

Article

Estimation of Soil Salt and Ion Contents Based on Hyperspectral Remote Sensing Data: A Case Study of Baidunzi Basin, China

Libing Wang¹ , Bo Zhang^{1,*}, Qian Shen², Yue Yao², Shengyin Zhang³, Huaidong Wei^{1,4}, Rongpeng Yao¹ and Yaowen Zhang¹

- ¹ College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, China; 2019212381@nwnu.edu.cn (L.W.); weihdgs@126.com (H.W.); yrpskyscraper@163.com (R.Y.); 19909311218@163.com (Y.Z.)
 - ² Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100049, China; shenqian@radi.ac.cn (Q.S.); yaoyue@aircas.ac.cn (Y.Y.)
 - ³ Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; zhangseepage@126.com
 - ⁴ State Key Laboratory Breeding Base of Desertification and Aeolian Sand Disaster Combating, Gansu Desert Control Desert Research Institute, Lanzhou 730070, China
- * Correspondence: zhangbo@nwnu.edu.cn



Citation: Wang, L.; Zhang, B.; Shen, Q.; Yao, Y.; Zhang, S.; Wei, H.; Yao, R.; Zhang, Y. Estimation of Soil Salt and Ion Contents Based on Hyperspectral Remote Sensing Data: A Case Study of Baidunzi Basin, China. *Water* **2021**, *13*, 559. <https://doi.org/10.3390/w13040559>

Academic Editors: Nicolo Colomabini and Thomas Meixner

Received: 23 December 2020

Accepted: 18 February 2021

Published: 22 February 2021

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



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Soil salinity due to irrigation diversion affects regional agriculture, and the development of soil composition estimation models for the dynamic monitoring of regional salinity is important for salinity control. In this study, we evaluated the performance of hyperspectral data measured using an analytical spectral device (ASD) field spec standard-res hand-held spectrometer and satellite sensor visible shortwave infrared advanced hyperspectral imager (AHSI) in estimating the soil salt content (SSC). First derivative analysis (FDA) and principal component analysis (PCA) were applied to the data using the raw spectra (RS) to select the best model input data. We tested the ability of these three groups of data as input data for partial least squares regression (PLSR), principal component regression (PCR), and multiple linear regression (MLR). Finally, an estimation model of the SSC, Na⁺, Cl⁻, and SO₄²⁻ contents was established using the best input data and modeling method, and a spatial distribution map of the soil composition content was drawn. The results show that the soil spectra obtained from the satellite hyperspectral data (AHSI) and laboratory spectral data (ASD) were consistent when the SSC was low, and as the SSC increased, the spectral curves of the ASD data showed little change in the curve characteristics, while the AHSI data showed more pronounced features, and this change was manifested in the AHSI images as darker pixels with a lower SSC and brighter pixels with a higher SSC. The AHSI data demonstrated a strong response to the change in SSC; therefore, the AHSI data had a greater advantage compared with the ASD data in estimating the soil salt content. In the modeling process, RS performed the best in estimating the SSC and Na⁺ content, with the R² reaching 0.79 and 0.58, respectively, and obtaining low root mean squared error (RMSE) values. FDA and PCA performed the best in estimating Cl⁻ and SO₄²⁻, while MLR outperformed PLSR and PCR in estimating the content of the soil components in the region. In addition, the hyperspectral camera data used in this study were very cost-effective and can potentially be used for the evaluation of soil salinization with a wide range and high accuracy, thus reducing the errors associated with the collection of individual samples using hand-held hyperspectral instruments.

Keywords: soil salinization; remote sensing; numerical modelling; digital soil mapping; arid regions

1. Introduction

Soil salinization is a global problem and a long-term threat to agricultural production, especially in arid and semi-arid areas. The risk of soil salinization will become more prominent in the future [1–4]. Agriculture in irrigated areas is of great significance to regional development, which relieves people's survival pressure in arid and semi-arid

areas. With the growing population, people develop more land to meet their development needs, and this is accompanied by soil compaction, fertility declines, acid–base imbalances, soil degradation, and other consequences caused by soil salinization [5]. With the increase in the salt content in soil, plant growth becomes more difficult [6,7]. From the perspective of soil salinization control, the amount of salt in soil is directly proportional to the difficulty of soil recovery. Therefore, the dynamic monitoring of soil salt content-related indicators and a timely quantitative grasp of the soil salt content are of great significance for the rational development and utilization of land resources and for maintaining the ecological sustainable development of irrigated areas.

Remote sensing as an effective technology to monitor regional soil salinization has been recognized in many studies [8,9]. Using a remote sensing method to monitor the regional soil salinization not only promotes the management of saline land but also helps to simulate the process of soil salinization, as well as provide data support for revealing the mechanism of soil salinization. The core of the quantitative evaluation of soil salinization by remote sensing data is to find the relationship between the content of related indicators and remote sensing data, and thus the purpose of using remote sensing technology to study soil salinization is to establish a reliable estimation model.

For the remote sensing data source, hyperspectral data have more spectral bands to choose from compared with multispectral data. Therefore, to establish a more reliable soil component content model, hyperspectral data were selected as the data source in this study. Generally speaking, hyperspectral remote sensing data can be divided into laboratory-measured spectral data, field-measured spectral data, and hyperspectral data obtained by an imaging spectrometer [10]. Both laboratory spectral measurements and field spectral measurements belong to the method of near hyperspectral remote sensing data acquisition. For example, Xiaoguang Zhang used the spectral data measured in the laboratory as the basic data to evaluate the ability of PLSR and PCR in estimating the soil salt content [11]. R.L. Dehaan used the HyMap hyperspectral imaging spectrometer to evaluate soil salinization in the Murray Darling Basin, Australia. However, in the process of collecting soil samples, the original soil moisture, soil particle size, and other properties of the samples change; therefore, the feasibility of using the spectrum data measured in the laboratory as the basic data for estimating soil salt content must be analyzed.

Mapping the distribution of regional salinization is the primary work of salinization control, which is of great significance for the rational development and utilization of land resources and the maintenance of ecological sustainable development in irrigation areas [12]. Scholars have done a great deal of research on this hot spot [13–15]. Compared with the spatial interpolation method [16] and the method of using a portable spectrometer to measure the spectral data, using the data measured by a satellite-borne imaging spectrometer to estimate soil salinization-related indicators [17] has the advantages of a small workload, strong timeliness, and a large spatial range. At present, most soil salinization mapping studies have been based on multispectral satellite image data [8,18,19], drone remote sensing data, and proximal hyperspectral remote sensing data [9,14,15]. The multispectral satellite data have fewer wavebands, which presents certain limitations in reflecting soil information. A hyperspectral imaging spectrometer has many bands, which can be used to reflect the fine characteristics of soil, and this has the characteristics of spectral integration. It is common to use a hyperspectral imaging spectrometer to study soil salinization using drones [20] or satellites [13] as the platform of the sensors. However, to use a drone as a platform, it is often difficult to carry out large-scale soil monitoring due to the limitations of the image width. A satellite hyperspectral imaging spectrometer typically has a large width and a certain revisit period, which provides a possibility for the dynamic evaluation of soil salinization across a large range. However, there are few cases using satellite hyperspectral imaging spectrometer data [21] because the available hyperspectral satellite data sources are scarce.

It is a common method to use the soil salt content (SSC) measured in the laboratory as an index to evaluate the degree of soil salinization [14]. Since most salts, especially

sodium chloride, strongly reflect solar energy, and this information will be recorded by satellites [15], the reflectance data of satellite images are related to the soil composition. Studies have also investigated the relationship between the ionic content of saline soils and spectral data [22]. The saline soil ion content is important for the study of the formation and evolution of salinity and the management of salinization. As soil ions are a more microcosmic concept compared with soil salinity, it is difficult to estimate the soil ion content in the study of soil salinization. At present, it is challenging to use spectral data to estimate the ion content of saline soil.

In view of the above problems, the research question of this paper is to determine whether the satellite hyperspectral data have more advantages in estimating the content of soil components than the spectral data obtained in the laboratory. What is the performance of common spectral data transformations and modeling methods for soil component content estimation in the study area? In this context, the objectives of this paper are to (1) compare and analyze the difference between the spectral data measured in the laboratory using an analytical spectral device (ASD) and the spectral data obtained by the satellite sensor advanced hyperspectral imager (AHSI) and to discuss the feasibility of the two in estimating the composition of the surface soil; (2) develop an estimation model suitable for this study area using the spectral data most suitable for estimating soil composition and to evaluate the performance of PLSR, PCR, and MLR.; (3) use the best modeling method and the best spectral transformation method to estimate the soil salt content and soil ion content and to use the satellite hyperspectral data to draw the spatial distribution map of the soil composition content; and (4) explore the ability of hyperspectral data to estimate the soil ion content.

2. Materials and Methods

2.1. Study Area

The Baidunzi basin of the Jingdian Irrigation Area was selected as the study area, and this basin has a prominent soil salinization problem. The irrigation area has a windy spring, a hot summer, a long frost-free period throughout the year, and sandy wind, which is typical of a continental climate. In the arid climate of the irrigation district, strong evaporation easily conducts the salts of the subsurface to the soil surface, resulting in the formation of surface salt, which is the phenomenon of salinization. The seasonal distribution of the irrigation area is clear.

The sunshine duration is long in the summer, short in the winter, and moderate in the autumn. The annual average temperature in the irrigation area is 8.3 °C, and the average precipitation is 185.7 mm, which is concentrated from May to September. The average annual evaporation is 2516 mm, the maximum annual evaporation is 3566 mm, and the minimum annual evaporation is 2227 mm [23]. The rise of the groundwater level caused by irrigation is the main factor leading to soil salinization in the Jingdian Irrigation Area. Due to the restriction of topography, the irrigated water carries salt into the Baidunzi basin (Figure 1), forming a tail-end lake [23].

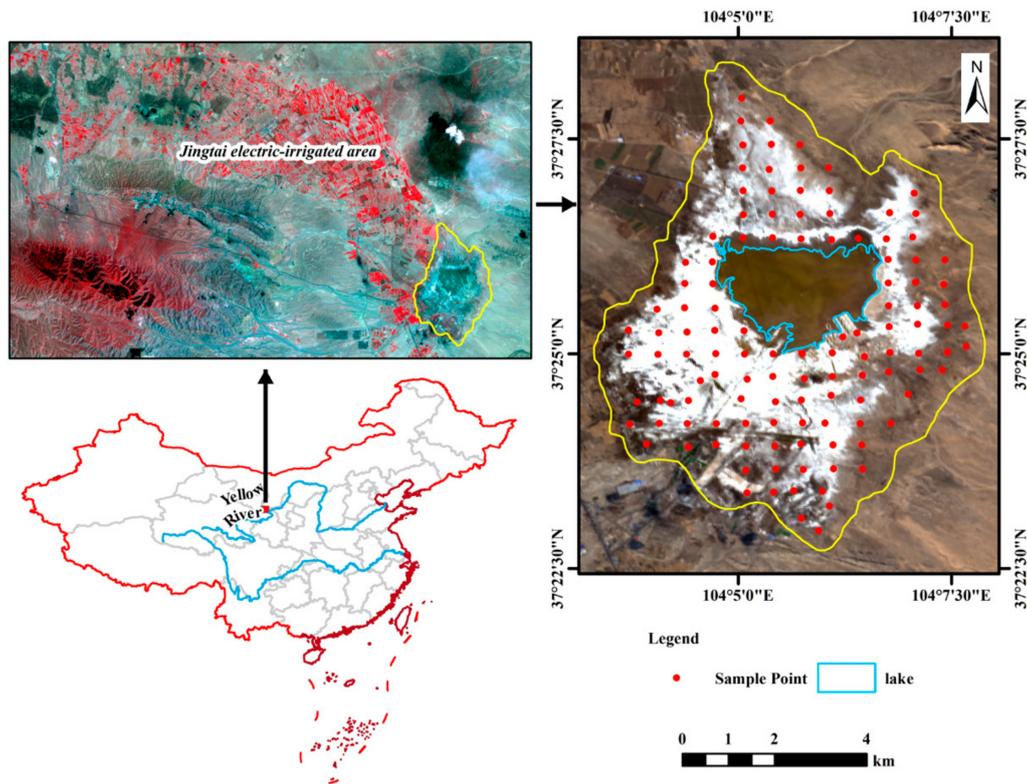


Figure 1. Location of the study area in China and in the Jingdian Irrigation Area, and the locations of collected samples.

2.2. Data Acquisition and Processing

2.2.1. Collection of Soil Data

Considering the transit time of the satellite, the field investigation was carried out on 5 December 2019. At this time, the study area was in winter, and there was almost no vegetation, which reduced the error of the salt content estimation caused by vegetation. A total of 119 soil samples were collected from the surface (0–10 cm). We took the samples back to the laboratory for analysis. The samples were first ground and sieved to eliminate the effect of the soil particle size, then the pH value of each sample was determined, and the SSC was determined using the water bath-evaporation method. Finally, ion chromatography and an inductively coupled plasma optical emission spectrometer were used to determine the cation and anion contents in the soil samples.

The soil spectra were measured in the laboratory in this study. A portable spectroradiometer (ASD) was used, and the spectral range was from 350 to 2500 nm, with a resolution of 3 nm in the visible near-infrared and 10 nm in the short-wave infrared. The spectroscopic measurements were performed in a dark room, where the prepared samples were placed in a 20-cm-diameter plastic Petri dish, and the surface of each sample was kept flat.

A tungsten–quartz–halogen filament lamp was used as the light source, and the angle of illumination was adjusted so that the brightest light was located in the center of the sample; the field of view of the sensor probe was 5° , and the distance to the soil surface was 15 cm, so that the area where the probe received the soil spectrum was much smaller than the Petri dish area to ensure that the probe received the reflected spectrum of the sample. Five spectral curves were collected for each soil sample, and the arithmetic mean value was taken as the actual reflectance of the soil sample.

2.2.2. Satellite Data Processing

We used the hyperspectral remote sensing images obtained by the visible-shortwave infrared advanced hyperspectral imager (AHSI) on board China's Gaofen-5 satellite, which

has a 60-km-wide image width and a 30 m spatial resolution and can simultaneously acquire spectral and spatial information of 330 consecutive spectral bands in the range of 400–2500 nm, with a spectral resolution of 5 nm in the visible near-infrared band and 10 nm in the short-wave infrared. The relative radiometric calibration accuracy is less than 3%, and the absolute radiometric calibration accuracy is less than 5%. Compared with the instrument on Hyperion, the signal-to-noise ratio of the hyperspectral camera is three to four times higher [24].

The images can be obtained from the website of the China Resource Satellite Application Center (<http://www.cresda.com/CN/> (accessed on 20 September 2020)). The images obtained were imaged on 27 November 2019. The AHSI images were geometrically corrected using the previously corrected images with a correction accuracy of 0.5 pixels. Atmospheric correction of the remote sensing images was then performed using the Fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) atmospheric correction model. The images were smoothed using a Savitzky–Golay Filter (SG) filter to reduce their system noise. The image water vapor absorption bands (1359–1409 and 1822–1923 nm) were marked as bad bands, and the bands with more noise due to being at the edge of the sensor (390–415 and 2285–2513 nm) were removed, resulting in 270 bands.

2.3. Method

To investigate the applicability of the laboratory-measured soil spectral data and AHSI data in estimating the soil content, the laboratory-measured spectra were first resampled into 270 bands of AHSI data. The responses for the different soil salinity levels in the two datasets were analyzed comparatively. The applicability of the two datasets was determined from the trend between the SSC and the reflectance of the two spectral datasets, and then the model was built based on the more applicable data. Spectral data come in many forms, such as raw spectral data and spectral transformed data, with different spectral transformations expressing soil salinity information differently from the suppression of other information [9].

The first derivative has been used extensively by researchers because this can amplify subtle information regarding the soil composition in spectral curves, thereby increasing the sensitive bands of the subject in the spectral data and obtaining good results [9,25]. In this paper, we performed a first-derivative analysis to expand the modeling data. Previous studies have shown that the first principal component of spectral data increases with increasing salt content [26,27]. Principal component analysis allows the hyperspectral data to be downsampled to obtain the main features of the data and reduces the correlation between the components, which is beneficial for modeling; thus, in this study, the data were subjected to principal component analysis, and the top ten principal components were selected.

The raw spectra of the hyperspectral data, the first derivative data of the spectra, and the spectral principal component data were selected as the raw datasets from which to select the modeling data. This stage entails the selection of feature data from the raw dataset that is sensitive to the salt data, which is key to modeling. When using spectral data, there are two options available: one is to use all spectral bands as modeling data, which uses a large amount of data, has significant co-linearity problems, and builds models with little generalizability.

The second approach is the use of feature bands, which are commonly used; this approach entails the selection of bands that are closely associated with the soil composition for modeling by correlation analysis [28,29]. In this study, we considered both the correlation between features and salinity data and the autocorrelation between features. We used the

correlation coefficient as an indicator for evaluating sensitive features (Equation (1)), and these features were selected at different visible near infrared and shortwave infrared bands.

$$r_i = \frac{\text{Cov}(R, S)}{\sqrt{D(R) \cdot D(S)}} = \frac{\sum_{i=1}^n (R_{ij} - \bar{R}_j)(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_{ij} - \bar{R}_j)^2 \sum_{i=1}^n (S_i - \bar{S})^2}} \quad (1)$$

where r_i is the correlation coefficient between the spectral reflectance and the soil components, R is the spectral reflectance, S is the soil component content, n is the total number of soil samples, j is the number of remote sensing data bands, i is the number of soil samples, R_{ij} is the spectral reflectance of the band, \bar{R}_j is the average value of soil reflectance in all bands of spectral data, S_i is the content of the soil component in each sample, and \bar{S} is the average value of the soil component content of each sample.

To increase the usefulness of the model, the three features that are optimal for modeling were selected as the modeling data. We used 70% of the total number of samples for model development and 30% as the model validation dataset. Based on the hyperspectral data, which had rich wavelengths, the number of variables based on these data is usually greater than the number of soil samples, and thus the ordinary least-squares model is no longer sufficient to build the model. Partial least squares regression (PLSR) is one of the most widely used methods to estimate the soil composition content.

Many scholars have established a good evaluation model based on this method [30,31]. In certain studies, its performance was even better than that of artificial neural networks [32]. Principal component regression (PCR) is an excellent method for hyperspectral data with a high correlation among different bands [11]. It is an ideal choice to use the partial principal components, which retain the main information of the image and have low correlation to the model for estimating the content of soil components.

Multiple linear regression (MLR) is more effective in the estimation of the soil component content by using multiple sensitive bands and has achieved good results in certain studies [33]. In this study, the model was developed using PLSR, PCR, and MLR using the coefficient of determination (R^2) and root mean squared error (RMSE) to evaluate the model (Equations (2) and (3)). The larger the R^2 value, the better the accuracy of the model, and the smaller the RMSE, the better the accuracy of the model. Finally, the model was cross validated.

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

$$\text{RMSE} = \sqrt{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)^2}{n}} \quad (3)$$

where n is the sample size, Y_i is the measured component content of the soil sample, \hat{Y}_i is the model prediction value of the soil composition content, and \bar{Y}_i is the average value of the soil composition content in the sample.

3. Results

3.1. Analysis of Soil Spectral Characteristics

The following data were obtained by measuring the soil samples (Table 1). The results showed that the average value of the soil pH was 8.38, and the average value of the soil salinity was 58.02 g/kg, which indicated that the soil salinity in the study area was severe. In this study, the spectral data of the soil were collected from the laboratory and satellite spectral data were obtained for a similar time period; thus, we were able to evaluate the ability of the laboratory and satellite data in establishing the soil content by analyzing the spectral characteristics of the salinized soil.

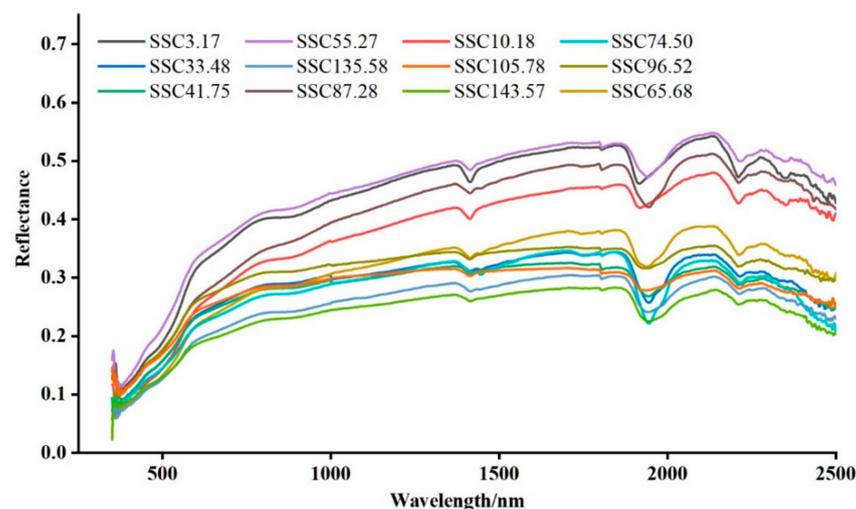
Table 1. Descriptive statistical parameters of the soil pH, SSC, and the contents of Na⁺, Ca²⁺, Mg²⁺, K⁺, Cl⁻, and SO₄²⁻.

	pH	SSC	Na ⁺	Ca ²⁺	Mg ²⁺	K ⁺	Cl ⁻	SO ₄ ²⁻
Unit	1	g/kg	ppm	ppm	ppm	ppm	ppm	ppm
Min	8.14	3.17	75.00	0.00	0.00	63.40	0.69	1.97
Max	8.98	143.57	11,960.00	6241.00	4142.00	693.40	3935.70	4541.90
Mean	8.38	58.02	3723.90	1361.50	960.16	286.23	1273.64	1506.53
SD	0.15	37.40	2796.85	1373.11	976.29	125.91	824.75	985.81

Figure 2 shows the laboratory measured soil spectra for different soil salinities, where the overall soil reflectance did not increase with the increasing soil salinity and vice versa. For example, a soil sample with a salinity of 143.57 g/kg had the lowest reflectance of the samples compared, while the sample with the highest reflectance was at the mean soil salinity level. This is consistent with the results of another study [21]. For the AHSI data, the same analysis was done, as in (Figure 3). The results showed that the soil reflectance increased as the soil salt content increased, although there were some exceptions, such as the reflectance of soil samples with a soil salt content of 143.57 g/kg was lower than the reflectance of samples with a salt content of 135.58 g/kg; however, on the whole, this trend was evident.

Figure 4 shows a comparison of the spectral data of soil samples measured in the laboratory using an ASD spectrometer with the AHSI data, which showed that the spectral data measured in the laboratory had better agreement with the AHSI spectral data when the soil salinity was low (e.g., the SSC was 3.17, 33.48, and 41.75 g/kg), and when the SSC was increasing compared to the laboratory measured spectra, the AHSI-measured reflectance had a significant increase in the visible NIR band, which also reflects the greater potential for estimating soil salinity using the visible NIR band. The AHSI reflectance curves also showed that the reflection peaks in the range of 550–650 nm became increasingly pronounced as the salt content increased, which is a characteristic not shown in the ASD data.

Figure 5 shows the correlation between the ASD data and AHSI data under different SSCs. The results showed that when the SSC was less than 33.48 g/kg, there was a good correlation between the ASD data and AHSI data. When the SSC was greater than 33.48 g/kg, the correlation gradually decreased. When the SSC reached 74.5 g/kg or greater than 74.5 g/kg, there was a negative correlation between the ASD data and AHSI data. In general, the AHSI data better reflected the real characteristics of saline soils; therefore, the AHSI data were chosen as the basis for the soil salinity estimation model.

**Figure 2.** Spectral curves of soil samples corresponding to different soil salinity levels measured in the laboratory using an analytical spectral device (ASD).

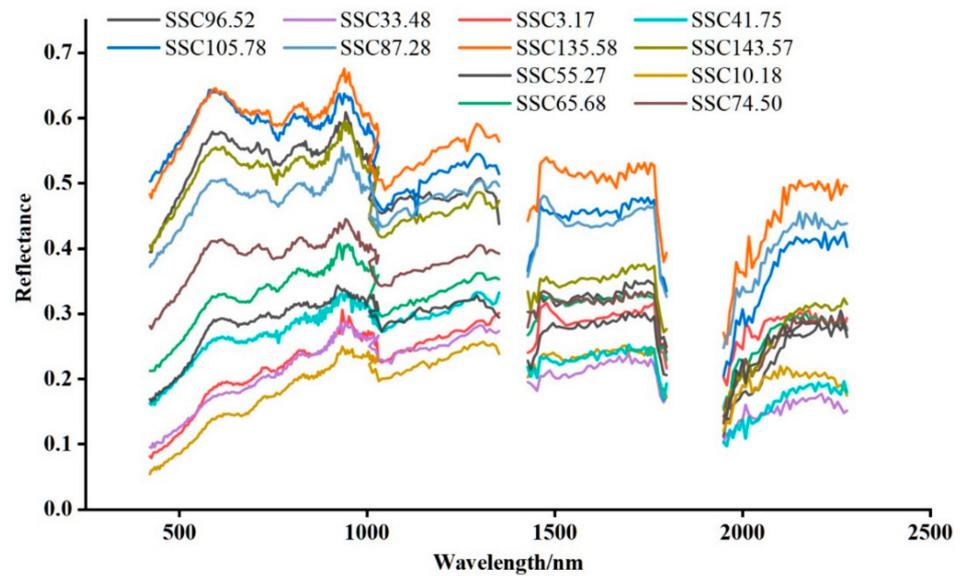


Figure 3. Spectral curves of soil samples on images obtained from the advanced hyperspectral imager (AHSI) sensor.

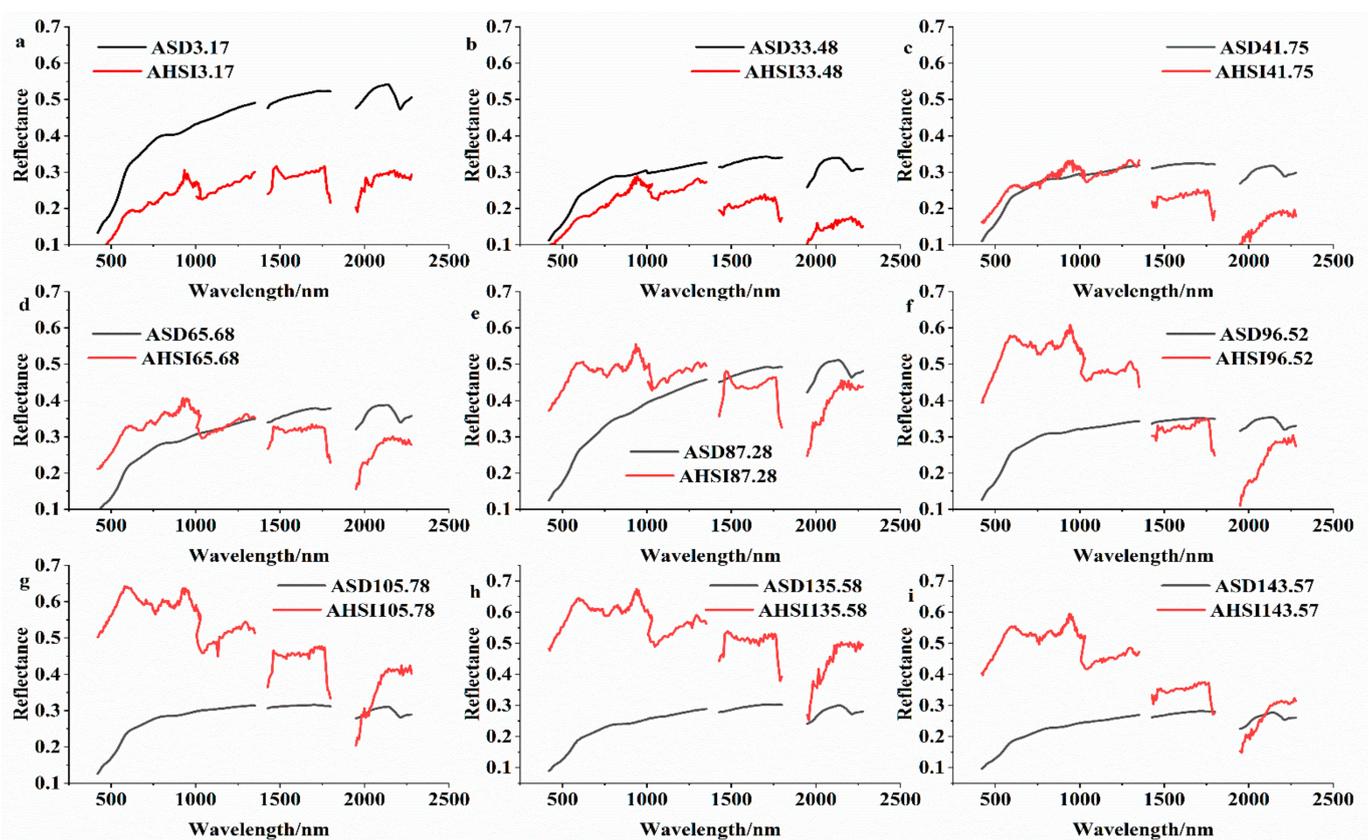


Figure 4. Resampling of the spectral curves obtained by laboratory instruments (ASD) to spectral curves obtained by satellite sensors (AHSI) and comparison with the AHSI data. (a–i) shows the comparison between ASD data and AHSI data under soil salt content of 3.17 g/kg, 33.48 g/kg, 41.75 g/kg, 65.68 g/kg, 87.28 g/kg, 96.52 g/kg, 105.78 g/kg, 135.58 g/kg and 143.57 g/kg.

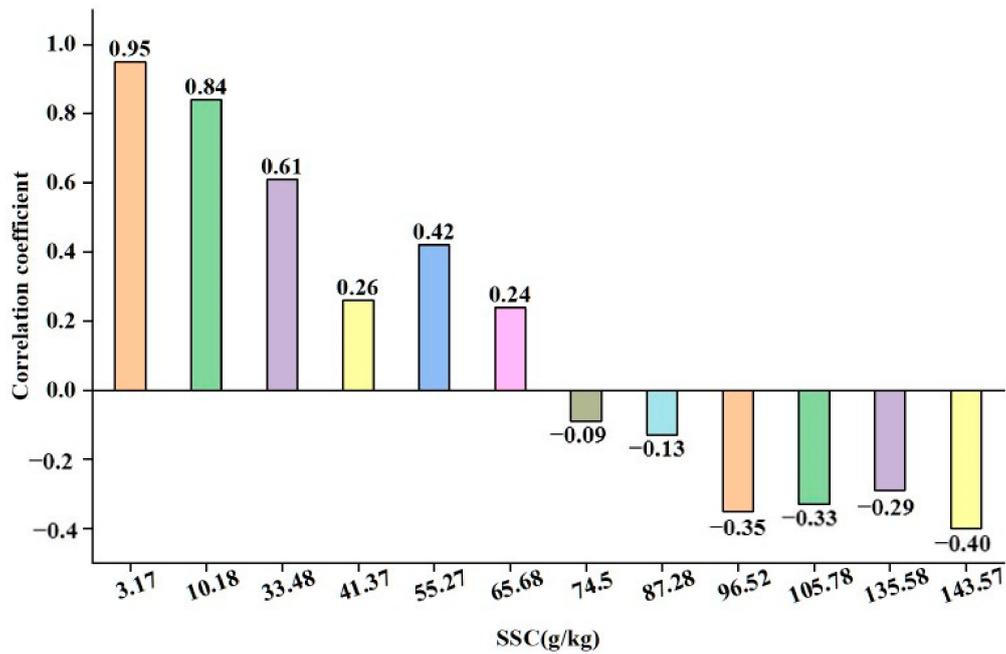


Figure 5. Correlation between ASD data and AHSI data under different SSCs.

3.2. Feature Band Selection

We performed correlation analysis of the raw spectral data, spectral first derivative data, and spectral principal component data with soil salinity data to summarize the maximum correlation coefficients between the three spectral data and soil composition data (Table 2). The correlation coefficient between the SSC and the raw spectrum was the best, the maximum correlation coefficient was 0.86, and the maximum correlation coefficient was obtained for the 424 nm band. The correlation coefficients of the Na⁺, Cl⁻, and SO₄²⁻ contents with the first derivative data were the best, and the maximum correlation coefficients were -0.68, -0.58, -0.52, and -0.54, respectively.

Table 2. Statistical table of the maximum correlation coefficient between the raw spectra (RS), first-derivative analysis (FDA), principal component analysis (PCA) and SSC, as well as Na⁺, Ca²⁺, Mg²⁺, K⁺, Cl⁻, and SO₄²⁻ contents.

Spectral Data Type	Index	SSC (g/kg)	Na ⁺ (ppm)	Ca ²⁺ (ppm)	Mg ²⁺ (ppm)	K ⁺ (ppm)	Cl ⁻ (ppm)	SO ₄ ²⁻ (ppm)
RS	Wavelength (nm)	424	424	591	2007	428	424	463
	Correlation coefficient	0.86 *	0.66 *	-0.08	-0.34 *	0.30 *	0.51 *	0.54 *
FDA	Wavelength (nm)	625	736	1105	1965	2167	1948	1956
	Correlation coefficient	-0.85 *	-0.68 *	0.24 *	-0.58 *	0.37 *	-0.52 *	-0.54 *
PCA	Component	1	1	6	2	2	2	2
	Correlation coefficient	0.80 *	0.60 *	-0.19 *	-0.52 *	-0.34 *	-0.46 *	-0.53 *

* The correlation is significant at the 0.05 level, using a two-tailed significance test.

The maximum correlation coefficients of Ca²⁺, Mg²⁺, and K⁺ were 0.24, -0.58, and 0.37, respectively. As the correlation coefficients were too small to establish a credible model to estimate the contents, this study did not carry out content estimation. We selected the characteristic bands of remote sensing data for soil components with high correlation coefficients, after which it was necessary to plot the correlation between each band or component of SSC and the AHSI raw data, first derivative data, and principal component data to find sensitive bands or sensitive band intervals.

(Figure 6a) shows the correlation of each soil component for the raw spectral data, and the correlation coefficients of the various soil component data changed similarly, with a decreasing trend from 420 nm, a slight increase after 1500 nm, and an increase after 2000 nm. In general, the correlations in the visible NIR band were better than those in the short-wave infrared. The correlation between the first-derivative data and the soil composition (Figure 6b) was very similar, and the overall trend was unevenly distributed. The first derivative amplified the details on the raw spectral data, resulting in improved correlations in the short-wave infrared band where the raw spectral correlation was lower, with higher correlation coefficients at 450–518, 779–788, 903, 1072, 1164, 1257, 1527, 1679, 2091, and 2235–2260 nm. Figure 6c shows the correlation between each principal component of the spectral data and the soil components, with the first, third, fourth, and sixth principal components being positively correlated with all soil data and the second and tenth principal components being negatively correlated with all soil data. The overall more strongly correlated principal components were the first, second, third, fourth, and tenth. Combined with the autocorrelation among the bands, the results of the sensitive bands screened in this paper are shown in (Table 3), and the three bands that were optimal for the modeling results were selected.

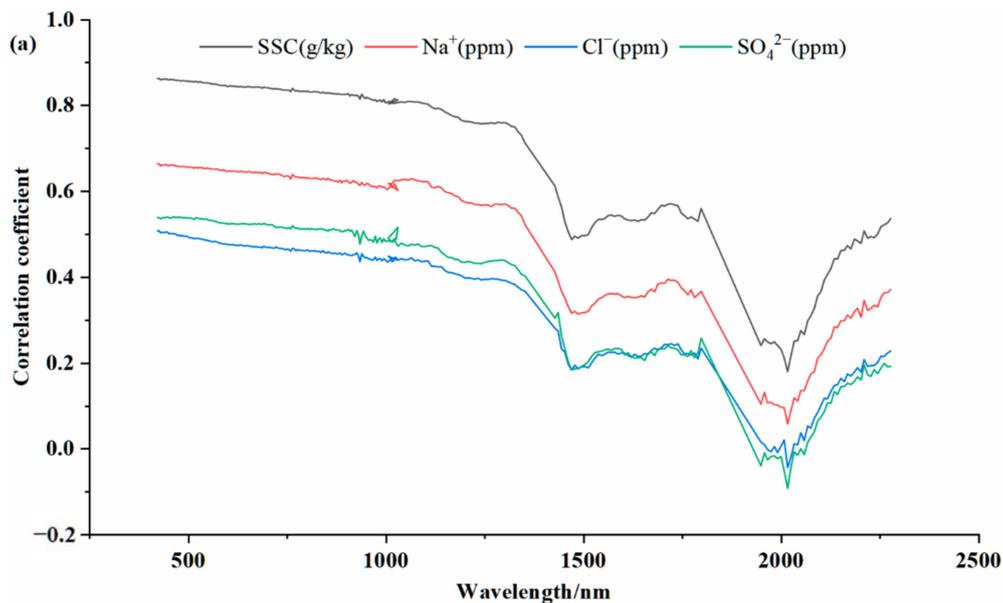


Figure 6. Cont.

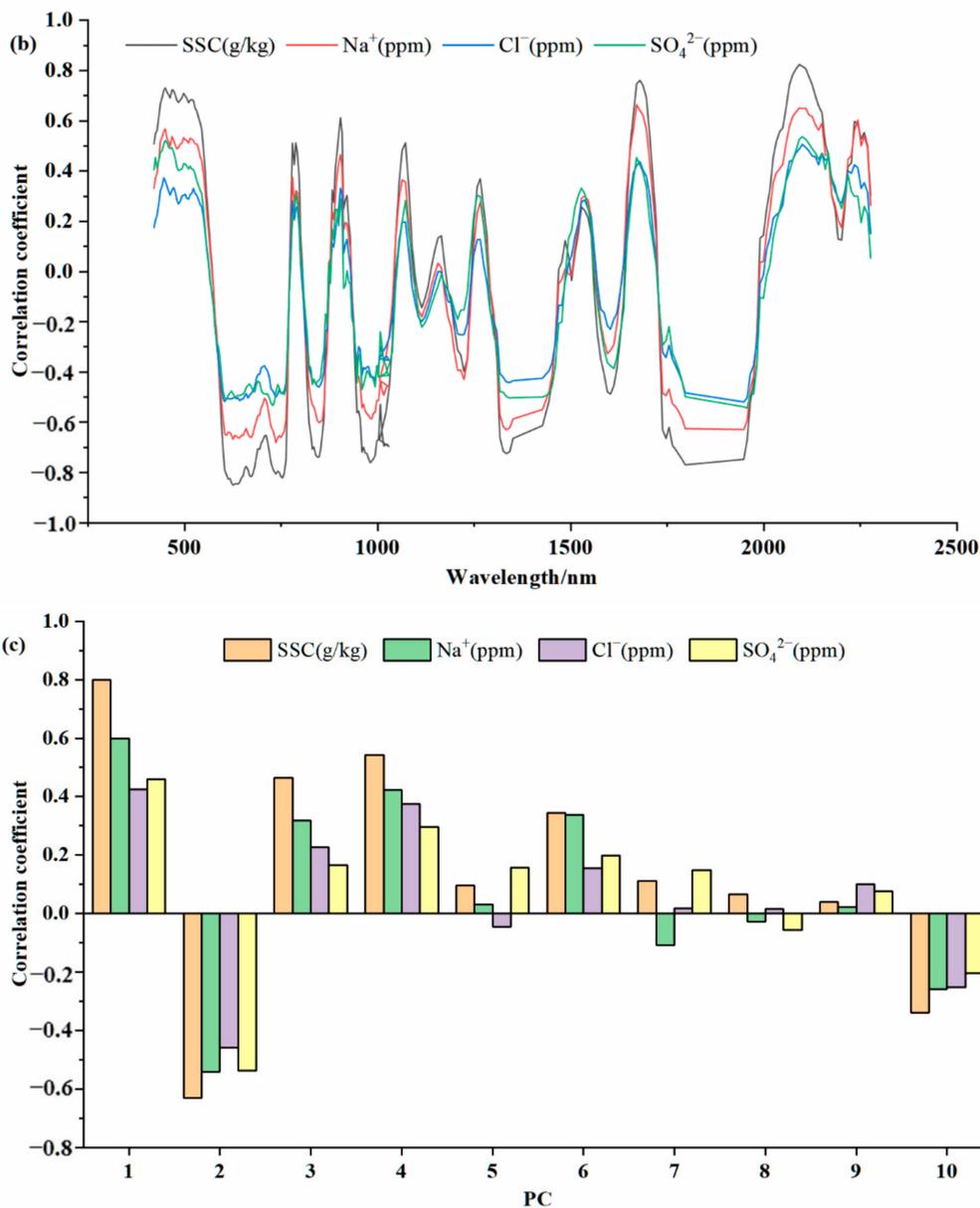


Figure 6. Correlation graph between the SSC, Na⁺, Cl⁻, and SO₄²⁻ contents and each band of RS data (a), graph of the correlation with each band of FDA data (b), graph of the correlation with each component of PCA (c).

Table 3. Based on the RS, FDA, and PCA data, the bands with the highest correlation with the SSC, Na⁺, Cl⁻, and SO₄²⁻ contents were extracted as the characteristic bands sensitive to soil components.

Spectral Data Type	SSC Feature Selection	Na ⁺ Feature Selection	Cl ⁻ Feature Selection	SO ₄ ²⁻ Feature Selection
RS	424, 441, 471 nm	424, 441, 471 nm	424, 441, 475 nm	463, 475, 420 nm
FDA	625, 612, 2091 nm	736, 625, 1670 nm	1948, 604, 642 nm	1956, 2100, 728 nm
PCA	1,2,4	1,2,4	1,2,4	2,1,4

3.3. Model Establishment and Verification

In this study, three methods, PLSR, PCR, and MLR, were used to model the estimation of the soil constituent contents. The results showed that the use of MLR yielded the largest R² for modeling the content of each constituent under different data transformations and obtained the smallest RMSE value at the maximum extracted value. Compared to PCR,

PLSR performed better in estimating the soil constituent contents under different data transformations. The raw spectral data showed better performance in estimating the SSC and Na⁺ contents with R² values of 0.79 and 0.58, respectively. The first-derivative transformation and principal component transformation performed optimally in estimating the Cl⁻ and SO₄²⁻ contents, respectively (Figure 6). Table 4 shows the soil composition content estimation models developed using the sensitive band and the optimal data transformations as well as the optimal method. (Figure 7) shows the results of cross validation of the developed models. Using the developed model, we obtained a final graph of the results of each soil content estimation (Figure 8), and it should be noted that we excluded the lake range from the soil content estimation.

Table 4. The model expression and evaluation parameters of data transformation and modeling method for estimating the soil composition.

Spectral Data Type	Method	Models	RMSE (g/kg)	R ²
RS	MLR	$SSC = 721.177*b424 + 1273*b441 - 1618*b471 + 16.46$	16.4	0.79
RS	MLR	$Na^+ = 68860*b424 + 172700*b441 - 2127*b471 + 2191.99$	1773.825	0.589
FDA	MLR	$Cl^- = -0.878*b1948 - 3.283*b604 - 1.041*b642 + 406.21$	563.41	0.46
PCA	MLR	$SO_4^{2-} = -0.06265*b2 + 0.0123*b1 + 0.256*b4 + 877.77$	670.01	0.41

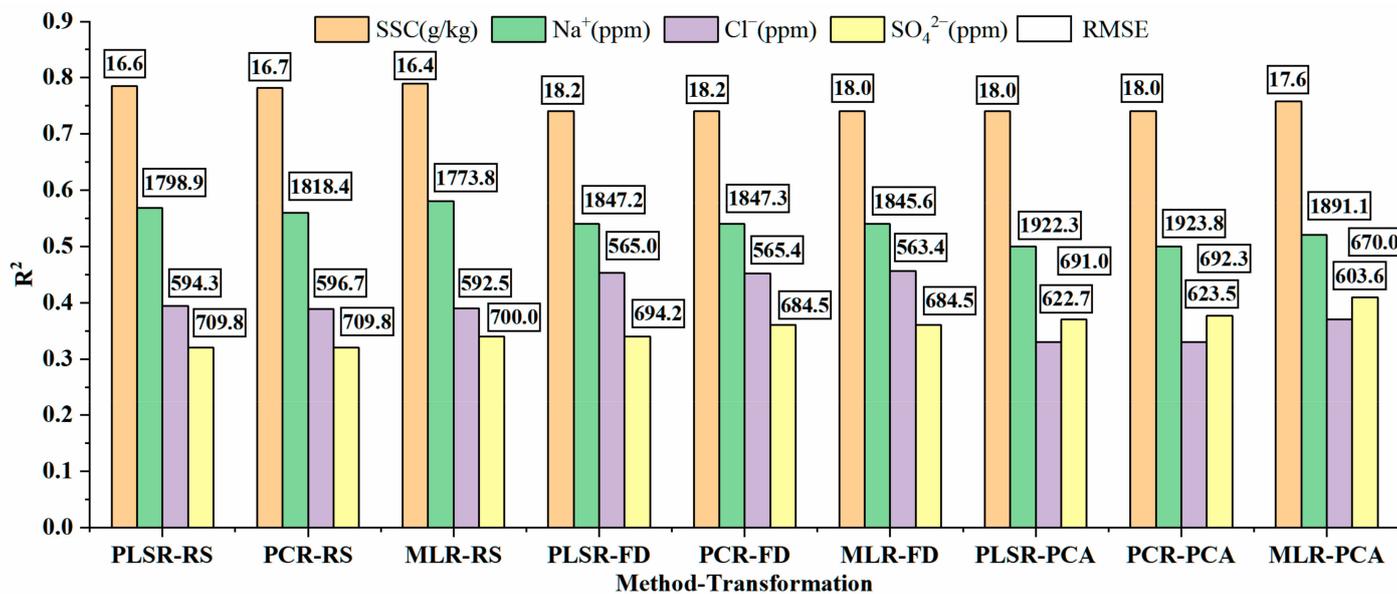


Figure 7. The performance of each soil component estimation model was established by combining different modeling data (RS, FDA, and PCA data) with different modeling methods (PLSR, PCR, and MLR). R² and RMSE were used to evaluate the performance of the models.

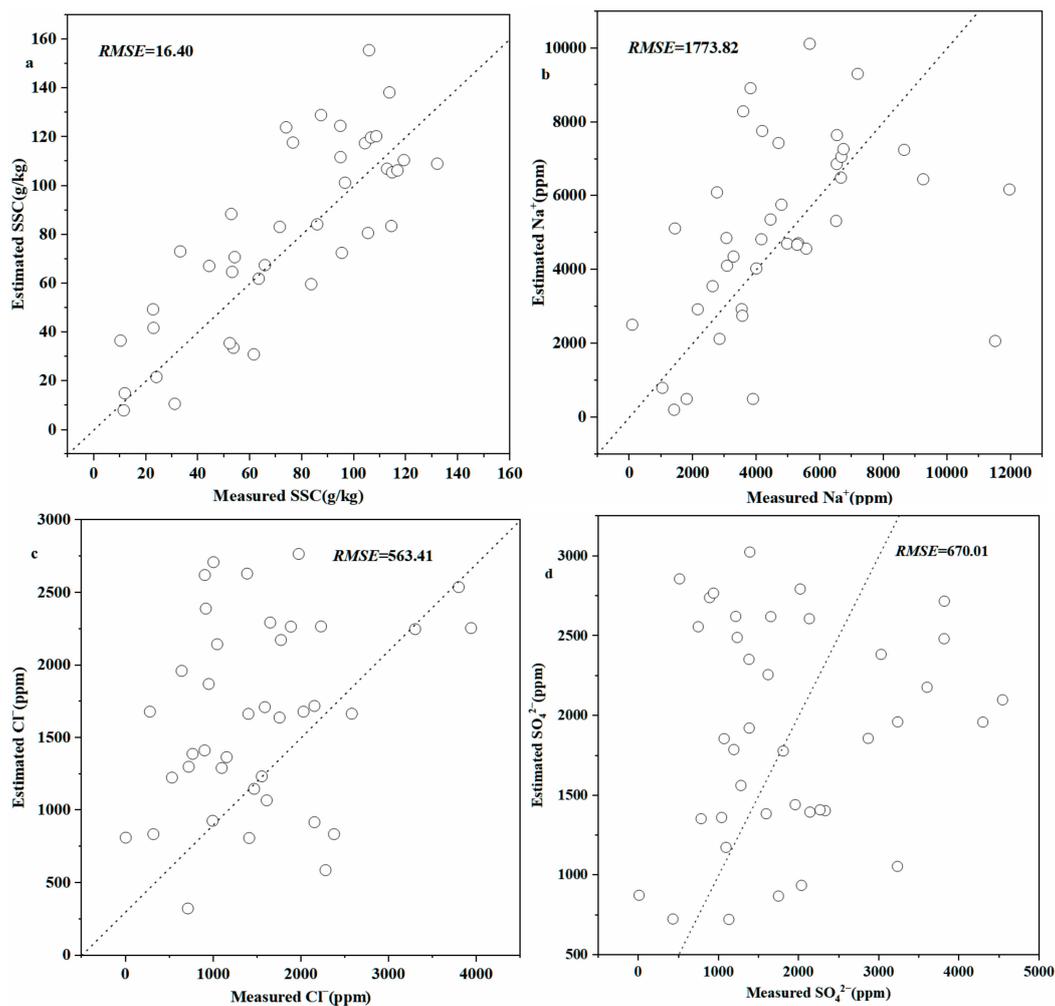


Figure 8. Relationship between the results for each soil component estimated using optimal data transformation and the best modelling methods and results for the actual measured soil components. (a) The relationship between measured SSC and estimated SSC; (b) the relationship between measured Na^+ content and estimated Na^+ content; (c) the relationship between measured Cl^- content and estimated Cl^- content; (d) the relationship between measured SO_4^{2-} content and estimated SO_4^{2-} content.

4. Discussion

We evaluated the ability of laboratory spectral data and AHSI data in estimating the soil composition, and used the AHSI data to model the estimation using the best modeling approach. We selected SSC, Na^+ , Cl^- , and SO_4^{2-} for modeling and mapping due to the limitation of low correlation between some of the ions and the spectral data.

Surface reflectance is affected by many factors, and the changes in reflectance with changing conditions as well as minimizing or eliminating these effects is key to improving the accuracy of estimates [32]. Satellite remote sensing data record information about the surface at a particular moment in time, and this is consistent for all samples collected simultaneously but is also affected by factors, such as the spectral and spatial resolution, image correction accuracy, and sensor characteristics, which can lead to errors between the measured and actual reflectance [34].

In contrast, a laboratory is able to control the interference of external factors to guarantee the measurements; however, there are also factors, such as the soil roughness, moisture, light, and changes in the field of view conditions, which are important factors that cause changes in the soil reflectivity [15]. It is therefore necessary to analyze the capabilities of both types of data when used as base data for estimating the soil composition content. A comparison of the ASD data with the AHSI data showed that as the soil salinity

increased, the reflectivity of the AHSI data increased, while the ASD data did not show this characteristic, which indicates that the ASD data were subject to a greater number of interfering factors.

From Figure 5, when the soil salt content was low, the ASD data and the AHSI data had a good correlation, which indicates that the laboratory spectral measurement data and satellite data estimated the soil in the area with a low soil salt content. The content of the components had good consistency, but as the soil salt content increased, the correlation between the ASD data and the AHSI data gradually decreased and even showed a negative correlation. When soil has a high salt content, the salt in the soil will precipitate out of the ground as the soil moisture changes, which will increase the soil reflectivity measured by the sensor, rather than decrease it as measured by the ASD data. This shows that the ASD data were subject to greater interference factors. Therefore, AHSI data have the potential to estimate the content of soil components.

The analysis of the AHSI reflectance data showed that the soil spectrum formed a reflection peak in the yellow and orange bands, which became increasingly pronounced as the salt content of the soil increased. This can be seen in Figures 3 and 4. The AHSI data had significant data noise in the 900–1000 nm range, and thus the spectral features that appeared are not characteristic of saline soils. For the soil compositions evaluated in this paper, the correlation coefficients of the raw soil spectra were not significantly different in the visible NIR (420–1006 nm), which is consistent with the results of several studies [21], and some of the short-wave infrared spectral bands (1006–1325 nm) also have the potential to estimate the salt content, as a significant decrease in the correlation coefficients was only observed starting at 1325 nm (Figure 6a).

The first-derivative transformation amplified the fine features of the spectral data, resulting in a significant increase in the correlation between each soil component and the spectral data from visible near infrared to short wave infrared, a result recognized in many studies [35,36], while the derivative transformation also reduced the correlation between sensitive bands, which is useful for modeling. The principal component transformation achieved good performance in estimating the SO_4^{2-} content.

Principal component transformations reduced the correlation between components and resulted in dimensionality reduction, which is a great source of inspiration for future methods of estimating the soil component content. MLR realized the best results in estimating the soil component content, and Abliz et al. also concluded that MLR was superior to PLSR in estimating the soil salinity in the Keriya Oasis [37]. However, there is no one method that can be applied to all regions [38], and regional differences are important factors that researchers should consider. At the same time, we can conclude that linear models are very effective for estimating soil composition, and their simple expressions have the advantage of being easy to use.

Soil brightness is affected by the degree of soil salinity [15,27]. Studies have shown that the soil salt content is positively correlated with the soil brightness, which is a characteristic of NaCl, and the results of this paper also showed such a feature. However, not all ions have such a correlation. The study by [39] illustrated that the brightness of MgCl_2 and CaCl_2 decreased when the salt concentration increased; however, unfortunately, the correlations of Mg^{2+} and Ca^{2+} with the AHSI spectral data were too low to develop a relatively credible model to assess their contents. Soil ion content estimation is currently a research challenge [40], mainly because most of the ion contents occur in trace amounts, and the spectral information is often masked by the spectral information of other substances. New and improved modelling methods will be the focus of future research.

From the spatial distribution maps of each soil component, we concluded that Na^+ , Cl^- , SO_4^{2-} , and SSC had good spatial agreement, reflecting the same salinity situation as reflected in the true color images, with the SSC values being higher for the whitened pixels and lower for the darkened pixels in the true color images. This is consistent with our expected results and reflects the credibility of the results. For regional salinity management, it is crucial to establish a highly accurate salinity evaluation model suitable for the region

to achieve dynamic monitoring of the salinity conditions [19,41]. We demonstrated that the AHSI had reliable performance in soil salinity evaluation, with more wavelengths and larger widths than sensors of the same type [24], and its graphical unity provided more favorable conditions for the accurate estimation of regional soil composition contents (Figure 9).

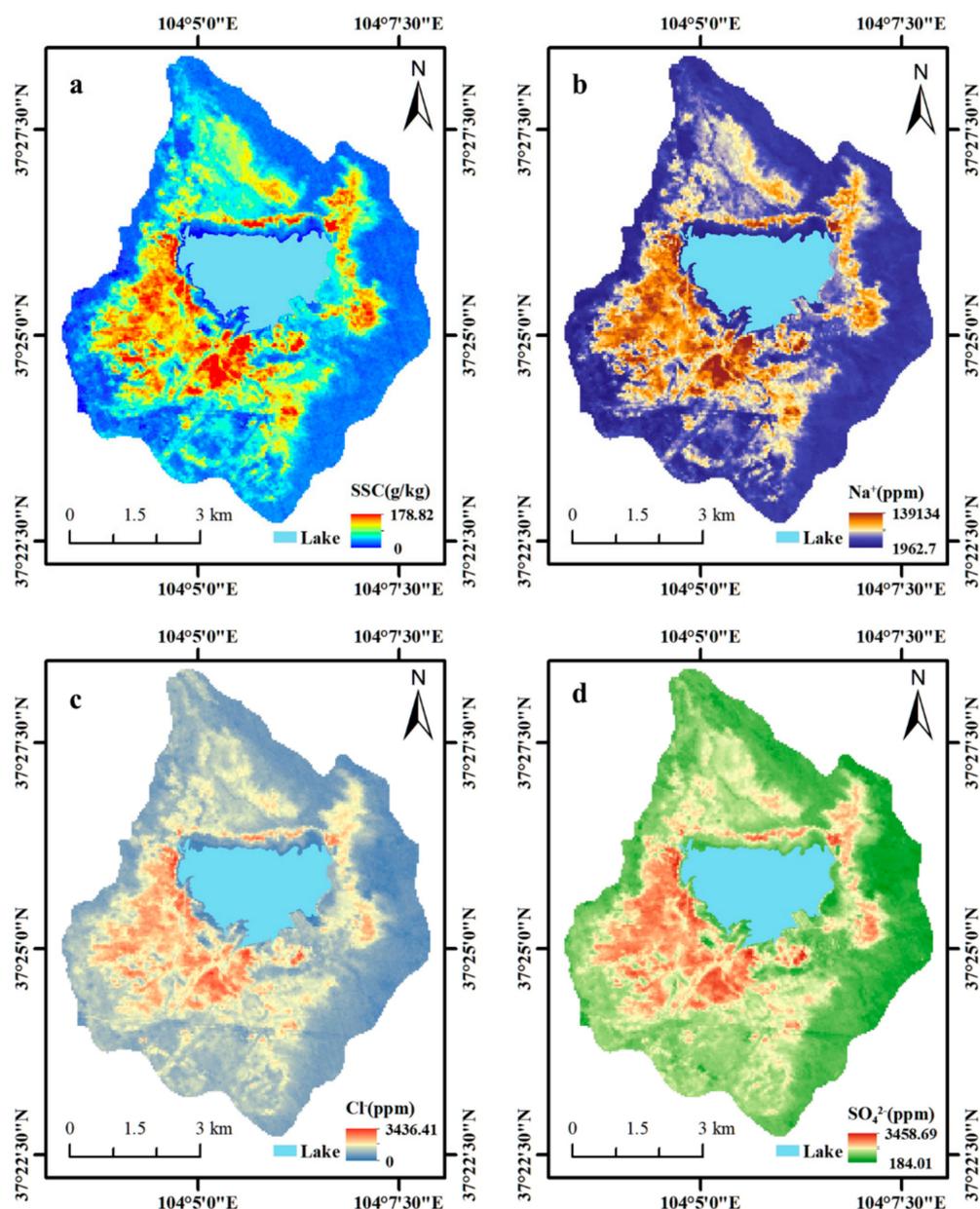


Figure 9. Estimation of the spatial distribution of the content of each soil component in the study area using the best available model. (a) Spatial distribution map of SSC in the study area; (b) spatial distribution map of Na⁺ content in the study area; (c) spatial distribution map of Cl⁻ content in the study area; (d) spatial distribution map of SO₄²⁻ content in the study area.

5. Conclusions

In this study, the correlation between the hyperspectral data measured in the laboratory and satellite hyperspectral data and soil salt content was compared. Finally, using AHSI satellite hyperspectral data as the basic data and using a variety of methods as comparison, an estimation model of the soil component content in this area was established, and the following conclusions were drawn.

- (a) With the increase in the soil salt content, there was a positive correlation between the satellite hyperspectral data and the soil salt content. When the soil salt content was less than 33.48 g/kg, the hyperspectral data measured by laboratory and satellite had a good correlation, but with an increase in the soil salt content, the correlation between the two gradually decreased. When the soil salt content was higher than 74.5 g/kg, the two had a negative correlation. Therefore, when the soil salt content is high, it is not recommended to use the spectral data measured in the laboratory as the basic data to establish the soil salt content estimation model.
- (b) By analyzing the spectral characteristics of saline soil reflected in AHSI data, with the increase in the soil salt content, the soil spectrum formed an obvious reflection peak in the yellow and orange light bands, which was beneficial to the establishment of the soil salt content estimation model based on the spectral characteristics. The first-derivative transformation increased the correlation between the soil composition and spectral data in the visible near infrared to short wave infrared bands and reduced the correlation between characteristic bands.
- (c) The best data transformation methods are different for estimating different soil components. When the original spectrum was used as the input data of the estimation model, the effect of estimating the soil salt content and sodium ion content was ideal, R^2 was 0.79 and 0.58, respectively, and a small RMSE was obtained. In the aspect of the modeling method, multiple linear regression achieved good results. At the same time, this study showed that AHSI data performed better in estimating the soil salt content.
- (d) The improvement of the soil ion content estimation accuracy is still the focus of future research, as it is difficult to establish a reliable soil ion content estimation model. In this study, we used satellite hyperspectral data to establish a soil ion content estimation model. We measured the concentrations of common ions in soil samples, namely, Na^+ , Ca^{2+} , Mg^{2+} , K^+ , Cl^- , and SO_4^{2-} , and calculated the correlation between the concentration of each ion and the satellite hyperspectral data. Unfortunately, only Na^+ , Cl^- , and SO_4^{2-} demonstrated a high correlation with the hyperspectral data. Therefore, to have a certain reference value, in this paper, we only established an estimation model with Na^+ , Cl^- , and SO_4^{2-} ions. This also shows that the hyperspectral satellite sensor data used in this paper have advantages for regional high-precision soil salinization monitoring.

Author Contributions: Data curation, L.W.; writing—original draft preparation, L.W.; writing—review and editing, L.W. and B.Z.; software, Q.S. and Y.Y.; validation, H.W.; R.Y. and Y.Z.; investigation, S.Z.; resources, B.Z.; funding acquisition, L.W. and H.W. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (31360204) and the Scientific Research Project of Northwest Normal University (2019KYZZ012058).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Available on request from corresponding author.

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

References

1. Scudiero, E.; Skaggs, T.H.; Corwin, D.L. Comparative regional-scale soil salinity assessment with near-ground apparent electrical conductivity and remote sensing canopy reflectance. *Ecol. Indic.* **2016**, *70*, 276–284. [[CrossRef](#)]
2. Daliakopoulos, I.N.; Tsanis, I.K.; Koutroulis, A.; Kourgialas, N.N.; Varouchakis, A.E.; Karatzas, G.P.; Ritsema, C.J. The threat of soil salinity: A European scale review. *Sci. Total Environ.* **2016**, *573*, 727–739. [[CrossRef](#)]
3. Rasool, S.; Hameed, A.; Azooz, M.M.; Muneeb u, R.; Siddiqi, T.O.; Ahmad, P. Salt Stress: Causes, Types and Responses of Plants. In *Ecophysiology and Responses of Plants under Salt Stress*; Ahmad, P., Azooz, M.M., Prasad, M.N.V., Eds.; Springer: New York, NY, USA, 2013; pp. 1–24.

4. Mashimbye, Z.E.; Cho, M.A.; Nell, J.P.; De Clercq, W.P.; Van Niekerk, A.; Turner, D.P. Model-Based Integrated Methods for Quantitative Estimation of Soil Salinity from Hyperspectral Remote Sensing Data: A Case Study of Selected South African Soils. *Pedosphere* **2012**, *22*, 640–649. [[CrossRef](#)]
5. Singh, A. Soil salinization and waterlogging: A threat to environment and agricultural sustainability. *Ecol. Indic.* **2015**, *57*, 128–130. [[CrossRef](#)]
6. Hamzeh, S.; Naseri, A.A.; AlaviPanah, S.K.; Mojaradi, B.; Bartholomeus, H.M.; Clevers, J.G.P.W.; Behzad, M. Estimating salinity stress in sugarcane fields with spaceborne hyperspectral vegetation indices. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 282–290. [[CrossRef](#)]
7. Periasamy, S.; Shanmugam, R.S. Multispectral and Microwave Remote Sensing Models to Survey Soil Moisture and Salinity. *Land Degrad. Dev.* **2017**, *28*, 1412–1425. [[CrossRef](#)]
8. Gorji, T.; Yildirim, A.; Hamzehpour, N.; Tanik, A.; Sertel, E. Soil salinity analysis of Urmia Lake Basin using Landsat-8 OLI and Sentinel-2A based spectral indices and electrical conductivity measurements. *Ecol. Indic.* **2020**, *112*, 106173. [[CrossRef](#)]
9. Wang, S.; Chen, Y.; Wang, M.; Zhao, Y.; Li, J. SPA-Based Methods for the Quantitative Estimation of the Soil Salt Content in Saline-Alkali Land from Field Spectroscopy Data: A Case Study from the Yellow River Irrigation Regions. *Remote Sens.* **2019**, *11*, 967. [[CrossRef](#)]
10. 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](#)]
11. Zhang, X.; Huang, B. Prediction of soil salinity with soil-reflected spectra: A comparison of two regression methods. *Sci. Rep.* **2019**, *9*, 5067. [[CrossRef](#)] [[PubMed](#)]
12. Bakker, D.M.; Hamilton, G.J.; Hetherington, R.; Spann, C. Salinity dynamics and the potential for improvement of waterlogged and saline land in a Mediterranean climate using permanent raised beds. *Soil Tillage Res.* **2010**, *110*, 8–24. [[CrossRef](#)]
13. Hu, J.; Peng, J.; Zhou, Y.; Xu, D.; Zhao, R.; Jiang, Q.; Fu, T.; Wang, F.; Shi, Z. Quantitative Estimation of Soil Salinity Using UAV-Borne Hyperspectral and Satellite Multispectral Images. *Remote Sens.* **2019**, *11*, 736. [[CrossRef](#)]
14. 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](#)] [[PubMed](#)]
15. Rocha Neto, C.O.; Teixeira, D.A.; Leão, A.R.; Moreira, C.L.; Galvão, S.L. Hyperspectral Remote Sensing for Detecting Soil Salinization Using ProSpecTIR-VS Aerial Imagery and Sensor Simulation. *Remote Sens.* **2017**, *9*, 42. [[CrossRef](#)]
16. Yao, R.J.; Yang, J.S.; Gao, P.; Shao, H.B.; Liu, G.M.; Yu, S.P. Comparison of statistical prediction methods for characterizing the spatial variability of apparent electrical conductivity in coastal salt-affected farmland. *Environ. Earth Sci.* **2014**, *71*, 233–243. [[CrossRef](#)]
17. Hively, W.D.; McCarty, G.W.; Reeves, J.B.; Lang, M.W.; Oesterling, R.A.; Delwiche, S.R. Use of Airborne Hyperspectral Imagery to Map Soil Properties in Tilled Agricultural Fields. *Appl. Environ. Soil Sci.* **2011**, *2011*, 358193. [[CrossRef](#)]
18. Wei, Y.; Shi, Z.; Biswas, A.; Yang, S.; Ding, J.; Wang, F. Updated information on soil salinity in a typical oasis agroecosystem and desert-oasis ecotone: Case study conducted along the Tarim River, China. *Sci. Total Environ.* **2020**, *716*, 135387. [[CrossRef](#)] [[PubMed](#)]
19. Masoud, A.A.; Koike, K.; Atwia, M.G.; El-Horiny, M.M.; Gemail, K.S. Mapping soil salinity using spectral mixture analysis of landsat 8 OLI images to identify factors influencing salinization in an arid region. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *83*, 101944. [[CrossRef](#)]
20. Ben-Dor, E.; Patkin, K.; Banin, A.; Karnieli, A. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data—a case study over clayey soils in Israel. *Int. J. Remote Sens.* **2002**, *23*, 1043–1062. [[CrossRef](#)]
21. Jiang, H.; Shu, H.; Lei, L.; Xu, J. Estimating soil salt components and salinity using hyperspectral remote sensing data in an arid area of China. *J. Appl. Remote Sens.* **2017**, *11*, 016043. [[CrossRef](#)]
22. Zovko, M.; Romić, D.; Colombo, C.; Di Iorio, E.; Romić, M.; Buttafuoco, G.; Castrignanò, A. A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia. *Geoderma* **2018**, *332*, 60–72. [[CrossRef](#)]
23. Zhang, M.; Zeng, Z.; Zhou, Q. The Both Causes of Formation and Prevention of Secondary Salinization of Soil in Jingtai Irrigation Area. *J. Lanzhou Univ. (Nat. Sci.)* **1990**, *26*, 148–153.
24. Liu, Y. Visible-shortwave Infrared Hyperspectral Imager of GF-5 Satellite. *Spacecr. Recovery Remote Sens.* **2018**, *39*, 25–28.
25. Xiaoping, W.; Fei, Z.; Hsiang-te, K.; Haiyang, Y. Spectral response characteristics and identification of typical plant species in Ebinur lake wetland national nature reserve (ELWNNR) under a water and salinity gradient. *Ecol. Indic.* **2017**, *81*, 222–234. [[CrossRef](#)]
26. Unger, D.; Bowes, C.; Farrish, K.; Hung, I.K. Mapping oilfield brine-contaminated sites with mid-spatial resolution remotely sensed data. *Gisci. Remote Sens.* **2013**, *50*, 623–632. [[CrossRef](#)]
27. Moreira, L.C.J.; Teixeira, A.d.S.; Galvão, L.S. Potential of multispectral and hyperspectral data to detect saline-exposed soils in Brazil. *Gisci. Remote Sens.* **2015**, *52*, 416–436. [[CrossRef](#)]
28. Kumar, S.; Gautam, G.; Saha, S.K. Hyperspectral remote sensing data derived spectral indices in characterizing salt-affected soils: A case study of Indo-Gangetic plains of India. *Environ. Earth Sci.* **2015**, *73*, 3299–3308. [[CrossRef](#)]
29. Allbed, A.; Kumar, L.; Aldakheel, Y.Y. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma* **2014**, *230–231*, 1–8.

30. Peng, J.; Biswas, A.; Jiang, Q.; Zhao, R.; Hu, J.; Hu, B.; Shi, Z. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* **2019**, *337*, 1309–1319. [[CrossRef](#)]
31. Sidike, A.; Zhao, S.; Wen, Y. Estimating soil salinity in Pingluo County of China using QuickBird data and soil reflectance spectra. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 156–175. [[CrossRef](#)]
32. Farifteh, J.; Van der Meer, F.; Atzberger, C.; Carranza, E.J.M. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sens. Environ.* **2007**, *110*, 59–78. [[CrossRef](#)]
33. Racetin, I.; Krtalic, A.; Srzic, V.; Zovko, M. Characterization of short-term salinity fluctuations in the Neretva River Delta situated in the southern Adriatic Croatia using Landsat-5 TM. *Ecol. Indic.* **2020**, *110*, 105924. [[CrossRef](#)]
34. Lagacherie, P.; Baret, F.; Feret, J.-B.; Madeira Netto, J.; Robbez-Masson, J.M. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sens. Environ.* **2008**, *112*, 825–835. [[CrossRef](#)]
35. Huang, Z.; Turner, B.J.; Dury, S.J.; Wallis, I.R.; Foley, W.J. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sens. Environ.* **2004**, *93*, 18–29. [[CrossRef](#)]
36. Qian, T.; Tsunekawa, A.; Peng, F.; Masunaga, T.; Wang, T.; Li, R. Derivation of salt content in salinized soil from hyperspectral reflectance data: A case study at Minqin Oasis, Northwest China. *J. Arid Land* **2019**, *11*, 111–122. [[CrossRef](#)]
37. Abliz, A.; Tiyip, T.; Ding, J.; Sawut, M.; Hou, Y.; Nurmemet, I. Estimating Soil Salt Content in the Keriya Oasis Using Hyperspectral Slope Index. *Nat. Environ. Pollut. Technol.* **2017**, *16*, 141–146.
38. Rahmati, M.; Hamzehpour, N. Quantitative remote sensing of soil electrical conductivity using ETM+ and ground measured data. *Int. J. Remote Sens.* **2017**, *38*, 123–140. [[CrossRef](#)]
39. Zhang, T.; Shao, Y.; Gong, H.; Li, L.; Wang, L. Salt Content Distribution and Paleoclimatic Significance of the Lop Nur “Ear” Feature: Results from Analysis of EO-1 Hyperion Imagery. *Remote Sens.* **2014**, *6*, 7783–7799. [[CrossRef](#)]
40. Wang, J.; Liu, Y.; Wang, S.; Liu, H.; Fu, G.; Xiong, Y. Spatial distribution of soil salinity and potential implications for soil management in the Manas River watershed, China. *Soil Use Manag.* **2020**, *36*, 93–103. [[CrossRef](#)]
41. Kasim, N.; Tiyip, T.; Abliz, A.; Nurmemet, I.; Sawut, R.; Maihemuti, B. Mapping and Modeling of Soil Salinity Using WorldView-2 Data and EM38-KM2 in an Arid Region of the Keriya River, China. *Photogramm. Eng. Remote Sens.* **2018**, *84*, 43–52. [[CrossRef](#)]