

Article

Soil Salinity Retrieval from Advanced Multi-Spectral Sensor with Partial Least Square Regression

Xingwang Fan ^{1,2}, Yuanbo Liu ^{1,*}, Jinmei Tao ³ and Yongling Weng ⁴

¹ Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, No. 73 East Beijing Road, Nanjing 210008, China; E-Mail: xwfan1989@163.com

² University of Chinese Academy of Sciences, No. 19 Yuquan Road, Beijing 100049, China

³ Jiangsu Provincial Bureau of Surveying and Mapping, No. 75 West Beijing Road, Nanjing 210013, China; E-Mail: phoebe_tao@sina.com

⁴ Surveying and Mapping Department, School of Transportation, Southeast University, No. 2 Sipailou, Nanjing 210096, China; E-Mail: wengyongling@seu.edu.cn

* Author to whom correspondence should be addressed; E-Mail: ybliu@niglas.ac.cn; Tel.: +86-25-8688-2164; Fax: +86-25-8688-2167.

Academic Editors: Arnon Karnieli and Prasad S. Thenkabail

Received: 21 May 2014 / Accepted: 23 December 2014 / Published: 6 January 2015

Abstract: Improper use of land resources may result in severe soil salinization. Timely monitoring and early warning of soil salinity is in urgent need for sustainable development. This paper addresses the possibility and potential of Advanced Land Imager (ALI) for mapping soil salinity. *In situ* field spectra and soil salinity data were collected in the Yellow River Delta, China. Statistical analysis demonstrated the importance of ALI blue and near infrared (NIR) bands for soil salinity. A partial least square regression (PLSR) model was established between soil salinity and ALI-convolved field spectra. The model estimated soil salinity with a R^2 (coefficient of determination), RPD (ratio of prediction to deviation), bias, standard deviation (SD) and root mean square error (RMSE) of 0.749, 3.584, 0.036 g·kg⁻¹, 0.778 g·kg⁻¹ and 0.779 g·kg⁻¹. The model was then applied to atmospherically corrected ALI data. Soil salinity was underestimated for moderately (soil salinity within 2–4 g·kg⁻¹) and highly saline (soil salinity >4 g·kg⁻¹) soils. The underestimates increased with the degree of soil salinization, with a maximum value of ~4 g·kg⁻¹. The major contribution for the underestimation (>80%) may result from data inaccuracy other than model ineffectiveness. Uncertainty analysis confirmed that improper atmospheric correction contributed to a very conservative uncertainty of 1.3 g·kg⁻¹. Field sampling within remote

sensing pixels was probably the major source responsible for the underestimation. Our study demonstrates the effectiveness of PLSR model in retrieving soil salinity from new-generation multi-spectral sensors. This is very valuable for achieving worldwide soil salinity mapping with low cost and considerable accuracy.

Keywords: soil salinization; advanced land imager; partial least square regression; Yellow River Delta

1. Introduction

Rapid population growth and economic development demand effective use of land resources. Excessive or improper land use may result in severe soil degradation [1]. Soil salinization is one of the major problems occurring in irrigated drylands [2,3]. It may trigger other problems such as soil dispersion, erosion or desertification [4,5], which are harmful to crop yields [6,7] and human health [8]. As a result, monitoring of soil salinity is urgently needed for salinization control [9,10].

Remote sensing possesses unique advantages over conventional proximal approaches in monitoring regional soil salinity [11–13]. It provides an inexpensive means for mapping large-scale soil salinity and a synoptic overview of soil salinization at remote regions. To date, multi-spectral remote sensing data are widely used in soil salinity studies. Since the 1970s, the Landsat Multi-Spectral Scanner (MSS)/Thematic Mapper (TM)/Enhanced Thematic Mapper (ETM+) have acquired over four decades of global data, invaluable for mapping salt-contaminated areas [14,15]. The Earth Observing (EO)-1 Advanced Land Imager (ALI) was launched to validate and demonstrate new techniques for Landsat 8 Operational Land Imager (OLI) [16,17]. The advanced multi-spectral sensors provide new opportunities for soil salinity mapping.

Surface reflectivity may vary with different levels of soil salinization [18], which allows image interpretation [19,20] or classification [21,22]. Statistical analysis can be also used to quantify soil salinity with multi-spectral remote sensing data. In most cases, multiple bands are transformed into an index more sensitive to soil salinity than a single band [23,24]. The most frequently used indices include Normalized Difference Vegetation Index (NDVI) and various salinity indices [25–31]. For single band, Shrestha [25] compared six Landsat ETM+ bands and concluded that the short-wave infrared (SWIR) band was most closely correlated with soil salinity. Soil salinity was positively correlated with satellite signals at ETM+ band 7 (Table 1). The values of correlation coefficient (r) and coefficient of determination (R^2) were only 0.484 and 0.234 between predicted and measured soil salinity. In general, the predictive accuracy was unfavorable for these models with only one independent variable. As a result, a large number of studies selected band combinations to improve salinity detection. Douaoui *et al.* [26] demonstrated that soil salinity was moderately ($r = 0.43 \sim 0.50$) correlated with a combination of red and green bands, but poorly ($r = 0.00 \sim 0.09$) with vegetation indices. The poor performance of vegetation index (VI) was also confirmed in Aldakheel [27]. Soil salinity was estimated with individual relationships for subareas ($R^2 = 0.20 \sim 0.50$), whereas there existed very weak correlation between soil salinity and NDVI for the whole study area [27]. In contrast, Allbed *et al.* [28] reported high correlations between soil salinity and VIs ($R^2 = 0.51 \sim 0.78$).

In general, soil salinity is negatively related with Vis. However, the accuracy for use of VIs may be case and site dependent. Besides a variety of VIs, Madani [29] also found a strong relationship between soil salinity and the difference of TM SWIR band and near-infrared (NIR) band. In addition, Bannari [30] declared that the index composed from SWIR bands was the most optimal indicator for soil salinity. It is likely that soil salinity may be also sensitive to SWIR bands, in addition to red and NIR bands which compose VIs. Inferentially, soil salinity could potentially be estimated from a combination of visible, NIR and SWIR bands. Odeh and Onus [31] also estimated soil salinity with a linear regression of two indicators. However, the R^2 value was only 0.32. In general, soil salinity can be quantitatively estimated from a variety of indices generated from multi-spectral data; however, the accuracy has not yet been well documented.

Since the 1980s, imaging spectroscopy has been applied for retrieval of soil salinity [32–38]. Salinized soil has distinct spectral features in reflective solar bands, related to water in hydrated minerals. The absorption features were detected at 505 nm, 920 nm, 1415 nm, 1915 nm and 2205 nm for saline soils, in addition to 680 nm, 920 nm and 1780 nm for soil scalds and highly salinized soils [12,33]. With statistical models, the dependence of soil salinity on diagnostic spectral features can be established. Ben-Dor *et al.* [34] performed a multiple regression analysis to predict soil salinity using 38 selected hyperspectral bands. Farifteh *et al.* [35] compared partial least square regression (PLSR) and artificial neural network (ANN) models to retrieve soil salinity from hyperspectral images, and concluded that the two models were comparable. Weng, *et al.* [36] also applied a PLSR model to EO-1 Hyperion images on a pixel-by-pixel basis. The PLSR model is effective in dealing with strong collinearity between independent variables [39], and has been successfully used in soil salinity retrievals from hyperspectral data [35,36,40]. Moreover, new technologies have been developed, for example, the use of sub-surface-to-surface correlation, frequency domain electromagnetic induction (FDEM) and ground penetrating radar (GPR) [41,42]. However, complicated atmospheric correction and high cost may preclude use of the hyperspectral remote sensing for mapping large-scale soil salinity [37,43]. Therefore, existing studies are generally confined to experimental analysis [44,45].

Hick and Russell [46] stated that multi-spectral remote sensing may not be optimal for soil salinity retrieval. To the best of our knowledge, only a few studies have correlated soil salinity with multi-spectral bands. For instance, Fan *et al.* [47] used Landsat TM and auxiliary elevation data to retrieve soil salinity. However, their models were established with moderate correlation coefficients of 0.504–0.736. Notably, recently launched sensors have several advantages over Landsat sensors, especially in spectral resolution and signal-to-noise ratio (Table 1). ALI has more bands in the blue, NIR and SWIR range of the electromagnetic spectrum. The improved quality may benefit high-accuracy soil salinity mapping. In our recent study, soil salinity was successfully retrieved from ALI data with a generalized regression neural network (GRNN) model [48]. In addition to the nonlinear model, a linear model is required to disclose explicitly the relationship between soil salinity and ALI bands.

This study was to investigate the possibility and potential in retrieving soil salinity from advanced multi-spectral sensor data with a linear regression model. For this purpose, ALI image was acquired over the Yellow River Delta (YRD), China, where long-standing soil salinization was reported due to both original and secondary salinization [21,49,50]. First, *in situ* data of soil spectra and salinity were statistically analyzed to find optimal spectral bands for soil salinity retrieval. Second, a PLSR model was established based on soil salinity measures and ALI-convolved field spectra. The model was then

applied to atmospherically corrected ALI data on a pixel-by-pixel basis. At last, errors associated with salinity retrieval were evaluated with a quantified sensitivity analysis. Our study may provide an inexpensive but reliable approach for quantitative monitoring of soil salinization at a large scale. The paper was organized as follows: Section 2 describes study materials and methodology; Section 3 reports the main results; Section 4 discusses on our findings and Section 5 concludes.

2. Materials and Methodology

2.1. Study Area

The Yellow River Delta is located at northeast of Shandong Province, China (Figure 1). The average annual precipitation is 530–630 mm, of which 70% occurs in the summer season (May–July). The average potential evaporation is 1900–2400 mm. This is approximately three times the precipitation needed for natural soil salinization. In addition, the groundwater is shallow with an average depth of 1.14 m [36]. According to Food and Agriculture Organization (FAO) classification, dominant soils include Gleyic Solonchaks, Calcaric Fluvisol and Salic Fluvisols [51]. Major mineral salts are NaCl and MgCl₂ [33]. Natural vegetation consists of salt tolerant herbs, grasses and shrubs. A part of YRD was selected as our study area, corresponding to an ALI scene (Figure 1). It covers a spatial extent within 37.2°N–38.2°N and 118.7°E–119.3°E, with a total area of about 3200 km². The western part is primarily agricultural croplands, especially along the Yellow River, and the eastern part is the offshore of Bohai Sea. There exist some oil fields in the North and some salt pans in the South.

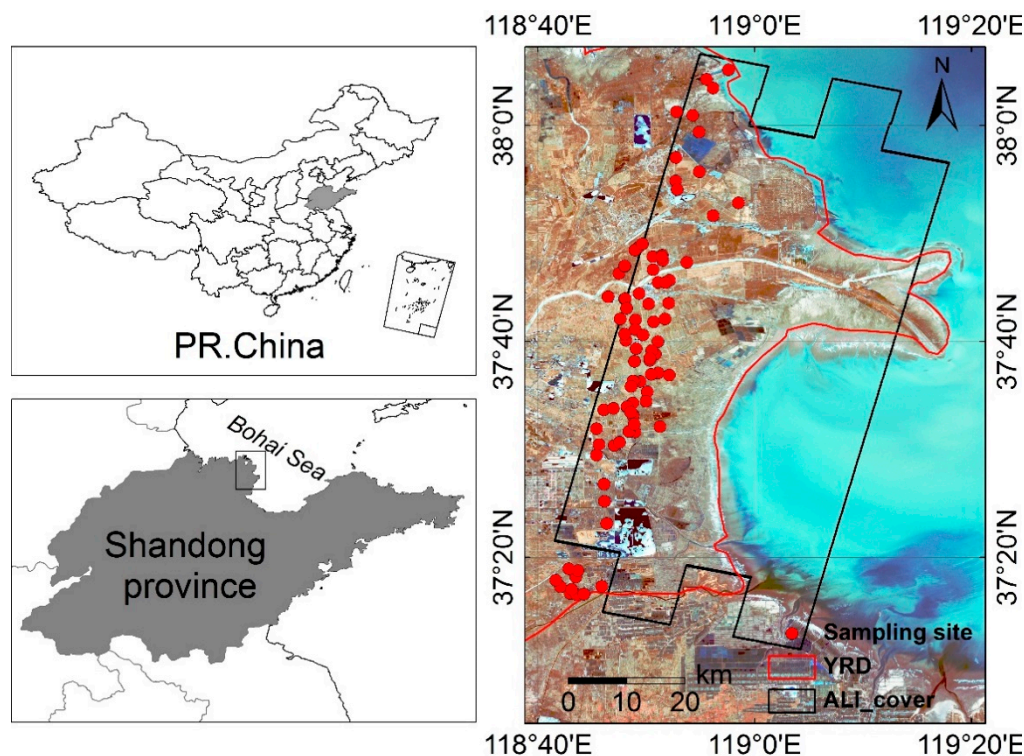


Figure 1. Location of study area. The study area is a part of the Yellow River Delta in Shandong province, China. The black frame confines our target area for soil salinity retrieval, corresponding to an ALI scene. The red line delineates the boundary of the Delta, and the red solid circles denote sampling sites.

2.2. Data Processing

2.2.1. Remote Sensing Data

An EO-1 ALI Level 1B multi-spectral image was used for soil salinity retrieval. It was acquired on 14 April 2005 from the Data Acquisition Requests (DAR) System at [52]. The sensor has nine multi-spectral bands (30 m) at solar reflective bands. Table 1 shows spectral bands of Landsat 7 ETM+, EO-1 ALI and Landsat 8 OLI sensors [16,17] to highlight unique spectral advantages of the ALI. The ALI bands are centered at wavelengths of 0.443 μm (Band 1p), 0.483 μm (Band 1), 0.565 μm (Band 2), 0.662 μm (Band 3), 0.790 μm (Band 4), 0.868 μm (Band 4p), 1.250 μm (Band 5p), 1.650 μm (Band 5) and 2.215 μm (Band 7). It has three more spectral bands (Bands 1p, 4p and 5p) and narrower spectral bandwidths over the ETM+. Compared to the OLI, the ALI has slightly broader spectral bandwidths, with an additional NIR band (Band 4).

Table 1. Comparison of Landsat 7 ETM+, EO-1 ALI and Landsat 8 OLI reflective spectral bands. The band names follow the ALI naming conventions. NIR = near-infrared band, SWIR = short-wave infrared band.

Band	Wavelength (μm)		
	Landsat 7 ETM+	EO-1 ALI	Landsat 8 OLI
Blue 1p	--	0.43–0.45	0.43–0.45
Blue 1	0.45–0.52	0.45–0.52	0.45–0.51
Green 2	0.52–0.60	0.53–0.61	0.53–0.59
Red 3	0.63–0.69	0.63–0.69	0.64–0.67
NIR 4	0.77–0.90	0.78–0.81	--
NIR 4p	--	0.85–0.89	0.85–0.88
SWIR 5p	--	1.20–1.30	1.36–1.38
SWIR 5	1.55–1.75	1.55–1.75	1.57–1.65
SWIR 7	2.09–2.35	2.08–2.35	2.11–2.29

ALI L1B data are provided in digital number (DN). The DN values were converted to top-of-atmospheric (TOA) reflectance with calibration coefficients provided in [53] for nine spectral bands. The conversion can be described as follows:

$$\rho_{TOA} = G \times DN + B \quad (1)$$

where ρ_{TOA} denotes the TOA reflectance, DN denotes the ALI digital number, G and B denote the gain and offset coefficients for sensor calibration.

TOA reflectance was an input for atmospheric correction. The TOA NDVI can be obtained from the following equation:

$$NDVI = \frac{\rho_{TOA,4} - \rho_{TOA,3}}{\rho_{TOA,4} + \rho_{TOA,3}} \quad (2)$$

where $NDVI$ denotes the normalized difference vegetation index, $\rho_{TOA,3}$ and $\rho_{TOA,4}$ denote the TOA reflectance for ALI bands 3 and 4.

Non-soil ALI pixels were removed to ensure a reasonable soil salinity retrieval. Water surface, vegetation cover and artificial objects were discarded with an NDVI-based decision tree method [54].

In most cases, water surface had a negative NDVI value, and it was removed with $NDVI < 0.03$. The threshold was set positive to remove shallow water area. Most artificial objects were spectrally flat with NDVI value close to zero, and they were removed with a conservative threshold of $NDVI < 0.05$. Vegetation had a high NDVI value, and it was removed with $NDVI > 0.14$. The remaining pixels were mainly comprised of bare soils at different saline levels. The average TOA reflectance was 16.83%, 15.54%, 15.71%, 16.58%, 19.84%, 19.89%, 20.63%, 19.76% and 15.79% for the nine bands. These values were used for simulation described in Section 2.3.4.

Auxiliary remote sensing data include the TERRA Moderate Resolution Imaging Spectroradiometer (MODIS) atmosphere products. TERRA overpassed the study area near-simultaneously with the EO-1, allowing use of the products for atmospheric correction of ALI image. The data include aerosol (MOD04) and atmosphere profile (MOD07) products, available at [55]. MOD04 was extracted for aerosol optical thickness (AOT) at 550 nm (AOT550), and MOD07 for total ozone concentration (TOC) and total precipitable water vapor content (TPW). To match ALI spatial resolution, these data were converted to vector form and interpolated to 30 m using inverse distance weighting (IDW) algorithm. For each $30\text{ m} \times 30\text{ m}$ grid, the ambient four MODIS product values were used for the interpolation. The interpolated values and the MODIS product values were compared with R^2 and root mean square error (RMSE) as defined in Equation (6) and Equation (10). Soil salinity uncertainties related to these errors were discussed in Section 4.3.

2.2.2. Field Campaign and Laboratory Analysis

A total of 68 topsoil (depth $< 5\text{ cm}$) samples were collected during a field campaign between 11 and 17 April 2005. Meteorological conditions were favorable during the campaign with low atmospheric humidity, low wind speed and long hours of sunshine (Table 2). Sampling sites were determined with a false color ($R = \text{band } 4$, $G = \text{band } 3$, $B = \text{band } 2$) image of Landsat TM acquired on 5 March 2005 (Figure 1). According to our knowledge and former studies, the strip-like transect covered soil salinity varying from higher level in south and north to lower level in between. The number of soil samples in each salinization level was balanced. At each sampling site, spectral reflectance was measured with an ASD Fieldspec Pro FR Spectroradiometer [56]. The spectroradiometer's probe was set within 15° field of view angle, and held vertically to ground surface. Before each measurement, a Spectralon[®] reference panel was measured. Soil reflectance was defined as a ratio of radiation reflected from soil surface to that from the reference panel. Within each $90\text{ m} \times 90\text{ m}$ square (3×3 ALI pixels), surface reflectance was measured at four homogenous targets, and averaged to represent the spectra of each sampling site. The homogenous targets were determined through a cluster analysis of the TM image and *in situ* visual verification. Correspondingly, topsoil samples were taken and sealed in plastic bags. At the same time, a handheld GPS was used to record the coordinates for each sampling site.

Field spectra data were provided at 1 nm intervals within 350–2500 nm. Due to strong atmospheric absorption, data quality was affected at spectral bands within 1355–1410 nm, 1810–1940 nm and 2451–2500 nm. The reflectance values deviated from the normal range of 0%–100%, and were omitted. As a result, a total of 1914 bands remained for further analysis. Since ALI spectral bands covered none of the discarded wavelengths, the omission had no impact on ALI-based salinity modelling. To reduce high-frequency noise caused by either atmospheric effect or instrument response,

a five-point smoothing algorithm was applied to the spectra data. Subsequently, the smoothed data were resampled in accordance to ALI spectral band responses. The resampled spectra were used as x -variables for soil salinity modelling. Standard laboratory method [57] was used to measure soil salinity in soil samples. The concentrations of Cl^- , SO_4^{2-} , CO_3^{3-} , HCO_3^- , K^+ , Na^+ , Ca^{2+} , and Mg^{2+} were measured using 1:5 soil-water mixtures. Soil salinity content (SSC) was defined as the total amount of the above eight ions. It was used as y -variable for modelling salinity–reflectance relationship. In terms of SSC value, soil salinization level was classified into non-saline, slightly saline, moderately saline and highly saline. The ranges were $0 < \text{SSC} < 1 \text{ g}\cdot\text{kg}^{-1}$, $1 < \text{SSC} < 2 \text{ g}\cdot\text{kg}^{-1}$, $2 < \text{SSC} < 4 \text{ g}\cdot\text{kg}^{-1}$ and $\text{SSC} > 4 \text{ g}\cdot\text{kg}^{-1}$ according to the coastal saline soil classification in China [58]. Field spectra were averaged within each saline level to find relations between soil salinity and soil reflectance.

Table 2. Meteorological conditions for the study area during the field campaign.

Statistical Indicators	Meteorological Factors			
	Temperature (°C)	Relative Humidity (%)	Wind Speed (m/s)	Sunshine Duration (h)
Mean	13.5	39.4	2.8	9.6
SD	3.5	8.5	0.9	2.5

Note: The data are available at [59] from a meteorological station at Dongying city, close to sampling sites.

2.3. Methodology

2.3.1. Atmospheric Correction

Remote sensing data are generally affected by the atmosphere along the path between ground target and satellite sensor. For a solar reflective band, main atmospheric effects include atmospheric absorption and scattering. Solar radiation is partially absorbed by ozone in visible bands, and by water vapor in NIR bands. Aerosol imposes impact on the entire solar reflective bands, with significant impacts on short wavelength. To retrieve surface reflectance, TOA reflectance data need to be corrected for ozone, water vapor and aerosol effects. A general atmospheric correction for a Lambertian and uniform surface can be written as follows [60]:

$$\rho_{TOA} = T_g \left(\rho_a + \frac{\rho T_s T_v}{1 - \rho S} \right) \quad (3)$$

where ρ_{TOA} denotes the TOA reflectance, ρ_a denotes the atmospheric path radiance due to molecular scattering and aerosol scattering, ρ denotes the surface reflectance, T_g denotes the atmospheric absorption due to ozone and water vapor, T_s and T_v respectively denote the downward and upward atmospheric scattering due to atmospheric molecules and aerosols, and S denotes the atmospheric spherical albedo.

Variables in Equation (3) can be related to sensor calibration and atmospheric parameters. For their detailed descriptions, see [58]. Sensor calibration coefficients are available from [53] and atmospheric parameters from MODIS atmosphere products [61]. With these known parameters, surface reflectance can be obtained by solving Equation (3). A mathematic form of the surface reflectance is as follows:

$$\rho = \frac{\rho_{TOA} - T_g \rho_a}{T_g T_s T_v + S(\rho_{TOA} - T_g \rho_a)} \quad (4)$$

2.3.2. Partial Least Square Regression (PLSR) Model and Performance Assessment

The PLSR model has been widely used for soil salinity retrieval due to its strong ability in processing of data with high collinearity [35,36,39,40]. To identify significant x -variables (wavelengths) for soil salinity retrieval, a PLSR model was developed based on Martens' Uncertainty Test [62]. The test was a significance method, now implemented in The Unscrambler 9.7 software [63]. The test can identify dominant sources of instability in regression modelling and quantify the stability of the regression results. It also allows identification of perturbing variables and selection of significant variables. In this study, all 68 samples were injected for the uncertainty test, with soil salinity as y -variable and the corresponding 1914-band spectra data as x -variables. The output was the important wavelengths for soil salinity retrieval.

The PLSR model can be established for soil salinity retrieval based on ALI-convolved field spectra data. The model takes a form as follows:

$$SSC = \sum_{i=1}^{N=9} w_i \times \rho_i + w_0 \quad (5)$$

where SSC denotes soil salinity content, ρ denotes the ALI-convolved spectral reflectance, w denotes model coefficient for the i th band, w_0 denotes the constant item, and N denotes the number of spectral bands. In this study, two-thirds (45) of the 68 samples were used for modelling, and the others for validation. The x -variables were nine-band ALI convolved spectra data, and the y -variable was soil salinity.

Model performance can be evaluated with R^2 , ratio of prediction to deviation (RPD), Bias, standard deviation (SD) and RMSE. The R^2 and RPD indicate the strength of statistical correlation between measured and predicted values. The model can be accurate for $R^2 > 0.91$ /RPD > 2.5 , good for $0.82 < R^2 < 0.90$ /RPD > 2 , moderate for $0.66 < R^2 < 0.81$ /RPD > 1.5 , and poor for $0.5 < R^2 < 0.65$ [35]. The Bias measures the mean difference of predicted *versus* measured value, and the SD represents the random component of total uncertainty (RMSE) [64]. The metrics can be described as follows:

$$R^2 = 1 - \frac{\sum (\hat{y}_i - y_i)^2}{\sum (y_i - \bar{y})^2} \quad (6)$$

$$RPD = \frac{\left\{ \sum \hat{y}_i - \left[\left(\sum \hat{y}_i \right)^2 / N / (N-1) \right] \right\}^{1/2}}{\left\{ \sum (\hat{y}_i - y_i)^2 - \left[\sum (\hat{y}_i - y_i) \right]^2 / N / (N-1) \right\}^{1/2}} \quad (7)$$

$$Bias = \frac{\sum (\hat{y}_i - y_i)}{N} \quad (8)$$

$$SD = \left[\frac{\sum (\hat{y}_i - y_i - Bias)^2}{N-1} \right]^{1/2} \quad (9)$$

$$RMSE = \left[\frac{\sum (\hat{y}_i - y_i)^2}{N} \right]^{1/2} \quad (10)$$

where y denotes the measures with a mean value of \bar{y} , \hat{y} denotes the predicted values, and N denotes the number of samples.

2.3.3. Soil Salinity Mapping and Error Analysis

Soil salinity can be mapped by applying Equation (5) and substituting x -variables with corresponding ALI-band surface reflectance data. Retrieved soil salinity can be extracted with GPS coordinates of sampling sites, and compared with measured and modelled values. For convenience, the retrieved soil salinity was labeled as SSC_{image} . Similarly, the measured and modelled values were labeled as $SSC_{measure}$ and SSC_{model} . The total SSC error was subsequently determined as the difference between the retrieved and measured values ($SSC_{image} - SSC_{measure}$). The total error can be separated into model- and data-related errors. The model related error is the difference between the modelled and the measured values ($SSC_{model} - SSC_{measure}$), which indicates the performance of the PLSR model. The data related error is the difference between the retrieved and the modelled values ($SSC_{image} - SSC_{model}$), which indicates the discrepancy between the field measures and the ALI-retrieved values. The total errors can be plotted against the measured soil salinity to detect major contributors to the retrieval error in soil salinity.

2.3.4. Uncertainty Analysis

According to Equations (1) and (4), ALI surface reflectance is an analytic function of sensor calibration (CAL), TOC, TPW and AOT. Incorporating Equation (5), the ALI-derived soil salinity is also an analytic function of CAL, TOC, TPW and AOT, and can be expressed as follows:

$$SSC = f(CAL, TOC, TPW, AOT) \quad (11)$$

It indicates that uncertainty associated with soil salinity retrieval is related to the uncertainties in sensor calibration and ozone, water vapor and aerosol products. The uncertainty due to each component is a derivative of function (11) with regard to CAL, TOC, TPW and AOT. Assuming that each component be independent, the total uncertainty can be derived as follows:

$$U_{SSC} = \left\{ \underbrace{\sum_{i=1}^N \left[w_i \times \frac{\partial f(TOA, i)}{\partial TOA} \times TOA_i \times re_{CAL_i} \right]^2}_{\text{Calibration related}} \right\}^{1/2} + \underbrace{\left[\sum_{i=1}^N w_i \times \frac{\partial f(TOC, i)}{\partial TOC} \times TOC \right] \times re_{TOC}}_{\text{Ozone related}} \quad (12)$$

$$+ \underbrace{\left[\sum_{i=1}^N w_i \times \frac{\partial f(TPW, i)}{\partial TPW} \times TPW \right] \times re_{TPW}}_{\text{Water vapor related}} + \underbrace{\left[\sum_{i=1}^N w_i \times \frac{\partial f(AOT, i)}{\partial AOT} \times AOT \right] \times re_{AOT}}_{\text{Aerosol related}}$$

where U_{SSC} denotes the soil salinity uncertainty, re_{CAL} , re_{TOC} , re_{TPW} and re_{AOT} denote relative uncertainty of CAL, TOC, TPW and AOT, and f denotes Equation (11).

Equation (12) can be divided into four components: CAL-, TOC-, TPW- and AOT-related uncertainty. For uncertainty analysis, TOA reflectance was set as the mean of TOA soil reflectance.

TOC, TPW and AOT were set as the mean values of corresponding MODIS atmosphere products. For sensitivity analysis, TOA reflectance varied by $\pm 5\%$ (in absolute amount) for each band. The threshold can be determined with variation range of field spectra data. Low reflectance stands for slightly saline soil, and high reflectance for highly saline soil. TOC varied by $\pm 0.1 \text{ cm}\cdot\text{atm}^{-1}$, and TPW by $\pm 0.5 \text{ g}\cdot\text{cm}^{-2}$, and AOT by ± 0.5 . High TPW and AOT values represent humid and turbid air, and low TPW and AOT values represent dry and clear air. The relative uncertainty of each component was set within 0%–20% to represent a low-to-high degree of uncertainty.

3. Results

3.1. Statistical Descriptions of Field Spectra Data

Figure 2 shows a relationship between soil salinity and field spectra data. The correlation coefficient ranged between -0.534 and 0.322 , with maximum positive r value at 350 nm and maximum negative r value at 1354 nm. Soil salinity was positively correlated with spectral reflectance for wavelengths within 350 nm and 523 nm, and the correlation coefficient decreased with wavelength. For longer spectral wavelength, soil salinity was negatively correlated with spectral reflectance. The correlation coefficient increased from almost zero at 523 nm to about -0.5 at approximately 1000 nm. It remained stable for spectral wavelengths within 1000 nm and 2450 nm. Figure 2 also illustrates spectral positions of the ALI spectral bands. ALI bands 1p and 1 were located at blue band, and soil spectral reflectance at these bands was positively correlated with soil salinity. Other bands were located at positions with longer wavelengths, and soil reflectance was negatively correlated with soil salinity.

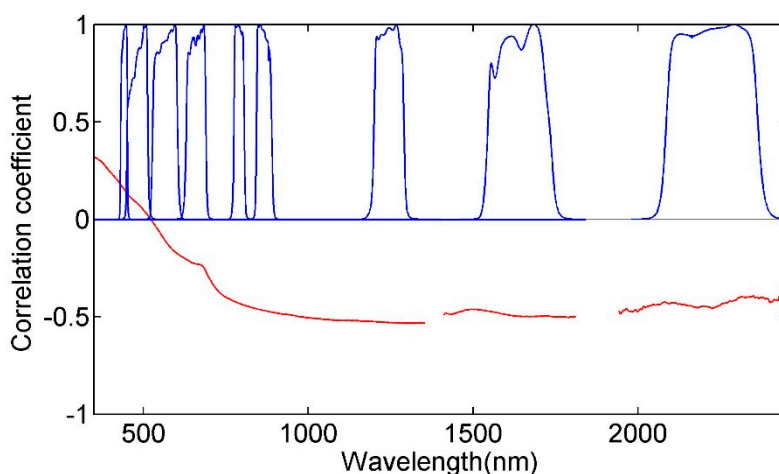


Figure 2. Correlation relationship between soil salinity and field spectra data (red lines). ALI relative spectral functions were superimposed to illustrate the positions of ALI spectral bands (blue lines).

Figure 3 shows the mean spectral reflectance for soil at different saline levels. There were 19, 14, 21 and 14 samples for SSC within $0\text{--}1 \text{ g}\cdot\text{kg}^{-1}$, $1\text{--}2 \text{ g}\cdot\text{kg}^{-1}$, $2\text{--}4 \text{ g}\cdot\text{kg}^{-1}$ and $\text{SSC} > 4 \text{ g}\cdot\text{kg}^{-1}$. In general, soil reflectance decreased with the degree of soil salinization. The trend was more significant for infrared bands than for visible bands. Spectral reflectance was similar to each other over a wide range of spectral wavelength for non-saline and slightly saline soils, indicating that it is generally difficult to

discriminate them. Notably, their differences became larger at spectral wavelengths longer than approximately 1900 nm. Highly saline soil had higher spectral reflectance than moderately saline soils at shorter wavelengths, and the reflectance was comparable with that of non-saline and slightly saline soils. These spectral behaviors provided a base for quantitative retrieval of soil salinity.

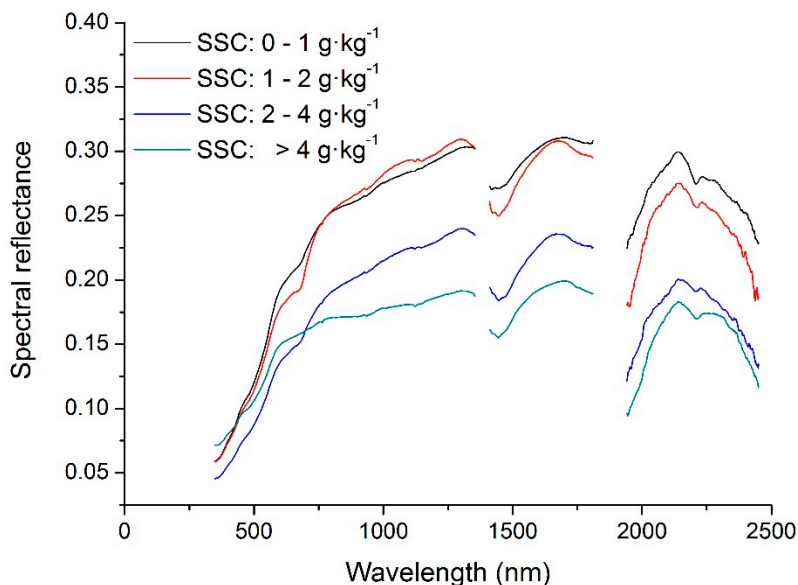


Figure 3. Mean spectral reflectance for soils at different saline levels. Non-saline, slightly saline, moderately saline and highly saline soils are defined in terms of soil salinity value as $0 < \text{SSC} < 1 \text{ g}\cdot\text{kg}^{-1}$, $1 < \text{SSC} < 2 \text{ g}\cdot\text{kg}^{-1}$, $2 < \text{SSC} < 4 \text{ g}\cdot\text{kg}^{-1}$ and $\text{SSC} > 4 \text{ g}\cdot\text{kg}^{-1}$.

3.2. Statistical Descriptions of ALI-Convolved Field Spectra Data

Soil salinity ranged between 0.39 and $6.25 \text{ g}\cdot\text{kg}^{-1}$, with a mean and SD value of $2.5 \text{ g}\cdot\text{kg}^{-1}$ and $1.6 \text{ g}\cdot\text{kg}^{-1}$. Table 3 shows descriptive statistics of soil salinity and ALI-convolved field spectra. For $\text{SSC} < 4 \text{ g}\cdot\text{kg}^{-1}$, spectral reflectance decreased with soil salinity for all nine bands. This may be caused by increased soil moisture, as the hygroscopic MgCl_2 was the major mineral salt in the soils [33]. Spectral reflectance increased at bands 1p-3 for high saline soils. A possible explanation may be the existence of soil crust that had a bright color in visible bands.

Table 4 shows correlation coefficients between soil salinity and ALI-convolved field spectra, and coefficients between ALI-convolved field spectral bands. Soil salinity had a low positive correlation coefficient with reflectance at bands 1p and 1 ($r = 0.144$ and 0.060), whereas it had a weak negative correlation coefficient at ALI green and red bands ($r = -0.134$ and -0.264). Soil salinity was moderately correlated with ALI NIR and SWIR bands ($r = -0.477$, -0.516 , -0.582 , -0.541 and -0.469). It was worth noting that band 1p was more sensitive to soil salinity than band 1 ($r = 0.144$ vs. 0.060). Similarly, band 4p also had increased sensitivity relative to band 4 ($r = -0.516$ vs. -0.477). Among others, the reflectance at band 5p had the largest negative correlation coefficient ($r = -0.582$) with soil salinity. Moreover, it had relatively lower r values with other bands. This implied that band 5p could provide complementary information for soil salinity retrieval.

Table 3. Descriptive statistics of soil salinity and ALI-convolved field spectra for soils at different saline levels.

Level	Range (g·kg ⁻¹)	Mean (g·kg ⁻¹)	Mean Spectral Reflectance (%)									
			1p	1	2	3	4	4p	5p	5	7	
non-saline	0–1	0.72	9.6	11.6	17.2	21.3	25.7	26.5	30.5	31.4	28.8	
slight	1–2	1.50	9.1	10.8	15.7	19.0	25.3	26.5	30.5	29.9	25.2	
moderate	2–4	2.91	7.1	8.5	12.3	15.2	19.4	20.4	24.2	23.8	19.7	
high	>4	4.80	12.2	13.3	16.6	18.8	20.2	20.3	22.1	22.6	20.4	

Table 4. Correlation coefficients between soil salinity and ALI-convolved field spectra and coefficients between ALI-convolved field spectra.

SSC/Spectra	SSC	1p	1	2	3	4	4p	5p	5	7
SSC	1.000	--	--	--	--	--	--	--	--	--
1p	0.144	1.000	--	--	--	--	--	--	--	--
1	0.060	0.994	1.000	--	--	--	--	--	--	--
2	-0.134	0.937	0.966	1.000	--	--	--	--	--	--
3	-0.264	0.864	0.910	0.981	1.000	--	--	--	--	--
4	-0.477	0.694	0.755	0.858	0.899	1.000	--	--	--	--
4p	-0.516	0.631	0.697	0.800	0.851	0.992	1.000	--	--	--
5p	-0.582	0.520	0.595	0.712	0.785	0.955	0.981	1.000	--	--
5	-0.541	0.581	0.657	0.780	0.857	0.937	0.945	0.962	1.000	--
7	-0.469	0.642	0.711	0.822	0.891	0.865	0.852	0.859	0.958	1.000

3.3. Important Spectral Wavelengths for Soil Salinity Retrieval

Figure 4 shows regression coefficients for the PLSR model obtained from total soil samples and field spectra data. The coefficients were positive for blue and green bands (350–566 nm), and negative for red and NIR bands (567–1034 nm). The higher end wavelength of 1034 nm was determined because regression coefficient at this wavelength was less than 0.01. Regression coefficient was very close to zero for wavelengths within 1034–1451 nm, and positive for wavelengths within 1452–1810 nm. For longer wavelength, regression coefficient was negative for wavelength shorter than 1988 nm, positive within 1988–2116 nm, and negative for wavelength longer than 2116 nm. The regression coefficient varied slightly for those bands without water vapor absorption. For wavelengths near 1400 nm, 1900 nm and thereafter, the coefficient varied greatly even for neighboring wavelengths.

Marten's Uncertainty Test showed that the most important wavelengths for soil salinity retrieval were located at 350–549 nm, 695–1001 nm, 1475–1681 nm, 1941–1949 nm, 2015–2051 nm, 2077–2078 nm, 2169–2183 nm, 2241–2301 nm, 2385–2391 nm, 2399–2413 nm, 2417–2423 nm, 2433–2436 nm and 2446–2450 nm. Among these spectral ranges, NIR-SWIR band (695–1001 nm) had the widest spectral interval of 307 nm, followed by SWIR band (1475–1681 nm) of 207 nm, and blue-green band (350–549 nm) of 200 nm. The ALI bands 1p, 1, 4, 4p and 5 were the important predictors for soil salinity retrieval. The central wavelength of ALI band 7 was located between two clusters of important wavelengths. This band had important implications due to its wide spectral coverage (Figure 2).

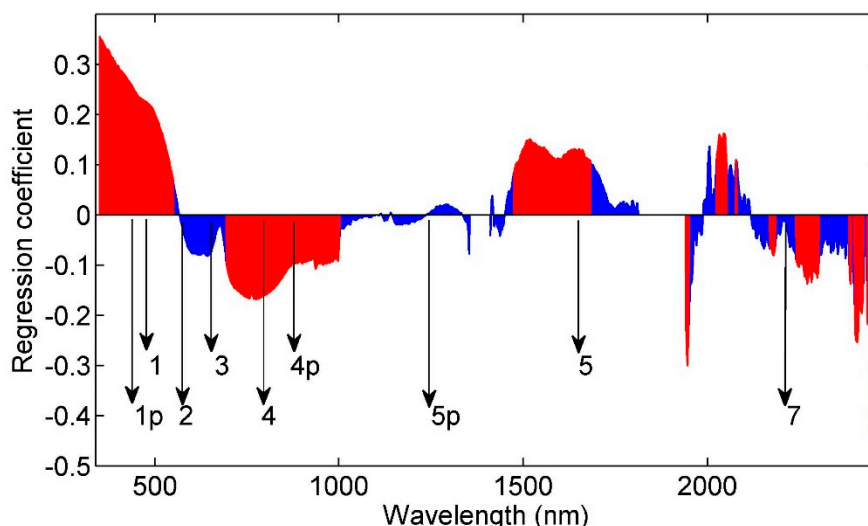


Figure 4. Regression coefficients for the PLSR model obtained from total soil samples and field spectra data. Important wavelengths for soil salinity retrieval are marked in red color. Central wavelengths of ALI bands were denoted with arrows and band numbers.

3.4. Partial Least Square Regression (PLSR) Model for Soil Salinity Retrieval

Measured soil salinity and ALI-convolved field spectra data yielded a linear model, described as follows:

$$SSC = 30.5 \times R_{1p} + 23.2 \times R_1 - 3.8 \times R_2 - 16.4 \times R_3 - 14.9 \times R_4 - 9.0 \times R_{4p} - 0.9 \times R_{5p} + 11.3 \times R_5 - 11.7 \times R_7 + 6.1 \quad (13)$$

where SSC denotes the modelled soil salinity, R denotes the ALI-convolved field spectra data, and subscripts 1p-7 denote sequential order of the ALI spectral bands.

Equation (13) highlighted the importance of bands 1p and 1 for soil salinity retrieval. Figure 5 compared measured and modelled soil salinity with collected samples. With respect to modelling samples, soil salinity values had a R^2 and RMSE of 0.797 and 0.779 $\text{g}\cdot\text{kg}^{-1}$. For validation samples, they were correlated with a R^2 and RMSE of 0.689 and 0.940 $\text{g}\cdot\text{kg}^{-1}$. In general, the model estimated soil salinity with a considerable accuracy; however, it may underestimate soil salinity according to less-than-unity slope values (0.748 and 0.757).

Table 5 provides statistics for evaluating the constructed salinity–reflectance relationship. For modelling samples, R^2 and RPD values were 0.749 and 3.584, indicating a moderate prediction ability as defined in [35]. Bias, SD and RMSE values were 0.036, 0.778 and 0.779, indicating a considerable accuracy. For validation samples, R^2 and RPD values were 0.689 and 2.838. The Bias, SD and RMSE values were -0.211 , 0.937 and 0.940. These values were comparable with those for modelling samples. For total samples, all these metric values lay between those for modelling and validation samples. R^2 and RPD values were 0.724 and 3.155. The Bias, SD and RMSE values were -0.048 , 0.842 and 0.837. The slope values were 0.748, 0.757 and 0.772 for modelling, validation and total samples, which may indicate underestimation of soil salinity.

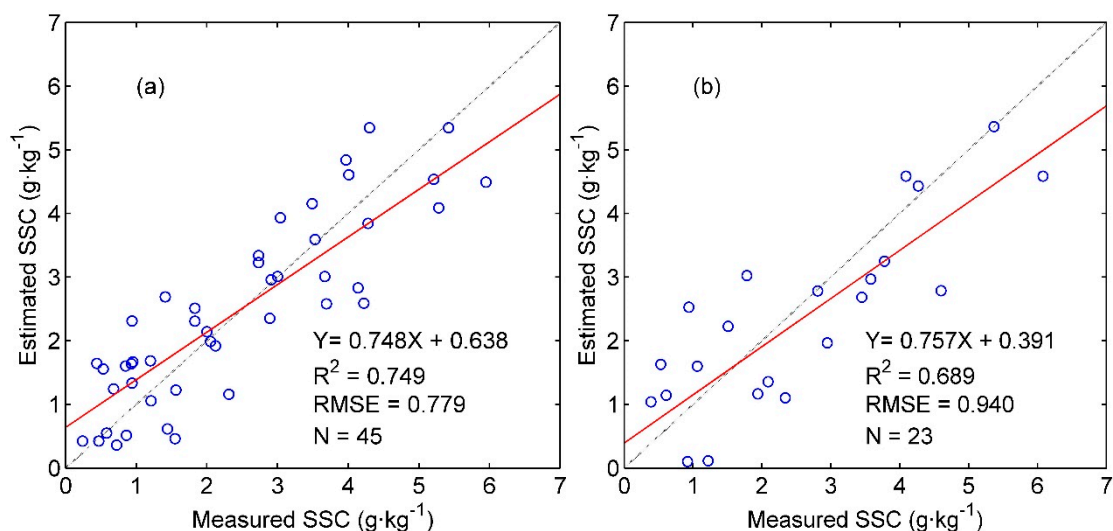


Figure 5. Comparison of measured and modelled soil salinity with modelling samples (a) and validation samples (b). The horizontal axis denotes measured soil salinity, and the vertical axis denotes modelled soil salinity. Slope and intercept, R^2 , RMSE and sample numbers are provided for the comparison.

Table 5. Performance assessment of the PLSR model. The statistics describes model performance for modelling samples, validation samples and total samples (modelling + validation), including R^2 , bias, SD (standard deviation), RMSE (root mean square error), ratio of prediction to deviation (RPD), number of samples (N), slope (a) and intercept (b).

Samples	R^2	RPD	Bias	SD	RMSE	N	Linear Regression	
							a	b
Modelling	0.749	3.584	0.036	0.778	0.779	45	0.748	0.638
Validation	0.689	2.838	-0.211	0.937	0.940	23	0.757	0.391
Total	0.724	3.155	-0.048	0.842	0.837	68	0.772	0.506

3.5. Salinity Mapping Using Partial Least Square Regression (PLSR) Model

Figure 6 shows the distribution of TOC, TPW and AOT interpolated from MODIS atmosphere products. TOC varied from 0.38–0.48 $\text{cm}\cdot\text{atm}^{-1}$, TPW from 0.74–1.49 $\text{g}\cdot\text{cm}^{-2}$, and AOT from 0.48–0.77. The mean values were 0.43 $\text{cm}\cdot\text{atm}^{-1}$, 1.08 $\text{g}\cdot\text{cm}^{-2}$ and 0.58. Comparison showed that R^2 and RMSE values were 0.733 and 0.01 $\text{cm}\cdot\text{atm}^{-1}$ for TOC, 0.840 and 0.07 $\text{g}\cdot\text{cm}^{-2}$ for TPW, and 0.951 and 0.02 for AOT. In a relative manner, the relative RMSE values were 2.87%, 6.41% and 3.58% for TOC, TPW and AOT, indicating a high accuracy of the interpolations. Figure 6a shows that ozone concentration was higher in the northern part. Figure 6b,c illustrates that the atmosphere was more humid and turbid in the southern part of study area. The southern part was close to urban area, which is the major cause of increased AOT values.

Figure 7 shows a synoptic soil salinity map over the study area and detailed salinity distributions over three typical agricultural lands. For the whole study area, soil salinity ranged between 0.00 and 9.95 $\text{g}\cdot\text{kg}^{-1}$, with a mean and SD of 1.98 $\text{g}\cdot\text{kg}^{-1}$ and 0.69 $\text{g}\cdot\text{kg}^{-1}$. In general, salinization was intensified from the inland toward the coastal area. The cultivated area along the Yellow River was slightly

affected by salinization, whereas the northern part (oil fields), the southern part (salt pans) and the eastern part (estuary) were moderately saline regions. Statistics showed that 10.5%, 36.4%, 53.0% and 0.1% of the entire area were covered with non-saline, slightly saline, moderately saline and highly saline soils.

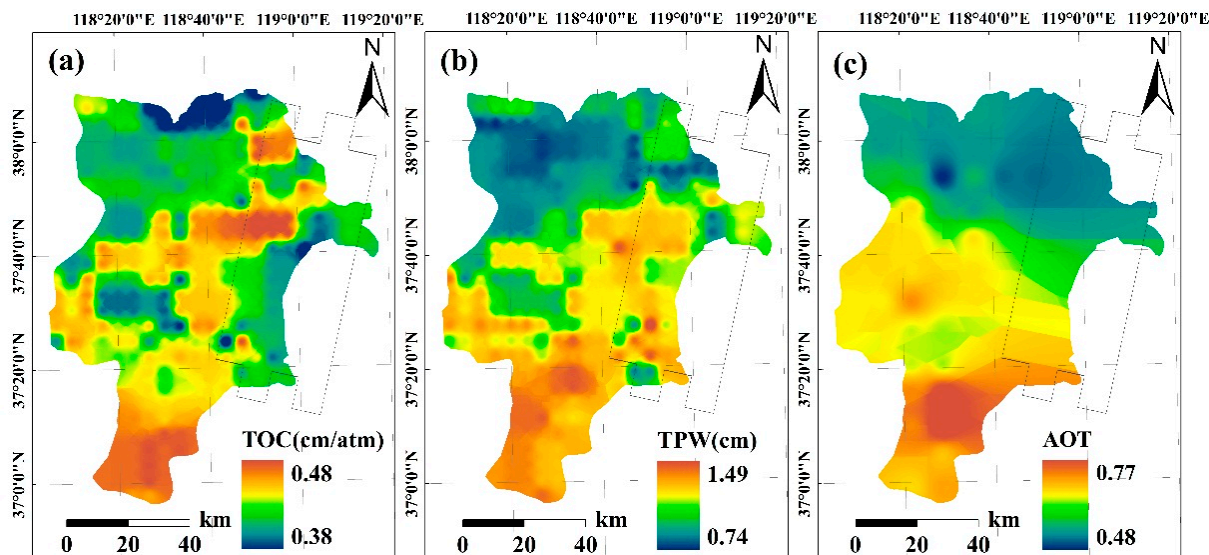


Figure 6. Spatial distribution of ozone concentration (a); water vapor content (b) and aerosol optical thickness at 550 nm (c) within the Yellow River Delta. The 30 m-resolution values were interpolated from MODIS atmosphere products.

Three agricultural lands were zoomed in for a detailed examination of soil salinity mapping. The land (a) was located in the northern part with slight-to-moderately saline soils; The land (b) was located along the Yellow River with non-saline and slightly saline soils; The land (c) was located in the southern part with moderate-to-highly saline soils. The detailed maps showed that cropland was likely consistent in soil salinity for all the $30\text{ m} \times 30\text{ m}$ pixels within the land. Moreover, salinity remained close to neighboring agricultural lands. Since cultivation practice was similar for agricultural lands, these retrievals were logical and reasonable.

3.6. Quantification of Error Related to Soil Salinity Retrieval

Figure 8a shows a scatterplot of measured soil salinity *versus* total retrieval errors. The total errors were highly correlated ($R^2 = 0.848$) with the measured soil salinity. The slope and intercept values were -0.701 and 0.823 , indicating an overall underestimate of soil salinity. For non-saline ($\text{SSC} < 1\text{ g}\cdot\text{kg}^{-1}$) and slightly saline ($1 < \text{SSC} < 2\text{ g}\cdot\text{kg}^{-1}$) soils, the errors were generally well within $1\text{ g}\cdot\text{kg}^{-1}$. For moderately ($2 < \text{SSC} < 4\text{ g}\cdot\text{kg}^{-1}$) and highly saline ($\text{SSC} > 4\text{ g}\cdot\text{kg}^{-1}$) soils, the errors increased linearly with soil salinity. The maximum error could be up to $-4\text{ g}\cdot\text{kg}^{-1}$. Figure 8b shows a weak correlation between measured soil salinity and model-related error. The error ranged between -2 and $2\text{ g}\cdot\text{kg}^{-1}$ with a fitting slope and intercept of -0.213 and 0.483 . This implied that the model might overestimate low salinity and underestimate high salinity. Figure 8c shows a moderate correlation ($R^2 = 0.467$) between measured soil salinity and data-related error, with a slope and intercept of -0.487 and 0.340 . The error increased with soil salinity, and the maximum error could be larger than

$-3 \text{ g}\cdot\text{kg}^{-1}$. In overall, both the model- and the data-related error resulted in the underestimation of high salinity. For $2 < \text{SSC} < 4 \text{ g}\cdot\text{kg}^{-1}$, the mean model- and data-related errors were $-0.21 \text{ g}\cdot\text{kg}^{-1}$ and $-1.06 \text{ g}\cdot\text{kg}^{-1}$. While for $\text{SSC} > 4 \text{ g}\cdot\text{kg}^{-1}$, the errors were $-0.48 \text{ g}\cdot\text{kg}^{-1}$ and $-2.10 \text{ g}\cdot\text{kg}^{-1}$. Thus, the data-related error accounted for 84% and 81% of the total error for moderately and highly saline soils.

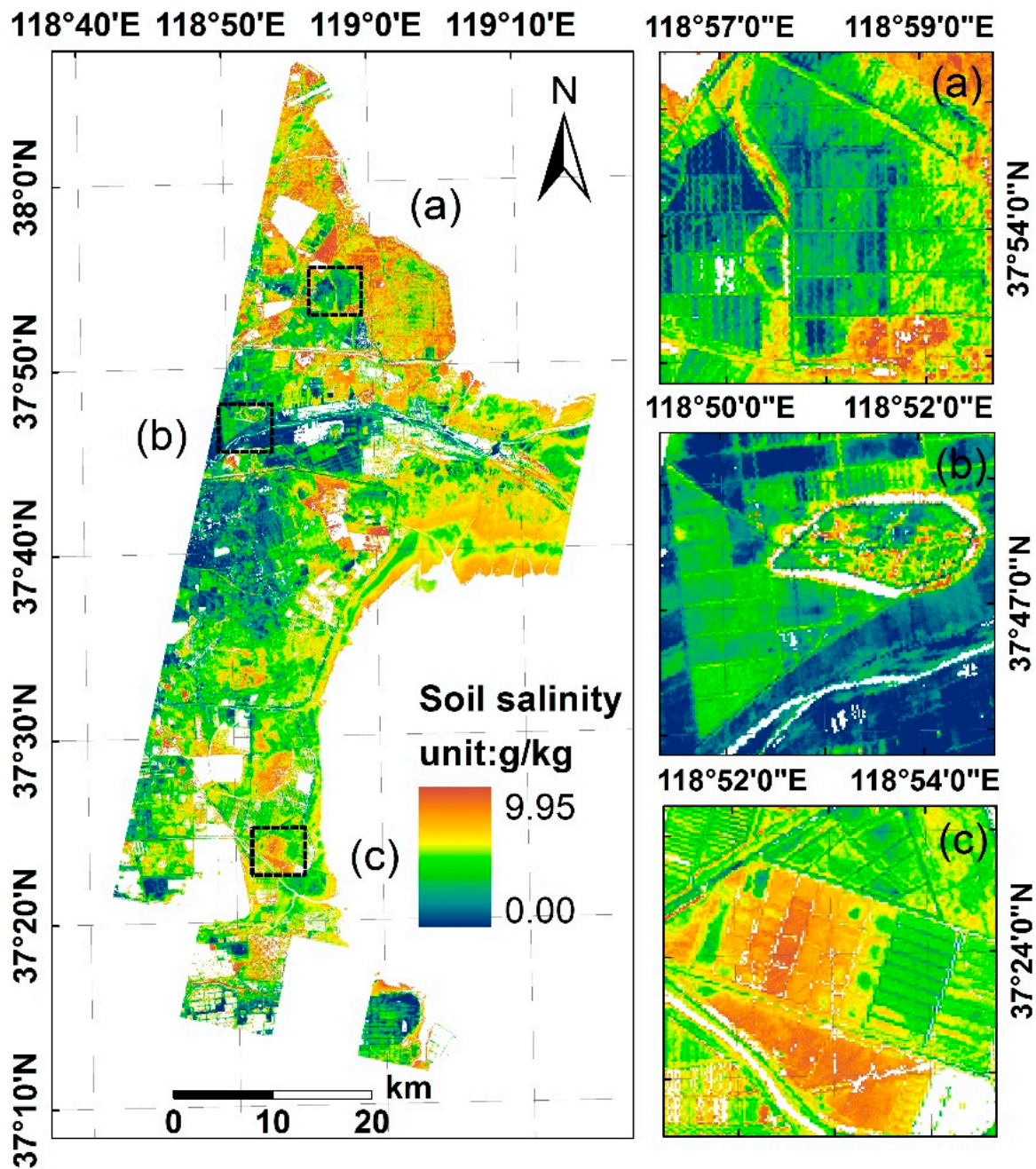


Figure 7. Soil salinity map over study area and detailed soil salinity distribution at three agricultural lands in the north with slight-moderately saline soils (a); along the Yellow River with non-saline and slightly saline soils (b) and in the south with moderate-to-highly saline soils (c).

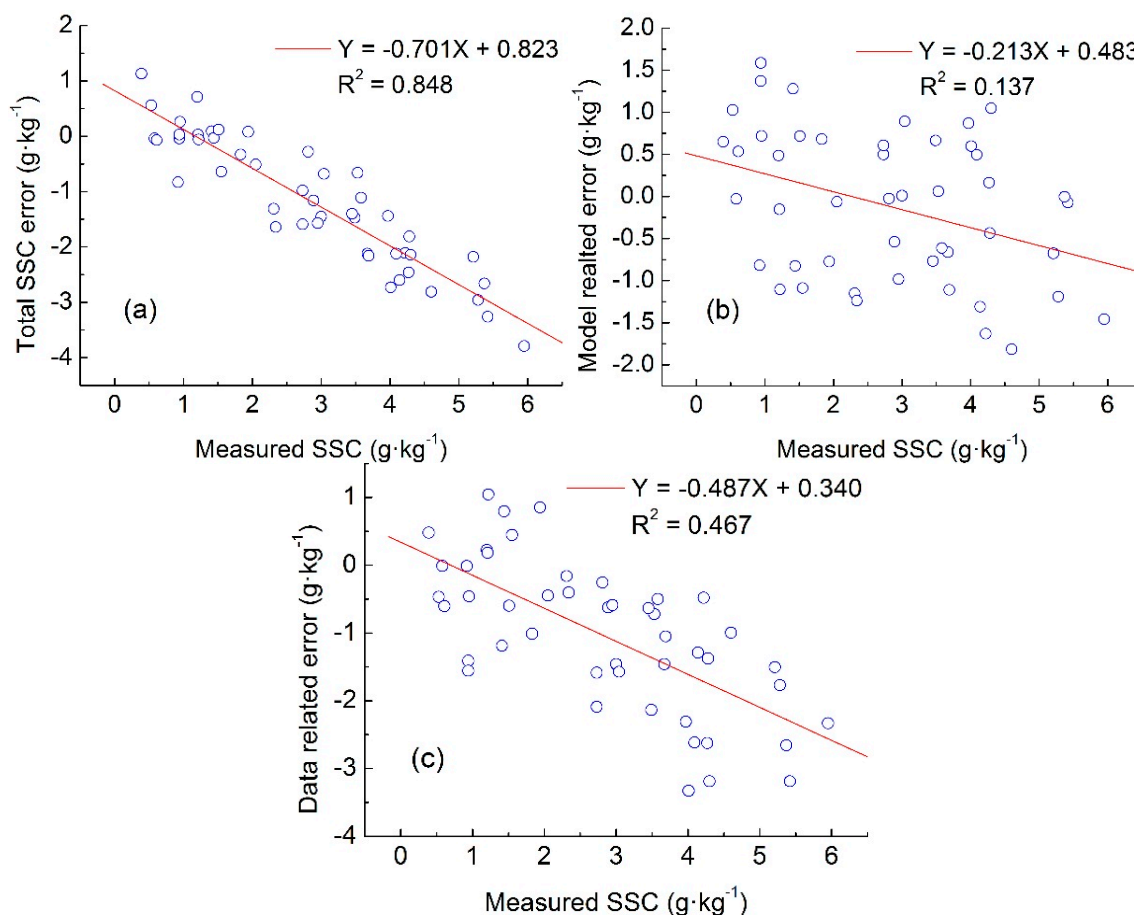


Figure 8. Scatterplots of measured soil salinity *versus* total retrieval error (a); model related error (b) and data related error (c). Slope and intercept values for linear regression and R^2 value are given for each scatterplot.

4. Discussion

4.1. Potential Use of Advanced Multi-Spectral Sensor for Soil Salinity Retrieval

Advanced multi-spectral sensors offer more opportunities for soil salinity quantification over a wide area of terrestrial surface. As for the ALI sensor, band 1p shows increased sensitivity to soil salinity than the blue band in Landsat series (Figure 2). Band 1p is of importance to the retrieval for its higher regression coefficient in the salinity–reflectance relationship. Figure 4 also demonstrates the importance of band 4p to soil salinity retrieval. Moreover, band 5p provides additional information (Table 4). These features constitute the fundamentals for soil salinity retrieval from the ALI multi-spectral data [30,48]. In general, multispectral remote sensing data owns comparable spatial resolution relative to satellite hyperspectral data (e.g., Hyperion), yet covers a wider area of terrestrial surface. The platforms enable repeated observations over a target area with a short revisit time. Nevertheless, insufficient spectral resolution (number of spectral band and spectral width) may be a major obstacle for its use in quantitative soil remote sensing. As demonstrated in this work, the key spectral bands of the new-generation satellite sensors may greatly improve soil salinity retrieval at a large scale from multi-spectral remote sensing data.

The ALI was developed for demonstrative validation of new technologies and strategies to improve Earth observations. It provides a prototype for Landsat 8 OLI. Relative to the ALI, the OLI retains bands 1p, 4p and 5p, while obsoletes band 4 (Table 1). Since spectral reflectance is highly correlated between bands 4 and 4p (0.992 as in Table 4), the absence of band 4 may have a minor impact on soil salinity retrieval. In addition, the OLI bands are spectrally comparative and are narrower than the corresponding ALI bands (Table 1). With a narrower spectral band pass, atmospheric absorption can be further reduced, and the OLI sensor has the potentials for worldwide soil salinity mapping.

4.2. Soil Salinity Retrieval from Multi-Spectral Sensor Data Based on Regression Models

Instead of all spectral bands, a majority of existing studies perform statistical analysis based on selected bands [25] and/or indices generated from band combinations. The commonly used indices can be generated from green/red [26], red/NIR [28], NIR/SWIR [29] and SWIR/SWIR bands [30]. This demonstrates from a different angle that the process of soil salinization may change soil surface reflectance at most spectral bands. Soil salinity variation can be detected with a single spectral band, dual-band or tri-band combination [26]. Nevertheless, these indices are generally weakly or moderately correlated with soil salinity. In most cases, the correlation coefficient may be less than 0.50 [25,26,30]. This paper incorporated all spectral bands for statistical analysis, and demonstrated the usefulness of the PLSR model for soil salinity retrieval with the ALI-convolved field spectra (Figure 6). The salinity–reflectance relationship was established with $R^2 > 0.7$ and RPD > 3 (Table 4). The bias was very close to zero, and the SD and the RMSE were less than $1.0 \text{ g}\cdot\text{kg}^{-1}$. These statistics demonstrated much improved retrieval accuracy with the linear regression models based on all-band satellite data.

4.3. Physical Explanations for Index Based Soil Salinity Estimation

Equation (13) not only shows a relationship for soil salinity estimation at YRD, but also implies physical explanations for index based soil salinity studies. The regression coefficients are -3.8 , -16.4 , -14.9 , 11.3 and -11.7 for green, red, NIR and two SWIR bands. Inferred from the coefficients, soil salinity may be insensitive to green band but sensitive to red band (coefficient of -3.8 vs. -16.4). This may be the reason why a combination of green and red bands has been extensively used for soil salinity mapping [26,28,31]. However, the sensitivity of soil salinity is similar to red and NIR bands (-16.4 vs. -14.9). This may partly account for why VIs can be used successfully in some studies whereas unsatisfactorily in other studies [26–28]. The general weak negative correlation between soil salinity and VIs may be contributable to the presence of vegetation in low saline soils. The coefficient is negative for NIR band yet positive for SWIR band 5 (-14.9 vs. 11.3). Moreover, the coefficient is positive for SWIR band 5 yet negative for SWIR band 7 (11.3 vs. -11.7). The contrasting coefficients may demonstrate the effectiveness of indices composed of NIR and SWIR bands [29] as well as two SWIR bands [30]. The use of dual-SWIR bands have been also demonstrated with hyperspectral data. By using Hyperion surface reflectance data, Weng *et al.* [33] have discovered a strong correlation between soil salinity and index constructed using reflectances at 2052 nm and 2203 nm.

In general, dual-band index can be composed from sensitive/insensitive bands (e.g., green and red bands), sensitive yet not contrastive bands (e.g., red and NIR bands) and contrasting bands (e.g., NIR

and SWIR bands, dual SWIR bands). The object of index is to highlight the information related to soil salinity whereas suppressing noise and background information. In this sense, our relationship can largely suppress unrelated information and thereby yield high accuracy.

4.4. Primary Uncertainty Associated with Soil Salinity Retrieval

Figure 8a shows a strong negative relationship between measured soil salinity and total retrieval error. It indicates that soil salinity was notably underestimated for moderately and highly saline soils. The underestimation is common in existing remote sensing studies. For example, Douaoui *et al.* [26] reported decreased correlation coefficients between soil salinity and several indices for highly saline soils. Farifteh *et al.* [35] also underestimated soil salinity for highly saline soils. The reason can be inferred from Figure 4 in which spectral reflectance varies differently for highly saline soils. Figure 8b reveals negative model-related errors for moderately and highly saline soils. However, the model-related error only accounted for a small proportion (<20%) of total retrieval error. The major error may result from data-related error as shown in Figure 8c.

Ben-Dor *et al.* [38] suggested that favorable atmospheric correction should be mandatory for imaging spectroscopy, whereas in multi-spectral remote sensing atmospheric effects are minor. However, quantification of the atmospheric effects is needed to verify this point. For multi-spectral remote sensing, atmospheric correction requires data of sensor calibration, ozone concentration, water vapor content and aerosol optical depth. Figure 9 illustrates a quantitative description for soil salinity uncertainty resulting from these parameters. The horizontal axes denote parameter and uncertainty, and the vertical axis denotes soil salinity uncertainty at different levels of the parameter and uncertainty. According to simulation results, sensor calibration, ozone, water vapor and aerosol may induce retrieval uncertainty of 0.7–0.9 g·kg⁻¹, <0.1 g·kg⁻¹, <0.2 g·kg⁻¹ and <1.0 g·kg⁻¹. Neglecting the sensor calibration uncertainty, the total retrieval error may not exceed 1.3 g·kg⁻¹. The uncertainty increases sharply at high TPW and/or AOT values for humid and turbid days. In most cases, remote sensing imageries are preferred in dry and clear days rather than in humid and turbid days. As a result, the total uncertainty of 1.3 g·kg⁻¹ is a conservative estimate. This also demonstrates that atmospheric correction is not the major source for retrieval error.

Data-related error may arise from scale difference between point sampling and remote sensing pixel. A high degree of surface heterogeneity may weaken the representativeness of point sampling for a 30 m × 30 m ALI pixel. The problem is one of the limitations for soil salinity mapping based on imaging spectroscopy [38]. As shown in Figure 8c, data-related error is moderately correlated with measured soil salinity. It was within 1 g·kg⁻¹ for non-saline and slightly saline soils, whereas increased with the degree of soil salinization for moderately and highly saline soils. Non-saline and slightly saline soils were generally distributed in agricultural lands (Figure 7). Uniform agricultural practice enabled similar spectral characteristics in neighboring agricultural lands. Thus, the sampling sites should be well representative within several hundred meters. However, moderately and highly saline soils unlikely had a homogeneous distribution within a large area. Salt crust and salt-tolerant vegetation may produce heterogeneous surfaces. As a result, a comprehensive field campaign and validation are necessary for quantitative remote sensing.

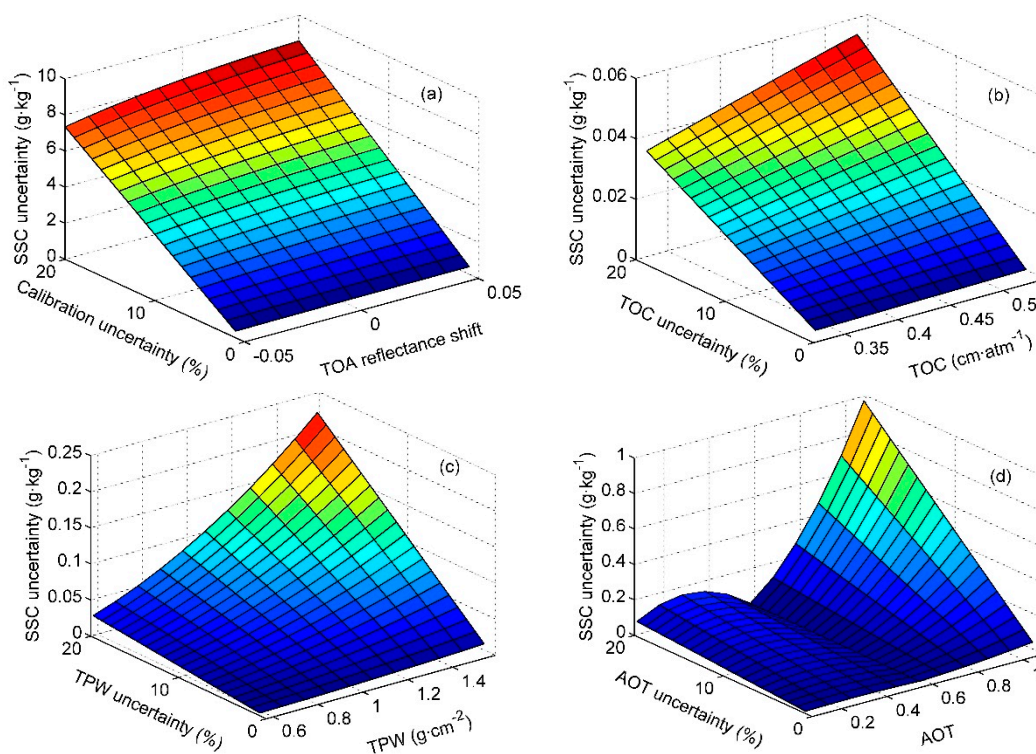


Figure 9. Simulated uncertainty in soil salinity retrievals due to ALI sensor calibration (a); ozone data (b); water vapor data (c); and aerosol data (d). The horizontal axes denote parameter and uncertainty, and the vertical axis denotes retrieval error at a given parameter and uncertainty.

5. Conclusions

Soil salinization has been a global concern, triggering secondary soil degradation and posing great threats to sustainable development. Timely detection and early warning of soil salinization is in urgent need. In this context, multi-spectral remote sensing data have been extensively used for soil salinity mapping. However, the coarse spectral resolution poses a major obstacle for quantitative retrieval. In virtue of improved spectral resolution of the ALI sensor, this paper applied a PLSR model to construct a relationship between measured soil salinity and ALI-convolved field spectra, with an examination in the Yellow River Delta, China. The model estimated soil salinity with a R^2 and RPD of 0.749 and 3.584. The estimates were almost unbiased in terms of SD and RMSE less than $1.0 \text{ g}\cdot\text{kg}^{-1}$. The results demonstrated the usefulness of the ALI bands 1p, 4p and potentially band 5p in mapping soil salinity. Landsat 8 OLI would be an important candidate for soil salinity mapping at a large scale. The PLSR model was applied to mapping soil salinity from atmospherically corrected ALI data. Error analysis showed that soil salinity was relatively accurate for low saline soils, and underestimated for moderately and highly saline soils. The underestimation may result from the poor representativeness of soil sampling, not from the model ineffectiveness or atmospheric correction.

Our study provides comprehensive spectral analysis based on ALI-convolved field spectra. It confirms the effectiveness of ALI bands in soil salinity retrieval. We also demonstrate the potential use of regression models for retrieving salinity from advanced multi-spectral sensors, and explain the

implications of our study for index based soil salinity estimation. The results and implications would be very valuable for inexpensive but accurate soil salinity mapping at a large scale.

Acknowledgments

We would like to express our thanks to those at School of Geographic and Oceanic Sciences, Nanjing University for their help in field campaign. We also thank staffs with Institute of Soil Sciences, Chinese Academy of Sciences (ISSAS) for their assistances in physio-chemical analysis of soil samples. This work was in part supported by Natural Science Foundation of China (NSFC) under Grant 41171272 and in part by another NSFC under Grant 41471352. We highly appreciate the anonymous reviewers for their constructive comments on an early draft of the manuscript.

Author Contributions

The first author Xingwang Fan proposed the main idea, and was responsible for paper writing. The second author Yuanbo Liu provided theoretical guidance on radiometric and atmospheric correction, and offered invaluable suggestions for paper revision. The third author Jinmei Tao made careful analysis of field spectra data and physio-chemical data. The last author Yongling Weng provided the original idea for soil salinity retrieval using multi-spectral data.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Greenland, D.J. Soil management and soil degradation. *Eur. J. Soil Sci.* **1981**, *32*, 301–322.
2. Rengasamy, P. World salinization with emphasis on Australia. *J. Exp. Bot.* **2006**, *57*, 1017–1023.
3. Lambers, H. Introduction, dryland salinity: A key environmental issue in southern Australia. *Plant Soil* **2003**, *257*, 5–7.
4. Thomas, D.S.G.; Middleton, N.J. Salinization: New perspectives on a major desertification issue. *J. Arid Environ.* **1993**, *24*, 95–105.
5. Metternicht, G.I.; Zinck, J.A. Remote sensing of soil salinity: Potentials and constraints. *Remote Sens. Environ.* **2003**, *85*, 1–20.
6. Childs, S.W.; Hanks, R.J. Model of soil salinity effects on crop growth. *Soil Sci. Soc. Am. J.* **1975**, *39*, 617–622.
7. Rozema, J.; Flowers, T. Crops for a salinized world. *Science* **2008**, *322*, 1478–1480.
8. Jardine, A.; Speldewinde, P.; Carver, S.; Weinstein, P. Dryland salinity and ecosystem distress syndrome: Human health implications. *EcoHealth* **2007**, *4*, 10–17.
9. Corwin, D.L.; Lesch, S.M. Apparent soil electrical conductivity measurements in agriculture. *Comput. Electron. Agric.* **2005**, *46*, 11–43.
10. Triantafilis, J.; Odeh, I.O.A.; McBratney, A.B. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. *Soil Sci. Soc. Am. J.* **2001**, *65*, 869–878.

11. Farifteh, J.; Farshad, A.; George, R.J. Assessing salt-affected soils using remote sensing, solute modelling, and geophysics. *Geoderma* **2006**, *130*, 191–206.
12. Mulder, V.L.; de Bruin, S.; Schaepman, M.E.; Mayr, T.R. The use of remote sensing in soil and terrain mapping—A review. *Geoderma* **2011**, *162*, 1–19.
13. Barnes, E.M.; Sudduth, K.A.; Hummel, J.W.; Lesch, S.M.; Corwin, D.L.; Yang, C.H.; Daughtry, C.S.T.; Bausch, W.C. Remote- and ground-based sensor techniques to map soil properties. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 619–630.
14. Dwivedi, R.S.; Rao, B.R.M. The selection of the best possible Landsat TM band combination for delineating salt-affected soils. *Int. J. Remote Sens.* **1992**, *13*, 2051–2058.
15. Howari, F.M. The use of remote sensing data to extract information from agricultural land with emphasis on soil salinity. *Aust. J. Soil Res.* **2003**, *41*, 1243–1253.
16. Ungar, S.G.; Pearlman, J.S.; Mendenhall, J.A.; Reuter, D. Overview of the earth observing one (EO-1) mission. *IEEE Trans. Geosci. Remote* **2003**, *41*, 1149–1159.
17. Roy, D.; Wulder, M.; Loveland, T.; Allen, R.; Anderson, M.; Helder, D.; Irons, J.; Johnson, D.; Kennedy, R.; Scambos, T. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172.
18. Mougnot, B.; Pouget, M.; Epema, G.F. Remote sensing of salt affected soils. *Remote Sens. Rev.* **1993**, *7*, 241–259.
19. Golovina, N.; Minskiy, D.Y.; Pankova, Y.I.; Solov'yev, D. Automated air photo interpretation in the mapping of soil salinization in cotton-growing zones. *Mapp. Sci. Rem. Sens.* **1992**, *29*, 262–268.
20. Dwivedi, R.S.; Sreenivas, K.; Ramana, K.V. Inventory of salt-affected soils and waterlogged areas: A remote sensing approach. *Int. J. Remote Sens.* **1999**, *20*, 1589–1599.
21. Guan, Y.X.; Liu, G.H.; Liu, Q.S.; Ye, Q.H. The study of salt-affected soils in the Yellow River delta based on remote sensing. *J. Remote Sens.* **2001**, *5*, 46–52. (in Chinese)
22. Elnaggar, A.A.; Noller, J.S. Application of remote-sensing data and decision-tree analysis to mapping salt-affected soils over large areas. *Remote Sens.* **2009**, *2*, 151–165.
23. Fernandez-Buces, N.; Siebe, C.; Cram, S.; Palacio, J.L. Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. *J. Arid Environ.* **2006**, *65*, 644–667.
24. Dwivedi, R.S.; Sreenivas, K. Image transforms as a tool for the study of soil salinity and 33 alkalinity dynamics. *Int. J. Remote Sens.* **1998**, *19*, 605–619.
25. Shrestha, R.P. Relating soil electrical conductivity to remote sensing and other soil properties for assessing soil salinity in northeast Thailand. *Land Degrad. Dev.* **2006**, *17*, 677–689.
26. Douaoui, A.E.K.; Nicolas, H.; Walter, C. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* **2006**, *134*, 217–230.
27. Aldakheel, Y.Y. Assessing NDVI spatial pattern as related to irrigation and soil salinity management in Al-Hassa Oasis, Saudi Arabia. *J. Indian Soc. Remote* **2011**, *39*, 171–180.
28. 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.

29. Madani, A.A. Soil salinity detection and monitoring using Landsat data: A case study from Siwa Oasis, Egypt. *Gisci. Remote Sens.* **2005**, *42*, 171–181.
30. Bannari, A.; Guedon, A.M.; El-Harti, A.; Cherkaoui, F.Z.; El-Ghmari, A. Characterization of slightly and moderately saline and sodic soils in irrigated agricultural land using simulated data of Advanced Land Imaging (EO-1) sensor. *Commun. Soil Sci. Plan.* **2008**, *39*, 2795–2811.
31. Odeh, I.O.A.; Onus, A. Spatial analysis of soil salinity and soil structural stability in a semiarid region of New South Wales, Australia. *Environ. Manag.* **2008**, *42*, 265–278.
32. Goetz, A.F.H. Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sens. Environ.* **2009**, *113*, S5–S16.
33. Weng, Y.L.; Gong, P.; Zhu, Z.L. Reflectance spectroscopy for the assessment of soil salt content in soils of the Yellow River Delta of China. *Int. J. Remote Sens.* **2008**, *29*, 5511–5531.
34. 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.
35. Farifteh, J.; van der Meer, F.D.; 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.
36. Weng, Y.L.; Gong, P.; Zhu, Z.L. Soil salt content estimation in the Yellow River delta with satellite hyperspectral data. *Can. J. Rem. Sens.* **2008**, *34*, 259–270.
37. Ben-Dor, E.; Taylor, R.G.; Hill, J.; Dematte, J.A.M.; Whiting, M.L.; Chabrillat, S.; Sommer, S. Imaging spectrometry for soil applications. *Adv. Agron.* **2008**, *97*, 321–392.
38. Ben-Dor, E.; Chabrillat, S.; Dematte, J.A.M.; Taylor, G.R.; Hill, J.; Whiting, M.L.; Sommer, S. Using imaging spectroscopy to study soil properties. *Remote Sens. Environ.* **2009**, *113*, S38–S55.
39. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemometr. Intell. Lab.* **2001**, *58*, 109–130.
40. Goldshleger, N.; Livne, I.; Chudnovsky, A.; Ben-Dor, E. New results in integrating passive and active remote sensing methods to assess soil salinity: A case study from Jezre'el Valley, Israel. *Soil Sci.* **2012**, *177*, 392–401.
41. Goldshleger, N.; Ben-Dor, E.; Lugassi, R.; Eshel, G. Soil degradation monitoring by remote sensing: Examples with three degradation processes. *Soil Sci. Soc. Am. J.* **2010**, *74*, 1433–1445.
42. Ben-Dor, E.; Metternicht G.; Goldshleger N.; Mor E.; Mirlas V.; Basson U. Review of remote sensing-based methods to assess soil salinity. In *Remote Sensing of Soil Salinization: Impact on Land Management*; Metternicht, G., Zinck, J.A., Eds.; CRC Press: Boca Raton, FL, USA, 2008; pp. 39–60.
43. Lass, L.W.; Prather, T.S.; Glenn, N.F.; Weber, K.T.; Mundt, J.T.; Pettingill, J. A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and babysbreath (*Gypsophila paniculata*) with a hyperspectral sensor. *Weed Sci.* **2005**, *53*, 242–251.
44. Moreira, L.C.J.; dos Santos Teixeira, A.; Galvão, L.S. Laboratory salinization of Brazilian alluvial soils and the spectral effects of gypsum. *Remote Sens.* **2014**, *6*, 2647–2663.
45. Pang, G.; Wang, T.; Liao, J.; Li, S. Quantitative model based on field-derived spectral characteristics to estimate soil salinity in Minqin County, China. *Soil Sci. Soc. Am. J.* **2014**, *78*, 546–555.

46. Hick, P.T.; Russell, W.G.R. Some spectral considerations for remote sensing of soil salinity. *Aust. J. Soil Res.* **1990**, *28*, 417–431.
47. Fan, X.; Pedroli, B.; Liu, G.; Liu, Q.; Liu, H.; Shu, L. Soil salinity development in the yellow river delta in relation to groundwater dynamics. *Land Degrad. Dev.* **2012**, *23*, 175–189.
48. Tao, J.M.; Fan, X.W.; Weng, Y.L. Soil salt content retrieval from ALI multi-spectral image based on GRNN (in Chinese). *Mod. Surv. Mapp.* **2012**, *35*, 10–12.
49. Yao, R.J.; Yang, J.S.; Liu, G.M.; Zou, P. Spatial variability of soil salinity in characteristic field of the Yellow River Delta. *Trans. Chin. Soc. Agric. Eng.* **2006**, *22*, 61–66. (in Chinese)
50. Weng, Y.L.; Gong, P.; Zhu, Z.L. A spectral index for estimating soil salinity in the Yellow River delta region of China using EO-1 Hyperion data. *Pedosphere* **2010**, *20*, 378–388.
51. Fang, H.; Liu, G.; Kearney, M. Georelational analysis of soil type, soil salt content, landform, and land use in the Yellow River Delta, China. *Environ. Manag.* **2005**, *35*, 72–83.
52. USGS EO-1 Website. Available online: <http://eo1.usgs.gov/> (accessed on 4 January 2015).
53. Chander, G.; Markham, B.L.; Helder, D.L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* **2009**, *113*, 893–903.
54. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409.
55. Goddard Space Flight Center. Available online: <http://ladsweb.nascom.nasa.gov/> (accessed on 4 January 2015).
56. Milton, E.J.; Schaepman, M.E.; Anderson, K.; Kneubuhler, M.; Fox, N. Progress in field spectroscopy. *Remote Sens. Environ.* **2009**, *113*, S92–S109.
57. Van Reeuwijl, L.P. *Procedures for Soil Analysis*, 3rd ed.; International Soil Reference and Information Centre (ISRIC): Wageningen, The Netherlands, 1992.
58. Wang, Z.Q.; You, W.R.; Zhu, S.Q. *Saline-Alkali Soil of China*; Science Publishing: Beijing, China, 1993; pp. 312–390. (in Chinese)
59. China Meteorological Administration. Available online: <http://www.cma.gov.cn/> (accessed on 4 January 2015).
60. Fan, X.W.; Liu, Y.B. Quantifying relationship between inter-sensor images in solar reflective bands: Implications for intercalibration. *IEEE Trans. Geosci. Remote* **2014**, *52*, 7727–7737.
61. Ju, J.C.; Roy, D.P.; Vermote, E.; Masek, J.; Kovalsky, V. Continental-scale validation of MODIS-based and LEDAPS Landsat ETM plus atmospheric correction methods. *Remote Sens. Environ.* **2012**, *122*, 175–184.
62. Martens, H.; Martens, M. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual. Prefer.* **2000**, *11*, 5–16.
63. Unscrambler 9.7 Software. Available online: <http://www.camo.com/> (accessed on 4 January 2015).
64. Nagol, J.R.; Vermote, E.F.; Prince, S.D. Effects of atmospheric variation on AVHRR NDVI data. *Remote Sens. Environ.* **2009**, *113*, 392–397.