

## Article

# Possibility of Zhuhai-1 Hyperspectral Imagery for Monitoring Salinized Soil Moisture Content Using Fractional Order Differentially Optimized Spectral Indices

Yasenjiang Kahaer <sup>1,2</sup>, Nigara Tashpolat <sup>1,2,\*</sup>, Qingdong Shi <sup>1,2,\*</sup> and Suhong Liu <sup>3</sup>

<sup>1</sup> College of Resources and Environment Sciences, Xinjiang University, Urumqi 830046, China; ysj\_0801@163.com

<sup>2</sup> Ministry of Education Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi 830046, China

<sup>3</sup> Beijing Key Laboratory of Environmental Remote Sensing and Digital City, Beijing Normal University, Beijing 100875, China; liush@bnu.edu.cn

\* Correspondence: ngr.t@xju.edu.cn (N.T.); shiqd@xju.edu.cn (Q.S.)

Received: 10 September 2020; Accepted: 26 November 2020; Published: 30 November 2020



**Abstract:** The possibility of quantitative inversion of salinized soil moisture content (SMC) from Zhuhai-1 hyperspectral imagery and the application effect of fractional order differentially optimized spectral indices were discussed, which provided new research ideas for improving the accuracy of hyperspectral remote sensing inversion. The hyperspectral data from indoor and Zhuhai-1 remote sensing imagery were resampled to the same spectral scale. The soil hyperspectral data were processed by fractional order differential preprocessing method and optimized spectral indices method, and the Pearson correlation coefficient (PCC/ $r$ ) analysis was made with SMC data. The sensitive optimized spectral indices were used to establish the ground hyperspectral estimation model, and a variety of modeling methods were used to select the best SMC inversion model. The results were as follows: the maximum one-dimensional  $r$  between SMC and the 466–938 nm band was  $-0.635$ , the maximum one-dimensional  $r$  with the 0.5-order absorbance spectrum was  $0.665$ , and the maximum two-dimensional  $r$  with the difference index (DI) calculated by the 0.5-order absorbance spectrum was  $\pm 0.72$ . The maximum three-dimensional  $r$  with the triangle vegetation index (TVI) calculated from the 0.5-order absorbance spectrum reached  $0.755$ , which exceeded the one-dimensional  $r$  extreme value of 400–2400 nm. The TreeNet gradient boosting machine (TGBM) regression model had the highest modeling accuracy, with a calibration coefficient of determination ( $R^2_C$ ) =  $0.887$ , calibration root mean square error ( $RMSE_C$ ) =  $2.488\%$ , standard deviation (SD) =  $6.733\%$ , and  $r = 0.942$ . However, the partial least squares regression (PLSR) model had the strongest predictive ability, with validation coefficient of determination ( $R^2_V$ ) =  $0.787$ , validation root mean square error ( $RMSE_V$ ) =  $3.247\%$ , and relative prediction deviation (RPD) =  $2.071$ . The variable importance in projection (VIP) method could not only improve model efficiency but also increased model accuracy.  $R^2_C$  of the optimal PLSR model was  $0.733$ ,  $RMSE_C$  was  $3.028\%$ ,  $R^2_V$  was  $0.805$ ,  $RMSE_V$  was  $3.100\%$ , RPD was  $1.976$ , and Akaike information criterion (AIC) was  $151.050$ . The three-band optimized spectral indices with fractional differential pretreatment could to a certain extent break through the limitation of visible near-infrared spectrum in SMC estimation due to the lack of shortwave infrared spectra, which made it possible to quantitatively retrieve saline SMC on the basis of Zhuhai-1 hyperspectral imagery.

**Keywords:** Zhuhai-1 hyperspectral imagery; salinized soil moisture content; fractional order differential; optimized spectral indices; Ugan–Kuqa Oasis

## 1. Introduction

Soil moisture is an important factor that affects plant growth and development, and it is also a key indicator for evaluating soil quality and judging farmland moisture [1]. However, soil moisture content (SMC) is also one of the most easily changed and contaminated indicators in various physical and chemical properties of soil, and thus it is urgent to explore an efficient and reliable large-area observation method [2]. Remote sensing technology has unique advantages in large-area observations. It contains two of the most promising directions for remote sensing disciplines: microwave remote sensing and hyperspectral remote sensing [3]. Microwave remote sensing has a better effect on SMC inversion, and researchers have also carried out a lot of studies that can be fully reflected when using CiteSpace software [4,5] to analyze the data obtained in the Web of Science (WoS) website with “remote sensing retrieval of soil moisture” as the keyword. As shown in Figure 1, three of the top six keywords are related to microwave remote sensing, namely, “soil moisture and ocean salinity (SMOS)”, “surface roughness”, and “radiative transfer model”.

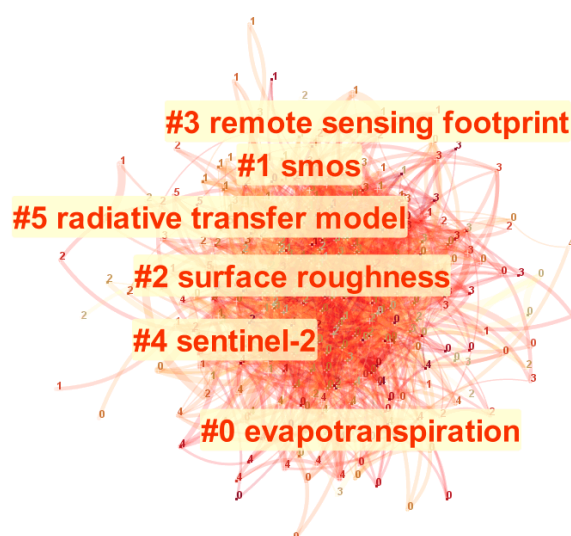


Figure 1. Keyword clustering knowledge map.

As shown in Figure 2, the current research hotspots in the field of remote sensing inversion of SMC are mainly based on SMOS, moderate-resolution imaging spectroradiometer (MODIS), and other images to carry out algorithmic research. As for the retrieval of SMC by hyperspectral remote sensing, due to the lack of hyperspectral images and low spatial resolution of images, there are fewer studies based on hyperspectral satellite images, and large-scale studies are even rarer. At present, researchers mostly carry out preliminary research on the basis of the ground-measured hyperspectral data. In the coming year, with the development of hyperspectral sensor manufacturing technology and the emergence of unmanned aerial vehicle (UAV) platforms, more and more hyperspectral images are available, such as GF-5, Zhuhai-1, PRISMA, Nano-Hyperspec, and Cubert S185, making the quantitative remote sensing research of small- and medium-scale soil properties based on hyperspectral images gradually enter the fast lane. Lu et al. compared the accuracy of quantitative inversion of partial least squares regression (PLSR) and stepwise multiple linear regression (SMLR) models based on indoor hyperspectral data, Hyperion hyperspectral images, and simulated Hyperion band hyperspectral data. The results were based on the accuracy of the quantitative inversion of soil organic carbon (SOC), total phosphorus (TP), PH, and cation exchange capacity (CEC). Hyperion image models for SOC, TP, and PH showed moderate accuracy ( $R^2_V > 0.6$ , relative prediction deviation (RPD)  $> 1.5$ ) and were generally better than other models established by indoor hyperspectral data and simulated Hyperion band hyperspectral data, which proved the feasibility and potential of Hyperion hyperspectral images in soil attribute quantitative remote sensing [6]. Tiwari et al. established an artificial neural network (ANN) model

for quantitative inversion of SOC based on Hyperion hyperspectral images and ground-measured hyperspectral data, and concluded that the ANN model was a powerful tool for inversion of SOC in farmland areas [7]. Gomez et al. conducted a study on the uncertainty analysis in the construction of the quantitative inversion model of soil clay content (SCC) based on airborne hyperspectral images acquired by the AISA-DUAL airborne imaging spectrometer. The research results provided certain guidance for improving the accuracy of the model [8,9]. At the same time, the sensitivity of carrying hyperspectral images to atmospheric and scale effects in the hyperspectral quantitative inversion of SCC had been also studied by Gomez et al., and the research results were expected to provide a theoretical basis for the development and application of related hyperspectral sensors [10]. Scholars such as Vaudour and Ouerghemmi used airborne hyperspectral images (AISA-Eagle images and HyMap images) to carry out relevant studies on the quantitative inversion of SOC ( $RMSE_V = 1.44$  g/kg) and SCC ( $R^2_V = 0.61$ ,  $RPD = 1.5$  when normalized difference vegetation index (NDVI) < 0.55), which achieved good results by the PLSR model [11,12]. Peon et al. combined airborne hyperspectral imagery acquired by the airborne hyperspectral scanner (AHS) and Hyperion satellite hyperspectral imagery to carry out a quantitative inversion study of SOC content, and finally concluded that the SMLR modeling method based on the spectral indices can achieve better inversion results ( $R^2_V = 0.60 - 0.62$  for AHS and  $R^2_V = 0.49 - 0.61$  for Hyperion) [13]. Liu et al. initially discussed the feasibility of applying the transfer learning method based on convolutional neural network to the quantitative inversion of SOC in HyMap imagery; its accuracy was  $R^2_V = 0.601$ ,  $RMSE_V = 8.62$ , and  $RPD = 1.54$  [14]. Ge et al., on the basis of the method of optimized spectral indices combined with machine learning modeling, initially carried out research on the monitoring of the SMC of farmland with UAV hyperspectral imagery acquired by the Headwall Nano-Hyperspec airborne hyperspectral imaging spectrometer. The inversion model had high accuracy ( $R^2_V = 0.907$ ,  $RMSE_V = 1.477$ , and  $RPD = 3.396$ ), which fully proved the potential of the SMC inversion by UAV hyperspectral remote sensing [15].

The optimized spectral indices method can fully extract the band combination that has the greatest  $r$  with the soil attribute content, achieve the rapid optimization effect in the complex hyperspectral parameters, and deeply mine the hyperspectral data to further improve the accuracy of the soil attribute content hyperspectral estimation model. The advantages of estimation accuracy and reduction of the influence of environmental factors on modeling have been paid more and more attention by relevant scientific researchers, and have become recent research hotspots. In the beginning, the optimized spectral indices method was initially applied in the field of agronomy and received good application results. Li et al. [16] conducted a systematic, comprehensive, and in-depth study on the optimized spectral indices method, believing that different optimized spectral indices also had different performances in the estimation of winter wheat canopy nitrogen contents. In general, two-band optimized spectral indices were better than the existing specific spectral indices, and three-band optimized spectral indices were even better than the two-band optimized spectral indices. It was predicted that the optimized spectral indices would be applied to hyperspectral remote sensing images in the future to achieve very good inversion results, potentially aiding in the design of a high-precision diagnosis system of plant nitrogen content and active remote sensing sensor of specific bands. Subsequently, the optimized spectral index was gradually introduced into the study of soil science, and the main soil parameters involved were salt, moisture, electrical conductivity, organic matter, and heavy metal content [17–25]. Wang et al. [26] used difference spectral index (DSI), ratio spectral index (RSI), and normalized difference spectral index (NDSI) band optimization algorithms to screen out the best sensitive band combination of soil salinity in the measured ground hyperspectral and image spectra, and then established a Bootstrap-BP neural network soil salt content (SSC) prediction model. The  $R^2_V$  of the model based on ground hyperspectral data was 0.95, the  $RMSE_V$  was 4.38 g/kg, and the  $RPD$  was 3.36. The estimation model based on the remote sensing image spectrum also had good accuracy, and the  $R^2_V$  was 0.91,  $RMSE_V$  was 4.82 g/kg, and  $RPD$  was 3.32. Yansenjiang et al. [27] used the optimized spectral indices to estimate the salt content of three soils with temporal and spatial heterogeneity, and revealed the effectiveness of the optimized spectral indices to eliminate heterogeneity to a certain

extent and improve universality. Li Chen et al. [28] carried out active explorations in the quantitative estimation of SMC on the basis of the optimized spectral indices method. The characteristics of SMC and soil spectral curve were comprehensively analyzed by obtaining field-measured SMC data and corresponding indoor soil surface hyperspectral reflectance data, and the accuracy of the traditional spectral modeling method and the optimized spectral indices method in the estimation of SMC was compared. It was considered that the SMC estimation model based on RSI ( $R_{1407}$ ,  $R_{1459}$ ) can more accurately estimate the SMC of coastal saline soil. Wang [29] et al. used the fractional differential method to preprocess the original hyperspectral data, and then used the optimized spectral indices method to calculate the  $r$  between the hyperspectral data and soil organic matter (SOM). It was found that the fractional differential pretreatment can also improve the  $r$  to a certain extent. The highest  $r$  was 0.52 at 1.2-order differential, while the combination of fractional differential and optimized spectral indices had a more significant effect on improving the  $r$ . The  $r$  can be increased to 0.86, which played an important role in improving the accuracy of the model. In some other studies, it was also concluded that optimized spectral indices combined with spectral transformation could achieve better modeling results and had good inversion feasibility [30,31]. Nijat et al. used the multi-dimensional modeling method to quantitatively retrieve the SSC on the basis of the WorldView-2 multi-spectral imagery [32], and the results obtained were consistent with the views of Li.

### Top 25 Keywords with the Strongest Citation Bursts



Figure 2. Analysis of emergent words.

Soil hyperspectral technology is characterized by having a large amount of information, being fast and easy, being non-destructive, and being non-polluting, among other factors, and is widely used



in SMC estimation research [33,34]. By analyzing the spectral shape characteristics of the tested soil samples in different soil types and comparing the  $r$ , researchers have found that the absorption peaks near 1400 nm and 1900 nm are the most significant bands that vary with SMC. The response law of soil hyperspectral curve to SMC change in this band was revealed, and the quantitative estimation model of SMC was constructed, and a relatively complete theoretical and methodological system was gradually formed [35–37]. Although hyperspectral technology has become an important method for predicting SMC, due to the large amount of redundant and invalid information contained in the soil hyperspectral spectrum, not only is the hyperspectral estimation model of SMC highly complex, but also the prediction accuracy of the model is affected. Therefore, the application of hyperspectral remote sensing in the retrieval of SMC is still very limited. However, the way in which to improve the prediction accuracy of remote sensing inversion models has always been a hot issue that researchers have been committed to solving in hyperspectral and even various quantitative remote sensing methods (such as microwave remote sensing, lidar remote sensing, and thermal infrared remote sensing) [38–40].

In summary, the quantitative inversion of SMC based on simulated satellite band hyperspectral data (400–1000 nm) combined with spectral fractional differential preprocessing and the three-dimensional optimized spectral indices method is still very rare and requires further research. Therefore, in this study, the indoor measured hyperspectral data and the newly emerged Zhuhai-1 hyperspectral image data were resampled to the same spectral scale. The fractional differential preprocessing method and optimized spectral indices method were used to process the simulated soil hyperspectral data, and the Pearson correlation coefficient (PCC) analysis was determined between soil hyperspectral data and measured SMC data. The variable optimization method was applied to screen out the sensitive spectral indices for the establishment of the ground hyperspectral estimation model. A variety of modeling methods were employed to select the best SMC inversion model. The main difficulties in this study were as follows: (i) 400–1000 nm is not the sensitive band of SMC in traditional methods, and it was also difficult to retrieve SMC through the 400–1000 nm band by the optimized spectral indices method; (ii) as a new method, the combination of fractional differential spectral preprocessing and three-dimensional optimized spectral indices faced challenges in the process of theory and implementation.

In this way, the major objectives of this study were to discuss the application effect of fractional differential and optimized spectral indices in hyperspectral estimation of SMC, and to verify the possibility of quantitative inversion of salinized SMC from Zhuhai-1 hyperspectral imagery. This paper provided a new research idea and practical example for improving the inversion accuracy of hyperspectral remote sensing.

## 2. Materials and Methods

### 2.1. Study Area and Sampling Sites

As shown in Figure 3, the Ugan River–Kuqa River Delta Oasis (hereinafter referred to as Ugan–Kuqa Oasis) is located in the north of the Taklimakan Desert in Xinjiang, China. It is a typical oasis in the arid area. Its research representativeness is reflected in the aspects of landform, hydrology, climate, soil, and vegetation [41–45]. Ugan–Kuqa Oasis is an alluvial plain oasis formed between Qiulitage Mountain and Tarim Basin. Its terrain slopes from the northwest to the southeast, with an average elevation of about 900 m. The geomorphological landscape is represented by Gobi, oasis, and desert. Under the influence of this topography, the Kuqa River and the Ugan River, which are formed by the melting water of the alpine ice and snow, have become the two main river systems flowing through the Ugan–Kuqa Oasis. The runoff lengths are about 126 and 452 km for Kuqa River and Ugan River, respectively, and become the main source of agricultural production and domestic water. The Ugan–Kuqa Oasis has the climatic characteristics of dryness and little rain, large temperature difference between day and night, and distinct seasons. It is characterized by a typical mid-latitude warm temperate continental arid climate with an average annual temperature of 10.5–11.4 °C. The temperature distribution characteristics in the region are opposite to the geomorphic

characteristics, and the overall situation is low in the north and high in the south, with an average annual precipitation of 51.6 mm, an average annual evaporation of 2123.7 mm, and a drought index between 17.3 and 21.8. There are many types of soil in the Ugan–Kuqa Oasis, mainly including saline-alkali soil, meadow soil, desert soil, sandy loam, fluvo-aquic soil, swamp soil, and brown calcium soil, and due to the influence of topography and climate, the phenomenon of soil salinization is very remarkable. Obviously, it is precisely for this reason that natural halophytes are widespread, with the more common ones being *Phragmites communis*, *Tamarix chinensis*, *Populus diversifolia*, *Apocynum venetum*, *Alhagi sparsifolia* Shap., *Halostachys caspica*, and *Glycyrrhiza uralensis*.

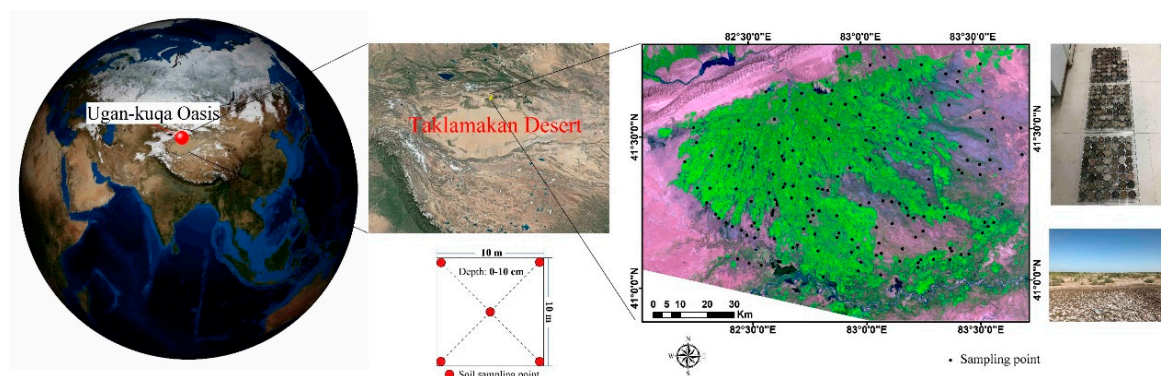


Figure 3. Maps of the study area.

## 2.2. Data Collection

### 2.2.1. Field Sampling and Spectral Measurements

Field sampling was conducted in August 2018. The main work of the field investigation was to collect surface soil samples. From half a month before the field investigation to the end of the investigation, the weather conditions were good and there was no rain. Prior to the field investigation, the location and number of the planned sampling points had been marked in the Ovitallmap software in the laboratory on the basis of previous sampling experience, and a preliminary sampling route plan had been formulated. As shown in Figure 3, the total number of sample points was 171, which were evenly distributed. They were located in the periphery of the oasis, in the oasis–desert interlace zone, and in the interior of the oasis, and covered different land use types. In the actual investigation, some samples were fine-tuned according to the actual conditions such as road accessibility. According to the 5-point sampling method, we used a ring knife to collect 5 soil samples of the 0–10 cm section of the soil in a 10 × 10 m square, and placed them in the aluminum boxes and sampling bags marked with the corresponding numbers. In the record book, the specific coordinates of the sampling point, the weight of the aluminum box filled with soil, the land use, the degree of soil salinization, the vegetation type and coverage, etc., were recorded. Finally, the landscape photos around the sampling point were taken by camera. After the soil sample was brought back to the laboratory, the spectrum of the soil in the aluminum box was first measured indoors. The indoor spectrum was measured using a ASD FieldSpec 3 spectrometer and was observed with a contact probe. Each soil sample was repeatedly observed 10 times and the arithmetic mean was taken as the reflectance value, and the whiteboard calibration was performed before each measurement [27]. Then, the soil in the aluminum box was dried in an oven and weighed immediately after being taken out of the oven to obtain the SMC of the soil sample. In this way, we could obtain the true SMC of 171 field soil samples and their corresponding soil hyperspectral data.

On the other hand, in order to acquire SMC control experiment data, we selected a representative neutral sandy loam sample area (the main soil type of the Ugan–Kuqa Oasis is sandy loam) during the field investigation, and approximately 20–30 kg of soil sample was collected and brought back to the laboratory. First, the soil sample was dried naturally to remove plants, plastics, and other impurities.

Next, the soil sample was washed with distilled water, with this fully removing the salt and other water-soluble substances in the soil sample. After each washing, the PH value of the filtrate was measured by the PHS-3C acidity meter to ensure the neutralization of the soil sample. Then, the soil sample was naturally dried. After this, the soil sample was grinded and sieved. Finally, the soil sample was used for the indoor water control experiment. In order to obtain soil samples that were close to the natural soil state that had different gradients of water content, we added 20 mL of distilled water to 100 aluminum boxes that had recorded their own weight, and then we added 50 g of neutral soil and let stand for half an hour. In order to ensure that the soil was fully saturated with water, we then placed all aluminum boxes in a drying oven set at 35 °C. The SMC of the soil samples collected in the field was between 0 and 37%. In addition, the SMC data measured by the control experiment should be of a random normal distribution, and thus 10 aluminum boxes in the oven should be taken out at certain appropriate intervals and quickly weighed to determine the SMC, with the corresponding hyperspectral data then being obtained by spectrometer.

The soil hyperspectral data measured by the ASD FieldSpec 3 hyperspectrometer was processed and exported by ViewSpec Pro software. In order to reduce the influence of noise, we removed the edge bands (350–399 nm and 2401–2500 nm) with low signal noise. The hyperspectral data (400–2400 nm) of soil samples were smoothed and denoised by the Savitzky–Golay filtering method [24].

### 2.2.2. Zhuhai-1 Hyperspectral Imagery

Zhuhai-1 hyperspectral Satellite (OrbitaHyperSpectral, OHS) is the second group of satellites in the Zhuhai-1 satellite constellation. It was successfully launched and operated normally on 26 April 2018. Four hyperspectral satellites (OHS-2A, OHS-2B, OHS-2C, and OHS-2D) are distributed on the same orbital surface; the orbit height is 510 km, the spatial resolution of each Zhuhai-1 hyperspectral image is 10 m, the push-broom imaging method is adopted and the width is 150 km, the signal-to-noise ratio is better than 300, the spectral range is between 400 and 1000 nm and it has 32 bands (the center wavelength is shown in Table 1), the spectral resolution is 2.5 nm, and the revisit period of the 4 Zhuhai-1 hyperspectral satellite network is 2 days [46,47].

**Table 1.** Center wavelength of detection band of Zhuhai-1 hyperspectral imagery/nm.

Channel	Center	Channel	Center	Channel	Center	Channel	Center
B01	466	B09	596	B17	716	B25	836
B02	480	B10	610	B18	730	B26	850
B03	500	B11	626	B19	746	B27	866
B04	520	B12	640	B20	760	B28	880
B05	536	B13	656	B21	776	B29	896
B06	550	B14	670	B22	790	B30	910
B07	566	B15	686	B23	806	B31	926
B08	580	B16	700	B24	820	B32	940

The band range of the spectrometer is 400–2400 nm, and the band interval is the same, which is 1 nm, while the band range of the Zhuhai-1 hyperspectral data is 400–1000 nm, which only contains 32 bands, and the band interval is not the same (Table 1). A necessary condition for the use of the fractional differential method in subsequent research is that the band interval is a constant, and thus further processing of the Zhuhai-1 hyperspectral image is required. Subsequent research can use the linear interpolation method based on Interactive Data Language (IDL) software to obtain the interpolated Zhuhai-1 imagery that meets the band conditions. By repeatedly testing, it is found that the difference in spectral reflectance curves before and after interpolation can be minimized, keeping the spectral information of the image free from distortion when the band interval is 8 nm and the interpolated Zhuhai-1 image contains 60 bands (Table 2). In addition, in order to enable the ground spectrum band to correspond to the center band of the satellite spectrum, we can select the corresponding band from the ground measured spectrum for the construction of the inversion model.

**Table 2.** Center wavelength of the interpolated Zhuhai-1 imagery/nm.

Channel	Center	Channel	Center	Channel	Center	Channel	Center
B01	466	B16	586	B31	706	B46	826
B02	474	B17	594	B32	714	B47	834
B03	482	B18	602	B33	722	B48	842
B04	490	B19	610	B34	730	B49	850
B05	498	B20	618	B35	738	B50	858
B06	506	B21	626	B36	746	B51	866
B07	514	B22	634	B37	754	B52	874
B08	522	B23	642	B38	762	B53	882
B09	530	B24	650	B39	770	B54	890
B10	538	B25	658	B40	778	B55	898
B11	546	B26	666	B41	786	B56	906
B12	554	B27	674	B42	794	B57	914
B13	562	B28	682	B43	802	B58	922
B14	570	B29	690	B44	810	B59	930
B15	578	B30	698	B45	818	B60	938

### 2.3. Fractional Order Differential Method

Fractional order differential was first proposed by the mathematician Gottfried in 1695. In the course of more than 300 years of development, there have been many definitions and there is no uniformity. In this article, the widely used Grünwald–Letnikov (GL) expression is used [48–50]. It is generalized from the definition of integer order differential, and its definition is

$$d^{\nu} f(x) = \lim_{h \rightarrow 0} \frac{1}{h^{\nu}} \sum_{n=0}^{\frac{t-k}{h}} (-1)^n \frac{\Gamma(\nu+1)}{n! \Gamma(\nu-n+1)} f(x-nh) \quad (1)$$

where the function  $f(x)$  is the reflectance of the spectral curve, the wavelength range is  $[k, t]$ ,  $\nu$  is the order,  $h$  is the step size,  $n$  is a constant, and  $\Gamma(\nu)$  is the Gamma function, and its expression is

$$\Gamma(\nu) = \int_0^{\infty} \exp(-u) u^{\nu-1} du = (\nu-1)! \quad (2)$$

From Equations (1) and (2), the difference expression of the fractional differential of the unary function  $f(x)$  can be derived as

$$\begin{aligned} \frac{d^{\nu} f(x)}{dx^{\nu}} &\approx \frac{1}{h^{\nu}} f(x) + \frac{(-\nu)}{h^{\nu}} f(x-h) \\ &+ \frac{(-\nu)(-\nu+1)}{2h^{\nu}} f(x-2h) \\ &+ \frac{(-\nu)(-\nu+1)(-\nu+2)}{6h^{\nu}} f(x-3h) \\ &+ \dots \\ &+ \frac{(-\nu)(-\nu+1)(-\nu+2) \dots (-\nu+n-1)}{n! h^{\nu}} f(x-nh) \end{aligned} \quad (3)$$

In order to make the formula easier to understand, Equation (3) is further simplified as

$$\begin{aligned} \frac{d^{\nu} f(x)}{dx^{\nu}} &\approx \frac{1}{h^{\nu}} [f(x) + (-\nu) f(x-h) \\ &+ (-\nu) \frac{(-\nu+1)}{2} f(x-2h) \\ &+ (-\nu) \frac{(-\nu+1)}{2} \frac{(-\nu+2)}{3} f(x-3h) \\ &+ \dots \\ &+ (-\nu) \frac{(-\nu+1)}{2} \frac{(-\nu+2)}{3} \dots \frac{(-\nu+n-1)}{n} f(x-nh)] \end{aligned} \quad (4)$$

When  $\nu = 1, 2$ , the expressions are

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (5)$$

$$\begin{aligned} f''(x) &= \lim_{h \rightarrow 0} \frac{f'(x+h) - f'(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2} \end{aligned} \quad (6)$$

According to the integer-order differential theory, Equation (5) and Equation (6) are, respectively, the first-order and second-order differentials of the derivable function  $f(x)$ .

At present, fractional differentiation has been widely used in signal filtering, pattern recognition, fractal theory, etc., and has achieved better results than integer-order differentiation. However, its application in remote sensing and spectroscopy is still lacking. Research and application of fractional differentiation are ignored. On the other hand, most of the current studies on spectral differentiation generally take  $h$  as 1, ignoring the situation where  $h$  is other values, and there are few studies involving different spectral scales. In this study, according to the actual situation of the satellite band,  $h$  is taken as 8. The calculation of the fractional differential of the spectral curve is realized by using a small piece of software written in JAVA language, and the calculation of the fractional differential of the remote sensing image can be processed by writing code in the IDL software.

#### 2.4. Optimized Spectral Indices Method

##### 2.4.1. Two-Band Spectral Indices

In light of the fact that the spectral index defined by the predecessors has clear theoretical significance, it still inevitably has a certain regionality, and thus the blind search method of optimized spectral indices is used to redefine the spectral indices suitable for the study area. The two-band spectral indices band optimization algorithm can calculate all possible spectral indices formed by the pairwise combination of different bands in the spectrum. This research comprehensively selects the following common two-band spectral indices formula for optimization, which are NDSI, RSI, difference index (DI), normalized perpendicular drought index (NPDI) [20], chlorophyll index (CI), soil index 2 (SI-2), and soil index 4 (SI-4) [17]. Its algorithm expression are

$$NDSI(i, j) = (R_i - R_j) / (R_i + R_j) \quad (7)$$

$$RSI(i, j) = R_i / R_j \quad (8)$$

$$DI(i, j) = R_i - R_j \quad (9)$$

$$NPDI(i, j) = (R_i + R_j) / R_j \quad (10)$$

$$CI(i, j) = (1/R_i - 1/R_j) \times R_j \quad (11)$$

$$SI-2(i, j) = \sqrt{R_i \times R_j} \quad (12)$$

$$SI-4(i, j) = \sqrt{R_i^2 \times R_j^2} \quad (13)$$

where  $R$  is the spectral reflectance value, and the subscripts ( $i$  nm and  $j$  nm) are wavelengths in nanometers.

##### 2.4.2. Three-Band Spectral Indices

In addition, the three-band spectral indices band optimization algorithm can calculate all possible spectral indices formed by the combination of different three bands in the spectrum. This research also comprehensively selected some common three-band spectral indices formulae for optimization,



these being soil index 1 (SI-1), soil index 3 (SI-3), normalized perpendicular drought index 3 (NPDI-3) [16], three band index 1 (TBI-1) [17], three band index 2 (TBI-2), three band index 3 (TBI-3), modified simple ratio index 1 (MSRI-1), modified simple ratio index 2 (MSRI-2), triangle vegetation index (TVI), modified triangle vegetation index (MTVI), modified normalized difference vegetation index (MNDVI), and health index (HI). The algorithm is defined as follows

$$SI-1(i, j, n) = R_i \times R_j / R_n \quad (14)$$

$$SI-3(i, j, n) = \sqrt{R_i \times R_j \times R_n} \quad (15)$$

$$NPDI-3(i, j, n) = (R_i / R_j - 1) / [(R_i - R_n) / (R_i + R_n)] \quad (16)$$

$$TBI-1(i, j, n) = R_i / (R_j + R_n) \quad (17)$$

$$TBI-2(i, j, n) = (R_i - R_j + 2 \times R_n) / (R_i + R_j - 2 \times R_n) \quad (18)$$

$$TBI-3(i, j, n) = (R_i - R_j + 2 \times R_n) / (R_i + R_j - R_n) \quad (19)$$

$$MSRI-1(i, j, n) = (R_i - R_j) / (R_n + R_j) \quad (20)$$

$$MSRI-2(i, j, n) = (R_i - R_j) / (R_n - R_j) \quad (21)$$

$$TVI(i, j, n) = 0.5 \times [120 \times (R_i - R_j) - 200 \times (R_n - R_j)] \quad (22)$$

$$MTVI(i, j, n) = 1.2 \times [1.2 \times (R_i - R_j) - 2.5 \times (R_n - R_j)] \quad (23)$$

$$MNDVI(i, j, n) = (R_i - R_j) / (R_i + R_j - 2 \times R_n) \quad (24)$$

$$HI(i, j, n) = (R_i - R_j) / (R_i + R_j) - 0.5 \times R_n \quad (25)$$

where  $R$  is the spectral reflectance value, and the subscripts ( $i$  nm,  $j$  nm, and  $n$  nm) are wavelengths in nanometers.

The calculation of the two-band optimized spectral indices and the three-band optimized spectral indices are respectively realized by the two-band spectral indices optimization software and the three-band spectral indices optimization software, independently developed on the basis of JAVA language.

## 2.5. Modelling Strategies

Considering that the modeling data and verification data need to be able to fully reflect the actual status of the SMC in the study area, we arranged 171 soil samples in ascending order of SMC, and 129 pieces of data were extracted at equal intervals for the calibration set and 42 pieces of data for the validation set. Sets are used for model establishment and accuracy verification [27].

The models are mainly divided into traditional statistical models and machine learning models, with this research comparing the application effects of the two model strategies. In light of the characteristics of the two models, when performing statistical modeling, the independent variable selection criterion  $|r|$  was greater than or equal to 0.750, whereas when performing machine learning modeling, the independent variable selection criterion  $|r|$  was greater than or equal to 0.249 (171 samples passed the 0.001 significance level test).

### 2.5.1. Optimal Variable Selection Method

#### (1) PCC method

In soil spectroscopy, it is generally believed that selecting some sensitive bands or spectral indices to establish an estimation model is beneficial to the improvement of the model accuracy. Moreover,

the PCC method is usually used to screen the sensitive bands or spectral indices that have a high  $r$  with the SMC by setting the  $r$  threshold. In statistics,  $r$  is defined as

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (26)$$

In this study,  $X$  represents the spectral reflectance value of a certain band,  $Y$  represents the measured SMC value, and the  $r$  value is between  $-1$  and  $1$ . The absolute value is closer to  $1$ , indicating the difference between  $X$  and  $Y$  is better.  $r$  is also calculated on the basis of the spectral indices optimization software independently developed by JAVA language.

## (2) VIP method

The variable importance criterion was first proposed by Wold [51]. The VIP value represents the degree to which the independent variable fits the model. The independent variable with a VIP value less than  $1$  has a small contribution to the dependent variable and can be considered to be eliminated. Therefore, the threshold of the independent variable VIP in this study was greater than or equal to  $1$ . The VIP value was calculated in Xlstat 2019 software, and its calculation formula is as follows

$$VIP_j = \sqrt{\frac{\sum_{f=1}^F W_{jf}^2 \times SSY_f \times J}{SSY_{total} \times F}} \quad (27)$$

where  $W_{jf}$  is the weight value of the  $j$  variable and the  $f$  component,  $SSY_f$  is the sum of squares of the variances of the  $f$  component and the  $J$  variables,  $SSY_{total}$  is the total sum of squares of the explanatory dependent variable, and  $F$  is the total number of components.

## 2.5.2. Modeling Approaches

A total of 8 models were used in this study, namely, multiple linear regression (MLR) model, Poisson regression (PR) model, nonlinear regression (NR) model, PLSR model, random forest (RF) regression model, multiple adaptive regression spline (MARS) model, classification and regression trees (CART) model, and regression model with TreeNet gradient boosting machine (TGBM). Among them, this article completed the modeling process of MLR in Unscrambler X 10.4 software, and completed the modeling process of PR, NR, and PLSR in Xlstat 2019 software. In Salford Predictive Modeler v8.2 software, the modeling process of RF, MARS, CART, and TGBM were completed [52]. A brief introduction of several models is provided below.

Regression model is used to establish the functional expression of the regression relationship (called regression equation) between the dependent variable(s) and independent variable(s) on the basis of mastering a large amount of observed data. In particular, machine learning regression model basically has no model assumptions, can process almost all kinds of data, and has no special requirements for input data.

## (1) MLR model

The MLR model assumes that multiple independent variables and single dependent variables have a linear relationship, which can be expressed by linear function. Therefore, it is necessary to input multiple independent variables and single dependent variable observation data when using this model.

## (2) PR model

The PR model is a kind of regression analysis that is used to model categorical data and contingency tables. The PR model assumes that the dependent variable is Poisson distribution and the logarithm of its expected value can be modeled by linear combination of unknown parameters. The PR model is sometimes called the log-linear model, especially when used as a contingency table model.

## (3) NR model

The NR model is a regression model in which the regression function has a nonlinear structure with respect to the unknown regression coefficients. The main assumption of NR model is that the dependent variable is more than one-order functional form of the independent variable. In this study, the single variable nonlinear regression model was used, and only the optimized spectral index with the highest  $r$  with SMC was selected as the independent variable. This article used a total of 18 different regression functions, and finally used the best regression function to represent the NR model.

## (4) PLSR model

Partial least squares (PLS) is a mathematical optimization technique that finds the best function match for a set of data by minimizing the sum of squares of errors [32].  $PLSR \approx MLR + \text{canonical correlation analysis (CCA)} + \text{principal component analysis (PCA)}$ . The PLSR method is also called the bilinear factor model. It mainly solves the regression problem of multiple dependent variables to multiple independent variables. It is especially suitable when the prediction matrix has more variables than the observations, and when there is multicollinearity in the value of independent variables. The general multivariate underlying model of PLSR is

$$\begin{aligned} X &= TP^T + E \\ Y &= UQ^T + F \end{aligned} \quad (28)$$

where  $X$  is a prediction matrix of  $n \times m$ ,  $Y$  is a response matrix of  $n \times p$ .  $T$  and  $U$  are matrices of  $n \times l$ , which are the projection of  $X$  and  $Y$ , respectively.  $P$  and  $Q$  are, respectively, the orthogonal load matrix of  $m \times l$  and  $p \times l$ , and the matrices  $E$  and  $F$  are error terms.

## (5) CART model

CART is a learning method to output the conditional probability distribution of random variable  $Y$  under the condition of given input random variable  $X$ . It can handle continuous and category data. CART algorithm is a binary recursive segmentation technology that divides the current sample into two sub-samples, so that each non-leaf node generated has two branches, and thus the decision tree generated by the CART algorithm is a binary tree with a simple structure. The CART model uses post-pruning, that is, pruning according to the verification data.

## (6) RF regression model

RF is an ensemble learning algorithm that belongs to bagging type. By combining multiple weak classifiers, the final result is voted or averaged, which makes the results of the whole model have high accuracy and generalization performance. In other words, the bagging method that uses CART decision trees as weak classifiers is generally called RF. Random is the core of RF. The correlation between decision trees is reduced by randomly selecting samples and features.

## (7) MARS model

The MARS model is a regression method specifically used for high-dimensional data and has a strong generalization ability. The regression method uses the tensor product of the spline function

as the basis function, and the determination of the basis function and the number of basis functions are automatically determined by the data. No manual selection is required. It is divided into three steps: forward process, backward pruning process, and model selection. Its advantage lies in its ability to process data with large data volume and high dimensionality, as well as fast calculation and accurate model.

#### (8) TGBM regression model

Boosting is an idea, and gradient boosting is a method of implementing boosting. Its main idea is that every time a model is established, the gradient descent direction of the model loss function is established before. It can consistently generate extremely accurate models. The algorithm typically generates thousands of small decision trees built in a sequential error-correcting process to converge to an accurate model. The TreeNet modeling engine's level of accuracy is usually not attainable by single models or by ensembles such as bagging or conventional boosting. As opposed to neural networks, the TreeNet methodology is not sensitive to data errors and needs no time-consuming data preparation, pre-processing, or imputation of missing values.

#### 2.5.3. Comparison of Model Accuracy

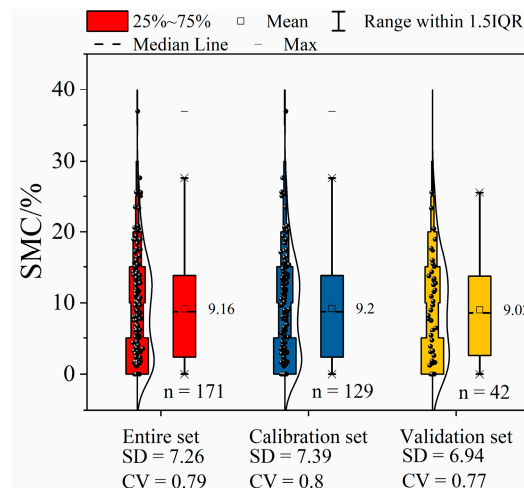
The optimal model was selected by comparing the accuracy parameters of each model for the inversion of the SMC in the study area. The accuracy parameters included the  $R^2_C$ ,  $R^2_V$ ,  $RMSE_C$ ,  $RMSE_V$ , RPD, and Akaike information criterion (AIC) [24].  $R^2_C$  and  $R^2_V$  determine the stability of the model. The closer the  $R^2_C$  and  $R^2_V$  is to 1, the better the stability of the model.  $RMSE_C$  and  $RMSE_V$  are used to characterize the accuracy of the model. The closer the  $RMSE_C$  and  $RMSE_V$  is to 0, the higher the accuracy of the model. The RPD threshold is divided into three categories. When  $a < 1.4$ , the model's estimation ability is low, when  $b \geq 1.4$ ,  $b < 2.0$ , the model's estimation ability is feasible. When  $c \geq 2.0$ , the model's estimation ability is good. In addition, AIC, as an index to evaluate the efficiency of the model can weigh the complexity of the estimated model and the goodness of the model fitted to the data. The smaller the value of AIC, the better the model can explain the data with the least free parameters, making it better at avoiding overfitting [32].

### 3. Results

#### 3.1. Analysis of SMC and Hyperspectral Characteristics

##### 3.1.1. Descriptive Statistical Analysis of SMC

It can be seen from Figure 4 that the structures of the three datasets were basically the same—the mean values were 9.16%, 9.20%, and 9.02%, and the SDs were 7.26%, 7.39%, and 6.94% for the entire set, calibration set, and validation set, respectively. The coefficient of variation (CV) corresponding to the calibration set and the validation set were 0.80 and 0.77, respectively. CV of SMC at all sampling points in the study area was 0.79, which was between the calibration set and the validation set. The degree of data dispersion was strong ( $CV > 0.36$ ). In addition, the maximum soil salinity in the study area was 60.37 mg/g, the minimum was 0.08 mg/g, and the average was 15.50 mg/g. The soil is generally strongly saline soil.

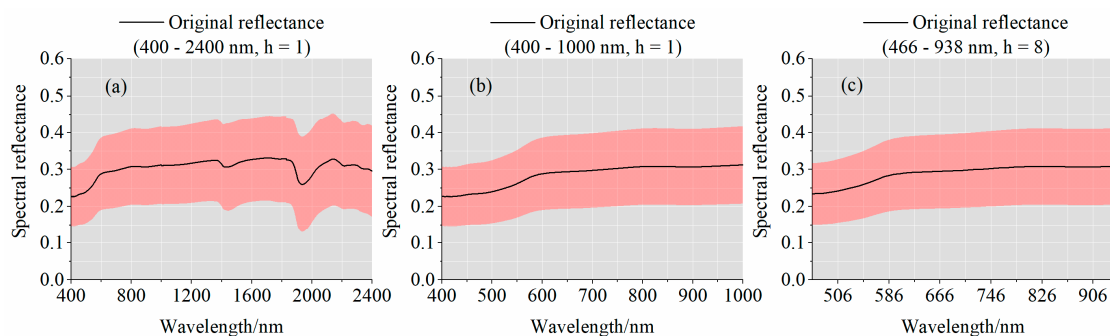


**Figure 4.** Statistical box diagram of soil moisture content (SMC) in different datasets.

### 3.1.2. Soil Hyperspectral Characteristics Analysis

#### (1) Soil average spectral curve analysis

Figure 5 shows the average spectral curves of 171 soil samples in the 400–2400 nm band with 1 nm intervals, the average spectral curves in the 400–1000 nm band with 1 nm intervals, and the average spectral curves in the 466–938 nm band with 8 nm intervals. Figure 5b is a part of Figure 5a. The difference between Figure 5b,c is that the band interval was different, that is, the scale was different. From the curve shape, there was almost no difference between them, indicating that the 8 nm spectral scale can also reflect the hyperspectral characteristics of soil, which laid the foundation for the following research.



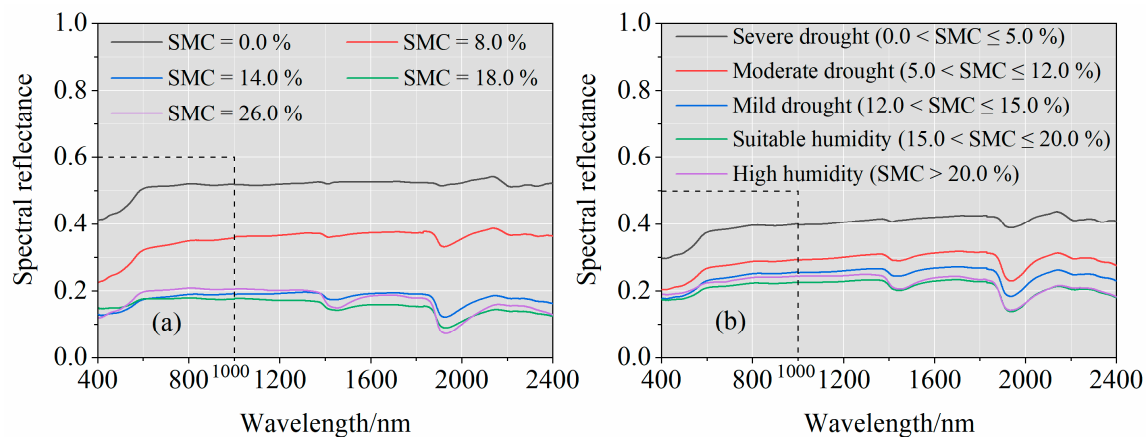
**Figure 5.** Soil average spectral curve. (a) the average spectral curves of 171 soil samples in the 400–2400 nm band with 1 nm intervals; (b) the average spectral curves in the 400–1000 nm band with 1 nm intervals; (c) the average spectral curves in the 466–938 nm band with 8 nm intervals.

#### (2) Analysis of soil spectral characteristics under different humidity conditions

Figure 6a is a partial soil spectral curve selected from the indoor soil moisture control experimental spectral data. Figure 6b is a spectral curve obtained after averaging 171 soil samples according to the soil moisture grading standard. Figure 6a was compared with Figure 6b, and it can be seen that the SMC in Figure 6a was within the corresponding SMC classification interval in Figure 6b. The basic spectral morphological characteristics of the two were the same, and both conformed to the general law of the influence of SMC on the soil spectrum. In the SMC range, the spectral reflectance decreased with the increase of SMC. On the other hand, there were many subtle differences between the two. Because the control experiment eliminated the interference of many factors on the soil spectrum, such as soil type, salinity, particle size, and surface roughness, the SMC was one-dimensionally related



to the original spectrum. The performance can reach  $-0.926$ , but in the actual situation, the undisturbed soil in the field will have heterogeneity due to various complex factors. This was the main reason for the difference in spectral characteristics. The optimized spectral indices method was used to eliminate the heterogeneity to the greatest extent. This was one of the core goals of this article.



**Figure 6.** Soil spectrum curves under different humidity conditions. (a) spectral curves selected from the indoor soil moisture control experimental spectral data; (b) spectral curves obtained after averaging 171 soil samples according to the soil moisture grading standard.

According to China's soil moisture grading standards [53], as shown in Table 3, SMC was divided into the following five categories:

**Table 3.** Grading table of soil drought degree.

Degree of Drought	SMC
High humidity	>20%
Suitable humidity	15–20%
Mild drought	12–15%
Moderate drought	5–12%
Severe drought	<5%

### (3) Analysis of soil spectral characteristics of different fractional differential orders

According to the fractional differential formula, we calculated the original spectral curves of the 171 soil samples. The average spectral curve, as shown in Figure 7, was obtained with 0.25 order intervals. It can be seen from the figure that the 0.00 order differential was the original spectral curve itself (Figure 5b), and the 1.00 and 2.00 order differentials were consistent with the normal results, indicating the correctness of the fractional order differential calculation process. In addition, it can also be found that the 0.25–0.75 order differential curve and the 1.25–1.75 order differential curve were obviously different in terms of the curve shape. In comparison, the 0.25–0.75 order differential curve was smoother.

### 3.2. PCC Analysis of SMC and Spectrum

#### 3.2.1. One-Dimensional PCC Analysis of SMC and Spectrum

When the traditional hyperspectral processing method was used to model the hyperspectral SMC, the sensitive band was usually determined by the PCC analysis between the SMC and the hyperspectral reflectance on a one-dimensional level. The higher the  $r$ , the higher the sensitivity of the waveband, and the better the effect of spectral modeling estimation. It can be seen from Figure 8b that the maximum  $r$  value in the 400–1000 nm band was  $-0.635$ , which was the same as the maximum  $r$  value

in the 466–938 nm band, and the position of the sensitive band was also roughly the same (Figure 8c). To a certain extent, the viewpoints in Section 3.1.2 were also verified. While in the 400–2400 nm band (Figure 8a), due to the presence of some short-wave infrared bands, the sensitive bands mainly appeared in the moisture absorption band around 2000 nm, and the  $r$  extreme value was much higher than the visible near-infrared band (400–1000 nm), reaching  $-0.737$ , which showed from one side that the SMC based on visible near-infrared spectroscopy had certain disadvantages.

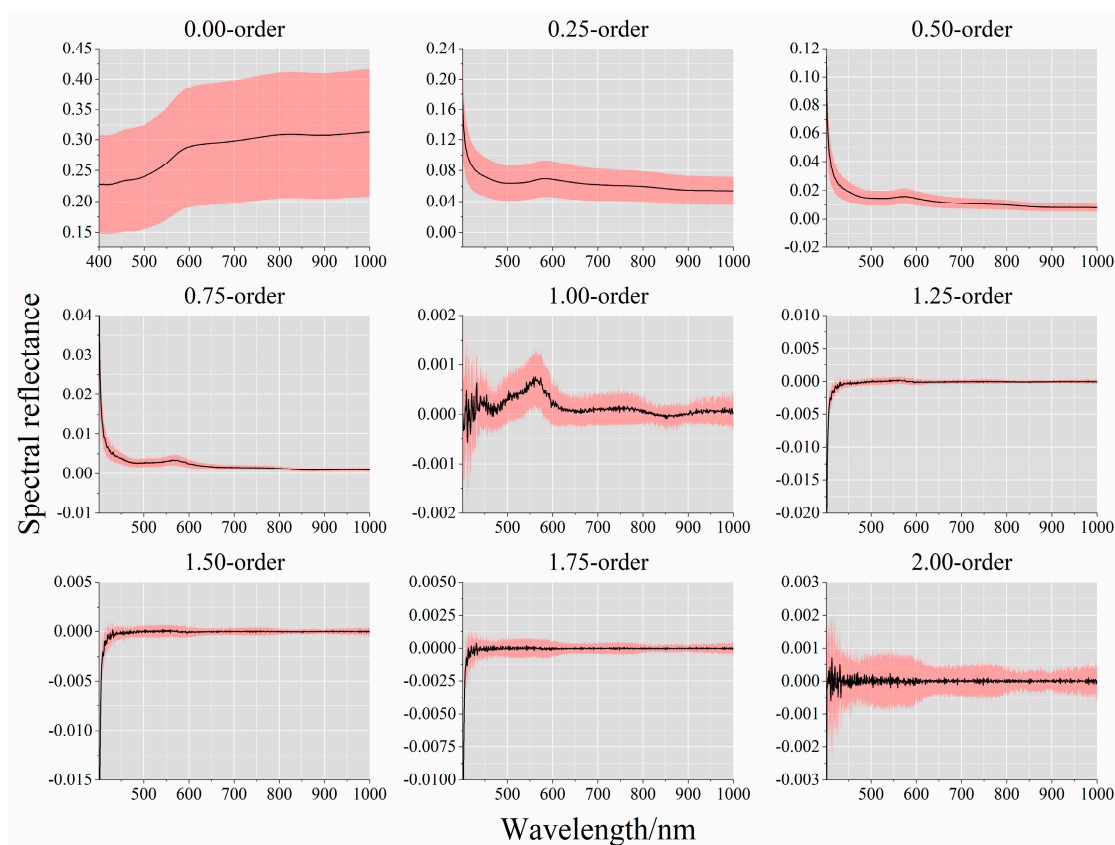


Figure 7. Soil spectrum curves of different fractional differential orders.

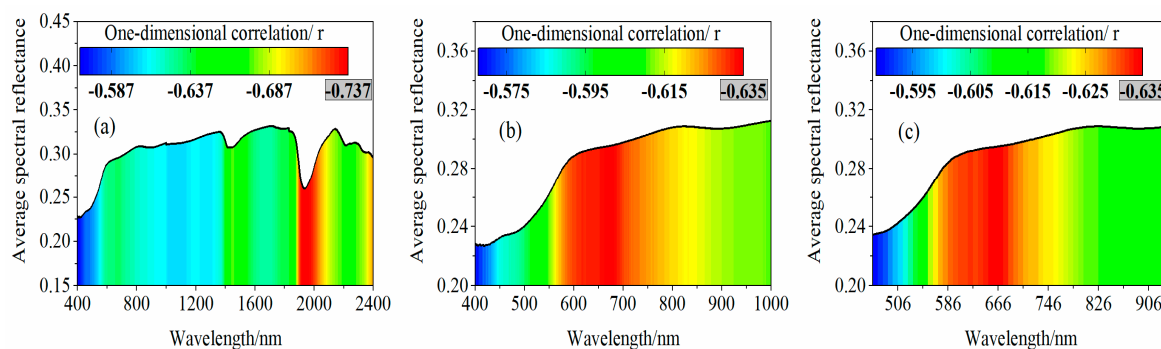


Figure 8. One-dimensional  $r$  diagram between SMC and original spectrum. (a)  $r$  diagram between SMC and 400–2400 nm spectra; (b)  $r$  diagram between SMC and 400–1000 nm spectra; (c)  $r$  diagram between SMC and 466–938 nm spectra.

The purpose of pretreatment of hyperspectral reflectance was to improve the  $r$  between soil water, salt content, and spectral reflectance, as well as to further improve the accuracy of the prediction model. Therefore, this study used several forms of commonly used nonlinear mathematical transformation

processing and fractional differential processing on the original spectrum, and the analysis results are shown in Figure 9. Traditional nonlinear mathematical transformations included square root transformation, reciprocal transformation, logarithmic transformation, reciprocal logarithmic transformation, and reciprocal logarithmic transformation, among which logarithmic transformation is also called absorbance spectrum. It was not difficult to see from the graph that the absolute value of the  $r$  extreme value increased to 0.653 after spectral preprocessing, indicating that the mathematical transformation method played a certain role. In particular, the maximum value was generated by logarithmic transformation and reciprocal logarithmic transformation. In the fractional differential preprocessing, the promotion effect of 0.5 order differential was the best. On the other hand, it can also be found that not all preprocessing methods can promote the effect. For example, the fractional differential processing with 0.75 order weakened the  $r$ .

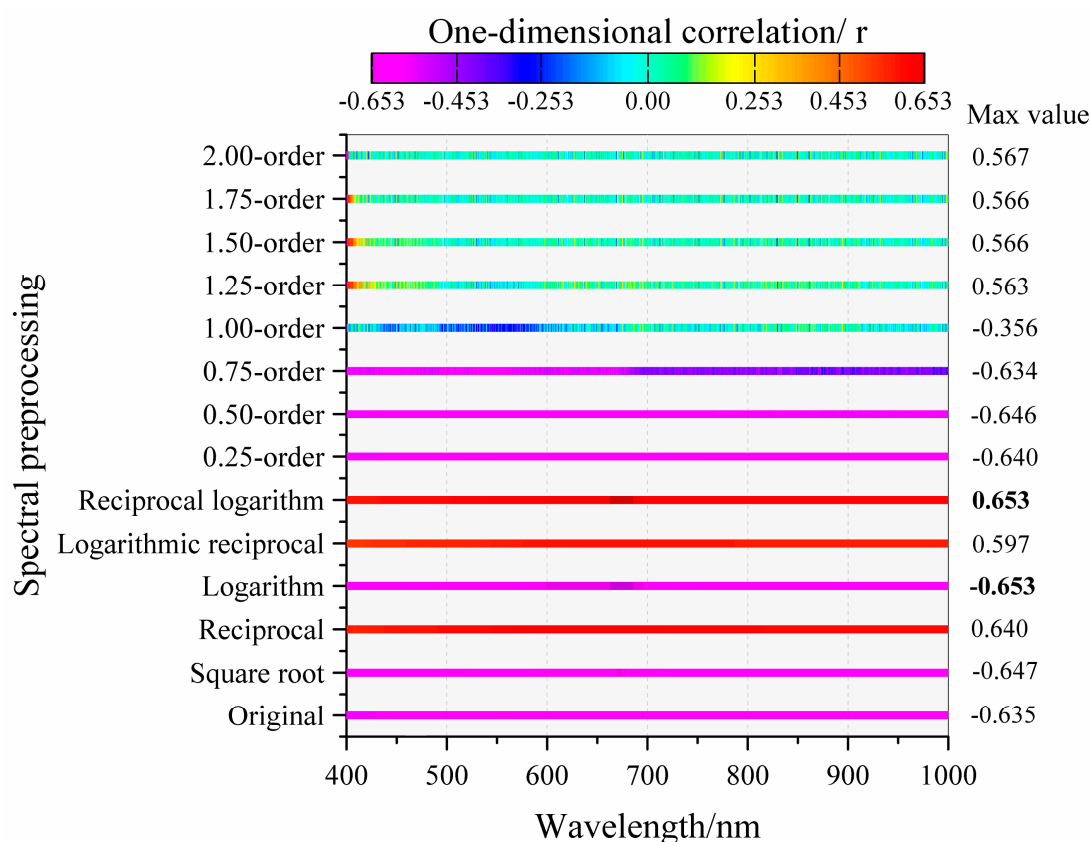


Figure 9. One-dimensional  $r$  diagram of SMC and transform spectrum.

The reciprocal logarithmic transformation and 0.5-order differential transformation that had the best effect at 400–1000 nm was selected from traditional mathematical transformations and fractional differential transformations to preprocess the original spectrum of 466–938 nm in turn, in order to obtain 0.5-order absorbance spectra. The  $r$  results are shown in Figure 10. It can be seen that the  $r$  increased steadily after combining the two mathematical transformations, reaching 0.665. Therefore, in the following two-dimensional and three-dimensional analysis, we used the 0.5-order absorbance spectrum in the calculation.

### 3.2.2. Two-Dimensional PCC Analysis of SMC and Spectrum

The two-dimensional  $r$  refers to the PCC analysis result between the SMC and the two-band optimized spectral indices. In the wavelength range of 400–2400 nm, seven two-band optimization algorithms were used to calculate the optimal spectral indices of all possible two-band combinations on the basis of the original hyperspectral reflectance of the soil sample, as well as to perform PCC

analysis with the SMC. The two-dimensional  $r$  heat map made in Origin Pro 2020 software is shown in Figure 11. It can be seen from the figure that the RSI, CI, and NPDI produced the same  $r$  extreme value, namely,  $-0.875$ , which was greatly improved in comparison with the one-dimensional  $r$  extreme value  $-0.737$  shown in Section 3.2.1. It showed that the two-dimensional  $r$  calculation based on the optimized spectral indices had a significant effect on improving the  $r$ , which fully proved the effectiveness of the optimized spectral indices. However, not all optimized spectral indices were useful, and some may even reduce the  $r$ . Figure 11 shows that the two-dimensional maximum  $r$  of the SI-4 was  $-0.673$ ; on the contrary, it was lower than the maximum one-dimensional  $r$ , which had the effect of weakening the  $r$ . The two-dimensional maximum  $r$  of the SI-2 was the same as the maximum one-dimensional  $r$ , which had no effect. It was precisely because of this possibility that it was necessary to try to optimize as many spectral indices as possible, as well as to find the index with the best effect to participate in the subsequent SMC estimation model.

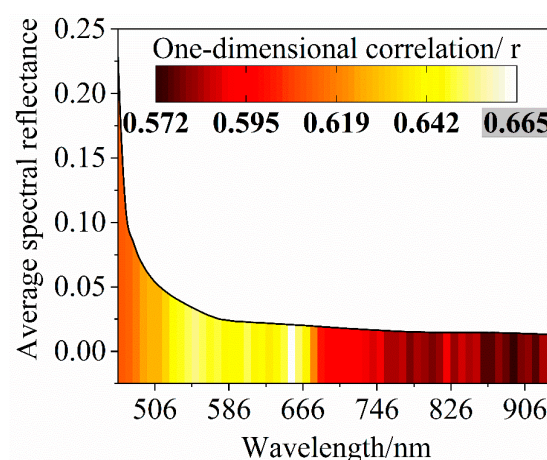


Figure 10. One-dimensional  $r$  diagram between SMC and 0.5-order absorbance spectrum.

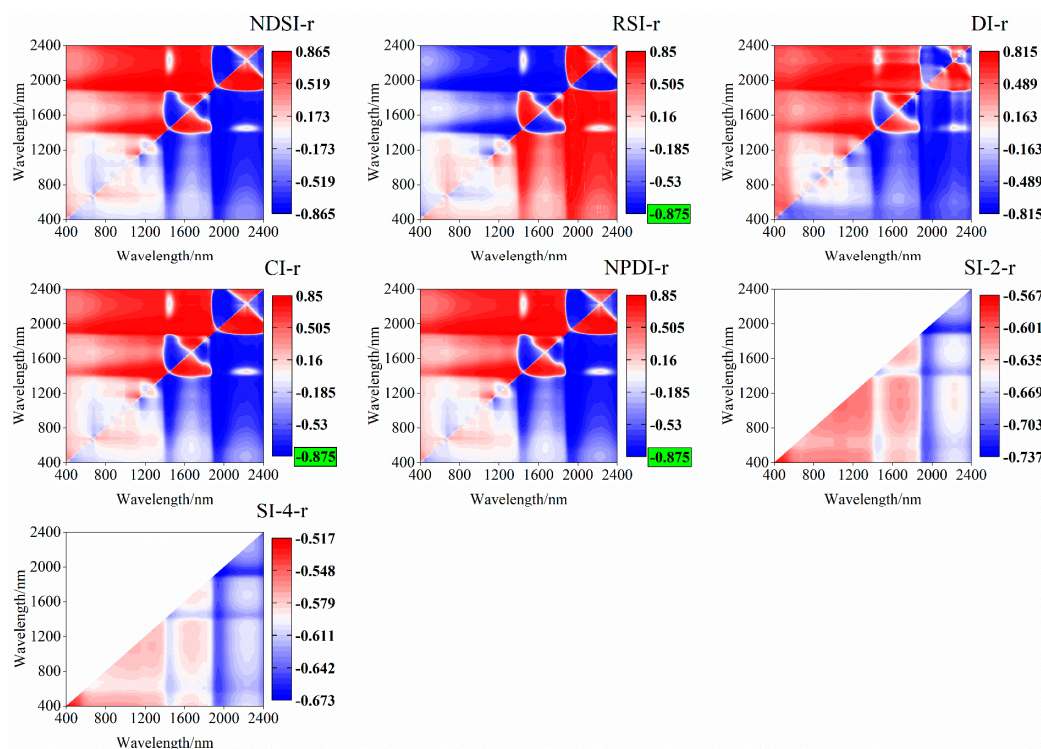
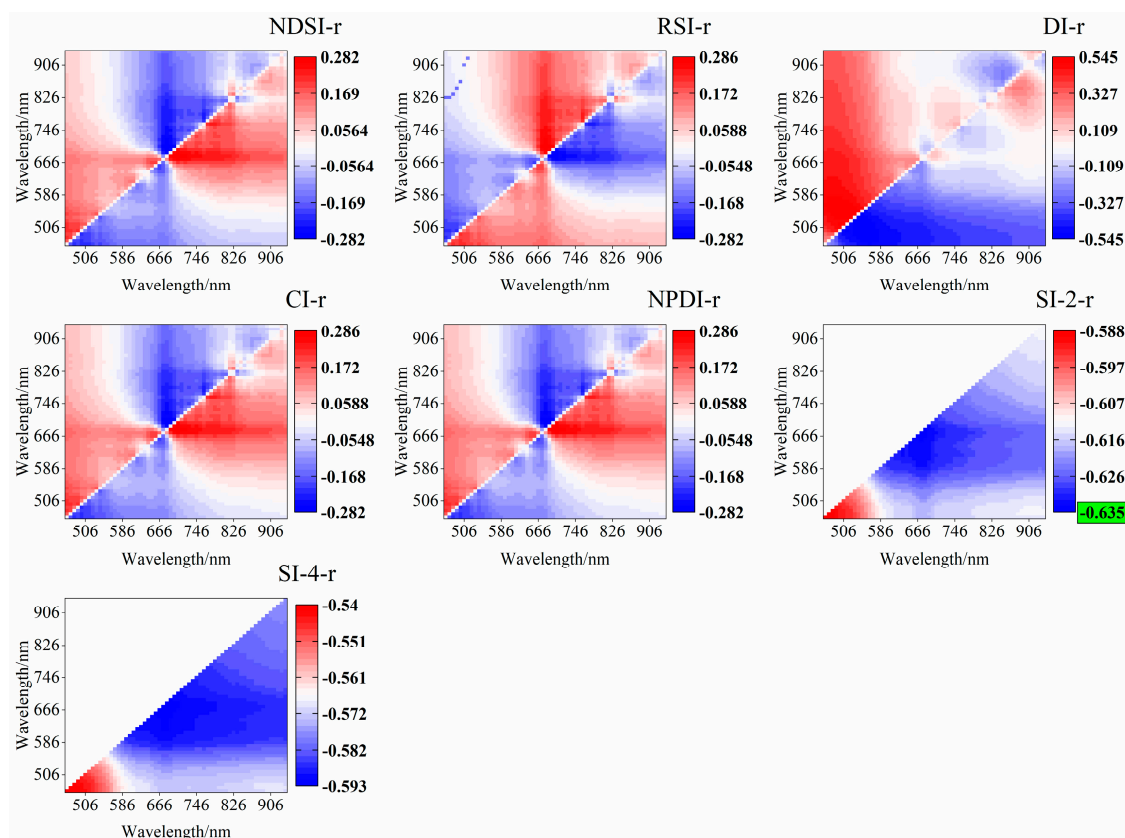


Figure 11. Two-dimensional  $r$  diagram of SMC and original spectrum of 400–2400 nm.



Figure 12 shows the two-dimensional PCC analysis results of seven optimized spectral indices and SMC calculated on the basis of the original spectrum of 466–938 nm. The extreme value of the  $r$  in the figure appeared in the SI-2 calculation process, and the extreme value was  $-0.635$ , which was the same as the extreme value in Section 3.2.1. Here, the method of optimized spectral indices seemed to fail. Even some optimized spectral indices had a significant drop in  $r$ . This may have been due to the fact that the original soil spectrum curve in the 466–938 nm band was relatively flat and there was no obvious difference. This also verified the point of view expressed in Section 3.2.1 to a certain extent. In other words, it was difficult to retrieve SMC based on visible and near-infrared spectroscopy.



**Figure 12.** Two-dimensional  $r$  diagram of SMC and original spectrum of 466–938 nm.

The original spectral curve in the 466–938 nm band was relatively flat and there was no obvious difference, and thus it was difficult to improve the  $r$ . After preprocessing the spectral reflectance by mathematical transformation, we found that the  $r$  can be improved. This may be because preprocessing of mathematical transformations such as fractional differentiation can increase the difference of the spectrum, thereby increasing the extreme value of the  $r$ . It can be found from Figure 13 that because the 0.5-order absorbance spectrum highlighted the spectral difference to the greatest extent, it can better help in optimizing the spectral indices to find the sensitive band and improve the  $r$ . At this time, the maximum  $r$  value calculated by the DI was  $\pm 0.72$ , which reached a higher level of  $r$  in comparison with Figure 12.

### 3.2.3. Three-Dimensional PCC Analysis of SMC and Spectrum

On the basis of two-dimensional PCC analysis, this study further explored the analysis process of the three-dimensional  $r$  based on the three-band optimized spectral indices. Taking into account the length of the article, the PCC analysis results of all 12 three-band optimized spectral indices are not shown, but the first two three-band optimized spectral indices with the best application effect on the  $(i, j)$  plane were selected for display, i.e., SI-1 and TVI. It can be seen from Figure 14 that both of



them had  $r$  extremes at the  $n = 794$  nm level, being 0.709 and 0.755, respectively. In particular, the TVI was the best performing index in the three-dimensional PCC analysis. Its  $r$  extreme value exceeded the one-dimensional  $r$  extreme value of 400–2400 nm (Figure 8). To a large extent, it had broken through the limitation of visible near-infrared spectroscopy in the estimation of SMC due to the lack of shortwave-infrared spectrum, and its utility was significantly better than the two-band DI (Figure 13).

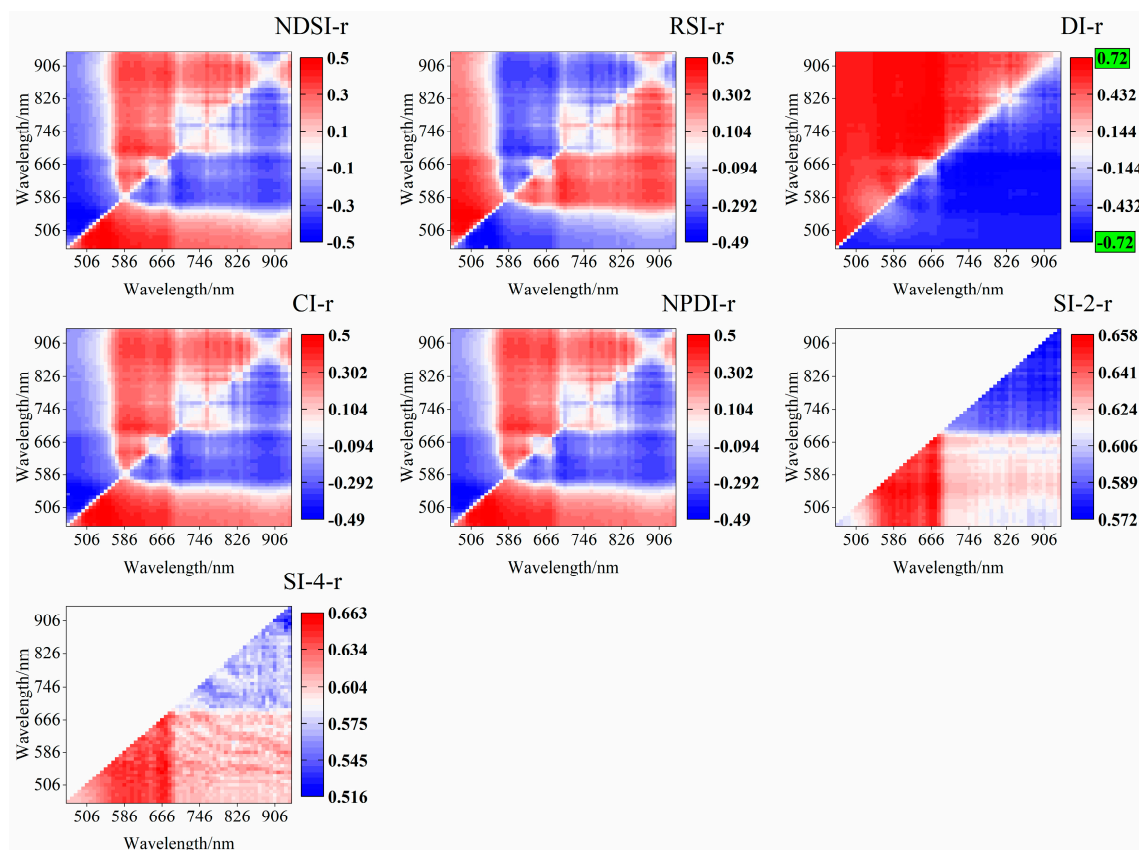


Figure 13. Two-dimensional  $r$  diagram between SMC and 0.5-order absorbance spectrum.

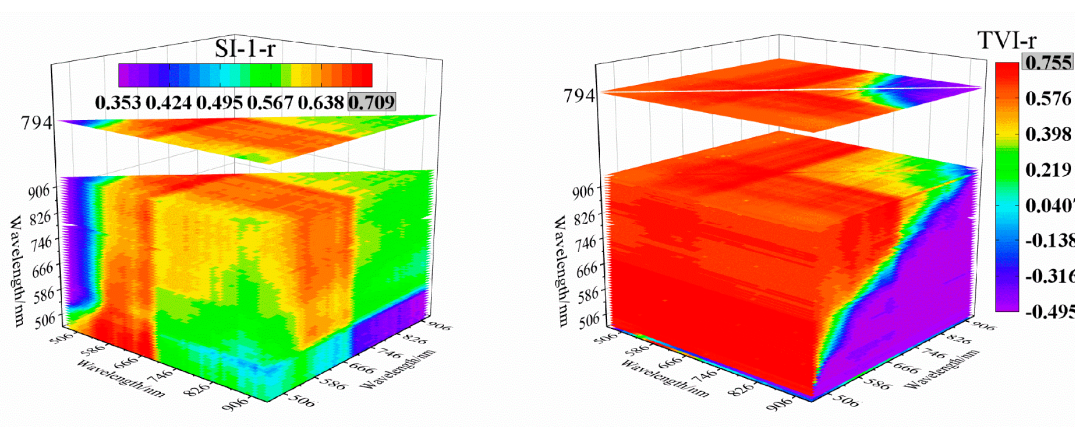


Figure 14. Three-dimensional  $r$  diagram of SMC and 0.5-order absorbance spectrum.

Since the TVI showed good utility on the  $(i, j)$  plane, this article further explored its effect on the  $(i, n)$  and  $(j, n)$  planes. It can be seen from Figure 15 that it had a very important effect on the  $r$ . The enhancement effect was consistent with the  $(i, j)$  plane—both were 0.755, but the plane where the extreme value appeared had changed, with  $j = 834$  nm and  $i = 666$  nm.

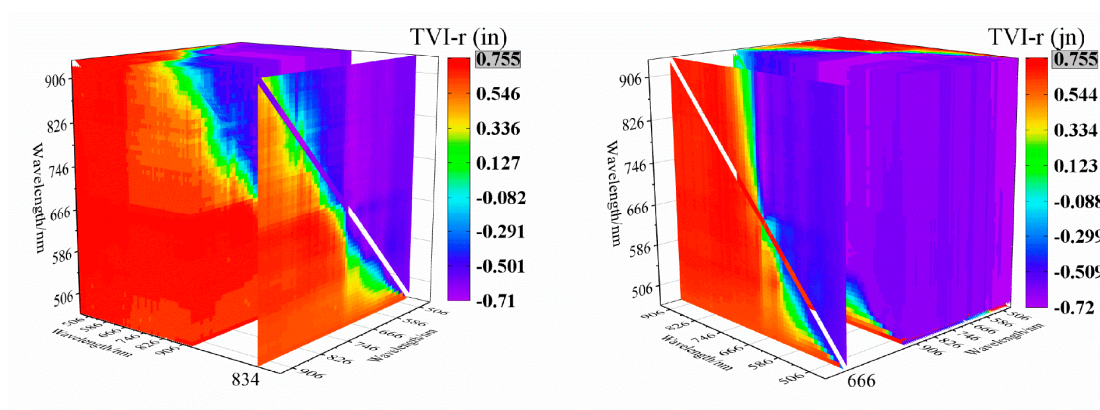


Figure 15. Three-dimensional  $r$  diagram of SMC and triangle vegetation index (TVI).

### 3.3. Model

The last chapter mainly conducted a more comprehensive analysis from the perspective of the extreme value of the  $r$ , and concluded that the TVI based on the 0.5-order absorbance spectrum was the best method. Therefore, we used the sensitive bands selected on the basis of this method to construct a ground-measured spectrum estimation model, used various model methods to establish the model, and finally selected the optimal model.

Figure 16 is a standardized Taylor diagram of the calibration set model. The evaluation indicators mainly included SD,  $RMSE_C$ ,  $r$ , and  $R^2_C$ . As can be seen from the graph, among all the eight models, the TGBM model had the best overall performance, with a  $R^2_C$  of 0.887, a  $RMSE_C$  of 2.488%, an SD of 6.733%, and a  $r$  of 0.942, while MLR, NR, RF, and CART had low accuracy. In addition, PLSR had the best comprehensive performance in conventional statistical models.

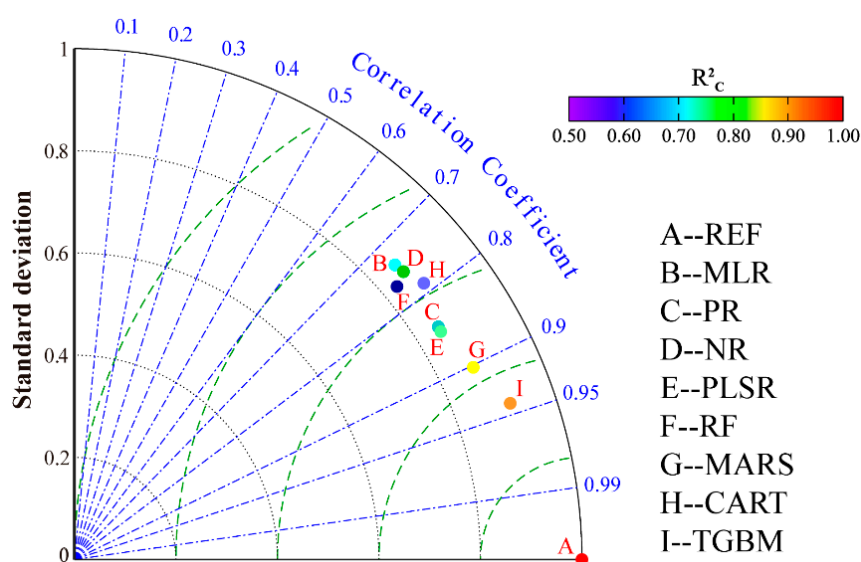


Figure 16. Taylor diagram of model accuracy evaluation of calibration set.

When constructing the model, the machine learning algorithm showed a strong fitting ability, but when verifying the model (Table 4), it was found that the PLSR model generated in actual application had the strongest predictive ability, and its  $R^2_V$  was 0.787,  $RMSE_V$  was 3.247%, and RPD was 2.071. The second was the MARS model with a  $R^2_V$  of 0.774,  $RMSE_V$  of 3.084%, and RPD of 1.972. Therefore, in this paper, the PLSR model was used as the inversion model of SMC. This was because of its high accuracy, and because its model expression was practical.

**Table 4.** Validation set model accuracy evaluation.

Model	$R^2_V$	RMSE <sub>V</sub> (%)	RPD
MLR	0.694	3.181	1.888
PR	0.684	3.870	1.495
NR	0.726	3.631	1.518
PLSR	0.787	3.247	2.071
RF	0.662	3.771	1.744
MARS	0.774	3.084	1.972
CART	0.544	4.922	1.245
TGBM	0.707	3.525	1.879

When applying the PLSR model, in order to minimize the number of model-independent variables and further enhance the ease of use of the model, this paper used the VIP method to filter model independent variables and improve model efficiency. The VIP method is actually a step of the PLSR model, which is encapsulated in the PLSR model. In continuing to build the PLSR model after excluding independent variables with VIP values less than 1, it is necessary to select the independent variables that have a positive effect on the accuracy of the model through the VIP method in order to achieve a suitable number of independent variables. In this process, the VIP method plays a role in the selection of model independent variables. As shown in Table 5 and Figure 17, 27 independent variables were selected from 54 independent variables with  $|r|$  greater than or equal to 0.750, and the PLSR model constructed subsequently had a certain improvement over the previous model.

**Table 5.** Accuracy evaluation of the variable importance in projection (VIP) process.

Number of Variables	$R^2_C$	RMSE <sub>C</sub> (%)	$R^2_V$	RMSE <sub>V</sub> (%)
54	0.727	3.061	0.796	2.545
27	0.733	3.028	0.805	3.100
9	0.738	3.002	0.761	2.627

Finally, the optimal PLSR model was selected. The model verification scatter diagram is shown in Figure 18. The model expression automatically output by the PLSR model was:

$$\begin{aligned} \text{SMC} = & 7.896 + 1.894 \times \text{B730/B674/B770} + 2.465 \times \text{B674/B730/B770} + 1.593 \times \text{B730/B666/B794} + 1.703 \\ & \times \text{B762/B666/B794} + 1.996 \times \text{B818/B666/B794} + 1.937 \times \text{B834/B658/B794} + 1.957 \times \text{B834/B666/B794} + 2.049 \\ & \times \text{B858/B666/B794} + 2.086 \times \text{B866/B666/B794} + 2.465 \times \text{B770/B674/B730} + 1.894 \times \text{B770/B730/B674} + 1.593 \\ & \times \text{B794/B730/B666} + 1.703 \times \text{B858B794/B762/B666} + 1.996 \times \text{B794/B818/B666} + 1.937 \times \text{B794/B834/B658} + \\ & 1.957 \times \text{B794/B834/B666} + 2.049 \times \text{B794/B858/B666} + 2.086 \times \text{B794/B866/B666} + 1.937 \times \text{B658/B794/B834} + \\ & 1.996 \times \text{B666/B794/B818} + 1.957 \times \text{B666/B794/B834} + 2.049 \times \text{B666/B794/B858} + 2.086 \times \text{B666/B794/B866} + \\ & 1.593 \times \text{B666/B794/B730} + 1.703 \times \text{B666/B794/B762} + 1.894 \times \text{B674/B770/B730} + 2.465 \times \text{B730/B770/B674}. \end{aligned}$$

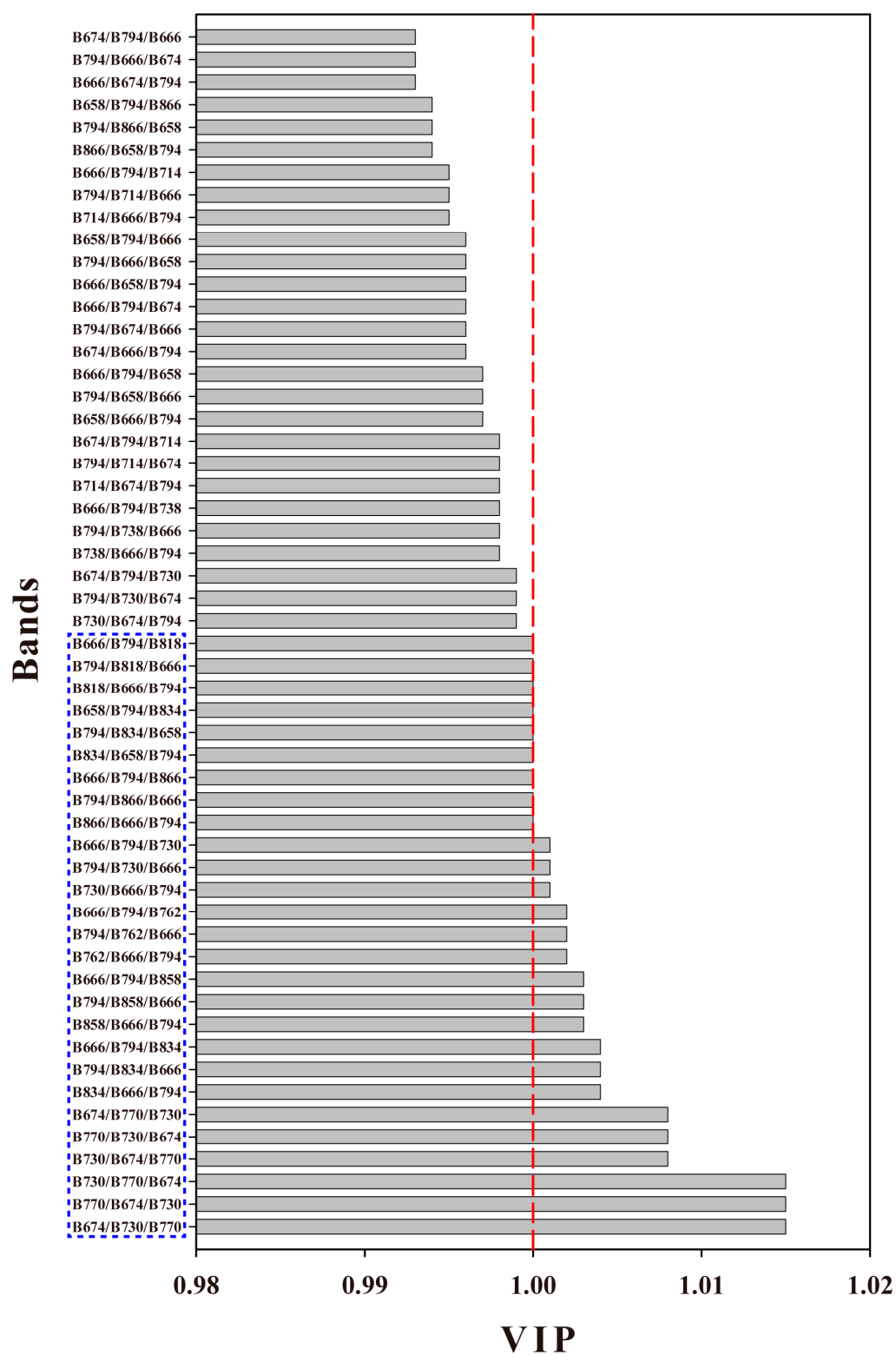
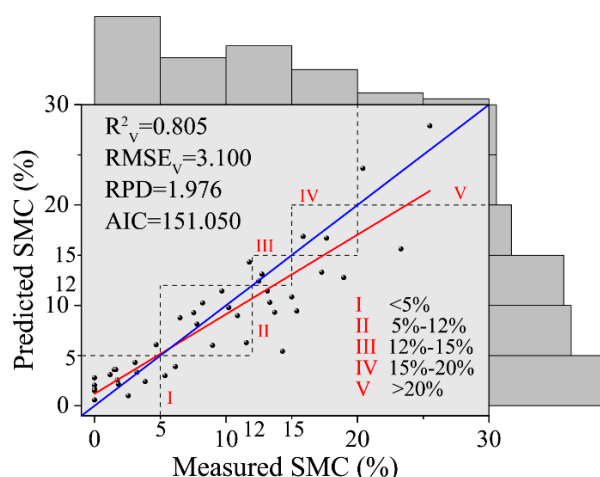


Figure 17. VIP screening diagram of model independent variables.



**Figure 18.** Scatter plot of optimal partial least squares regression (PLSR) model verification.

## 4. Discussion

### 4.1. Discussion on Optimized Spectral Indices

Remote sensing technology, especially hyperspectral remote sensing technology, mainly uses various methods and indicators based on spectral characteristics to achieve monitoring purposes. On the basis of this idea, many spectral indices have been developed to estimate soil water and salt content, which can extract useful information from complex spectral reflectance to eliminate the effects of soil heterogeneity, vegetation coverage, and weather conditions [54]. However, a spectral index developed on the basis of specific data in a specific area has obvious limitations, poor portability, and often cannot achieve good inversion results in other areas. Therefore, researchers have optimized the specific spectral index in order to improve the accuracy of quantitative prediction models.

The rapid development of computer software and hardware technology provides convenience for the realization of band optimization algorithms, and the rich band information of hyperspectral provides more possible combinations for band optimization algorithms, which can also effectively find more sensitive bands by increasing the band dimension. The combination of bands is beneficial to optimize the spectral index and further reduce the influence of heterogeneity, which has been verified in the research of other scholars [21]. However, this blind search method also has certain problems, that is, the spectral index defined by it has no clear theoretical significance compared with the spectral index defined by theoretical basis in traditional research. However, like many fields of computer science (especially the field of artificial intelligence), with the continuous development of informatization, this black box approach has gradually become a trend that can effectively discover the inherent patterns of data and further promote the development of the subject. Perhaps, just like the method used in this study, that is, to select the appropriate defined spectral index for optimization, this problem can be solved to a certain extent, and the optimization of the spectral index can be given a certain theoretical significance.

### 4.2. Discussion on Fractional Order Differential

Fractional order differentiation extends the order to non-integer order, generalizes the definition of differentiation, expands the order of operations, and organically unifies integer order and non-integer order differentiation in the definition, which shows the existence of fractional order differentiation is reasonable. In the numerical calculation process, the fractional order is not only related to the value of the point, but also related to the value of all points before the point. For fractional differentiation, according to the formula, the closer the point, the greater the weight given in the calculation, the greater the impact, the farther the point is, and the smaller the weight, and the impact will also decrease.



This is the well-known “memory” and “non-locality” of the fractional differential, and it is also the biggest difference between the fractional differential and the integer differential [55].

In this study, we may ask why the best result was achieved at the 0.5 level? Perhaps it is related to the slope and curvature of the curve. In the field of spectral analysis, some studies have pointed out that although the fractional differential has no clear physical meaning, the first differential of the spectrum curve is defined as the slope of the spectrum, and the second differential is defined as the curvature of the spectrum, corresponding to the inflection point and the extreme value. The fractional order can be considered as the sensitivity to the slope and curvature of the spectral line, that is, when the differential order increases from 0 to 1, the sensitivity of the differential result to the spectral reflectance decreases, and the sensitivity to the slope of the spectral line increases. When the order is increased from 1 to 2, the sensitivity of the differential result to the slope of the spectral line decreases, while the sensitivity to the curvature of the spectral line increases. However, since the shape of the soil spectral curve in this study area was relatively flat and the slope was dominant, the  $r$  extreme value appeared in the 0–1 order differential.

In the past, in the research of fractional order differentiation in spectroscopy and remote sensing, the step length was defaulted to 1. However, this research broke through the limitation of step length and contributed to the further expansion of fractional order differentiation in the field of remote sensing. As we all know, the satellite band interval is unlikely to be equidistant and is always 1 nm, which puts a large amount of constraints on the application of fractional differentiation. Even if the band is cut down, it will bring great uncertainty. It may bring a greater deviation from the actual situation. In this case, perhaps the method of equidistant satellite bands through linear interpolation in this article will be simpler and more practical.

#### 4.3. Discussion on Spectral Scale

In the linear interpolation of satellite bands, there will be a problem in ensuring that the interval step, that is, the scale, is appropriate, and the scale should not be too large or too small [56]. In this study, the interval between the 32 bands of the Zhuhai-1 hyperspectral satellite was 14–16 nm. Therefore, the 4, 8, and 16 nm scales were used to perform linear interpolation experiments and comparisons of the satellite spectrum curves. Finally, the 8 nm spectrum was found. This scale can better describe the shape of the original spectrum curve, while the spectrum curve shape information lost in the 4 nm and 16 nm spectrum scales was more and the error was larger. Of course, other methods of interpolation, such as non-linear interpolation, may be better, but this article did not perform a more in-depth exploration, with further discussion and verification needing to be done in future research.

#### 4.4. Discussion on Models

In this study, the PLSR model had higher prediction accuracy. In addition to the high practical application ability of the PLSR model itself, this fact may have been due to the limited number of dependent variables in this study, which limited the machine learning model to a certain extent. Because the general application scenarios of machine learning models have more dependent variables and fewer independent variables, only in this case can the advantages of machine learning models be fully demonstrated. This is being continuously verified in many related studies.

Near-infrared spectroscopy is generated by molecular vibration, with weak absorption and severe overlap of absorption peaks [57]. There must be redundant variables and invalid information variables that are not related to SMC in soil hyperspectral variables, which will increase the complexity of model operation and calculation. It even reduces the prediction accuracy of the model. Therefore, variable optimization becomes a key link in the modeling process of spectral analysis [22]. In this study, the VIP method was used to screen the best variables. Because the VIP method is embedded in the PLSR model, it can better combine the model to determine the variables.

#### 4.5. Discussion on Zhuhai-1 Hyperspectral Imagery

Microwave remote sensing, especially passive microwave remote sensing, is generally used for large-area and large-scale research because of its low spatial resolution and large coverage. Compared with microwave remote sensing, hyperspectral remote sensing has higher spatial resolution and image coverage. The area is small, suitable for small- and medium-scale research, while hyperspectral remote sensing sensors are affected by the mutual limitation of spectra and spatial resolution. Generally speaking, the spatial resolution is not particularly high. For example, the spatial resolution of Hyperion, GF-5, and HJ-1 HIS hyperspectral images are 30, 30, and 100 m, respectively. Compared with other hyperspectral satellites, the Zhuhai-1 hyperspectral satellite has the characteristics of higher spatial resolution and larger amplitude. However, due to the lack of short-wave infrared bands, and the wavelength of 400–1000 nm can only be selected under 32 bands. It is said that in many application scenarios, there will be more or less limitations, but method innovation can break through this limitation to a large extent, which has been fully verified in this research. In terms of comprehensive evaluation, Zhuhai-1 image data are of good quality, which makes up for the lack of hyperspectral image data to a certain extent [46]. With this higher spectral, spatial, and temporal resolution, we believed that more applications will be obtained in the field of quantitative remote sensing, especially in vegetation remote sensing.

### 5. Conclusions

This paper simulated the band settings of the Zhuhai-1 satellite by ground spectra, and analyzed the  $r$  between the fractional differential of different orders and the optimized spectral indices of different dimensions and the measured SMC, establishing a SMC estimation model. The degree of dispersion of soil moisture in the study area was strong. The 8 nm spectral scale can accurately reflect the hyperspectral characteristics of the soil. The hyperspectral characteristics of the soil with different SMC conformed to the general law, and the fractional differential calculation results were correct. The one-dimensional  $r$  between SMC and the 466–938 nm band < two-dimensional  $r$  < three-dimensional  $r$ , and the  $r$  of TVI calculated by the 0.5-order absorbance spectrum exceeded the one-dimensional  $r$  extreme value of 400–2400 nm.  $R^2_C$  of the optimal PLSR model was 0.733,  $RMSE_C$  was 3.028%,  $R^2_V$  was 0.805,  $RMSE_V$  was 3.100%, RPD was 1.976, and AIC was 151.050. The results showed that the fractional order differential three-band optimized spectral indices can to a certain extent break through the limitations of visible near-infrared spectroscopy in the estimation of SMC due to the lack of short-wave infrared spectra. We proved that it is possible to retrieve regional SMC using Zhuhai-1 hyperspectral imagery, which provides a new research idea for improving the accuracy of hyperspectral satellite remote sensing inversion.

**Author Contributions:** Conceptualization, Y.K. and N.T.; methodology, Y.K. and S.L.; software, Y.K. and S.L.; validation, N.T. and Q.S.; formal analysis, Y.K.; investigation, Y.K.; resources, Y.K. and N.T.; data curation, Y.K.; writing—original draft preparation, Y.K.; writing—review and editing, N.T., Q.S. and S.L.; visualization, Y.K.; supervision, N.T.; project administration, Q.S.; funding acquisition, N.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was carried out with the financial support provided by the National Natural Science Foundation of China “Study on the optimal scale for soil salinization water–salt remote sensing monitoring” (grant No.41761077).

**Acknowledgments:** The authors would like to thank those who assisted to collect and process field data. Furthermore, sincere thanks are also extended to the teachers and students of the Ministry of Education Key Laboratory of Oasis Ecology, Xinjiang University for their support and help. Finally, the authors wish to thank the referees for providing helpful suggestions for the improvement of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ebtehaj, A.; Bras, R.L. A physically constrained inversion for high-resolution passive microwave retrieval of soil moisture and vegetation water content in L-band. *Remote Sens. Environ.* **2019**, *233*, 111346. [[CrossRef](#)]
2. Bablet, A.; Viallefont, F.; Jacquemoud, S.; Fabre, S.; Briottet, X. High-resolution mapping of in-depth soil moisture content through a laboratory experiment coupling a spectroradiometer and two hyperspectral cameras. *Remote Sens. Environ.* **2020**, *236*, 111533. [[CrossRef](#)]
3. Aumann, H.H.; Chahine, M.T.; Gautier, C.; Goldberg, M.D.; Kalnay, E.; McMillin, L.M.; Revercomb, H.; Rosenkranz, P.W.; Smith, W.L.; Staelin, D.H.; et al. AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 253–264. [[CrossRef](#)]
4. Chen, C. Science mapping: A systematic review of the literature. *J. Data Inf. Sci.* **2017**, *2*, 1–40. [[CrossRef](#)]
5. Chen, C.; Hu, Z.; Liu, S.; Tseng, H. Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. *Expert Opin. Biol. Ther.* **2012**, *12*, 593–608. [[CrossRef](#)]
6. Lu, P.; Wang, L.; Niu, Z.; Li, L.; Zhang, W. Prediction of soil properties using laboratory VIS–NIR spectroscopy and Hyperion imagery. *J. Geochem. Explor.* **2013**, *132*, 26–33. [[CrossRef](#)]
7. Tiwari, S.K.; Saha, S.K.; Kumar, S. Prediction modeling and mapping of soil carbon content using artificial neural network, hyperspectral satellite data and field spectroscopy. *Adv. Remote Sens.* **2015**, *4*, 63–72. [[CrossRef](#)]
8. Gomez, C.; Drost, A.; Roger, J. Analysis of the uncertainties affecting predictions of clay contents from VNIR/SWIR hyperspectral data. *Remote Sens. Environ.* **2015**, *156*, 58–70. [[CrossRef](#)]
9. Gomez, C.; Gholizadeh, A.; Borůvka, L.; Lagacherie, P. Using legacy data for correction of soil surface clay content predicted from VNIR/SWIR hyperspectral airborne images. *Geoderma* **2016**, *276*, 84–92. [[CrossRef](#)]
10. Gomez, C.; Oltra-Carrió, R.; Bacha, S.; Lagacherie, P.; Briottet, X. Evaluating the sensitivity of clay content prediction to atmospheric effects and degradation of image spatial resolution using Hyperspectral VNIR/SWIR imagery. *Remote Sens. Environ.* **2015**, *164*, 1–15. [[CrossRef](#)]
11. Vaudour, E.; Gilliot, J.-M.; Bel, L.; Lefevre, J.; Chehdi, K. Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 24–38. [[CrossRef](#)]
12. Ouerghemmi, W.; Gomez, C.; Naceur, S.; Lagacherie, P. Semi-blind source separation for the estimation of the clay content over semi-vegetated areas using VNIR/SWIR hyperspectral airborne data. *Remote Sens. Environ.* **2016**, *181*, 251–263. [[CrossRef](#)]
13. Peón, J.; Recondo, C.; Fernández, S.; Calleja, J.J.F.; De Miguel, E.; Carretero, L. Prediction of Topsoil Organic Carbon Using Airborne and Satellite Hyperspectral Imagery. *Remote Sens.* **2017**, *9*, 1211. [[CrossRef](#)]
14. Liu, L.; Ji, M.; Buchroithner, M.F. Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. *Sensors* **2018**, *18*, 3169. [[CrossRef](#)] [[PubMed](#)]
15. Ge, X.; Wang, J.; Ding, J.-L.; Cao, X.; Zhang, Z.; Liu, J.; Li, X. Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring. *PeerJ* **2019**, *7*, e6926. [[CrossRef](#)] [[PubMed](#)]
16. Li, F.; Mistele, B.; Hu, Y.; Chen, X.; Schmidhalter, U. Optimising three-band spectral indices to assess aerial N concentration, N uptake and aboveground biomass of winter wheat remotely in China and Germany. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 112–123. [[CrossRef](#)]
17. Wang, J.; Ding, J.; Yu, D.; Ma, X.; Zhang, Z.; Ge, X.; Teng, D.; Li, X.; Liang, J.; Lizaga, I.; et al. Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China. *Geoderma* **2019**, *353*, 172–187. [[CrossRef](#)]
18. Ainiwaer, M.; Ding, J.; Kasim, N.; Wang, J.; Wang, J. Regional scale soil moisture content estimation based on multi-source remote sensing parameters. *Int. J. Remote Sens.* **2019**, *41*, 3346–3367. [[CrossRef](#)]
19. Liu, C.; Zhang, F.; Ge, X.; Zhang, X.; Chan, N.W.; Qi, Y. Measurement of total nitrogen concentration in surface water using hyperspectral band observation method. *Water* **2020**, *12*, 1842. [[CrossRef](#)]
20. Kahaer, Y.; Shengtian, Y.; Tashpolat, N.; Fei, Z. Hyperspectral estimation of soil electrical conductivity based on fractional order differentially optimized spectral indices. *Acta Ecol. Sin.* **2019**, *39*, 7237–7248. [[CrossRef](#)]
21. Zhang, Z.; Ding, J.-L.; Wang, J.; Ge, X. Prediction of soil organic matter in northwestern China using fractional order derivative spectroscopy and modified normalized difference indices. *Catena* **2020**, 185. [[CrossRef](#)]

22. Hong, Y.; Chen, S.; Zhang, Y.; Chen, Y.; Yu, L.; Liu, Y.; Liu, Y.; Cheng, H.; Liu, Y. Rapid identification of soil organic matter level via visible and near-infrared spectroscopy: Effects of two-dimensional correlation coefficient and extreme learning machine. *Sci. Total. Environ.* **2018**, *644*, 1232–1243. [[CrossRef](#)] [[PubMed](#)]
23. Zhu, C.; Zhang, Z.; Wang, H.; Wang, J.; Yang, S. Assessing soil organic matter content in a coal mining area through spectral variables of different numbers of dimensions. *Sensors* **2020**, *20*, 1795. [[CrossRef](#)] [[PubMed](#)]
24. Sawut, R.; Kasim, N.; Abliz, A.; Hu, L.; Yalkun, A.; Maihemuti, B.; Qingdong, S. Possibility of optimized indices for the assessment of heavy metal contents in soil around an open pit coal mine area. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 14–25. [[CrossRef](#)]
25. Hong, Y.; Shen, R.; Cheng, H.; Chen, Y.; Zhang, Y.; Liu, Y.; Zhou, M.; Yu, L.; Liu, Y.; Liu, Y. Estimating lead and zinc concentrations in peri-urban agricultural soils through reflectance spectroscopy: Effects of fractional-order derivative and random forest. *Sci. Total Environ.* **2019**, *651*, 1969–1982. [[CrossRef](#)]
26. Wang, X.; Zhang, F.; Ding, J.; Kung, H.-T.; Latif, A.; Johnson, V.C. Estimation of soil salt content (SSC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR), Northwest China, based on a Bootstrap-BP neural network model and optimal spectral indices. *Sci. Total Environ.* **2018**, *615*, 918–930. [[CrossRef](#)]
27. Kahaer, Y.; Tashpolat, N. Estimating salt concentrations based on optimized spectral indices in soils with regional heterogeneity. *J. Spectrosc.* **2019**. [[CrossRef](#)]
28. Li, C.; Zhang, G.-W.; Zhou, Z.; Zhao, W.-Q.; Meng, Y.-L.; Chen, B.; Wang, Y. Hyperspectral parameters and prediction model of soil moisture in coastal saline. *Chin. J. Appl. Ecol.* **2016**, *27*, 525–531. [[CrossRef](#)]
29. Wang, X.; Zhang, F.; Kung, H.-T.; Johnson, V.C. New methods for improving the remote sensing estimation of soil organic matter content (SOMC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR) in northwest China. *Remote Sens. Environ.* **2018**, *218*, 104–118. [[CrossRef](#)]
30. Jingzhe, W.; Jianli, D.; Xuankai, M.; Xiangyu, G.; Bohua, L.; Jing, L. Detection of soil moisture content based on UAV-derived hyperspectral imagery and spectral index in oasis cropland. *Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 164–172. [[CrossRef](#)]
31. Xianlong, Z.; Fei, Z.; Haiwei, Z.; Zhe, L.; Qing, H.; Lihua, C. Optimization of soil salt inversion model based on spectral transformation from hyperspectral index. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 110–117. [[CrossRef](#)]
32. Kasim, N.; Maihemuti, B.; Sawut, R.; Abliz, A.; Cui, D.; Abdumutallip, M. Quantitative estimation of soil salinization in an arid region of the Keriya Oasis based on multidimensional modeling. *Water* **2020**, *12*, 880. [[CrossRef](#)]
33. Xu, C.; Zeng, W.; Huang, J.; Wu, J.; van Leeuwen, W.J.D. Prediction of soil moisture content and soil salt concentration from hyperspectral laboratory and field data. *Remote Sens.* **2016**, *8*, 42. [[CrossRef](#)]
34. Oltra-Carrió, R.; Baup, F.; Fabre, S.; Fieuzal, R.; Briottet, X. Improvement of soil moisture retrieval from hyperspectral VNIR-SWIR data using clay content information: From laboratory to field experiments. *Remote Sens.* **2015**, *7*, 3184–3205. [[CrossRef](#)]
35. Yue, J.; Tian, J.; Tian, Q.; Xu, K.; Xu, N. Development of soil moisture indices from differences in water absorption between shortwave-infrared bands. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 216–230. [[CrossRef](#)]
36. Zhang, Y.; Tan, K.; Wang, X.; Chen, Y. Retrieval of soil moisture content based on a modified Hapke photometric model: A novel method applied to laboratory hyperspectral and Sentinel-2 MSI data. *Remote Sens.* **2020**, *12*, 2239. [[CrossRef](#)]
37. Haubrock, S.-N.; Kuhnert, M.; Hostert, P.; Kaufmann, H.; Chabrillat, S. Surface soil moisture quantification and validation based on hyperspectral data and field measurements. *J. Appl. Remote Sens.* **2008**, *2*, 023552. [[CrossRef](#)]
38. Schlerf, M.; Atzberger, C. Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data. *Sens. Environ.* **2006**, *100*, 281–294. [[CrossRef](#)]
39. Yuan, H.; Yang, G.; Li, C.; Wang, Y.; Liu, J.; Yu, H.; Feng, H.; Xu, B.; Zhao, X.; Yang, X. Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. *Remote Sens.* **2017**, *9*, 309. [[CrossRef](#)]
40. Lagacherie, P.; Arrouays, D.; Bourennane, H.; Gomez, C.; Martin, M.; Saby, N.P. How far can the uncertainty on a Digital Soil Map be known? A numerical experiment using pseudo values of clay content obtained from Vis-SWIR hyperspectral imagery. *Geoderma* **2019**, *337*, 1320–1328. [[CrossRef](#)]
41. Ainiwaer, M.; Ding, J.-L.; Wang, J.; Nasierding, N. Spatiotemporal dynamics of water table depth associated with changing agricultural land use in an arid zone oasis. *Water* **2019**, *11*, 673. [[CrossRef](#)]
42. Huang, S.; Ding, J.-L.; Liu, B.; Ge, X.; Wang, J.; Zou, J.; Zhang, J. The Capability of integrating optical and microwave data for detecting soil moisture in an oasis region. *Remote Sens.* **2020**, *12*, 1358. [[CrossRef](#)]

43. Huang, S.; Ding, J.-L.; Zou, J.; Liu, B.; Zhang, J.; Chen, W. Soil Moisture Retrieval Based on Sentinel-1 Imagery under Sparse Vegetation Coverage. *Sensors* **2019**, *19*, 589. [[CrossRef](#)] [[PubMed](#)]
44. Tashpolat, N.; Ding, J.-L.; Yu, D. Dielectric properties of saline soil based on a modified Dobson dielectric model. *J. Arid. Land* **2015**, *7*, 696–705. [[CrossRef](#)]
45. Wang, F.; Shi, Z.; Biswas, A.; Yang, S.; Ding, J. Multi-algorithm comparison for predicting soil salinity. *Geoderma* **2020**, *365*, 114211. [[CrossRef](#)]
46. Jiang, Y.; Wang, J.; Zhang, L.; Zhang, G.; Li, X.; Wu, J. Geometric processing and accuracy verification of Zhuhai-1 hyperspectral satellites. *Remote Sens.* **2019**, *11*, 996. [[CrossRef](#)]
47. Huadong, G.; Guang, L.; Dong, L.; Lu, Z.; Han, X. Progress of Earth observation in China. *Chin. J. Space Sci.* **2018**, *38*, 797–809. [[CrossRef](#)]
48. Yépez-Martínez, H.; Gómez-Aguilar, J. A new modified definition of Caputo-Fabrizio fractional-order derivative and their applications to the Multi Step Homotopy Analysis Method. *J. Comput. Appl. Math.* **2018**. [[CrossRef](#)]
49. Hong, Y.; Chen, S.; Liu, Y.; Zhang, Y.; Yu, L.; Chen, Y.; Liu, Y.; Cheng, H.; Liu, Y. Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy. *Catena* **2019**, *174*, 104–116. [[CrossRef](#)]
50. Hong, Y.; Liu, Y.; Chen, Y.; Liu, Y.; Yu, L.; Liu, Y.; Cheng, H. Application of fractional-order derivative in the quantitative estimation of soil organic matter content through visible and near-infrared spectroscopy. *Geoderma* **2019**, *337*, 758–769. [[CrossRef](#)]
51. Chen, X.; Huang, J.; Yi, M. Cost estimation for general aviation aircrafts using regression models and variable importance in projection analysis. *J. Clean. Prod.* **2020**, *256*, 120648. [[CrossRef](#)]
52. Mullet, T.C.; Gage, S.H.; Morton, J.M.; Huettmann, F. Temporal and spatial variation of a winter soundscape in south-central Alaska. *Landsc. Ecol.* **2016**, *31*, 1117–1137. [[CrossRef](#)]
53. Suon, S.; Li, Y.; Porn, L.; Javed, T. Spatiotemporal Analysis of Soil Moisture Drought over China during 2008–2016. *J. Water Resour. Prot.* **2019**, *11*, 700–712. [[CrossRef](#)]
54. Pôças, I.; Calera, A.; Campos, I.; Cunha, M. Remote sensing for estimating and mapping single and basal crop coefficients: A review on spectral vegetation indices approaches. *Agric. Water Manag.* **2020**, *233*, 106081. [[CrossRef](#)]
55. Wang, J.; Ding, J.; Abulimiti, A.; Cai, L. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS-NIR) spectroscopy, Ebinur Lake Wetland, Northwest China. *PeerJ* **2018**, *6*, e4703. [[CrossRef](#)] [[PubMed](#)]
56. Du, H.; Jiang, H.; Zhang, L.; Mao, D.; Wang, Z. Evaluation of spectral scale effects in estimation of vegetation leaf area index using spectral indices methods. *Chin. Geogr. Sci.* **2016**, *26*, 731–744. [[CrossRef](#)]
57. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).