Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China

Jingzhe Wang ^{a,b}, Jianli Ding ^{a,b,*}, Danlin Yu ^{c,d*}, Xuankai Ma ^a, Zipeng Zhang ^{a,b}, Xiangyu Ge ^{a,b}, Dexiong Teng ^{a,b}, Xiaohang Li ^{a,b}, Jing Liang ^{a,b}, Ivan Lizaga ^e, Xiangyue Chen ^{a,b}, Lin Yuan ^f, Yahui Guo ^g

^a Key Laboratory of Smart City and Environment Modelling of Higher Education Institute, College of Resources and Environment Science, Xinjiang University, Urumqi 800046, China

^b Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi 830046, China

^c School of Sociology and Population Studies, Renmin University of China, Beijing, 100872, China

^d Department of Earth and Environmental Studies, Montclair State University, Montclair, NJ, 07043, USA

^e Estación Experimental de Aula Dei, Spanish National Research Council (EEAD-CSIC), Avenida Montañana, 1005, 50059 Zaragoza, Spain

^f School of Architecture, Harbin Institute of Technology (Shenzhen), Shenzhen, 518055, China

^g Beijing Key Laboratory of Urban Hydrological Cycle and Sponge City Technology, College of Water Sciences, Beijing Normal University, Beijing 100875, China

Abstract: Soil salinization is one of the most important causes for land degradation and desertification and is an important threat to land management, farming activities, water quality, and sustainable development in arid and semi-arid areas. Soil salinization is often characterized with significant spatiotemporal dynamics. The salt-affected soil is predominant in the Ebinur Lake region in the Northwestern China. However, detailed local soil salinity information is ambiguous at the best due to limited monitoring techniques. Nowadays, the availability of Multi-Spectral Instrument (MSI) onboard Sentinel-2, offers unprecedented perspectives for the monitoring and mapping of soil salinity. The use of MSI data is an innovative attempt for salinity detection in arid land. We hypothesize that field observations and MSI data and MSI data-derived spectral indices using the partial least square regression (PLSR) approach will yield fairly accurate regional salinity map. Based on electrical conductivity of 1:5 soil:water extract (EC) of 72 ground-truth measurements (out of 116 sample sites)

and various spectral parameters, such as satellite band reflectance, published satellite salinity indices, red-edge indices, newly constructed two-band indices, and three-band indices from MSI data, we built a few inversion models in an attempt to produce the regional salinity maps. Different algorithms including Pearson correlation coefficient method (PCC), variable importance in projection (VIP), Gray relational analysis (GRA), and random forest (RF) were applied for variable selection. The results suggest that both the newly proposed normalized difference index (NDI) $[(B12-B7)/(B12+B7)]$ and three-band index (TBI4) $[(B12-B3)/(B3-B11)]$ show a better correlation with validation data and could be applied to estimate the soil salinity in the Ebinur Lake region. The established models were validated using the remaining 44 independent ground-based measurements. The RF-PLSR model performed the best across the five models with R^2_v , $RMSE_v$, and RPD of 0.92, 7.58 dS m⁻¹, and 2.36, respectively. The result from this model was then used to map the soil salinity over the study area. Our analyses suggest that soil salinization changes quite significantly in different seasons. Specifically, soil salinity in the dry season was higher than in the wet season, mostly in the lake area and nearby shores. We contend that the results from the study will be useful for soil salinization monitoring and land reclamation in arid or semi-arid regions outside the current study area.

Key words: Sentinel-2; Soil salinity; Red-edge; Spectral indices; Remote sensing

1. Introduction

Soil salinization is a serious global issue, being a considerable threat for the limited soil resources and ecosystem health in arid and semi-arid regions ([Metternicht and Zinck, 2003](#); [Nawar et al., 2014](#)). It is also one of the most important causes of land desertification and deterioration. Increasing salinity in soil has posed serious threat to farming activities, vegetation growth, bio-

diversity and sustainable development ([Aldabaa et al., 2015](#)). It is estimated that about 950 million ha of land are subjected to soil salinization at the global scale, and the affected area is expanding at a speed of two million ha yr⁻¹ ([Abbas et al., 2013](#); [Peng et al., 2019](#)). Apart from original high-water soluble salt content, soil resource in arid and semi-arid areas is also under severe threat of secondary salinization. This process is mainly a result of the great contrast of hydrologic cycle factors (scarce precipitation and intense evaporation), shallow groundwater levels, and irrational farming activities such as extensive agricultural operations and flood irrigation works ([Douaoui et al., 2006](#); [Ma et al., 2018](#)). Soil salinity is an effective evaluation indicator for soil salinization and varies extensively in temporal and spatial dimension ([Ding and Yu, 2014](#); [Gorji et al., 2015](#)). This suggests that only dynamic monitoring can provide sufficient understanding of the basis of current soil salinization situation so that effective soil restoration and land reclamation are possible ([Bannari et al., 2018](#)).

In practice, soil electric conductivity (EC) is often used to dynamically monitor soil salinization because of their strong correlation ([Peng et al., 2014](#)). It is found that traditional analyses of soil electrical conductivity (EC) are highly accurate, but are slow, discontinuous, financially expensive, and limited in space. In their study, [Ding and Yu \(2014\)](#) propose that remote sensing (RS) technology is a promising method for cost-effective monitoring and mapping of different soil properties in remote regions and environmental issues such as soil salinity. Over the past two decades, RS has demonstrated its ability in soil salinity monitoring from surface characteristics at various scales ([Allbed et al., 2014b](#); [Bell et al., 2001](#); [Kobayashi et al., 2015](#); [Verma et al., 1994](#); [Wang et al., 2018b](#)). In terms of the types of platform, they consist of four categories: multispectral sensor, hyperspectral sensor, synthetic aperture radar (SAR), and ground-based visible and near-infrared (VIS–NIR) instrument. Due to the capability of wide view field, real-time information acquisition, and periodic

surface coverage, multispectral RS makes monitoring and mapping soil salinity easier, especially in large-scale ([Fernández-Buces et al., 2006](#); [Lizaga et al., 2019b](#); [Wu et al., 2014](#)). Satellite data collected from various multispectral sensors, such as moderate resolution imaging spectroradiometer (MODIS), Landsat series, IKONOS, HuanJing satellite constellation-1 (HJ-1) and GaoFen-1 (GF-1) have been applied to extract salinization information at different spatial resolutions across the world. Using multivariate analysis, we are able to model the potential relationships that exist between RS data and salt content in soils. The combinations of multivariate regression and various multi spectral RS data can effectively monitor and map soil salinity ([Allbed et al., 2014a](#); [Bai et al., 2018](#); [Cao et al., 2016](#); [Farifteh et al., 2007](#); [Lobell et al., 2010](#); [Meng et al., 2016](#); [Wang et al., 2018b](#); [Zhang et al., 2015](#)). However, accuracy of multispectral analysis is often limited primarily because of large pixel size and broad bandwidths ([Davis et al., 2019](#)). The launch of Sentinel-2 satellite equipped with Multi-Spectral Instrument (MSI) in 2015 is an essential part of Global Monitoring for Environment and Security (GMES), which aimed at providing geo-information time series data ([Drusch et al., 2012](#); [Malenovsky et al., 2012](#)). The new generation on-board MSI measures the Earth's reflected radiance over 13 spectral bands from 443 nm to 2190 nm (covering VIS-NIR and the shortwave-infrared bands). It has a 10-day revisiting period cycle and a wide swath of 290 km ([ESA, 2015](#)). The delivered images have fine resolution (up to 10 m), novel spectral capabilities (i.e., three red-edge bands and two SWIR bands), offering unprecedented perspectives for the detection of useful information for a variety of land applications ([Lizaga et al., 2019a](#); [Sibanda et al., 2015](#); [Yang et al., 2018](#)).

Remote sensors use the electromagnetic energy reflected from the ground targets to collect information about saline soils ([Peng et al., 2019](#)). The spectral index is a simple composition of

different wavebands that can be used to establish the correlation between spectral data and specific targets. By generating optimal spectral indices using various combination of wavebands, we could detect more feature wavelengths and further enhance the correlations between specific properties and spectral characteristics of a target ([Ge et al., 2019](#)). Because of different spectral behavior in the Blue (B), Green (G), Red (R), and near infrared (NIR) spectrum of salinized soils on satellite images, various spectral indices were developed to detect and map soil salinity. These indices include Intensity Index (Int 1-2), normalized differential salinity index (NDSI), brightness index (BI), and salinity index (SI 1-5) ([Douaoui et al., 2006](#); [Khan et al., 2005](#)). The spectral soil salinity indices are relatively easy to calculate and reliable in certain conditions, which presents a scientific opportunity for soil salinity research ([Wang et al., 2018b](#); [Zhang et al., 2011](#)). Significant progress have been made in the field of soil salinity monitoring and mapping via the combination of various spectral soil salinity indices ([Allbed et al., 2014a](#); [Khan et al., 2005](#); [Triki Fourati et al., 2015](#); [Wang et al., 2018b](#)). However, these indices only consider the electromagnetic spectrum in the visible (VIS) and near-infrared (NIR) bands instead of the shortwave infrared (SWIR) and red-edge spectral bands. [Bannari et al. \(2018\)](#) applied the VIS-NIR and SWIR spectral bands of the Sentinel 2-MSI for soil salinity detection in the Kingdom of Bahrain, showing the potential of this data for soil salinity discrimination. The introduction of red-edge bands and SWIR bands provides the opportunity for improved characterization of soil salinity. It is predicted that the application of Sentinel-2 MSI (S2) data will greatly improve the estimation of soil salinity at regional scale.

Fig. 1 illustrates the soil salinity map of the world and Xinjiang Uyghur Autonomous Region (XUAR) of northwestern China. This is one of the severely salinized regions in China. Ebinur Lake is a representative ecological degradation region in Junggar basin, XUAR ([Ge et al., 2016](#); [Wang et](#)

al., 2018a). Currently, Ebinur Lake region has become a fragile environment threatened by severe soil salinization (Wang et al., 2017b). Local soil salinity has been examined with various RS data previously, but most of these studies have focused on a single period (Qing et al., 2005; Wang et al., 2011) instead of different seasons. Acquiring soil salinity maps in different seasons is crucial to understand the major mechanism underlying soil salinization and degradation.

To fill in this gap, we try to answer the following questions: Is the S2 data appropriate for dynamic soil salinity monitoring in the Ebinur Lake region? If so, how can we improve the estimating performance using extra spectral bands? What is the difference of soil salinity in different periods in the arid and semi-arid regions? These questions are unavoidable in RS-based monitoring of soil salinity and their answers are the primary objectives of the current study.

Figure 1 Soil salinity map of the world (figure reprinted from (Gorji et al., 2017))

Specifically, this study aims to (1) analyze the relationship between band information of the S2 data and soil salinity; (2) develop new S2 salinity indices and validate their capabilities for dynamic salinity monitoring; (3) construct an optimal information retrieving model by comparing the estimating accuracies of PLSR models based on various scenarios of input spectral covariates; and (4) monitor and map soil salinity in the dry and wet seasons in the Ebinur Lake region.

2. Materials and methods

2.1 Study Area

Ebinur Lake is the largest saltwater lake in Xinjiang Uyghur Autonomous Region (XUAR) ([Wang et al., 2019](#)). The study area is located in the southwestern region of the Junggar Basin (44°54' to 45°08' N and 82°35' to 83°10'E) and is mostly composed of salt-rich wetlands. The study area is situated at the lowest part of Ebinur Lake Catchment at 189 meters above sea level (m.a.s.l). Ebinur Lake region is characterized by a typical continental arid climate with scarce precipitation (< 150 mm/yr) and intense evaporation (> 2000 mm/yr). The region also has strong and consistent winds with speeds >8 m/s occurring on average 165 days per year. Based on the precipitation records and surface water quantity in this region, May and August are the beginning of the wet and dry seasons, respectively ([Ge et al., 2016](#)). The main land use and land cover (LU/LC) in the study area include water body, wetland (rare vegetation), desert, salinized land, and other types. In the dry and wet seasons, the differences of Ebinur Lake water surface area are considerable ([Ma et al., 2007](#)). Local prevalent soils types are mainly desert soil consisting of piedmont psephitic, salt, and gypsum. The soil texture is coarse with sand and sandy loam as the dominant types. Affected by the low and flat terrain, specific climatic features, and shallow groundwater level, soil salinity in the study area is severe with NaCl as the prevalent local salt minerals ([Wang et al., 2018a](#)). Under such conditions, most of the natural vegetation is halophytes, e.g., *Tamarix spp.*, *Haloxylon spp.*, *Halocnemum caspica*, and *Kalidium foliatum*. The vegetation cover is low, which considerably facilitates the estimation of soil salinity through remote sensing images.

Figure 2 Distribution of sampling sites in the study area.

2.2 Data collection

2.2.1 Field sampling and analysis

Considering the LU/LC, soil types, soil surface characteristics, previous field sampling experience, and accessibility to the potential sampling sites, a total of 146 representative sampling points (units) were selected with the same spatial resolution as the Sentinel-2 VIS-NIR bands (10m). In each sampling point, four samples were collected using a shovel and mixed in the field to create a representative composite sample (Figure 2). A portable GPS (Garmin GPS 72, accuracy < 10 m) was used to record the location of each sample. Soil samples were placed into sealed watertight bags and labelled for further chemical analysis. Field work was conducted on May ($n = 65$) and August ($n = 51$) 2018, during the wet and dry seasons, respectively. All samples were air-dried, ground, homogenised and sieved to ≤ 2 mm. For the specimen preparation, 100 ml distilled water was added into every 20 g soil sample and mixed adequately (1:5 soil-water extraction solution) for 30 mins. Then the leachate was extracted to measure EC via a digital multiparameter measuring apparatus (Multi 3420 Set B, WTW GmbH,

Germany) equipped with the composite electrode (TetraCon 925) at room temperature (25 °C). In addition, the research group has installed an integrated groundwater observation system consisted of HOBO Salt Water Salinity Data Logger (U24-002-C) and HOBO Water Level Data Logger (U20-001-02). These instruments were employed to automatically record long-term groundwater levels (m) and salinity (dS m^{-1}) in the study area (Figure 2).

2.2.2 S2 data acquisition and pre-processing

The Sentinel-2 MSI has 13 spectral bands: four VIS-NIR bands at 10 m, six red-edge and SWIR bands at 20 m, and three bands (Coastal aerosol, Water Vapor, and Cirrus) at 60 m spatial resolution (ESA, 2015). Specifications of spectral bands are provided in Table 1. In this study, two simultaneous cloud-free S2 Level-1C products in UTM/WGS84 projection were collected from the European Space Agency's Copernicus Open Access Hub (<https://scihub.copernicus.eu/>). Two S2 images captured on May 10th and August 28th, 2018 were regarded as the representative dry and wet seasons imageries in this study. The Sentinel Application Platform (SNAP) version 4.0.2. (<http://step.esa.int/main/toolboxes/snap/>) equipped with Sentinel-2 Toolbox and Sen2Cor plugin version 2.2.1 were applied for the atmospheric correction (Clevers and Gitelson, 2013; Louis et al., 2016). The algorithm parameters (Rural aerosol type and Mid-latitude summer atmosphere type) were chosen based on the location and environment type of the study area as well as the values recommended in the Sen2Cor configuration and user manual (Mueller-Wilm et al., 2017). It aimed at converting Sentinel-2 Level-1C images into Level-2A data, in other words, from Top Of Atmosphere (TOA) reflectance values to Bottom Of Atmosphere (BOA) reflectance values. In this step, bands in the 60 m-resolution (Band 1, 9 and 10) are dedicated primarily for detecting atmospheric features and therefore are not included in subsequent research. In addition, narrow-band

NIR (B8a) were employed instead of broadband NIR (B8), because the differences between them were not significant, which has been illustrated in (Fernández-Manso et al., 2016) and (Navarro et al., 2017). To retain as much information as possible, the mosaic of each pretreated spectral band was created at 10-m spatial resolution (Sibanda et al., 2018).

In these two seasons, the major scene components are water bodies and bare soils. The vegetation cover is low in both scenes, which considerably facilitates the estimation of soil salinity through RS images. In addition, water bodies were extracted using a supervised classification with the Maximum Likelihood Classification (MLC) method available in the ENVI® software version 5.1 (Exelis Visual Information Solutions, Boulder, CO, USA) from the two scenes.

Table 1 Spectral bands of MSI sensor.

Acronym	Band	Band center/nm	Band width/nm	Spatial resolution /m
B1	Coastal aerosol	443	20	60
B2	Blue	490	65	10
B3	Green	560	35	10
B4	Red	665	30	10
B5	Red-edge 1	705	15	20
B6	Red-edge 2	740	15	20
B7	Red-edge 3	783	20	20
B8	NIR	842	115	10
B8a	Narrow NIR	865	20	20
B9	Water Vapor	945	20	60
B10	Cirrus	1380	30	60

B11	SWIR1	1610	90	20
B12	SWIR2	2190	180	20

2.3 Spectral salinity indices

The application of spectral indices to investigate soil salinity is built upon the different spectral behavior associated in image pixels of the ground target (Peng et al., 2019). Subsequently, a series of spectral soil salinity indices were proposed for soil salinity assessments, especially in the arid regions. The relevance between the indices derived from broad-band RS data and the ground truth soil salinity data have been found to vary substantially (Allbed et al., 2014a). Some of the widely used soil salinity indices are listed in Table 2. Considering the novel spectral capabilities of three extra red-edge bands (Bannari et al., 2018; Sibanda et al., 2015), we used these three bands (B5-B7) instead of B4 and computed a variety of possible combinations of these new red-edge bands to produce potential soil salinity indices for estimating EC (Table3).

Table 2 Reference spectral salinity indices.

Acronym	Spectral index	Formula	Equation	Reference
Int1	Intensity Index1	$(G + R)/2$	$(B3 + B4)/2$	(Triki Fourati et al., 2015)
Int2	Intensity Index2	$(G + R + NIR)/2$	$(B3 + B4 + B8a)$	
NDSI	Normalized Difference	$(R - NIR)/(R + NIR)$	$(B4 - B8a)/(B4 + B8a)$	(Khan et al., 2005)
	Salinity Index			
S1	Salinity index I	B/R	$B2/B4$	(Khan et al., 2005)
S2	Salinity index II	$(B - R)/(B + R)$	$(B3 - B4)/(B3 + B4)$	
S3	Salinity index III	$(G \times R)/B$	$(B3 \times B4)/B2$	

S5	Salinity index V	$(B \times R)/G$	$(B2 \times B4)/B3$	
S6	Salinity index VI	$(R \times NIR)/G$	$(B4 \times B8a)/B3$	
SI	Salinity index	$(B + R)0.5$	$(B2 + B4)^{0.5}$	
SI1	Salinity index 1	$(G \times R)0.5$	$(B3 \times B4)^{0.5}$	(Douaoui et
SI2	Salinity index 2	$[(G)^2 + (R)^2 + (NIR)^2]^{0.5}$	$[(B3)^2 + (B4)^2 + (B8a)^2]^{0.5}$	al., 2006)
SI3	Salinity index 3	$[(R)^2 + (G)^2]^{0.5}$	$[(B4)^2 + (B3)^2]^{0.5}$	

214

215

216 **Table 3 The newly proposed red-edge spectral indices.**

Acronym	Spectral index	Formula	Equation
Int1 re1	Intensity Index 1 red-edge 1	$(G + \text{Red-edge 1})/2$	$(B3 + B5)/2$
Int1 re2	Intensity Index 1 red-edge 2	$(G + \text{Red-edge 2})/2$	$(B3 + B6)/2$
Int1 re3	Intensity Index 1 red-edge 3	$(G + \text{Red-edge 3})/2$	$(B3 + B7)/2$
Int2 re1	Intensity Index 2 red-edge 1	$(G + \text{Rededge-1} + NIR)/2$	$(B3 + B5 + B8a)$
Int2 re2	Intensity Index 2 red-edge 2	$(G + \text{Rededge-2} + NIR)/2$	$(B3 + B6 + B8a)$
Int2 re3	Intensity Index 2 red-edge 3	$(G + \text{Rededge-2} + NIR)/2$	$(B3 + B7 + B8a)$
NDSI re1	Normalized Difference	$(\text{Red-edge 1} - NIR)/(\text{Red-}$	$(B5 - B8a)/(B5 + B8a)$
	Salinity Index red-edge 1	$\text{edge 1} + NIR)$	
NDSI re2	Normalized Difference	$(\text{Red-edge 2} - NIR)/(\text{Red-}$	$(B6 - B8a)/(B6 + B9a)$
	Salinity Index red-edge 2	$\text{edge 2} + NIR)$	
NDSI re3	Normalized Difference	$(\text{Red-edge 3} - NIR)/(\text{Red-}$	$(B7 - B8a)/(B7 + B10a)$
	Salinity Index red-edge 3	$\text{edge 3} + NIR)$	

S1 re1	Salinity index I red-edge 1	$B/\text{Red-edge 1}$	$B2/B5$
S1 re2	Salinity index I red-edge 2	$B/\text{Red-edge 2}$	$B2/B6$
S1 re3	Salinity index I red-edge 3	$B/\text{Red-edge 3}$	$B2/B7$
S2 re1	Salinity index II red-edge 1	$(B - \text{Red-edge 1})/(B + \text{Red-edge 1})$	$(B3 - B5)/(B3+B5)$
S2 re2	Salinity index II red-edge 2	$(B - \text{Red-edge 2})/(B + \text{Red-edge 2})$	$(B3 - B6)/(B3+B6)$
S2 re3	Salinity index II red-edge 3	$(B - \text{Red-edge 3})/(B + \text{Red-edge 3})$	$(B3 - B7)/(B3+B7)$
S3 re1	Salinity index III red-edge 1	$(G \times \text{Red-edge 1})/B$	$(B3 \times B5)/B2$
S3 re2	Salinity index III red-edge 2	$(G \times \text{Red-edge 2})/B$	$(B3 \times B6)/B2$
S3 re3	Salinity index III red-edge 3	$(G \times \text{Red-edge 3})/B$	$(B3 \times B7)/B2$
S5 re1	Salinity index V red-edge 1	$(B \times \text{Red-edge 1})/G$	$(B2 \times B5)/B3$
S5 re2	Salinity index V red-edge 2	$(B \times \text{Red-edge 2})/G$	$(B2 \times B6)/B3$
S5 re3	Salinity index V red-edge 3	$(B \times \text{Red-edge 3})/G$	$(B2 \times B7)/B3$
S6 re1	Salinity index VI red-edge 1	$(\text{Red-edge 1} \times \text{NIR})/G$	$(B5 \times B8a)/B3$
S6 re2	Salinity index VI red-edge 2	$(\text{Red-edge 2} \times \text{NIR})/G$	$(B6 \times B8a)/B3$
S6 re3	Salinity index VI red-edge 3	$(\text{Red-edge 3} \times \text{NIR})/G$	$(B7 \times B8a)/B3$
SI re1	Salinity index red-edge 1	$(B + \text{Red-edge 1})^{0.5}$	$(B2 + B5)^{0.5}$
SI re2	Salinity index red-edge 2	$(B + \text{Red-edge 2})^{0.5}$	$(B2 + B6)^{0.5}$
SI re3	Salinity index red-edge 3	$(B + \text{Red-edge 3})^{0.5}$	$(B2 + B7)^{0.5}$
SI1 re1	Salinity index 1 red-edge 1	$(G \times \text{Red-edge 1})^{0.5}$	$(B3 \times B5)^{0.5}$

SI1 re2	Salinity index 1 red-edge 2	$(G \times \text{Red-edge } 2)^{0.5}$	$(B3 \times B6)^{0.5}$
SI1 re3	Salinity index 1 red-edge 3	$(G \times \text{Red-edge } 3)^{0.5}$	$(B3 \times B7)^{0.5}$
SI2 re1	Salinity index 2 red-edge 1	$[(G)^2 + (\text{Red-edge } 1)^2 + (\text{NIR})^2]^{0.5}$	$(B3^2 + B5^2 + B8a^2)^{0.5}$
SI2 re2	Salinity index 2 red-edge 2	$[(G)^2 + (\text{Red-edge } 2)^2 + (\text{NIR})^2]^{0.5}$	$(B3^2 + B6^2 + B8a^2)^{0.5}$
SI2 re3	Salinity index 2 red-edge 3	$[(G)^2 + (\text{Red-edge } 3)^2 + (\text{NIR})^2]^{0.5}$	$(B3^2 + B7^2 + B8a^2)^{0.5}$
SI3 re1	Salinity index 3 red-edge 1	$[(\text{Red-edge } 1)^2 + (G)^2]^{0.5}$	$[(B5)^2 + (B3)^2]^{0.5}$
SI3 re2	Salinity index 3 red-edge 2	$[(\text{Red-edge } 2)^2 + (G)^2]^{0.5}$	$[(B6)^2 + (B3)^2]^{0.5}$
SI3 re3	Salinity index 3 red-edge 3	$[(\text{Red-edge } 3)^2 + (G)^2]^{0.5}$	$[(B7)^2 + (B3)^2]^{0.5}$

Many scholars have regarded contour map of the R-values between EC and spectral bands as an effective approach to select sensitive bands for soil salinity information (Hong et al., 2019; Jin et al., 2016). In addition, the generated optimal band combination can be easily compared to those of previous studies (Jin et al., 2013). The contour maps of the R-values between EC and spectral bands can provide comprehensive information about the predictive ability of the combinations of the two separate wavelengths of RS image data (Hong et al., 2018; Wang et al., 2018c). In this study, two-band (2D) indices, namely, (1) ratio index (RI), (2) difference index (DI), and (3) normalized difference index (NDI) were calculated, and subsequently evaluated for their correlation with EC to extract practical salinity spectral indices for assessment of EC, as shown in Equations (1)-(3).

$$RI (R_{\lambda 1}, R_{\lambda 2}) = R_{\lambda 1} / R_{\lambda 2} \quad (1)$$

$$DI (R_{\lambda 1}, R_{\lambda 2}) = R_{\lambda 1} - R_{\lambda 2} \quad (2)$$

$$NDI(R_{\lambda 1}, R_{\lambda 2}) = (R_{\lambda 1} - R_{\lambda 2}) / (R_{\lambda 1} + R_{\lambda 2}) \quad (3)$$

In addition, following the methodology proposed in (Jackson, 1983) and (Tian et al., 2011), we attempted to add an extra band ($\lambda 3$) to construct several three-band indices (TBI), which were calculated according to the following formulae:

$$TBI_1(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = R_{\lambda 1} / (R_{\lambda 2} \times R_{\lambda 3}) \quad (4)$$

$$TBI_2(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = R_{\lambda 1} / (R_{\lambda 2} + R_{\lambda 3}) \quad (5)$$

$$TBI_3(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = (R_{\lambda 1} - R_{\lambda 2}) / (R_{\lambda 2} + R_{\lambda 3}) \quad (6)$$

$$TBI_4(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = (R_{\lambda 1} - R_{\lambda 2}) / (R_{\lambda 2} - R_{\lambda 3}) \quad (7)$$

$$TBI_5(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = (R_{\lambda 2} + R_{\lambda 3}) / R_{\lambda 1} \quad (8)$$

$$TBI_6(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = (R_{\lambda 1} - R_{\lambda 2}) / [(R_{\lambda 1} - R_{\lambda 2}) - (R_{\lambda 2} - R_{\lambda 3})] \quad (9)$$

$$TBI_7(R_{\lambda 1}, R_{\lambda 2}, R_{\lambda 3}) = (R_{\lambda 1} - R_{\lambda 2}) - (R_{\lambda 2} - R_{\lambda 3}) \quad (10)$$

Where $R_{\lambda 1}$, $R_{\lambda 2}$, and $R_{\lambda 3}$ represent the reflectance values from the random available spectral bands of S2 data (Band 2-7, 8a, 11 and 12). As described, the contour map of the correlation coefficients (R) was used to select the optimal spectral band combination. This specific method was used to reveal the possible combinations of different bands. Based on Equations (1) - (10), we generate the contour maps with MATLAB 2014a (The Mathworks Inc, Natick, MA, USA).

2.4 Modelling strategy

To ensure the generalization and robustness of the models, all soil samples (collected in dry and wet seasons) were treated as the entire set. The whole dataset ($n = 116$) was divided into two sections via Kennard-Stone (K-S) algorithm (Wang et al., 2018a): a calibration set of 72 soil samples and a validation set of 44 soil samples for cross-validation and independent validation, respectively.

In this study, we use the Partial Least Squares Regression (PLSR) to establish the relationship between EC and the remote sensing images generated spectral indices in order to evaluate which spectral indices are suitable for soil salinity monitoring and mapping. PLSR is a commonly used approach to build a multi-variable model, with the advantages of dimension reduction and handling of collinearities among various independent variables ([Sidike et al., 2014](#); [Wang et al., 2017a](#)). PLSR is an appropriate approach for calibrating models with severe collinearity in the independent variables, particularly in cases where the sample size is small. Many scholars applied this method to retrieve soil salinity information based on different RS data and found that its performance is reasonably satisfactory for estimating EC with remote sensing image ([Fan et al., 2016](#); [Farifteh et al., 2007](#); [Peng et al., 2016](#)). When building the PLSR models, it was also found that some non-predictive bands might lead to unsatisfactory results ([Peng et al., 2019](#)). Hence, band selection is crucial for the performance of the PLSR models. For that reason, a variety of selection strategies have been developed ([Genuer et al., 2010](#); [Zou et al., 2010](#)). We'll briefly discuss the commonly used Pearson's correlation coefficient (PCC), variable importance in projection (VIP), gray relational analysis (GRA), and random forest (RF) approaches below for band selection.

2.4.1 Pearson's correlation coefficient method (PCC)

In statistical analysis, the Pearson's correlation coefficient (PCC) is one of the most commonly used method to quantify correlation between two variables ([He et al., 2015](#)). Suppose that n distinct samples $\{(X_i, Y_i)\}_{i=1}^n$ are available, the corresponding PCC value (r) is calculated as follows:

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (10)$$

where X_i and Y_i represent the measured values of two input attributes (spectral data and measure

EC in this study, respectively), $\bar{X}$ and $\bar{Y}$ represent the mean values of them. The PCC value has a range of $[-1, 1]$, a greater absolute value of PCC represents higher correlation between the two variables, and vice versa.

2.4.2 Variable importance in projection (VIP)

Variable influence in projection scores (also known as ‘variable importance on projection’ scores or VIP scores) was first proposed by (Wold et al., 1983). VIP scores are effective in identifying predictor variables in the X space that best explain Y ’s variance, which is calculated as follows:

$$VIP = \sqrt{k \times \frac{\sum_{a=1}^A R_d(Y, t) W^2}{\sum_{a=1}^A R_d(Y, t)}} \quad (11)$$

where k is the total number of independent variables; A is the total number of components; t represents the ingredient; $R_d(Y, t)$ is the sum of squares of Y explained by component a ; W^2 is the importance of the variable in each component. In general, the VIP values greater than 1 represent the most relevant variables.

2.4.3 Gray Relational Analysis (GRA)

Gray relational analysis (GRA), a systematic analysis approach, can evaluate correlations between related factor sequences and characteristic data with the ultimate objective (Jin et al., 2013). Its performance of understanding complicated interrelationships between multi-factors and multi-variables have been widely evaluated (Morán et al., 2006; Wang et al., 2018c). A detailed description of GRA can be found in (Deng, 2002; Kuo et al., 2008). For this specific algorithm, a quantitative description of the effect of each featured factor (Gray correlation degree, GCD) is available. Suppose that $X_0 = \{x_0(t), t = 1, 2, \dots, n\}$ and $X_i = \{x_i(t), t = 1, 2, \dots, n\}$ are defined as the reference sequence and the comparative sequence, respectively. The GCD is given as:

$$\text{GCD} = \frac{1}{n} \sum_{t=1}^n \gamma(x_0(t), x_i(t)) \quad (12)$$

$$\gamma(x_0(t), x_i(t)) = \frac{\min_t |x_0(t) - x_i(t)| + \xi \max_t |x_0(t) - x_i(t)|}{|x_0(t) - x_i(t)| + \xi \max_t |x_0(t) - x_i(t)|}$$

where ξ is the distinguishing coefficient with a range of $[0, 1]$, in this study setting $\xi = 0.5$ (Wang et al., 2018c).

2.4.4 Random Forest (RF)

Random Forest (RF), developed by (Breiman, 2001), is capable of aggregating ideas as well as solving classification and regression problems. It is an ensemble machine learning algorithm and appears to be very effective hence increasingly popular for variable selection (Wang et al., 2018a). In the RF framework, variable importance is primarily affected by two main parameters: the size of the input variables subset (*mtry*), and the number of trees in the forest (*ntree*). Currently the performance of RF is mostly evaluated by the out-of-bag (OOB) error. In this study, the obtained RF variable importance is accuracy-based importance (MDA, Mean Decrease Accuracy). A detailed illustration of numerical and sensitivity experiments is available in (Genuer et al., 2010). The RF variable importance of X^i is calculated as:

$$VI(X^i) = \frac{1}{ntree} \sum_t (errO\tilde{O}B_t^i - errOOB_t) \quad (13)$$

Where t represents each tree in the forest; OOB_t represents the associated section not included in the bootstrap processes used to construct t ; $errOOB_t$ is the error of a specific tree t on the associated OOB_t sample; and $errO\tilde{O}B_t^i$ represents a perturbed sample affected by the permuted values of X^i . In this study, the size of the input variables subset was set to 23 based on repeated testing, and the number of trees was set to 500 according to the stabilizing feature of the $errO\tilde{O}B_t^i$.

2.5 Statistical analysis

In this study, executions of PCC, VIP, and GRA were implemented with the PLS_Toolbox (Version 8.0.2, Eigenvector Research, Inc., Wenatchee, USA) in MATLAB 2014a. The implementation of RF was conducted via the randomForest package (version 4.6-12) in the R platform (version 3.4.0) ([Liaw and Wiener, 2002](#); [R Development Core Team, 2018](#)). To determine the ideal number of latent variables when using PLSR, leave-one-out cross-validation (LOOCV) was conducted to prevent over- or under-fitting the data, which may produce models with poor performance ([Martens et al., 1992](#)).

To assess and compare the overall performance of the estimation model, three measures were used: (1) the coefficients of determination (R^2), (2) root mean square errors (RMSE), and (3) ratio of performance to deviation (RPD). According to inversion model classification criterion outlined in ([Chang et al., 2001](#)), the models could be classified to three categories: Level A ($RPD \geq 2.00$) is the most effective model with reliable predictive ability; Level B ($1.40 \leq RPD < 2.00$) is a good model with generally adequate outcomes; Level C ($RPD \leq 1.40$) is an unreliable model. In general, most effective models are characterized by high R^2 and RPD values and a low RMSE.

Additionally, based on the standard of different soil salinity levels given by the United States Department of Agriculture (USDA) ([Schoeneberger et al., 2002](#)), the soil salinity can be partitioned into five levels: □) non-saline ($0-2 \text{ dS m}^{-1}$), □) very slightly saline ($2-4 \text{ dS m}^{-1}$), □) slightly saline ($4-8 \text{ dS m}^{-1}$), □) moderately saline ($8-16 \text{ dS m}^{-1}$), and □) strongly saline ($>16 \text{ dS m}^{-1}$). When scatter point falls within a specific class boundary, this suggests the correct classification, and vice versa. The rate of correct classification was used to assess the estimation capability, which is simply the ratio operation between the number of correctly classified samples and that of the total samples ([Hong et al., 2018](#)).

3. Results

3.1 Chemical and spectral characteristics

In the study area, the mean value of EC in the wet season (May) was 14.96 dS m^{-1} , and approximately 45% less than the ones in the dry season (August, 27.39 dS m^{-1}) (Figure 3). The coefficients of variation of EC were 97.02% and 137.61% in the wet and dry periods, respectively. This result indicates that the topsoil is salinized in the Ebinur Lake region with large variabilities. Furthermore, this specific condition might contribute to the improvement of our models' predictive capabilities (Hong et al., 2019; Kuang and Mouazen, 2011). The mean and S.D. values of the calibration dataset were 19.45 dS m^{-1} and 18.17 dS m^{-1} , respectively, slightly lower than those of the validation dataset (Figure 3). The selected calibration and validation datasets had nearly identical wide range spanned from $0.11\text{--}67.51 \text{ dS m}^{-1}$ and $0.16\text{--}67.04 \text{ dS m}^{-1}$, respectively. These results indicate that soil salinity of both selected datasets adequately represent the entire dataset (Table 4). The calibrated models can then be used to generate satisfactory predictive results.

Figure 3 Violin plots showing the descriptive statistics of soil salinity for the wet, dry, calibration, and validation dataset (dS m^{-1}). S.D.: standard deviation.

Table 4 Distribution of soil salinity levels in different datasets.

Soil salinity	Dataset			
	Wet season	Dry season	Calibration	Validation
Non-saline ($0-2 \text{ dS m}^{-1}$)	15	3	11	7
Very slightly saline ($2-4 \text{ dS m}^{-1}$)	7	5	8	4
Slightly saline ($4-8 \text{ dS m}^{-1}$)	7	3	6	4
Moderately saline ($8-16 \text{ dS m}^{-1}$)	12	9	13	8
Strongly saline ($>16 \text{ dS m}^{-1}$)	24	31	34	21
Total	65	51	72	44

As presented in Fig. 4, soil samples with different salinity levels show very similar spectral shapes and amplitudes of the spectral reflectance curves. The reflectivity of soil samples increased with the increase in soil salinity ranging from non-saline (0.38 dS m^{-1} , lowest reflectance) to strongly saline (59.49 dS m^{-1} , highest reflectance) samples. Similar to published results of reflectance spectra, the spectrum curves were sharply increased in the range of VIS-NIR, especially from Blue to NIR band (B2–B4), reaching a maximum value at the SWIR1 (B11). However, the regularity between different EC levels and the corresponding spectral feature was less noticeable in the SWIR 2 band, which agrees with the findings in [El Harti et al. \(2016\)](#). This specific spectral characteristic could be attributed to the different chemical composition of salt in soil ([Farifteh et al., 2006](#)). It was noticed that the reflectance in Red-edge range (B5-B7) could discriminate the soil samples with different salinity levels, although the difference was not significant among bands. Generally, these initial analyses lay the foundation for the subsequent constructions of satellite spectral indices.

Figure 4 Reflectance spectra curves of soils with different EC in the study area.

3.2 Identifying optimal band of S2 data for the EC retrieval

To further analyze the potential of S2 data, the descriptive statistics of the spectral differences of soils with different EC values were calculated. Results showed various reflectance values in each spectral band of S2. Among these nine bands, Band 2 seemed to generate relatively stable performance with a narrower range while the spectral reflectance values of the other bands were broad and overlapping to varying degrees. Here, the single-band retrieval model was constructed to quantify local soil salinity (Table 5). The optimum single-band model produced an RPD of 1.30 (≤ 1.40), indicating the model is not ideally reliable. This model was constructed based on Blue band (B2) of S2 data with generated R^2_v and $RMSE_v$ of 0.70 and 13.8 dS m^{-1} , respectively. Hence, a single-band is not likely to have acceptable accuracy, and is not capable of effective dynamic monitoring of soil salinity in the study area. The feasibilities of published satellite spectral indices of RS imageries need to be further verified.

Figure 5 Box plot of each spectral band of S2 data of soils in with varying salinity degree.

Table 5 Relationships between S2 bands and EC data in the study area (dS m^{-1}).

Band (X)	Linear model	R^2_{CV}	$RMSE_{CV}$	R^2_V	$RMSE_V$	RPD
Blue (B2)	$y = 143.94x - 7.37$	0.52	16.97	0.70	13.86	1.30
Green (B3)	$y = 144.41x - 10.56$	0.54	16.72	0.70	13.89	1.29
Red (B4)	$y = 130.65x - 11.19$	0.56	16.55	0.67	14.39	1.25
Red-edge 1 (B5)	$y = 130.07x - 13.40$	0.57	16.41	0.66	14.53	1.24
Red-edge 2 (B6)	$y = 133.06x - 15.30$	0.57	16.36	0.65	14.68	1.22
Red-edge 3 (B7)	$y = 136.37x - 17.25$	0.59	16.10	0.66	14.65	1.23
NIR narrow (B8a)	$y = 131.32x - 17.06$	0.60	16.02	0.65	14.81	1.21
SWIR1 (B11)	$y = 108.99x - 15.84$	0.56	16.51	0.58	15.82	1.14
SWIR2 (B12)	$y = 88.75x - 7.15$	0.52	17.07	0.51	16.74	1.07

R^2_{CV} , determination coefficient for cross-validation; $RMSE_{CV}$, root mean square error of cross-validation
 R^2_V , determination coefficient for validation; $RMSE_V$, root mean square error of validation; RPD, the ratio of standard deviation.

3.3 Comparison of red-edge indices with the existing soil salinity indices

The relationships between existing soil salinity indices and the soil salinity data extracted in the

Ebinur Lake region were compared (Figure 6). All the published indices except for NDSI, S1 and S2 had a significant correlation with EC at the 0.01 level ($p < 0.01$). We will discuss individually these three indices. The PCCs spanned from 0.417 to 0.503. The increased relevance can be observed when replacing the red band by a red-edge band. The introduction of red-edge bands enhances the sensitivities of the indices to soil salinity, which might be attributed to the novel spectral information and higher signal-to-noise ratio in the red-edge region. The mean PCC values between EC and traditional salinity indices, re 1-, re 2-, and re 3-derived indices are 0.476, 0.480, 0.487, and 0.491, respectively, which can be used to extract soil salinity information. Among these 40 indices, the index of SI2 re3 performs better than the others with an optimal PCC of 0.508. With respect to the three less ideal (NDSI, S1, and S2), the improvement is not exceptional with the introduction of red-edge bands. Although they failed the significance test, these less predictive spectral covariates were not excluded from this study.

Figure 6 Relevance between EC and different soil salinity indices (existing indices and derived red-edge indices).

3.4 Constructions of 2D indices and TBI

The relationship between the 2D indices (RI, DI, and NDI) for the two random spectral bands selected between B2-B12 in S2 bands and soil salinity (EC) data are shown in Figure 7. The right-side color bar indicates the color of the PCC values. The colors of dark-red and dark-blue represent a relatively high PCC (red for positive and blue for negative) between the measured EC and the band combinations. Results showed that bands with high correlation of RI, DI, and NDI to EC mainly

concentrated in SWIR bands (B11 and B12). As shown in the diagram, the maximum absolute PCC was between EC and NDI (0.52), its explicit expression was $[(B12-B7)/(B12+B7)]$. While the PCC of DI (B12-B6) was 0.16, suggesting its poor correlation with soil salinity. The PCCs are -0.16 and 0.41 for the DI (B12-B6) and RI (B12/B7), respectively (Table 6).

The optimal band combinations of each TBI were extracted and their correlation coefficients with soil salinity were calculated (Figure 8). As illustrated in the slice maps, the combinations of SWIR2 Band (B12) and two other bands usually yield higher PCCs. Among all three-band indices, the TBI₄ $[(B12-B3)/(B3-B11)]$ has the best effect with a maximum PCC of 0.54. The combination of three random spectral bands with the highest PCC differed in each form of TBIs. Comparing to the two-band (2D) spectral indices, the TBI magnifies the threshold range of the correlation coefficient with soil salinity.

Table 6 The relationship between soil salinity and the various optimal spectral indices.

	Optimal combination bands	<i>R</i>
2D index		
RI	B4/B12	0.41
DI	B3-B7	-0.16
NDI	$(B12-B7)/(B12+B7)$	0.52
Three-band index		
TBI ₁	$B12/(B8a \times B8a)$	-0.43
TBI ₂	$B7/(B12+B12)$	-0.51
TBI ₃	$(B6-B12)/(B6+B12)$	0.50
TBI ₄	$(B12-B3)/(B3-B11)$	0.54

TBI_5	$(B12+B11)/B3$	0.53
TBI_6	$(B11-B3)/(B11-2\times B3+B12)$	0.44
TBI_7	$B11-2\times B12+B5$	-0.49

(a) RI (b) DI (c) NDI

Figure 7 Correlation coefficients between the soil electrical conductivity and the all potential two-band combinations based on nine spectral bands (B2, B3, B4, B5, B6, B7, B8a, B11, B12) of S2 data: (a) RI, (b) DI, and (c) NDI.

Note: The X-, Y-, and Z- axes represents λ_1 , λ_2 , and λ_3 , respectively. The upper and lower limits of the color bar represent the largest positive correlation coefficient and the largest negative correlation coefficient.

(a) horizontal slice map (b) vertical slice map (c) optimal slice map

Figure 8 Slice maps of the correlation coefficients between the soil electrical conductivity and

the all potential three-band combinations t based on nine spectral bands (B2, B3, B4, B5, B6, B7, B8a, B11, B12) of S2 data: (a) horizontal slice map, (b) vertical slice map, and (c) optimal slice map.

Note: The X-, Y-, and Z- axes represents λ_1 , λ_2 , and λ_3 , respectively. The upper and lower limits of the color bar represent the largest positive correlation coefficient and the largest negative correlation coefficient.

3.5 Model estimations and comparisons

PCC, VIP, GCD, and RF variable importance scores between the 67 spectral covariates (nine spectral bands, 12 existing satellite soil salinity indices, 36 red-edge indices, three 2D indices, and seven TBIs) and soil salinity are illustrated in Figure 9. We applied all 67 spectral covariates and the top 20-ranked spectral covariates selected by different algorithms to calibrate the PLSR model for salinity estimation. The performance of the PLSR models based on different covariates selecting strategy is illustrated in Table 7. The estimating performances of Raw-PLSR model with 67 spectral covariates shows the lowest values with an RPD value of 1.48. Once the models were calibrated based on the selected spectral covariates, their estimation accuracy and stability have been improved in varying degrees. Specifically, the performance of the RF-PLSR model to estimate EC was found to be the best. With reference to the three-level classifications of RPD shown in Section 2.5, the RF-PLSR model ($R^2_v = 0.92$, $RMSE_v = 7.58 \text{ dS m}^{-1}$ and $RPD = 2.36$) generates the most reliable estimation, whereas the other models could only yield generally adequate outcomes regardless of the variable selections implemented. Among them, VIP and PCC possess similar performance, which is to be expected since these two approaches follow identical principles. Box plots of residual analysis of validation results in different models are presented in Figure 10. With the widest range of residuals, the estimating performances of Raw-PLSR model are less satisfactory. However, the results showed

that the distribution of the residuals in the validation dataset of RF-PLSR model followed the standardized normal distribution, with the narrowest range. In addition, based on the performance on the cross-validation dataset and validation dataset, we can consider that there are neither over-fitting or under-fitting in RF-PLSR model. The scatter plots (Figure 11) of the measured versus estimated soil salinity are well distributed along the 1:1 line, suggesting high estimation performance. All these results suggest that the RF-PLSR model has the ideal performance to link spectral information and soil salinity.

Moreover, Figure 11 also illustrates the boundaries of different soil salinity levels (black dash lines), i.e., □) non-saline (0–2 dS m⁻¹), □) very slightly saline (2–4 dS m⁻¹), □) slightly saline (4–8 dS m⁻¹), □) moderately saline (8–16 dS m⁻¹), and □) strongly saline (>16 dS m⁻¹). Only 12 soil samples were mis-classified by the RF-PLSR model, suggesting 72.73% the classification correctness of the model, better than any other models (there are 20, 19, 18, and 14 misclassified soil samples for Raw-PLSR, PCC-PLSR, VIP-PLSR, and GRA-PLSR models, which corresponds to effective classification rates of 54.54%, 56.82%, 59.09%, and 68.18%, respectively).

Figure 9 PCC, VIP score, GCD, and RF variable importance of 67 spectral covariates (nine spectral bands, 12 existing satellite soil salinity indices, 36 red-edge indices, three 2D indices, and seven TBIs) for soil salinity estimation.

Acronym	Modelling strategy	Parameters	LVs	R^2_{cv}	$RMSE_{cv}$	R^2_v	$RMSE_v$	RPD
Raw-PLSR	-	67	4	0.85	10.75	0.78	12.12	1.48
PCC-PLSR	PCC	20	2	0.87	9.97	0.82	11.18	1.61
VIP-PLSR	VIP	20	2	0.85	10.43	0.82	11.13	1.61
GRA-PLSR	GRA	20	3	0.89	9.25	0.88	9.08	1.98
RF-PLSR	RF	20	4	0.93	8.35	0.92	7.58	2.36

LVs, the optimal number of latent variables; R^2_{CV} , determination coefficient for cross-validation; RMSE_{CV}, root mean square error of cross-validation R^2_V , determination coefficient for validation; RMSE_V, root mean square error of validation; RPD, the ratio of standard deviation.

Figure 10 Box plot of residual analysis of validation results of the different models.

Figure 11 Scatter plots of the measured versus estimated EC using different models with various modelling strategies: (a)Raw-PLSR, (b) PCC-PLSR, (c) VIP-PLSR, (d) GRA-PLSR, and (e)

RF-PLSR. Red and blue lines represent the 1: 1 line and fitting line between the measured and estimated EC values, respectively.

3.6 Soil salinity maps of dry and wet seasons

Because the RF-PLSR model produces the most ideal results, the study used the calibrated RF-PLSR model and two S2 scenes for the RS estimation of soil salinity. Furthermore, the estimated soil salinity values were divided into the five levels as aforementioned. From wet season to dry season, the extent of water bodies declined. The spatial distribution of soil salinity in the Ebinur Lake region in Figure 12 agrees with our prior knowledge of this specific region and field investigations. Considering the dynamics of water surface area of Ebinur Lake, we found that the proportions of water body, non-saline soil, very slightly saline soil, slightly saline soil, moderately saline soil, and strongly saline soil were 30.11%, 6.91%, 2.72%, 32.27%, 16.28%, and 11.71% in wet season, and 23.86%, 8.13%, 9.00%, 16.60%, and 21.28% in dry season. Overall, the salt-affected soil is predominant in this study area. From the maps (Figure 12), we can see that soil salinity decreases from the center to the edges of the study area, and there is difference between the trend of EC distribution maps in the wet and dry seasons. Affected by the topography, areas with the lowest degree of salinization are located mostly in the north of the study area (Alatao Mountain). Areas with the highest salinity occupied the southern and eastern parts of the Ebinur Lake region. The 10-m resolution EC map unfold more details of salinization than those of medium resolution.

To further check the difference between soil salinization in wet and dry seasons, we created a Sankey diagram of the soil salinity levels of both seasons in Figure 13. Sankey diagrams are useful tools for visualizing soil salinity levels change using comparative flows between categories from one time period to another. They display soil salinity levels change flows “from” and “to” various nodes,

and transition lines between pairs of soil salinity levels categories are used to display these flows. In this study, the stacked bars represent the proportion of each soil salinity levels, the width of the directional lines are proportional to the magnitude of the changes in flows. From the diagram, we can see that the increased extent of strongly saline soil occurring from wet season to dry season are attributed to the conversion of slightly and moderately salinized soils. The proportions of no-saline and moderately saline soils were relatively stable in the study area. Specifically, the soil salinity in the dry season was higher than in the wet season, mostly in the lake area and nearby shores, primarily due to the expansion of the strongly salinity level ($> 16 \text{ dSm}^{-1}$), and an associated decrease of the slightly saline level ($4\text{--}8 \text{ dS m}^{-1}$) and moderately saline level ($8\text{--}16 \text{ dS m}^{-1}$) soils in the lake shores.

Figure 12 Soil salinity (EC) maps in (a) dry and (b) wet seasons in the Ebinur Lake region.

Figure 13 Sankey diagram for comparison of soil salinization dynamics in dry and wet seasons in the Ebinur Lake region.

4. Discussion

4.1 Application of various spectral indices across different platforms

From the above analyses, we can see that the spectral indices are but simple combinations of reflectance values for different spectral bands. More importantly, our study suggests that we can establish rather strong correlation between the indices and the corresponding physical parameters (soil salinity in our study), and the spectral estimation can be operated subsequently ([Jackson, 1983](#)). In recent years, substantial soil salinity indices and vegetation spectral indices have been used for the soil salinization monitoring across different RS platforms, especially for the ground-based VIS–NIR ([Hong et al., 2018](#)). Despite the high spectral resolution, consecutive spectral information, it can only extract salinization information at single-point scale rather than large areas ([El Harti et al., 2016](#); [Wang et al., 2018b](#)). In addition, their universal applicability and generalization are relatively poor, though the vegetation indices are more sensitive to changes in soil salinity under higher vegetation coverage, due to the effects of salt stress in plants. The vegetation indices are of limited use in arid and semi-arid regions as in the current study.

The multi-spectral RS, on the other hand can meet the needs of monitoring and mapping of soil salinity in arid and semi-arid regions mainly because of the massive availability of data resources, less access difficulty and appropriate temporal resolution. This suggests that the satellite spectral indices are potentially more appropriate approaches. Based on the measured soil salinity and a Combined Spectral Response Index (COSRI), [Fernández-Buces et al. \(2006\)](#) mapped the salt-affected area in the former lake Texcoco with satisfactory accuracy ($R^2 = 0.75$). [Sidike et al. \(2014\)](#) also indicated that the spectral covariates, such as Int1, Int2, and soil salinity indices (SI1, SI2, SI3) were sensitive to soil salinity, which could help increase the mapping accuracy of soil salinity.

In this study, we revealed all possible combinations of two or three bands in the multispectral remote sensing images to construct 2D indices and TBI (Band 2-7, 8a, 11 and 12) for estimating EC in the Ebinur Lake region. The combined bands of 2D index and TBI are mainly concentrated in two zones, namely, red-edge bands and SWIR bands, the results are supported by those of ([Allbed et al., 2014b](#)) and ([Peng et al., 2019](#)). The red band performs the best at characterizing the patterns and features of soil salinity ([Bannari et al., 2018](#); [Hu et al., 2019](#)). This is because the red band has high correlation with EC ground measurements and the behavior of their spectral signatures for SWIR bands are controlled significantly by the chemical composition and the mineralogy of the salt in the soils. In this study, we found that the TBI had better performance than other spectral covariates. More importantly, however, there is no universal spectral index which can yield a single satisfactory outcome in any environmental conditions.

4.2 Estimation of salinity and the uncertainty analysis

For the RS-based monitoring of soil salinization, its accuracy is susceptible to vegetation coverage, soil moisture content (SMC) and soil texture ([Ding and Yu, 2014](#)). Specifically, the difference in vegetation coverage may obscures soil information and cause large error in the identifications of soils with various salinity levels. Higher SMC can cause decreased spectral reflectivity in the VIS-NIR region. Consequently, the optimal parameters should not be fixed parameters, instead they should be redetermined locally and temporally according to the varied environmental conditions.

In this study, we found our spectral covariates perform ideally. We contend this is likely because of: (1) the temporal consistency between field sampling and satellite passes, (2) the spatial consistence that most of the sampling sites are located in the pure soil pixels, which means the soil

reflectance are not polluted by non-soil targets, (3) minimum effect of SMC on soils because of extremely intense local evaporation conditions, (4) high concentration of accumulated salt within the topsoil, and (5) the combination of a variety of mutually complementary spectral covariates including satellite band reflectance, published satellite salinity index, red-edge index, newly constructed 2D index, and TBI.

This study shows how the PLSR approach, coupled with fine spatial resolution RS data and various spectral indices could successfully monitor and map soil salinity in arid land in dry and wet seasons. However, salinity is not the only factor that influences soil spectra characteristics. The RS-based estimation of soil salinity might be interfered by the adsorption features of massive soil components, such as soil irons, water-soluble salt ions, organic matter and other soil components ([Nawar et al., 2014](#); [Peng et al., 2016](#)). In addition, the terrain attributes and soil textures can also have impacts on the mapping of soil salinity. It is hence worth noting that any such work needs to be carefully cross examined with in situ verification.

4.3 Soil salinity in dry and wet seasons

Soil salinity is significantly affected by different environmental factors ([Gorji et al., 2015](#)). In Ebinur Lake region, during the wet seasons, the highest salinity levels were found in the lakeside. However, stormy waves can push the lake water to the nearby shores, which submerges the topsoil and decreases its salinity. This is one of the reasons why soil salinity levels are typically lower during the wet seasons than the dry seasons. The highest salinity level shifts to the eastern bordering region of the study area during the dry seasons, especially near the management stations of Shaqiu and Dongdaqiao.

In the Ebinur Lake region, precipitation is distributed unevenly in both time and space ([Wang et](#)

al., 2018a). In the wet seasons, rainwater leaches salt from the topsoil to deep soil, leading to the relative lower topsoil salinity. Meanwhile, vegetation enters into the growth period, which causes the decline of the groundwater table and high phreatic water salinity (Figure 14). Additionally, the region is dominated with sand and sandy loam soil with strong capillary effects (Ge et al., 2016). The strong winds from the Alataw Pass also accelerate regional evaporation rate. Affected by the capillarity effect, salt contents from groundwater could easily seep up to the topsoil during the dry seasons. In the dry seasons, water-soluble salt often accumulated in the top soil after evaporation.

Figure 14 The groundwater table depth and salinity in the study

4.4 Rehabilitation of eco-environment

Currently, how to take effective measures to avoid further salinization of moderately and lightly salinized regions, so as to rehabilitate the eco-environment in the dry and semi-dry regions has attracted scholarly and governmental attention. Severe soil salinization will ultimately lead to soil desertification and land degradation, which would eventually limit sustainable urban development and even endanger human habitation. There are about 140 kilometers of the highway runs along the Ebinur Lake, which is an essential part of the New Eurasian Continental Bridge (also known as the Second Eurasian Continental Bridge). Affected by the dust storms, both the subgrade and pavement

of National Highway 312 were severely damaged many times ([Wang et al., 2011](#)). Local prevailing northwestern winds with salt-rich dust are considered one of most severe constraints for the entire traffic network there. Ebinur Lake nearly dried up in 2005 due to excessive farming irrigation and reduced surface runoff, though water has returned to the dried up lake in the past two years ([Su et al., 2016](#)). [Wang et al. \(2019\)](#) has analyzed the annual and intra-annual variations of Ebinur Lake and found that the surface area of water body in 2017 reached an all-time high. Still, it is important to pay more attention to local salinization level and other environmental effects despite of the return of lake water. With the increase of water surface area of Ebinur Lake, the bare dry lakebed that comprises abundant loose saline sediments was submerged, and the dust was suppressed to a certain extent. Understandably, local soil salinity decreased significantly comparing to just a few years ago.

The ecological water conveyances play an important role in the regeneration of natural vegetation in the lower reaches of Tarim River ([Ling et al., 2017](#); [Liu et al., 2016](#)). This project provides valuable experience for the ecological restoration under similar condition, especially for the management of soil salinization in the Ebinur Lake region. For the protection and management of eco-environment, as well as supporting local sustainable development, it might be inevitable to implement inter-basin water transfer project. On the other hand, the water conveyance will likely expand the lake's water surface area and rise the lake's water level, which could result in considerable ineffective evaporation and inundation of the New Eurasian Continental Bridge if the lake's surface water area was not managed effectively. For this reason, it is crucial to determine a threshold of reasonable water area of Ebinur Lake. This topic, however, is beyond the scope of the current investigation and will be explored in the future.

5. Conclusion

In this study, using field collected soil salinity data, we investigated the predictive capability of Sentinel 2-MSI (S2) spectral bands and derived spectral indices in a salt-affected wetland, Ebinur Lake region. The performance of existing salinity indices was tested, and the spectral bands specifically sensitive to soil salinity were identified to develop various new soil salinity indices. The spectral index is an effective way to retrieve soil salinity, especially in arid and semi-arid regions. The introduction of red-edge bands enhances the sensitivities of the derived spectral indices to soil salinity. Significant relationships between soil salinity and 2D indices (RI, DI, and NDI) were obtained in this study. Among these indices, the NDI $[(B12-B7)/(B12+B7)]$ has a maximum absolute PCC of 0.52. For the TBIs, we found that the combinations of the SWIR2 Band (B12) and two other bands usually yield higher PCCs. Among all three-band indices, the TBI4 $[(B12-B3)/(B3-B11)]$ has the best effect with a maximum PCC of 0.54. Based on 10-m resolution S2 data and field-measured soil salinity data, the PLSR models of soil EC were constructed using a variety of spectral covariates obtained from satellite images, including satellite band reflectance, published satellite salinity index, red-edge index, newly constructed 2D index, and TBI. The estimation performance varies with different input spectral covariates, the RF-PLSR model performed the best across the five models with the R^2_v , $RMSE_v$, and RPD of 0.92, 7.58 $dS\ m^{-1}$, and 2.36, respectively. All models and our field data suggest that salt-affected soil is predominant in this study area. Not surprisingly, the soil salinity in the dry season was higher than in the wet season, mostly in the lake area and nearby shores. The model generated 10-m resolution EC map unfolds more details of soil salinization than those of medium resolution. This study is expected to be useful for dynamic soil salinization monitoring and land reclamation in arid or semi-arid regions outside the current study area, though necessary

adjustment of ideal spectral covariates is likely needed. In addition, though not implemented in our current study, we intend to do so in the future to adopt Sentinel-2 time series data for soil salinity mapping in different periods and believe we could further reduce the uncertainty and improve the accuracy of the predictions. Other than the inclusion of time series data in our future studies, we also intend to focus on conducting studies on the tasks for the rehabilitation of local eco-environment and approaches to support sustainable development in arid and semi-arid regions.

Acknowledgements

The authors wish to thank Mr. Xiangzhuo Liu and Dr. Tayierjiang Aishan for helping in providing helpful suggestions. We are especially grateful to the anonymous reviewers and editors for appraising our manuscript and for offering instructive comments. This study was supported by the National Natural Science Foundation of China (41771470). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Reference

- Abbas, A., Khan, S., Hussain, N., Hanjra, M.A., Akbar, S., 2013. Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Physics and Chemistry of the Earth, Parts A/B/C* 55-57, 43-52.
- Aldabaa, A.A.A., Weindorf, D.C., Chakraborty, S., Sharma, A., Li, B., 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. *Geoderma* 239-240, 34-46.
- Allbed, A., Kumar, L., Aldakheel, Y.Y., 2014a. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma* 230-231, 1-8.
- Allbed, A., Kumar, L., Sinha, P., 2014b. Mapping and Modelling Spatial Variation in Soil Salinity in the Al Hassa Oasis Based on Remote Sensing Indicators and Regression Techniques. *Remote Sensing* 6(2), 1137-1157.
- Bai, L., Wang, C., Zang, S., Wu, C., Luo, J., Wu, Y., 2018. Mapping Soil Alkalinity and Salinity in Northern Songnen Plain, China with the HJ-1 Hyperspectral Imager Data and Partial Least Squares Regression. *Sensors* 18(11), 3855.
- Bannari, A., El-Battay, A., Bannari, R., Rhinane, H., 2018. Sentinel-MSI VNIR and SWIR Bands Sensitivity Analysis for Soil Salinity Discrimination in an Arid Landscape. *Remote Sensing* 10(6), 855.
- Bell, D., Menges, C., Ahmad, W., van Zyl, J.J., 2001. The Application of Dielectric Retrieval Algorithms

- for Mapping Soil Salinity in a Tropical Coastal Environment Using Airborne Polarimetric SAR. Remote Sensing of Environment 75(3), 375-384.
- Breiman, L., 2001. Random Forests. Machine Learning 45(1), 5-32.
- Cao, L., Ding, J., Halik, Ü., Su, W., Ning, J., Miao, C., Li, H., 2016. Extraction and Modeling of Regional Soil Salinization Based on Data from GF-1 Satellite. Acta Pedologica Sinica 53(6), 1399-1409 (in Chinese).
- Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy—Principal Components Regression Analyses of Soil Properties Journal Paper no. J-18766 of the Iowa Agric. and Home Econ. Exp. Stn., Ames, IA. Soil Science Society of America Journal 65(2), 480-490.
- Clevers, J.G.P.W., Gitelson, A.A., 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. International Journal of Applied Earth Observation and Geoinformation 23, 344-351.
- Davis, E., Wang, C., Dow, K., 2019. Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: a case study of agricultural lands in coastal North Carolina. International Journal of Remote Sensing, 1-20.
- Deng, J., 2002. Grey System Theory. Press of Huazhong University of Science and Technology, Wuhan.
- Ding, J., Yu, D., 2014. Monitoring and evaluating spatial variability of soil salinity in dry and wet seasons in the Werigan–Kuqa Oasis, China, using remote sensing and electromagnetic induction instruments. Geoderma 235-236, 316-322.
- Douaoui, A.E.K., Nicolas, H., Walter, C., 2006. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. Geoderma 134(1), 217-230.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. Remote Sensing of Environment 120, 25-36.
- El Harti, A., Lhissou, R., Chokmani, K., Ouzemou, J.-e., Hassouna, M., Bachaoui, E.M., El Ghmari, A., 2016. Spatiotemporal monitoring of soil salinization in irrigated Tadla Plain (Morocco) using satellite spectral indices. International Journal of Applied Earth Observation and Geoinformation 50, 64-73.
- ESA, 2015. ESA Introducing Sentinel-2.
- Fan, X., Weng, Y., Tao, J., 2016. Towards decadal soil salinity mapping using Landsat time series data. International Journal of Applied Earth Observation and Geoinformation 52, 32-41.
- Farifteh, J., Farshad, A., George, R.J., 2006. Assessing salt-affected soils using remote sensing, solute modelling, and geophysics. Geoderma 130(3), 191-206.
- Farifteh, J., Van der Meer, F., Atzberger, C., Carranza, E.J.M., 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). Remote Sensing of Environment 110(1), 59-78.
- Fernández-Buces, N., Siebe, C., Cram, S., Palacio, J.L., 2006. Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. Journal of Arid Environments 65(4), 644-667.
- Fernández-Manso, A., Fernández-Manso, O., Quintano, C., 2016. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. International Journal of Applied Earth Observation and Geoinformation 50, 170-175.

- 715 Ge, X., Wang, J., Ding, J., Cao, X., Zhang, Z., Liu, J., Li, X., 2019. Combining UAV-based hyperspectral
716 imagery and machine learning algorithms for soil moisture content monitoring. *PeerJ* 7, e6926.
- 717 Ge, Y., Abuduwaili, J., Ma, L., Wu, N., Liu, D., 2016. Potential transport pathways of dust emanating
718 from the playa of Ebinur Lake, Xinjiang, in arid northwest China. *Atmospheric Research* 178-
719 179, 196-206.
- 720 Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. *Pattern*
721 *Recognition Letters* 31(14), 2225-2236.
- 722 Gorji, T., Sertel, E., Tanik, A., 2017. Monitoring soil salinity via remote sensing technology under data
723 scarce conditions: A case study from Turkey. *Ecological Indicators* 74, 384-391.
- 724 Gorji, T., Tanik, A., Sertel, E., 2015. Soil Salinity Prediction, Monitoring and Mapping Using Modern
725 Technologies. *Procedia Earth and Planetary Science* 15, 507-512.
- 726 He, Y.-L., Geng, Z.-Q., Zhu, Q.-X., 2015. Data driven soft sensor development for complex chemical
727 processes using extreme learning machine. *Chemical Engineering Research and Design* 102, 1-
728 11.
- 729 Hong, Y., Chen, S., Zhang, Y., Chen, Y., Yu, L., Liu, Y., Liu, Y., Cheng, H., Liu, Y., 2018. Rapid
730 identification of soil organic matter level via visible and near-infrared spectroscopy: Effects of
731 two-dimensional correlation coefficient and extreme learning machine. *Science of The Total*
732 *Environment* 644, 1232-1243.
- 733 Hong, Y., Liu, Y., Chen, Y., Liu, Y., Yu, L., Liu, Y., Cheng, H., 2019. Application of fractional-order
734 derivative in the quantitative estimation of soil organic matter content through visible and near-
735 infrared spectroscopy. *Geoderma* 337, 758-769.
- 736 Hu, J., Peng, J., Zhou, Y., Xu, D., Zhao, R., Jiang, Q., Fu, T., Wang, F., Shi, Z., 2019. Quantitative
737 Estimation of Soil Salinity Using UAV-Borne Hyperspectral and Satellite Multispectral Images.
738 *Remote Sensing* 11(7), 736.
- 739 Jackson, R.D., 1983. Spectral indices in N-Space. *Remote Sensing of Environment* 13(5), 409-421.
- 740 Jin, X., Du, J., Liu, H., Wang, Z., Song, K., 2016. Remote estimation of soil organic matter content in
741 the Sanjiang Plain, Northeast China: The optimal band algorithm versus the GRA-ANN model.
742 *Agricultural and Forest Meteorology* 218-219, 250-260.
- 743 Jin, X., Xu, X., Song, X., Li, Z., Wang, J., Guo, W., 2013. Estimation of Leaf Water Content in Winter
744 Wheat Using Grey Relational Analysis-Partial Least Squares Modeling with Hyperspectral
745 Data. *Agronomy Journal* 105(5), 1385-1392.
- 746 Khan, N.M., Rastoskuev, V.V., Sato, Y., Shiozawa, S., 2005. Assessment of hydrosaline land degradation
747 by using a simple approach of remote sensing indicators. *Agricultural Water Management* 77(1),
748 96-109.
- 749 Kobayashi, C., Lau, I.C., Wheaton, B., Bourke, L., Kakuta, S., Tachikawa, T., 2015. Mapping of soil
750 salinity using an airborne hyperspectral sensor in Western Australia, 2015 IEEE International
751 Geoscience and Remote Sensing Symposium (IGARSS), pp. 2684-2687.
- 752 Kuang, B., Mouazen, A.M., 2011. Calibration of visible and near infrared spectroscopy for soil analysis
753 at the field scale on three European farms. *European Journal of Soil Science* 62(4), 629-636.
- 754 Kuo, Y., Yang, T., Huang, G.-W., 2008. The use of grey relational analysis in solving multiple attribute
755 decision-making problems. *Computers & Industrial Engineering* 55(1), 80-93.
- 756 Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. *R News* 2(3), 18-22.
- 757 Ling, H., Zhang, P., Guo, B., Xu, H., Ye, M., Deng, X., 2017. Negative feedback adjustment challenges
758 reconstruction study from tree rings: A study case of response of *Populus euphratica* to river

- discontinuous flow and ecological water conveyance. *Science of The Total Environment* 574, 109-119.
- Liu, G., Yin, G., Kurban, A., Aishan, T., You, H., 2016. Spatiotemporal dynamics of land cover and their impacts on potential dust source regions in the Tarim Basin, NW China. *Environmental Earth Sciences* 75(23), 1477.
- Lizaga, I., Gaspar, L., Quijano, L., Dercon, G., Navas, A., 2019a. NDVI, 137Cs and nutrients for tracking soil and vegetation development on glacial landforms in the Lake Parón Catchment (Cordillera Blanca, Perú). *Science of The Total Environment* 651, 250-260.
- Lizaga, I., Quijano, L., Gaspar, L., Ramos, M.C., Navas, A., 2019b. Linking land use changes to variation in soil properties in a Mediterranean mountain agroecosystem. *CATENA* 172, 516-527.
- Lobell, D.B., Lesch, S.M., Corwin, D.L., Ulmer, M.G., Anderson, K.A., Potts, D.J., Doolittle, J.A., Matos, M.R., Baltes, M.J., 2010. Regional-scale Assessment of Soil Salinity in the Red River Valley Using Multi-year MODIS EVI and NDVI. *Journal of Environmental Quality* 39(1), 35-41.
- Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Müller-Wilm, U., Cadau, E., Gascon, F., 2016. SENTINEL-2 SEN2COR: L2A Processor for Users.
- Ma, L., Yang, S., Simayi, Z., Gu, Q., Li, J., Yang, X., Ding, J., 2018. Modeling variations in soil salinity in the oasis of Junggar Basin, China. *Land Degradation & Development* 29(3), 551-562.
- Ma, M., Wang, X., Veroustraete, F., Dong, L., 2007. Change in area of Ebinur Lake during the 1998–2005 period. *International Journal of Remote Sensing* 28(24), 5523-5533.
- Malenovsky, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R., Berger, M., 2012. Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sensing of Environment* 120, 91-101.
- Martens, H., Naes, T., Naes, T., 1992. Multivariate calibration. John Wiley & Sons, New York.
- Meng, L., Zhou, S., Zhang, H., Bi, X., 2016. Estimating soil salinity in different landscapes of the Yellow River Delta through Landsat OLI/TIRS and ETM+ Data. *Journal of Coastal Conservation* 20(4), 271-279.
- Metternicht, G.I., Zinck, J.A., 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sensing of Environment* 85(1), 1-20.
- Morán, J., Granada, E., Míguez, J.L., Porteiro, J., 2006. Use of grey relational analysis to assess and optimize small biomass boilers. *Fuel Processing Technology* 87(2), 123-127.
- Mueller-Wilm, U., Devignot, O., Pessiot, L., 2017. S2 MPC Sen2Cor configuration and user manual. European Space Agency: Paris, France.
- Navarro, G., Caballero, I., Silva, G., Parra, P.-C., Vázquez, Á., Caldeira, R., 2017. Evaluation of forest fire on Madeira Island using Sentinel-2A MSI imagery. *International Journal of Applied Earth Observation and Geoinformation* 58, 97-106.
- Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., 2014. Modeling and Mapping of Soil Salinity with Reflectance Spectroscopy and Landsat Data Using Two Quantitative Methods (PLSR and MARS). *Remote Sensing* 6(11), 10813-10834.
- Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., Shi, Z., 2019. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* 337, 1309-1319.
- Peng, J., Ji, W., Ma, Z., Li, S., Chen, S., Zhou, L., Shi, Z., 2016. Predicting total dissolved salts and soluble ion concentrations in agricultural soils using portable visible near-infrared and mid-infrared spectrometers. *Biosystems Engineering* 152, 94-103.

- Peng, J., Xiang, H.-y., Guo, Y., Shi, Z., 2014. Comparative study on hyperspectral inversion accuracy of soil salt content and electrical conductivity. *Spectroscopy and Spectral Analysis* 34(2), 510-514.
- Qing, K., Yu, R., Li, X., Deng, X., 2005. Application of spectral angle mapping model to rapid assessment of soil salinization in arid area, *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.*, pp. 2355-2357.
- R Development Core Team, 2018. R : A language and environment for statistical computing, R Foundation for Statistical Computing.
- Schoeneberger, P., Wysocki, D., Benham, E., Broderson, W., 2002. *Field Book for Describing and Sampling Soils, Version 2.0.*
- Sibanda, M., Mutanga, O., Dube, T., S Vundla, T., L Mafongoya, P., 2018. Estimating LAI and mapping canopy storage capacity for hydrological applications in wattle infested ecosystems using Sentinel-2 MSI derived red edge bands. *GIScience & Remote Sensing*, 1-19.
- Sibanda, M., Mutanga, O., Rouget, M., 2015. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS Journal of Photogrammetry and Remote Sensing* 110, 55-65.
- Sidike, A., Zhao, S., Wen, Y., 2014. Estimating soil salinity in Pingluo County of China using QuickBird data and soil reflectance spectra. *International Journal of Applied Earth Observation and Geoinformation* 26, 156-175.
- Su, X., Liu, Z., Wei, T., Wang, Y., Liu, Y., 2016. Change of Ebinur Lake Area and Its Response Characteristics of the Runoff Change. *Research of Soil and Water Conservation* 23(3), 252-256 (in Chinese).
- Tian, Y.C., Yao, X., Yang, J., Cao, W.X., Hannaway, D.B., Zhu, Y., 2011. Assessing newly developed and published vegetation indices for estimating rice leaf nitrogen concentration with ground- and space-based hyperspectral reflectance. *Field Crops Research* 120(2), 299-310.
- Triki Fourati, H., Bouaziz, M., Benzina, M., Bouaziz, S., 2015. Modeling of soil salinity within a semi-arid region using spectral analysis. *Arabian Journal of Geosciences* 8(12), 11175-11182.
- Verma, K.S., Saxena, R.K., Barthwal, A.K., Deshmukh, S.N., 1994. Remote sensing technique for mapping salt affected soils. *International Journal of Remote Sensing* 15(9), 1901-1914.
- Wang, H., Tiyp, T., XIE, X., FAN, Y., ZHANG, F., Sawut, M., 2011. Assessment of Soil Salinization Sensitivity for Different Types of Land Use in the Ebinur Lake Region in Xinjiang. *Progress in Geography* 30(5), 593-599 (in Chinese).
- Wang, J., Ding, J., Abulimiti, A., Cai, L., 2018a. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS-NIR) spectroscopy, Ebinur Lake Wetland, Northwest China. *PeerJ* 6, e4703.
- Wang, J., Ding, J., Li, G., Liang, J., Yu, D., Aishan, T., Zhang, F., Yang, J., Abulimiti, A., Liu, J., 2019. Dynamic detection of water surface area of Ebinur Lake using multi-source satellite data (Landsat and Sentinel-1A) and its responses to changing environment. *CATENA* 177, 189-201.
- Wang, J., Tiyp, T., Ding, J., Zhang, D., Liu, W., Wang, F., Tashpolat, N., 2017a. Desert soil clay content estimation using reflectance spectroscopy preprocessed by fractional derivative. *PLOS ONE* 12(9), e0184836.
- Wang, X., Zhang, F., Ding, J., Kung, H.-t., Latif, A., Johnson, V.C., 2018b. 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. *Science of The Total Environment* 615, 918-930.

- Wang, X., Zhang, F., Kung, H.-t., Johnson, V.C., 2018c. New methods for improving the remote sensing estimation of soil organic matter content (SOMC) in the Ebinur Lake Wetland National Nature Reserve (ELWNNR) in northwest China. *Remote Sensing of Environment* 218, 104-118.
- Wang, Y., Liu, Z., Yao, J., Bayin, C., 2017b. Effect of climate and land use change in Ebinur Lake Basin during the past five decades on hydrology and water resources. *Water Resources* 44(2), 204-215.
- Wold, S., Martens, H., Wold, H., 1983. The multivariate calibration problem in chemistry solved by the PLS method. *Matrix Pencils*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 286-293.
- Wu, W., Mhaimeed, A.S., Al-Shafie, W.M., Ziadat, F., Dhehibi, B., Nangia, V., De Pauw, E., 2014. Mapping soil salinity changes using remote sensing in Central Iraq. *Geoderma Regional* 2-3, 21-31.
- Yang, X., Qin, Q., Grussenmeyer, P., Koehl, M., 2018. Urban surface water body detection with suppressed built-up noise based on water indices from Sentinel-2 MSI imagery. *Remote Sensing of Environment* 219, 259-270.
- Zhang, T.-T., Qi, J.-G., Gao, Y., Ouyang, Z.-T., Zeng, S.-L., Zhao, B., 2015. Detecting soil salinity with MODIS time series VI data. *Ecological Indicators* 52, 480-489.
- Zhang, T.-T., Zeng, S.-L., Gao, Y., Ouyang, Z.-T., Li, B., Fang, C.-M., Zhao, B., 2011. Using hyperspectral vegetation indices as a proxy to monitor soil salinity. *Ecological Indicators* 11(6), 1552-1562.
- Zou, X., Zhao, J., Povey, M.J.W., Mel, H., Mao, H., 2010. Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta* 667(1), 14-32.