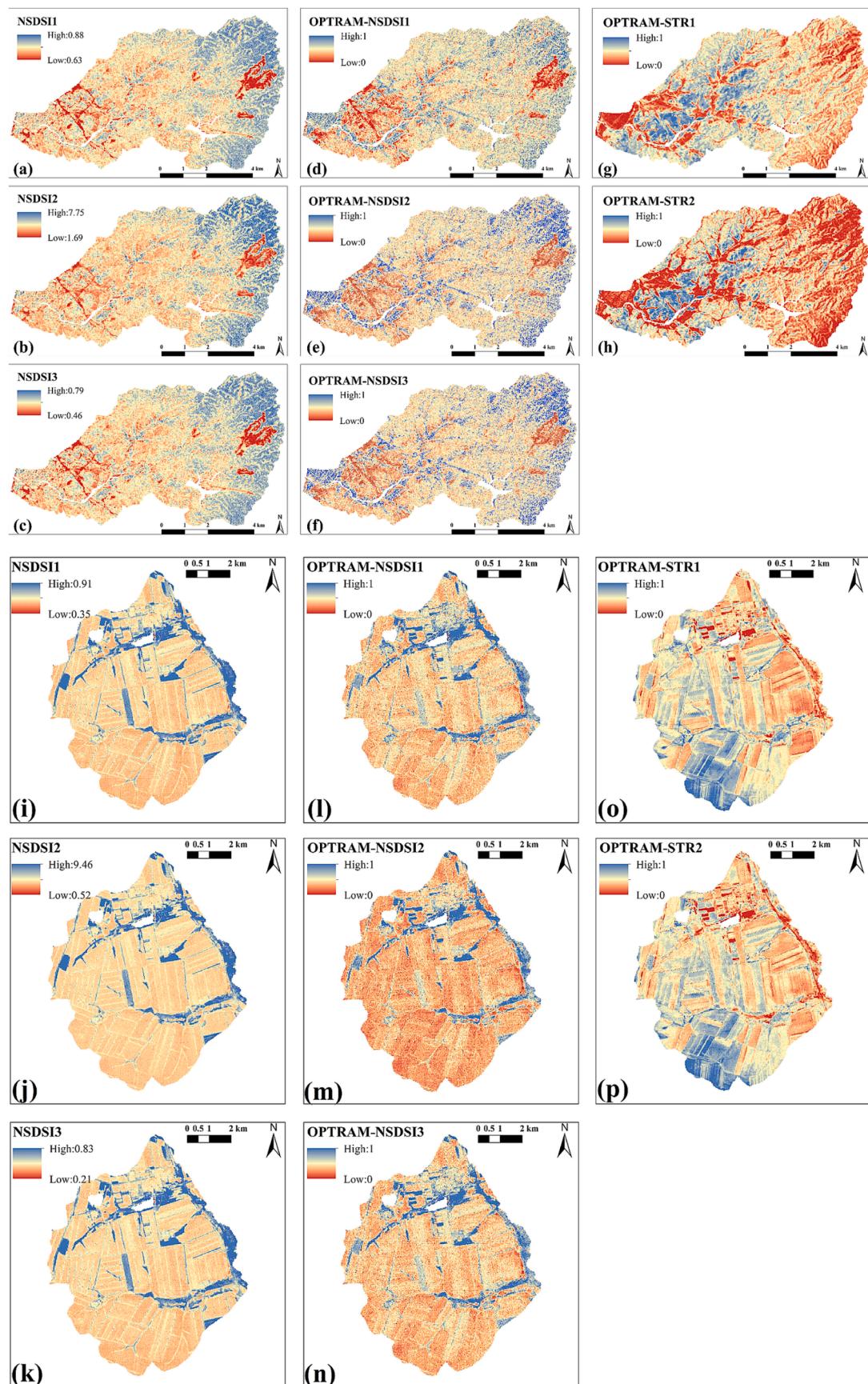


Fig. 6. The spatial distribution of the NSDSIs (a, b, c and i, j, k), OPTRAM-NSDSI estimated SM (d, e, f and l, m, n) and OPTRAM-STR estimated SM covariates (g, h and o, p) in Zhuxi and Heshan, NSDSI: normalized shortwave-infrared difference bare soil moisture indices, OPTRAM-NSDSI: normalized shortwave-infrared difference bare soil moisture indices with Optical TRApeZoid Model, OPTRAM-STR: shortwave infrared transformed reflectance with Optical TRApeZoid Model.

Table 3

The selected covariates for different covariates combinations in Zhuxi and Heshan of optimal date.

Study area	Combinations	The selected Covariates
Zhuxi	C	ELEV, SLP, TWI, SType, TEM, PER
	C + NDVI	ELEV, SLP, TWI, SType, TEM, PER, NDVI
	C + NSDSIs	ELEV, SLP, TWI, SType, TEM, PER, NSDSI1, NSDSI2, NSDSI3
	C + OPTRAM-NSDSI	ELEV, SLP, TWI, SType, TEM, PER, OPTRAM-NSDSI1, OPTRAM-NSDSI3
	C + OPTRAM-STR	ELEV, SLP, TWI, SType, TEM, PER, OPTRAM-STR1, OPTRAM-STR2
Heshan	C	ELEV, plan, prof
	C + NDVI	ELEV, plan, prof, NDVI
	C + NSDSIs	ELEV, plan, prof, NSDSI2, NSDSI3
	C + OPTRAM-NSDSI	ELEV, plan, prof, OPTRAM-NSDSI1, OPTRAM-NSDSI2, OPTRAM-NSDSI3
	C + OPTRAM-STR	ELEV, plan, prof, OPTRAM-STR1, OPTRAM-STR2

in Heshan, it is lower in valley areas and higher in farmland areas. Notably, the estimated SM values are similar across different bands with the same method.

4.3. The covariates selection for different covariate combinations

Table 3 shows the selected covariates for the four environmental covariates combinations in SOM mapping at the optimal date, respectively. In the “C” covariates combination, ELEV, SLP, TWI, SType, TEM and PER were selected in Zhuxi, ELEV, plan and prof were selected in Heshan. NDVI on May 13th for Zhuxi and May 29th for Heshan with the highest correlation with the measured SOM was used. In “C + NSDSIs”, “C + OPTRAM-NSDSI” and “C + OPTRAM-STR”, SM indices derived from different bands were selected for Zhuxi and Heshan, respectively.

4.4. SOM prediction accuracies

Fig. 7 illustrates the LCCC and R^2 values obtained from cross-validation for SOM prediction using different environmental combinations in Zhuxi and Heshan. Each result is determined by the optimal covariates selected within each environmental combination using the method outlined in [section 3.4](#). The results indicate that using either NSDSIs, OPTRAM-NSDSI or OPTRAM-STR estimated SM improved the mapping accuracies of Zhuxi and Heshan compared with using the commonly-used environmental covariates. The added value of NSDSIs and OPTRAM-NSDSI is higher than OPTRAM-STR in both Zhuxi and Heshan. Adding NDVI improved the accuracy of 5.8% in terms of LCCC compared with the “C” covariates combination in Zhuxi, while NDVI didn't improve the mapping accuracy in Heshan, thus we didn't compare the “C + NDVI” with those covariates combinations with SM variables.

Although not all the SM variables adding to the commonly-used covariates generated higher mapping accuracies than NDVI, several SM variables (i.e. Apr. 20th, Jun. 04th, Jun. 27th, Aug. 10th, and Nov. 22nd in Zhuxi) generated higher accuracies than NDVI. This indicates that SM variables at appropriate dates could be a more effective predictor than NDVI. The highest improvement of accuracy for Zhuxi occurs on Apr. 20th. At this date, “C + NSDSIs” and “C + OPTRAM-NSDSI” show an increase of 26.8% and 18.1% in LCCC compared to “C”, respectively, NSDSIs are more effective than OPTRAM-NSDSI estimated SM at this date. In the case of Heshan, the largest improvement of LCCC was observed on Jun. 06th, with an increase of 41.7% and 56.7% for “C + NSDSIs” and “C + OPTRAM-NSDSI”, respectively. At the nearby dates of May. 29th and Jun. 15th, OPTRAM-NSDSI estimated SM also obtained a large improvement in accuracy. In this case, OPTRAM-NSDSI estimated SM showed a more promising predictive power than NSDSIs. By examining the mapping accuracy and NDVI change curves along image

dates, it shows that the largest accuracy improvement in both Zhuxi and Heshan appears around the spring green-up period, not the bare soil or the densest vegetation period. It may indicate that images at this date are capable of showing the variation in SM over the study area which is related to the spatial variation of SOM. The average improvement in accuracy (6.6% with “C + NSDSIs” and 6.2% with “C + OPTRAM-NSDSI”) for Zhuxi is lower than that of Heshan (8.6% and 14.7%), probably due to more complex soil variation and soil-environment relationships in Zhuxi. The results of R^2 are similar to those of LCCC. The Scatterplot of measured SOM and predicted SOM using different covariates combinations in (C, C + NDVI, C + NSDSIs, C + OPTRAM-NSDSI and C + OPTRAM-STR) at Apr. 20th for Zhuxi and Jun. 06th for Heshan are shown in [Fig. S3](#).

4.5. The predicted SOM maps

Fig. 8 displays the predicted SOM maps generated using different combinations of covariates at the dates with their highest mapping accuracies. The results of SOM mapping at the other dates with SM covariates are presented in [Fig. S4](#) to [Fig. S9](#). The spatial distribution of SOM maps with SM variables generated from remote sensing images of different dates in both Zhuxi and Heshan holds a similar pattern. However, differences do exist, mainly due to the input SM variables at different dates. The spatial distribution of SOM predicted by each combination of covariates displays similar patterns for each study area ([Fig. 8](#)). The results show that in Zhuxi, the predicted SOM content is higher in the eastern high-elevation, western river valley and northwest flat region, in Heshan, the predicted SOM is higher in the river valley area with lower elevation. The addition of NDVI, NSDSIs, OPTRAM-NSDSI, or OPTRAM-STR covariates to the commonly-used variables results in more detailed spatial variation ([Fig. 8b, c, d and g, h, i](#)). The difference between the predicted SOM using the combination of “C + NSDSIs” ([Fig. 8b](#) and [Fig. 8g](#)) and “C + OPTRAM-NSDSI” ([Fig. 8c](#) and [Fig. 8h](#)) is illustrated in [Fig. 8e](#) and [Fig. 8j](#). Larger differences were observed in the mapping results between using the combinations “C + NSDSIs” and “C + OPTRAM-NSDSI” in Zhuxi than in Heshan. This may be related to that Zhuxi with a larger SOM content range has more complex soil variation characteristics and environmental conditions than Heshan.

5. Discussions

5.1. Applications of NSDSIs and OPTRAM-NSDSI for SOM mapping

In this study, it was observed that using both NSDSIs and the SM variables estimated by OPTRAM-NSDSI can improve the accuracy of SOM mapping. This is related to that SM plays a crucial role in influencing the vegetation growth and soil carbon turnover, and has a direct impact on the abundance and activity of soil microorganisms by regulating the availability of oxygen and water within soil ([Linn and Doran, 1984](#); [Shabtai et al., 2022](#); [Skopp et al., 1990](#); [Vanderborgh et al., 2024](#)). Furthermore, [Feng and Liu \(2015\)](#) suggest that agricultural irrigation activities and land cover have a significant impact on SM, in addition to precipitation, accordingly the spatial distribution of SM can indicate irrigation or other related agricultural practices.

We utilized the NSDSI indices and OPTRAM-NSDSI to estimate the SM. NSDSIs are calculated from SWIR1 and SWIR2 bands, while OPTRAM-NSDSI is calculated from NSDSIs combined with the Optical TRApezoid Model. SWIR is considered the optimal band for the inversion of SM in optical bands ([Sadeghi et al., 2015](#)), the reflectance of SWIR is linearly correlated with SM ([Sadeghi et al., 2017; Sadeghi et al., 2015; Yue et al., 2019](#)). Different from the single spectral calculation method like STR, NSDSIs involve the normalized transformation of two spectral reflectances, making them less susceptible to complex interferences ([Du et al., 2024](#)). This may be a reason that OPTRAM-NSDSI generated a larger improvement in prediction accuracy than OPTRAM-

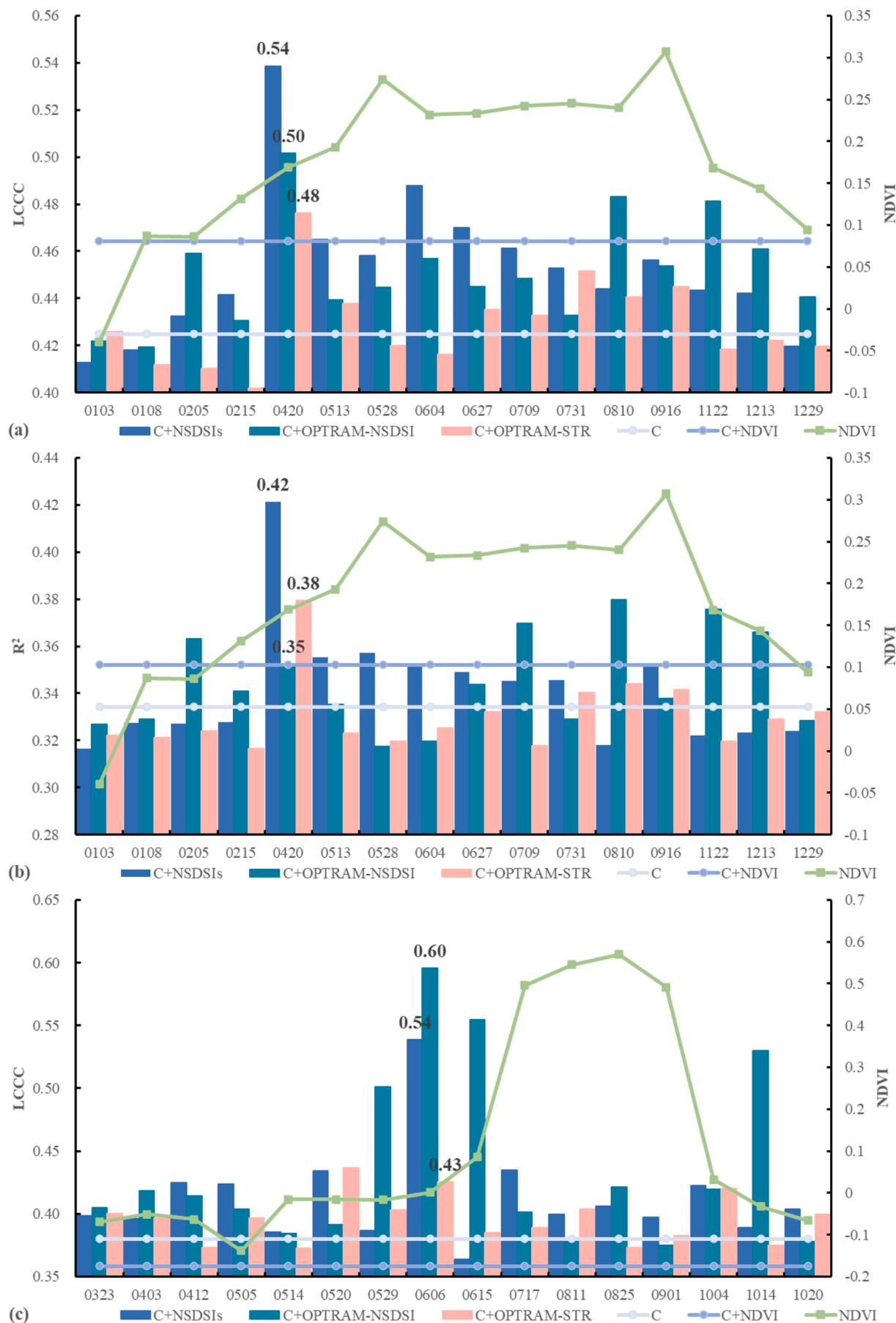


Fig. 7. The LCCC (a and c) and R^2 (b and d) of SOM predictions using different covariates combinations (C, C + NDVI, C + NSDSIs, C + OPTRAM-NSDSI and C + OPTRAM-STR) from different dates at the two study areas. (a) and (b) are for Zhuxi, (c) and (d) are for Heshan. The NDVI curve is obtained by averaging the values at the sample points. C: combination only has commonly-used environmental covariates; C + NDVI: combination composed of the commonly-used environmental covariates with Normalized Difference Vegetation Index; C + NSDSIs: combination composed of the commonly-used environmental covariates with Normalized shortwave-infrared difference bare soil moisture indices; C + OPTRAM-NSDSI: combination composed of the commonly-used environmental covariates with Normalized shortwave-infrared difference bare soil moisture indices with Optical TRApeZoid Model estimated SM. C + OPTRAM-STR: combination composed of the commonly-used environmental covariates with shortwave infrared transformed reflectance with Optical TRApeZoid Model estimated SM.

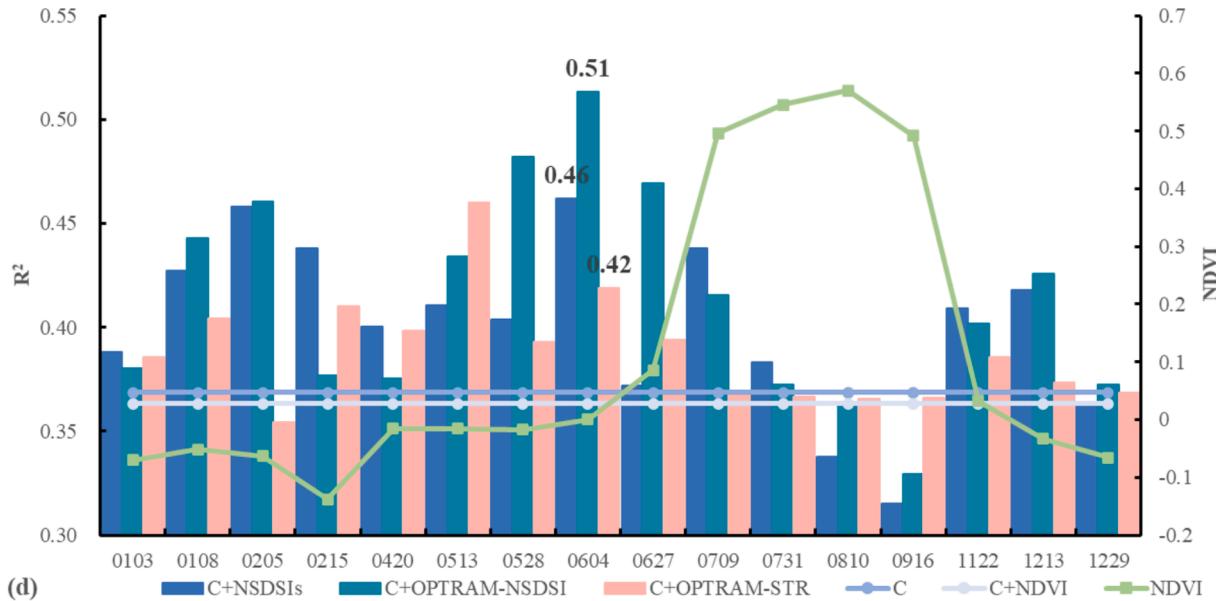


Fig. 7. (continued).

STR. This indicates that OPTRAM-NSDSI can be a potential alternative for the estimation of SM for an area with vegetation coverage in addition to OPTRAM-STR. Another advantage of OPTRAM is that the model doesn't require thermal bands or separate parameterization and atmospheric calibration of images for each distinct observation date. The absence of thermal bands allows for the utilization of OPTRAM-NSDSI on high spatial resolution satellites that only have optical bands, such as Sentinel-2 (Sadeghi et al., 2017). Compared with some SM products based on microwave remote sensing, OPTRAM combined with optical remote sensing has a higher spatial resolution to fulfill more detailed DSM requirements. It is important to determine the wet and dry edges to employ OPTRAM for calculating SM. In this paper, we adopted an automated method of Tang et al. (2010), which makes the parameter determination more objective than the visual interpretation way.

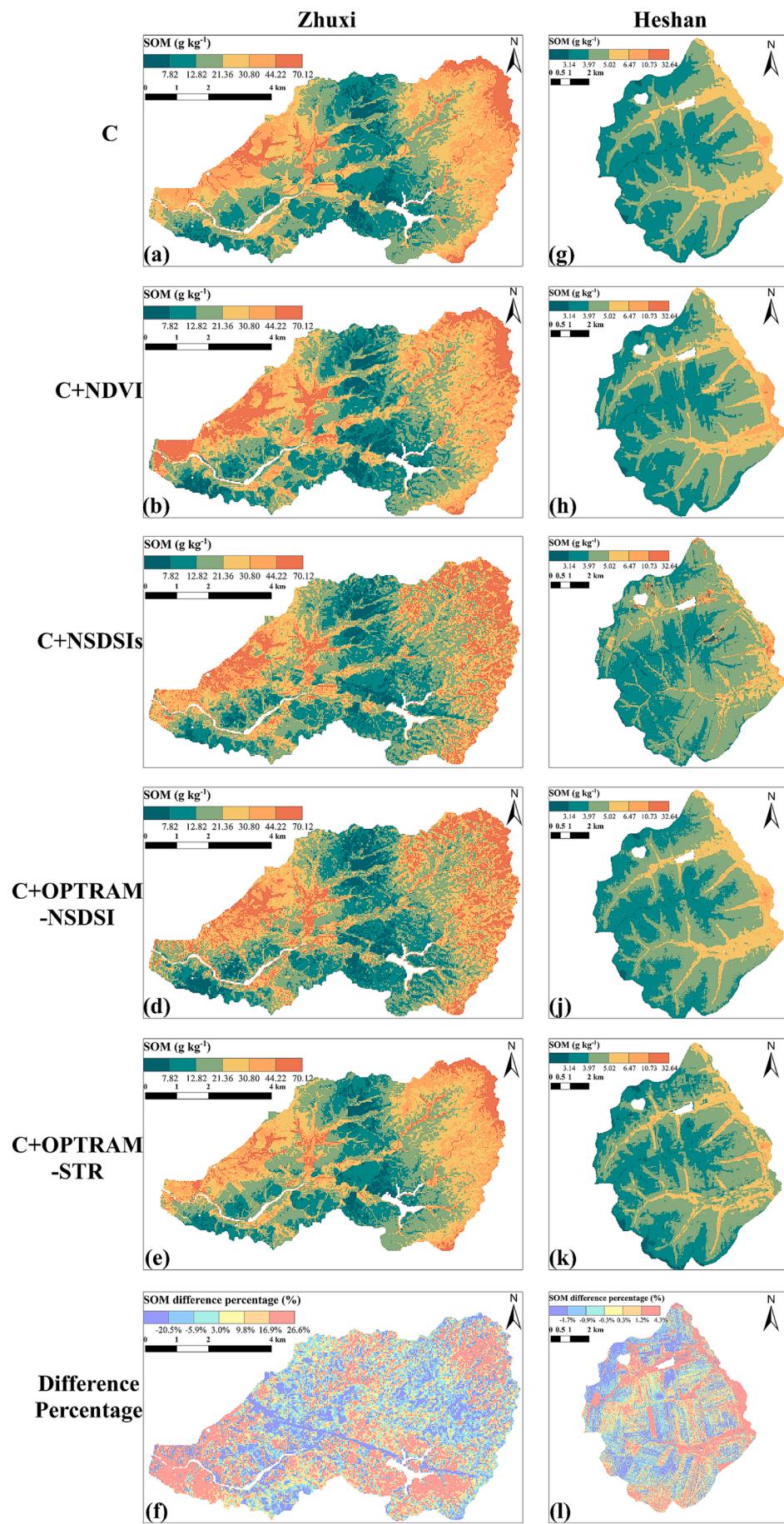
Choosing the appropriate date for remote sensing images is crucial for using NSDSIs and OPTRAM-NSDSI covariates. The SM variables at appropriate dates are more effective than NDVI at dates highly correlated with SOM. The dates of the SM variables with the highest prediction accuracies are different with the dates of the NDVI most correlated with SOM, indicating distinct mechanisms of how SM and vegetation relate to SOM. When we examine the prediction accuracies at different dates and the NDVI profiles along the same period in Fig. 7, we found that the optimal dates didn't appear at the densest vegetation stage or bare soil stage, but at the time when NDVI starts to increase, about the "green up" stage of vegetation growth. This may be associated with significant periods such as the onset of the initiation of vegetation growth, the commencement of leaf sprouting in trees, or human irrigation activity (Ryan et al., 2017). It is reported that vegetation greening can be very sensitive to SM (Yang et al., 2023b). Lan et al. (2024) and Li et al. (2023) suggested that vegetation consumes soil water through evapotranspiration during the greening stage, thus playing an important role in soil water availability dynamics. During the period of the "green up" phase, the stored soil moisture could be correlated to the vegetation development and the carbon input into the soil, thus changes in field

moisture content may highlight the spatial differentiation of SM or SOM. More research is necessary to explore the optimal image dates for estimating SM in other areas. Experiments on the relationship among the measured SM, the estimated SM, and SOM at different dates should also be conducted in the future to understand the complex processes behind SM relating to SOM.

Both NSDSIs and OPTRAM-NSDSI estimated SM generated higher prediction accuracies than OPTRAM-STR in the two study areas with different vegetation densities. However, the largest improvement of prediction accuracy obtained by OPTRAM-NSDSI estimated SM is higher than that of NSDSIs in Heshan, and NSDSIs generated a larger highest accuracy than OPTRAM-NSDSI estimated SM in Zhuxi. This may be related to that Heshan has a relatively uniform vegetation type (crops) with a similar growth period, and a high vegetation cover density during the growth season of crops. However, Zhuxi has more complex land use and a relatively lower vegetation cover density. The highest NDVI during the dense vegetation period is 0.57 vs. 0.27 in Heshan and Zhuxi, respectively. This may indicate that OPTRAM-NSDSI is more effective than NSDSI in areas with high vegetation densities, although either of these two variables is a potentially effective predictor for SOM in the two study areas with different vegetation coverage densities. The generated SM covariates provide a higher added value of SOM prediction in Heshan than Zhuxi. This is probably because Zhuxi has more complex soil-environment relationships with more land use types than Heshan.

5.2. Limitations and perspectives

The study shows the effectiveness of the remote sensing-based estimated SM for SOM mapping, however, there are some limitations and further perspectives in this study. Firstly, optical satellites are susceptible to interference from cloud cover, leading to a reduction in the quantity of available images. Consequently, this results in an uneven distribution of the time series of remote sensing images (Ambrosone et al., 2020; Zhang and Zhou, 2016). This may be addressed by



(caption on next page)

Fig. 8. The predicted SOM maps in Zhuxi and Heshan with different covariates combinations on Apr. 20th of Zhuxi and Jun. 06th of Heshan (“C”, “C + NDVI”, “C + NSDSIs”, “C + OPTRAM-NSDSIs” and “C + OPTRAM-STR”, a, b, c, d, f and h, i, j, k, l) and the difference between using C + NSDSIs and C + OPTRAM-NSDSI (subtract C + OPTRAM-NSDSI from C + NSDSIs and divided by C + NSDSI, f and l). C: combination only has commonly-used environmental covariates; C + NDVI: combination composed of the commonly-used environmental covariates with Normalized Difference Vegetation Index; C + NSDSIs: combination composed of the commonly-used environmental covariates with Normalized shortwave-infrared difference bare soil moisture indices; C + OPTRAM-NSDSIs: combination composed of the commonly-used environmental covariates with Normalized shortwave-infrared difference bare soil moisture indices with Optical TRApezioid Model estimated SM; C + OPTRAM-STR: combination composed of the commonly-used environmental covariates with shortwave infrared transformed reflectance with Optical TRApezioid Model estimated SM.

integrating multiple satellites, such as the MODIS series satellites (Araya et al., 2016; Babaeian et al., 2018; Chen et al., 2018). Secondly, the OPTRAM-NSDSIs is to estimate SM based on a linear fit at the wet and dry edges, which may be not optimal for acquiring precise SM information (Ambrosone et al., 2020; Ma et al., 2022). Non-linear fitting to wet and dry edge equations may be an option, such as exponential or quadratic fit (Ambrosone et al., 2020; Ma et al., 2022). Thirdly, NDVI improved the prediction accuracies in Zhuxi study area, as in many other studies (He et al., 2021; Yang et al., 2020). Combining both NDVI and SM indices may work in SOM prediction in some areas, which can be further investigated. Furthermore, we used the remote sensing images on different dates to estimate SM separately, while extracting more features from SM profiles over a long period may provide extra information on SM variation for SOM prediction. Previous studies that generated phenological variables or variables indicating crop periodic growth characteristics on NDVI time series data, and found that the inclusion of those variables improved the accuracy of SOC mapping (Yang et al., 2020; Yang et al., 2019). This may provide inspiration for extracting new variables from remote sensing-based SM indices over years. Further investigations can be conducted at this aspect.

6. Conclusion

Our study introduces a new approach, OPTRAM-NSDSI, using NSDSIs instead of STR within the trapezoid model, to estimate SM in areas with vegetation coverage, and conduct a comparative analysis between the OPTRAM-NSDSI estimated SM and NSDSIs, OPTRAM-STR to map SOM across two study areas with different vegetation densities and land use types. The results indicate that adding SM variables into the commonly-used environmental covariates improved SOM prediction, and NSDSIs or the SM variables generated using OPTRAM-NSDSI provide higher added value than the SM variable generated using OPTRAM-STR in either study area. The results also indicate that OPTRAM-NSDSI is more effective than NSDSIs in areas with higher vegetation densities (Heshan study area) while NSDSIs generated the highest prediction accuracy improvement in areas with lower vegetation densities (Zhuxi study area). The highest accuracy improvement of 26.8% in terms of LCCC was obtained by NSDSIs on Apr. 20th in Zhuxi, and the highest improvement of 56.7% in terms of LCCC was obtained by OPTRAM-NSDSI estimated SM on Jun. 6th in Heshan, respectively. The optimal dates are about the “green-up” stage of vegetation growth. This study examines the applicability of three remote sensing-based SM variables at different dates for SOM mapping in two study areas with different vegetation densities, which provides a reference for using SM information to improve SOM mapping in areas covered with vegetation.

CRediT authorship contribution statement

Chenconghai Yang: Writing – original draft, Software, Resources, Methodology, Investigation. **Lin Yang:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation. **Lei Zhang:** Writing – review & editing. **Feixue Shen:** Writing – review & editing. **Di Fu:** Validation, Writing – review & editing. **Shengfeng Li:** Data curation, Conceptualization. **Zhiqiang Chen:** Resources. **Chenghu Zhou:** Funding acquisition.

Declaration of competing interest

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

Acknowledgments

This study is supported by the National Natural Science Foundation of China (Project No. 42471468), the Special Fund of Jiangsu Province Carbon Peak and Carbon Neutral Technology Innovation (Grant No. BK20220037), Fundamental Research Funds for the Central Universities (Grant No. 0209-14380115 and 0209-14912210).

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhydrol.2025.132980>.

Data availability

Data will be made available on request.

References

- Abriha, D., Srivastava, P.K., Szabó, S., 2023. Smaller is better? Unduly nice accuracy assessments in roof detection using remote sensing data with machine learning and k-fold cross-validation. *Heliyon* 9 (3), e14045. <https://doi.org/10.1016/j.heliyon.2023.e14045>.
- Ambrosone, M., Matese, A., Gennaro, S., et al., 2020. Retrieving soil moisture in rainfed and irrigated fields using Sentinel-2 observations and a modified OPTRAM approach. *Int. J. Appl. Earth Obs. Geoinf.* 89, 102113. <https://doi.org/10.1016/j.jag.2020.102113>.
- Araya, S., Lyle, G., Lewis, M., Ostendorf, B., 2016. Phenologic metrics derived from MODIS NDVI as indicators for Plant Available Water-holding Capacity. *Ecol. Indic.*, 60, 1263–1272. <https://doi.org/10.1016/j.ecolind.2015.09.012>.
- Arrouays, D., Mulder, V.L., Richer-de Forges, A.C., 2021. Soil mapping, digital soil mapping and soil monitoring over large areas and the dimensions of soil security – A review. *Soil Secur.*, 5, 100018. <https://doi.org/10.1016/j.soilsec.2021.100018>.
- Arshad, M., Li, N., Bella, L., Triantafyllis, J., 2020. Field-scale digital soil mapping of clay: Combining different proximal sensed data and comparing various statistical models. *Soil Sci. Soc. Am. J.*, 84. <https://doi.org/10.1002/sssaj.20008>.
- Babaeian, E., Sadeghi, M., Franz, T.E., Jones, S., Tuller, M., 2018. Mapping soil moisture with the OPTical TRApezioid Model (OPTRAM) based on long-term MODIS observations. *Remote Sens. Environ.*, 211, 425–440. <https://doi.org/10.1016/j.rse.2018.04.029>.
- Carlson, T.N., Gillies, R.R., Perry, E.M., 1994. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sensing Reviews*, 9 (1–2), 161–173. <https://doi.org/10.1080/02757259409532220>.
- Carlson, T.N., Gillies, R.R., Schmugge, T.J., 1995. An interpretation of methodologies for indirect measurement of soil water content. *Agric. for. Meteorol.*, 77 (3), 191–205. [https://doi.org/10.1016/0168-1923\(95\)02261-U](https://doi.org/10.1016/0168-1923(95)02261-U).
- Carrasco, J.L., Phillips, B.R., Puig-Martinez, J., King, T.S., Chinchilli, V.M., 2013. Estimation of the concordance correlation coefficient for repeated measures using SAS and R. *Comput. Methods Programs Biomed.*, 109 (3), 293–304. <https://doi.org/10.1016/j.cmpb.2012.09.002>.
- Chen, H., 2011. Spatial Variability and Quality Evaluation of Soil Nutrient in the Small Watershed of Red Eroded Soil Area — A Case Study in the Small Watershed of Zhuxi (in Chinese). M Thesis, Fujian Normal University. doi:[10.7666/d.y1998992](https://doi.org/10.7666/d.y1998992).
- Chen, T., de Jeu, R.A.M., Liu, Y.Y., van der Werf, G.R., Dolman, A.J., 2014. Using satellite based soil moisture to quantify the water driven variability in NDVI: A case study over mainland Australia. *Remote Sens. Environ.*, 140, 330–338. <https://doi.org/10.1016/j.rse.2013.08.022>.
- Chen, Y., et al., 2018. Mapping croplands, cropping patterns, and crop types using MODIS time-series data. *Int. J. Appl. Earth Obs. Geoinf.*, 69, 133–147. <https://doi.org/10.1016/j.jag.2018.03.005>.

- Chen, S., Zha, X., Bai, Y., Wang, L., 2019. Evaluation of soil erosion vulnerability on the basis of exposure, sensitivity, and adaptive capacity: A case study in the Zhuxi watershed, Changting, Fujian Province, Southern China. *CATENA* 177, 57–69. <https://doi.org/10.1016/j.catena.2019.01.036>.
- Chen, M., Zhang, Y., Yao, Y., et al., 2020. Evaluation of the OPTRAM Model to Retrieve Soil Moisture in the Sanjiang Plain of Northeast China. *Earth and Space Science* 7 (6), e2020EA001108. <https://doi.org/10.1029/2020EA001108>.
- Du, R., et al., 2024. Combing transfer learning with the OPTical TRApeZoid Model (OPTRAM) to diagnosis small-scale field soil moisture from hyperspectral data. *Agric. Water Manage.*, 298, 108856. <https://doi.org/10.1016/j.agwat.2024.108856>.
- Feng, H., Liu, Y., 2015. Combined effects of precipitation and air temperature on soil moisture in different land covers in a humid basin. *J. Hydrol.*, 531, 1129–1140. <https://doi.org/10.1016/j.jhydrol.2015.11.016>.
- Galvão, L.S., Vitorello, I., 1998. Variability of Laboratory Measured Soil Lines of Soils from Southeastern Brazil. *Remote Sens. Environ.*, 63 (2), 166–181. [https://doi.org/10.1016/S0034-4257\(97\)00135-1](https://doi.org/10.1016/S0034-4257(97)00135-1).
- Gao, L., et al., 2020. Remote sensing algorithms for estimation of fractional vegetation cover using pure vegetation index values: A review. *ISPRS J. Photogramm. Remote Sens.*, 159, 364–377. <https://doi.org/10.1016/j.isprsjprs.2019.11.018>.
- Hassanpour, R., Zarehagi, D., Neyshabouri, M.R., Feizizadeh, B., Rahmati, M., 2020. Modification on optical trapezoid model for accurate estimation of soil moisture content in a maize growing field. *Journal of Applied Remote Sensing* 14 (3), 034519.
- He, X., et al., 2021. Soil organic carbon prediction using phenological parameters and remote sensing variables generated from Sentinel-2 images. *CATENA* 205, 105442. <https://doi.org/10.1016/j.catena.2021.105442>.
- Huang, H., et al., 2022. A review on digital mapping of soil carbon in cropland: progress, challenge, and prospect. *Environ. Res. Lett.*, 17 (12), 123004. <https://doi.org/10.1088/1748-9326/aca41e>.
- Huang, J., Malone, B.P., Minasny, B., McBratney, A.B., Triantafyllis, J., 2017. Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. *Sci. Total Environ.*, 609, 621–632. <https://doi.org/10.1016/j.scitotenv.2017.07.201>.
- Hutchinson, M.F., 1998. Interpolation of Rainfall Data with Thin Plate Smoothing Splines: II. Analysis of Topographic Dependence. *Journal of Geographic Information and Decision Analysis*, 2, 168–185.
- Jacquemoud, S., 1993. Inversion of the PROSPECT + SAIL canopy reflectance model from AVIRIS equivalent spectra: Theoretical study. *Remote Sens. Environ.*, 44 (2), 281–292. [https://doi.org/10.1016/0034-4257\(93\)90022-P](https://doi.org/10.1016/0034-4257(93)90022-P).
- Krull, E.S., Baldock, J.A., Skjemstad, J.O., 2003. Importance of mechanisms and processes of the stabilisation of soil organic matter for modelling carbon turnover. *Functional Plant Biology*, 30 (2), 207–222. <https://doi.org/10.1071/fpb02085>.
- Lal, R., 2004. Soil Carbon Sequestration Impacts on Global Climate Change and Food Security. *Science* 304 (5677), 1623–1627.
- Lal, R., 2020. Food security impacts of the “4 per Thousand” initiative. *Geoderma* 374, 114427. <https://doi.org/10.1016/j.geoderma.2020.114427>.
- Lin, X., et al., 2024. Vegetation greening accelerated hydrological drought in two-thirds of river basins over China. *J. Hydrol.*, 637, 131436. <https://doi.org/10.1016/j.jhydrol.2024.131436>.
- Li, N., Zare, E., Huang, J., Triantafyllis, J., 2018. Mapping Soil Cation-Exchange Capacity using Bayesian Modeling and Proximal Sensors at the Field Scale. *Soil Sci. Soc. Am. J.*, 82. <https://doi.org/10.2136/sssaj2017.10.0356>.
- Li, Y., et al., 2023. Warming and greening exacerbate the propagation risk from meteorological to soil moisture drought. *J. Hydrol.*, 622, 129716. <https://doi.org/10.1016/j.jhydrol.2023.129716>.
- Lawrence, I.K.L., 1989. A Concordance Correlation Coefficient to Evaluate Reproducibility. *Biometrics* 45 (1), 255–268. <https://doi.org/10.2307/2532051>.
- Li, Z.L., Leng, P., Zhou, C., et al., 2021. Soil moisture retrieval from remote sensing measurements: Current knowledge and directions for the future. *Earth Sci. Rev.* 218, 103673. <https://doi.org/10.1016/j.earscirev.2021.103673>.
- Linn, D.M., Doran, J.W., 1984. Effect of Water-Filled Pore Space on Carbon Dioxide and Nitrous Oxide Production in Tilled and Nontilled Soils. *Soil Sci. Soc. Am. J.*, 48 (6), 1267–1272. <https://doi.org/10.2136/sssaj1984.03615995004800060013x>.
- Longo-Minnolo, G., et al., 2022. A stand-alone remote sensing approach based on the use of the optical trapezoid model for detecting the irrigated areas. *Agric. Water Manage.*, 274, 107975. <https://doi.org/10.1016/j.agwat.2022.107975>.
- Ma, C., Johansen, K., McCabe, M.F., 2022. Combining Sentinel-2 data with an optical-trapezoid approach to infer within-field soil moisture variability and monitor agricultural production stages. *Agric. Water Manage.*, 274, 107942. <https://doi.org/10.1016/j.agwat.2022.107942>.
- McBratney, A., Mendonça Santos, M., Minasny, B., 2003. On Digital Soil Mapping. *Geoderma* 117, 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4).
- Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. *Geoderma* 264, 301–311. <https://doi.org/10.1016/j.geoderma.2015.07.017>.
- Mokhtari, A., Sadeghi, M., Afrasiabian, Y., Yu, K., 2023. OPTRAM-ET: A novel approach to remote sensing of actual evapotranspiration applied to Sentinel-2 and Landsat-8 observations. *Remote Sens. Environ.*, 286, 113443. <https://doi.org/10.1016/j.rse.2022.113443>.
- Moran, M.S., Clarke, T.R., Inoue, Y., Vidal, A., 1994. Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote Sens. Environ.* 49 (3), 246–263. [https://doi.org/10.1016/0034-4257\(94\)90020-5](https://doi.org/10.1016/0034-4257(94)90020-5).
- Oldeman, L.R., van Engelen, V.W.P., 1993. A world soils and terrain digital database (SOTER)—An improved assessment of land resources. *Geoderma* 60 (1), 309–325. [https://doi.org/10.1016/0016-7061\(93\)90033-H](https://doi.org/10.1016/0016-7061(93)90033-H).
- Price, J.C., 1990. Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE Trans. Geosci. Remote Sens.*, 28 (5), 940–948. <https://doi.org/10.1109/36.58983>.
- Robinson, D.A., et al., 2008. Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review. *Vadose Zone J.*, 7 (1), 358–389. <https://doi.org/10.2136/vzj2007.0143>.
- Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. Royal Stat. Soc. Ser. B (Stat. Methodol.)*, 71 (2), 319–392. <https://doi.org/10.1111/j.1467-9868.2008.00700.x>.
- Ryan, C.M., Williams, M., Grace, J., Woollen, E., Lehmann, C.E.R., 2017. Pre-Rain Green-up Is Ubiquitous across Southern Tropical Africa: Implications for Temporal Niche Separation and Model Representation. *213* (2), 625–633. <https://doi.org/10.1111/nph.14262>.
- Sadeghi, M., Babaeian, E., Tuller, M., Jones, S., 2017. The optical trapezoid model: A novel approach to remote sensing of soil moisture applied to Sentinel-2 and Landsat-8 observations. *Remote Sens. Environ.*, 198, 52–68. <https://doi.org/10.1016/j.rse.2017.05.041>.
- Sadeghi, M., Jones, S.B., Philpot, W.D., 2015. A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. *Remote Sens. Environ.*, 164, 66–76. <https://doi.org/10.1016/j.rse.2015.04.007>.
- Sanderman, J., Baldock, J.A., 2010. Accounting for soil carbon sequestration in national inventories: a soil scientist's perspective. *Environ. Res. Lett.*, 5 (3), 034003. <https://doi.org/10.1088/1748-9326/5/3/034003>.
- Sandholt, I., Rasmussen, K., Andersen, J., 2002. A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sens. Environ.*, 79 (2), 213–224. [https://doi.org/10.1016/S0034-4257\(01\)00274-7](https://doi.org/10.1016/S0034-4257(01)00274-7).
- Shabtai, I.A., et al., 2022. Soil organic carbon accrual due to more efficient microbial utilization of plant inputs at greater long-term soil moisture. *Geochim. Cosmochim. Acta* 327, 170–185. <https://doi.org/10.1016/j.gca.2022.04.028>.
- Skopp, J., Jawson, M.D., Doran, J.W., 1990. Steady-State Aerobic Microbial Activity as a Function of Soil Water Content. *Soil Sci. Soc. Am. J.*, 54 (6), 1619–1625. <https://doi.org/10.2136/sssaj1990.03615995005400060018x>.
- Susha Lekshmi, S.U., Singh, D.N., Shojaei Baghini, M., 2014. A critical review of soil moisture measurement. *Measurement* 54, 92–105. <https://doi.org/10.1016/j.measurement.2014.04.007>.
- Tang, R., Li, Z.-L., Tang, B., 2010. An application of the Ts–VI triangle method with enhanced edge determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: Implementation and validation. *Remote Sens. Environ.*, 114 (3), 540–551. <https://doi.org/10.1016/j.rse.2009.10.012>.
- Tian, J., Yue, J., Philpot, W.D., Dong, X., Tian, Q., 2021. Soil moisture content estimate with drying process segmentation using shortwave infrared bands. *Remote Sens. Environ.*, 263, 112552. <https://doi.org/10.1016/j.rse.2021.112552>.
- Twomey, S.A., Bohren, C.F., Mergenthaler, J.L., 1986. Reflectance and albedo differences between wet and dry surfaces. *Appl Opt.* 25 (3), 431–437. <https://doi.org/10.1364/ao.25.000431>.
- Vanderborgh, J., et al., 2024. Combining root and soil hydraulics in macroscopic representations of root water uptake. *Vadose Zone J.*, 23 (3), e20273. <https://doi.org/10.1002/vzj2.20273>.
- Wang, J., Zuo, R., 2021. Spatial modelling of hydrothermal mineralization-related geochemical patterns using INLA+SPDE and local singularity analysis. *Comput. Geosci.*, 154, 104822. <https://doi.org/10.1016/j.cageo.2021.104822>.
- Wang, L., Wang, X., Kooch, Y., Song, K., Wu, D., 2023a. Improvement of data imbalance for digital soil class mapping in Eastern China. *Comput. Electron. Agric.*, 214, 108322. <https://doi.org/10.1016/j.compag.2023.108322>.
- Wang, S., et al., 2023b. An independent framework-based evapotranspiration model (IFEM) for dual-source: From field to regional scale. *Remote Sens. Environ.*, 298, 113792. <https://doi.org/10.1016/j.rse.2023.113792>.
- Wang, X., et al., 2022. Prediction of soil organic matter using VNIR spectral parameters extracted from shape characteristics. *Soil Tillage Res.*, 216, 105241. <https://doi.org/10.1016/j.still.2021.105241>.
- Wu, W., 2021. Multiscale Bayesian spatial modelling and mapping of soil organic carbon with the INLA-SPDE approach (in Chinese). D Thesis, Zhejiang University. DOI: <https://doi.org/10.27461/d.cnki.gzdx.2021.000541>.
- Wu, Z., et al., 2021. Mapping farmland soil organic carbon density in plains with combined cropping system extracted from NDVI time-series data. *Sci. Total Environ.*, 754, 142120. <https://doi.org/10.1016/j.scitotenv.2020.142120>.
- Yan, Y., et al., 2023. Optimal soil organic matter mapping using an ensemble model incorporating moderate resolution imaging spectroradiometer, portable X-ray fluorescence, and visible near-infrared data. *Comput. Electron. Agric.*, 210, 107885. <https://doi.org/10.1016/j.compag.2023.107885>.
- Yang, C., Yang, L., Zhang, L., Zhou, C., 2023a. Soil organic matter mapping using INLA-SPDE with remote sensing based soil moisture indices and Fourier transforms decomposed variables. *Geoderma* 437, 116571. <https://doi.org/10.1016/j.geoderma.2023.116571>.
- Yang, J., et al., 2023b. Green-up and brown-down: Modelling grassland foliage phenology responses to soil moisture availability. *Agric. for. Meteorol.*, 328, 109252. <https://doi.org/10.1016/j.agrformet.2022.109252>.
- Yang, L., et al., 2020. Improving prediction of soil organic carbon content in croplands using phenological parameters extracted from NDVI time series data. *Soil Tillage Res.*, 196, 104465. <https://doi.org/10.1016/j.still.2019.104465>.
- Yang, L., et al., 2019. Predicting soil organic carbon content in croplands using crop rotation and Fourier transform decomposed variables. *Geoderma* 340, 289–302. <https://doi.org/10.1016/j.geoderma.2019.01.015>.
- Yang, L., et al., 2013. An integrative hierarchical stepwise sampling strategy for spatial sampling and its application in digital soil mapping. *Int. J. Geog. Inf. Sci.*, 27 (1), 1–23. <https://doi.org/10.1080/13658816.2012.658053>.

- Yoshida, H., et al., 2018. Changes in the chemical composition of soil organic matter including water-soluble component during incubation: A case study of coniferous and broadleaf forest soils. *CATENA* 171, 22–28. <https://doi.org/10.1016/j.catena.2018.06.032>.
- Yue, J., Tian, J., Tian, Q., Xu, K., Xu, N., 2019. Development of soil moisture indices from differences in water absorption between shortwave-infrared bands. *ISPRS J. Photogramm. Remote Sens.*, 154, 216–230. <https://doi.org/10.1016/j.isprsjprs.2019.06.012>.
- Zeng, C., et al., 2016. Mapping soil organic matter concentration at different scales using a mixed geographically weighted regression method. *Geoderma* 281, 69–82. <https://doi.org/10.1016/j.geoderma.2016.06.033>.
- Zhang, D., Zhou, G., 2016. Estimation of Soil Moisture from Optical and Thermal Remote Sensing: A Review. *Sensors* 16, 1308. <https://doi.org/10.3390/s16081308>.
- Zhang, G., Zhu, A.X., 2019. A representativeness heuristic for mitigating spatial bias in existing soil samples for digital soil mapping. *Geoderma* 351, 130–143. <https://doi.org/10.1016/j.geoderma.2019.05.024>.
- Zhang, L., et al., 2021. A self-training semi-supervised machine learning method for predictive mapping of soil classes with limited sample data. *Geoderma* 384, 114809. <https://doi.org/10.1016/j.geoderma.2020.114809>.
- Zhu, A.X., et al., 2010. Construction of membership functions for predictive soil mapping under fuzzy logic. *Geoderma* 155 (3), 164–174. <https://doi.org/10.1016/j.geoderma.2009.05.024>.