Chin. Geogra. Sci. 2019 Vol. 29 No. 5 pp. 784–797
https://doi.org/10.1007/s11769-019-1071-x

Springer Science Press
www.springerlink.com/content/1002-0063

Spatial Prediction of Soil Salinity in a Semiarid Oasis: Environmental Sensitive Variable Selection and Model Comparison

LI Zhen[1], LI Yong[1], XING An[1], ZHUO Zhiqing[1], ZHANG Shiwen[2], ZHANG Yuanpei[3], HUANG Yuanfang[1]

(1. *College of Resources and Environment Sciences*, *China Agricultural University*, *Beijing* 100193, *China*; 2. *School of Earth and Environment*, *Anhui University of Science and Technology*, *Huainan* 232001, *China*; 3. *Institute of Crop Sciences*, *Ningxia Academy of Agricultural and Forestry Sciences*, *Yinchuan* 750002, *China*)

Abstract: Timely monitoring and early warning of soil salinity are crucial for saline soil management. Environmental variables are commonly used to build soil salinity prediction model. However, few researches have been done to summarize the environmental sensitive variables for soil electrical conductivity (EC) estimation systematically. Additionally, the performance of Multiple Linear Regression (MLR), Geographically Weighted Regression (GWR), and Random Forest regression (RFR) model, the representative of current main methods for soil EC prediction, has not been explored. Taking the north of Yinchuan plain irrigation oasis as the study area, the feasibility and potential of 64 environmental variables, extracted from the Landsat 8 remote sensed images in dry season and wet season, the digital elevation model, and other data, were assessed through the correlation analysis and the performance of MLR, GWR, and RFR model on soil salinity estimation was compared. The results showed that: 1) 10 of 15 imagery texture and spectral band reflectivity environmental variables extracted from Landsat 8 image in dry season were significantly correlated with soil EC, while only 3 of these indices extracted from Landsat 8 image in wet season have significant correlation with soil EC. Channel network base level, one of the terrain attributes, had the largest absolute correlation coefficient of 0.47 and all spatial location factors had significant correlation with soil EC. 2) Prediction accuracy of RFR model was slightly higher than that of the GWR model, while MLR model produced the largest error. 3) In general, the soil salinization level in the study area gradually increased from south to north. In conclusion, the remote sensed imagery scanned in dry season was more suitable for soil EC estimation, and topographic factors and spatial location also play a key role. This study can contribute to the research on model construction and variables selection for soil salinity estimation in arid and semi-arid regions.

Keywords: soil salinity; environmental variable; random forest regression; geographic weighted regression; Yinchuan Plain irrigation oasis

Citation: LI Zhen, LI Yong, XING An, ZHUO Zhiqing, ZHANG Shiwen, ZHANG Yuanpei, HUANG Yuanfang, 2019. Spatial Prediction of Soil Salinity in a Semiarid Oasis: Environmental Sensitive Variable Selection and Model Comparison. *Chinese Geographical Science*, 29(5): 784–797. https://doi.org/10.1007/s11769-019-1071-x

1 Introduction

As a global environmental hazard, soil salinity adversely affects plant growth, crop production, soil quality and water quality, resulting in loss of farmland and soil degradation, especially in arid and semi-arid irrigated areas (Allbed and Kumar, 2013; El Harti et al., 2016; Ma et al., 2018). The Food and Agriculture Organization of the United Nations has estimated that about 397 million ha land covered by saline soil is distributed across most

Received date: 2018-12-20; accepted date: 2019-04-10
Foundation item: Under the auspices of National Natural Science Foundation of China (No. 41571217), National Program on Key Basic Research Project (No. 2016YFD0300801)
Corresponding author: HUANG Yuanfang. E-mail: yfhuang@cau.edu.cn

continents (Koohafkan and Stewart, 2008). It is necessary to map and monitor soil salinity to establish the areal extent of saline soil and track changes in salinity in order to formulate appropriate and timely management strategies for such soils (Shrestha, 2006).

Conventional methods of measuring soil salinity are based on field sampling and laboratory analysis, which are time-consuming, costly, and inefficient. Moreover, these methods are not feasible for frequent monitoring of large areas (Nanni and Demattê, 2006; Farifteh et al., 2007; Vermeulen and Van Niekerk, 2017). A commonly used alternative method is to use some easily available indices to indirectly reflect soil salinization. For example, five salinity indices, three vegetation indices, three spectral intensity indices, and seven spectral band reflectivity indices extracted from the MODIS Terra data were used to generate a predicted salinity (Bouaziz et al., 2011). Bannari et al. (2008) found that ASTER-SI which combines the SWIR1 and SWIR2 bands showed high correlation to electrical conductivity (EC), while the normalized difference salinity index (NDSI) calculated by the blue and red bands had low correlation with soil EC. Zhang et al. (2015) used the time series normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) to estimate the soil EC. Cai et al. (2010) found that the image texture traits were good indicators for saline soil classification. In addition, Peng et al. (2019) indicated that terrain conditions play an important role in soil salinity estimation. Elnaggar and Noller (2010) found a significant correlation between soil EC and elevation, slope and wetness indices. Vermeeulen et al. (2017) found significant correlations between soil EC and environmental factors such as the distance to drainage, profile curvature, slope and groundwater table depth. Taghizadeh-mehrjardi et al. (2016) found that the wetness indices, the valley bottom flatness index and the elevation were the most important predictors of soil salinity. Furthermore, some scholar considered the evapotranspiration (ET) as a good indicator of soil salination as high ET cause more salt to accumulate on the soil surface (Abou and Ali, 2018). In summary, most of the studies using indirect variables to predict soil salinity mainly choose environmental variables from the following two perspectives: 1) the spectral characteristics of the surface and the overlying cover affected by saline soil, including the spectral reflectance of the bare soil, the spectrum and coverage of plants in the crop growing season (or wet season); 2) the formation factors of the saline soil, such as topographical factors, soil properties and meteorological factors. However, few researches have been done to summarize and assess the environmental sensitive variables for soil (EC) estimation systematically.

Environmental variables are commonly integrated with various methods to predict soil salinity. On the one hand, Statistical analysis, especially linear regression, has great potential in soil EC modeling because it is simple and efficient. Moreover, it is commonly used to evaluate the effectiveness of independent variables in predicting soil EC as it can directly reflect the relationship between influencing factors and soil salinity. For example, Allbed et al. (2014) estimated the soil salinity using Multiple linear regression (MLR) and remote sensing indicators extracted from IKONOS images, and the coefficient of determination (R^2) for all models were higher than 0.65, signifying a high prediction accuracy. On the other hand, a large number of studies have shown that machine learning algorithms, such as artificial neural network (ANN), support vector regression (SVR), decision tree (DT), and especially random forest regression (RFR), can achieve high prediction accuracy in soil properties mapping (Lu et al., 2018). Numerous studies indicate that RFR performs well and can obtain robust estimations. For example, sea surface salinity estimates by ANN and RFR indicated that RFR model was usually stable with different parameters compared to the ANN model (Liu et al., 2013). Wang et al. (2018a) proposed that the RFR exhibited a better performance in predicting SOC stocks than the BRT regardless of input variables. Rodriguez-Galiano et al. (2015) compared the performance of ANN, regression trees, RFR, and SVR for mapping of mineral perspectivity and the results showed that RFR showed higher stability and robustness. Furthermore, some studies indicated that geographically weighted regression (GWR) can obtain accurate predictions by building the spatial non-stationary relationship between dependent and independent variables. Wu et al. (2016) found that GWR had a higher prediction accuracy than MLR and Co-kriging for soil salinity estimation in the Yellow River Delta. Unfortunately, the performance of RFR, MLR, and GWR for predicting soil salinity has not been examined.

The irrigation district in the northern Yinchuan Plain, China is one of the largest and oldest oases in the Yellow

River Basin. Like most irrigation districts in semi-arid areas, soil salinity is one of the primary agronomic constraints. Mapping and monitoring soil salinity are crucial for the development of local precision agriculture. In this study, we firstly summarized 64 environmental variables, including 11 terrain attributes, 12 salinity indices, 30 spectral band reflectivity and imagery texture indices in dry season and wet season, 3 vegetation indices, 3 spectral intensity indices, 3 spatial location factors, evapotranspiration, and the content of silt and clay. Then the optimal variables for the three models were selected by the correlation analysis between the 184 topsoil EC and 64 environmental variables. Lastly, the performance of MLR, GWR, and RFR model on soil salinity prediction was compared. Objectives of this study are to: 1) select the optimal environmental variables used to estimate soil salinity; 2) compare the performance of the MLR, GWR, and RFR on soil salinity prediction; 3) reveal the spatial distribution of soil salinity in the study area. This research can provide reference for similar study and the results may help the government to make appropriate and timely management strategies for saline land.

2 Materials and Methods

2.1 Study area

This study was carried out in the northern Yinchuan Plain irrigation oasis, in Ningxia Hui Autonomous Region, China. The study area encompasses two counties and five districts, Pingluo and Helan counties, Huinong District, Dawukou District, Jinfeng District, Xixia District, and Xingqing District. It lies between the latitudes of 38.27°N and 39.39°N and the longitudes of 105.85°E and 106.98°E, with an east-west width of approximately 51 km, north-south length of approximately 130 km and area of 6967.80 km^2 (Fig. 1). The climate is semi-arid with cold winters and long warm summer. The average annual precipitation and evaporation are 185 mm and 1825 mm, respectively, with a ratio of evaporation to precipitation close to 10 (Dou et al., 2011; Sidike et al., 2014). Due to the high evaporation to precipitation ratio, poor drainage conditions, geologic characteristics and some inappropriate management, soil salinity has always been the primary agronomic constraint to sustainable development in the study area.

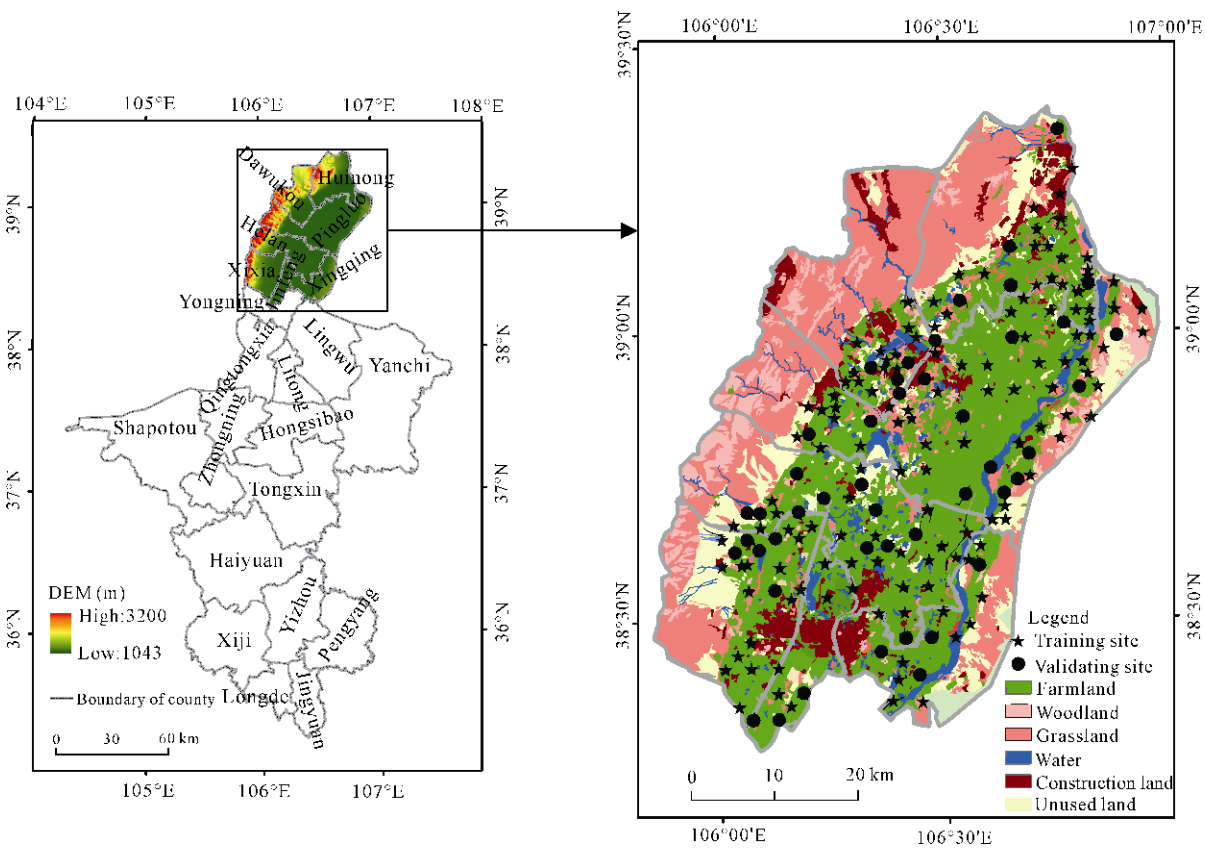

Fig. 1 Study area in the north of Yinchuan Plain irrigation district, Ningxia Hui Autonomous Region, China and spatial distribution of soil samples

2.2 Data source

In this study, the topsoil EC, Landsat 8 remote sensed image, digital elevation model, MOD16 global evapotranspiration product, and soil texture map were used to predict soil salinity.

Firstly, a total of 184 composite topsoil (0–20 cm) samples were collected from March 29 to April 6, 2017 (Fig. 1). Samples were collected using a 2.5 km × 2.5 km grid-sampling method combined with stratified sampling based on land use type in 2015 (The data set is provided by Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (RESDC), http://www.resdc.cn). Each sample location was recorded by the Global Positioning System. In order to reduce the errors resulting from spatial mismatch, each composite soil sample consisted of four sub-samples collected at a distance of 10 m to the east, west, south, and north of the sample location. With water to soil ratio of 5∶1, the soil EC was measured using a conductivity meter (Bao, 2001). The soil salt content (SSC) of 19 samples with different EC levels was measured using a Dezimalwaage after drying as described by Bao (2001).

Secondly, two Landsat 8 images scanned on February 10th and September 6th, 2017 were used to extract environmental variables for building soil salinity prediction model. According to previous studies, dry season (around March, a time when salt surface features are enhanced and the impact of soil salinity on surface attributes is intensified) is the optimum time to predict soil salinity from remote sensing imagery due to the extensive bare soil, while others suggested that soil salinity can be detected by vegetation indices in wet season, namely the plant growing season (Zhou et al., 2012; Zhang et al., 2015). Therefore, the two images scanned in dry season and wet season were selected and the corresponding Landsat path and row were 29 and 33, respectively. In order to remove atmospheric artifacts, atmospheric correction was performed on the image using the Flash model based on ENVI 5.3 software. In addition, the digital elevation model with a spatial resolution of 30 m were used to extract terrain attributes of the study area, and the data can be download from the website of www.usgs.gov.

Lastly, an 8-day composite MODIS MOD16 global evapotranspiration product (download from https://ladsweb.modaps.eosdis.nasa.gov/) acquired on September 6th, 2017 was used to obtain the ET of the study area and it was resampled to 30 m to keep a consistent spatial resolution with Landsat 8 remote sensing image. The soil texture map was provided by a study (Li et al., 2018).

2.3 Environmental variables for soil salinity prediction model construction

On the one hand, four kinds of environmental variables related to the formation of saline soil, including terrain attributes, spatial location, soil matrix, and meteorological factor, were summarized according related studies. To be specific, 11 terrain attributes extracted from the digital elevation model using SAGA GIS software were used to depict the terrain traits which affect soil salt movement and redistribution. The specific analysis refers to Conrad's research (Conrad et al., 2015). Three spatial location factors, namely the distance to the Yellow River and drainage ditch, the longitude and latitude coordinates of the projection coordinate system, were also considered as environment variables as the spatial distribution of soil EC is generally correlated. The content of silt and clay was used to signify the properties of soil matrix. The ET was used to represent the meteorological factor.

On the other hand, referring to previous study, five kinds of environmental variables signifying the spectral characteristic of saline land surface were summarized to build soil salinity prediction model, including the band spectral reflectivity, imagery textures, salinity indices, spectral intensity indices, and vegetation indices. Specifically, 14 bands spectral reflectivity of the two images with the spatial resolution of 30 m in visible and infrared wavelengths, including B1 (Coastal, 0.43–0.45 μm), B2 (Blue, 0.45–0.51 μm), B3 (Green, 0.53–0.59 μm), B4 (Red, 0.64–0.67 μm), B5 (NIR, 0.85–0.88 μm), B6 (SWIR1, 1.57–1.65 μm) and B7 (SWIR2, 2.11–2.29 μm) were used in this study. Sixteen imagery textures of the two images were derived by the grey-level co-occurrence matrix (GLCM) with window size of 3 × 3 pixels based on the first principal of the spectral bands (Lu et al., 2014). Based on the spectral band reflectivity of the remote sensed image in dry season, twelve salinity indices and three spectral intensity indices were calculated and extracted. Moreover, three vegetation indices were extracted from the remote sensed imagery in wet season.

All environmental variables used to build soil salinity estimation model and their abbreviations, calculation

formula, and reference were summarized in Table 1. For variable selection, the importance of the environmental variables was estimated using Spearman correlation coefficient (r), which is commonly used to measure the linear association between variables. The r was calculated using the analysis tool in SPSS 21 software.

Table 1 Environmental variables used to predict soil EC in this study

Covariates sets	Land surface parameters	Abbreviations	Formulations	References
Terrain attributes	Elevation	DEM		www.usgs.gov
	Aspect	AS		Conrad et al. (2015)
	Analytical hill shading	AH		Conrad et al. (2015)
	Channel network base level	CNBL		Conrad et al. (2015)
	Channel network distance	CND		Conrad et al. (2015)
	Closed depression	CD		Conrad et al. (2015)
	Convergence index	CI		Conrad et al. (2015)
	LS factor	LSF		Conrad et al. (2015)
	Relative slope position	RSP		Conrad et al. (2015)
	Topographic wetness index	TWI		Conrad et al. (2015)
	Valley depth	VD		Conrad et al. (2015)
Spatial location	Distance to drainage channel	DC		
	Longitude	LO		
	Latitude	LA		
Meteorological factor	Evapotranspiration	ET		https://ladsweb.modaps.eosdis.nasa.gov/
Soil matrix	Content of silt and clay particle	SC		Li et al. (2018)
Spectral reflectivity	Landsat 8 spectral bands in 10th February, 2017	B1_F; B2_F; B3_F; B4_F; B5_F; B6_F; B7_F;		https://www.usgs.gov/
	Landsat 8 spectral bands in 6th September, 2017	B1_S; B2_S; B3_S; B4_S; B5_S; B6_S; B7_S;		https://www.usgs.gov/
Imagery texture	Landsat 8 imagery texture in 10th February, 2017	Mean_F, Contrast_F, Correlation_F, Dissimilarity_F, Entropy_F, Homogeneity_F, Second Moment_F, Variance_F;		Lu et al. (2014)
	Landsat 8 imagery texture in 6th September, 2017	Mean_S, Contrast_S, Correlation_S, Dissimilarity_S, Entropy_S, Homogeneity_S, Second Moment_S, Variance_S		Lu et al. (2014)
Salinity index	Salinity index 1	SI1	$[B4\_F \times B3\_F]^{0.5}$	Jiang et al. (2019)
	Salinity index 2	SI2	$[B5\_F^2+B4\_F^2+B3\_F^2]^{0.5}$	Jiang et al. (2019)
	Salinity index 3	SI3	$[B4\_F^2+B3\_F^2]^{0.5}$	Jiang et al. (2019)
	Salinity index 4	SI4	$[B4\_F \times B2\_F]^{0.5}$	Khan et al. (2005)
	Salinity index 5	SI5	B2_F/B4_F	Abbas and Khan (2007)
	Salinity index 6	SI6	B6_F/B7_F	Bannari et al. (2008)
	Salinity index 7	SI7	(B3_F×B4_F) / B2_F	Abbas and Khan (2007)
	Salinity index 8	SI8	(B2_F-B4_F)/ (B2_F+B4_F)	Abbas and Khan (2007)
	Salinity index 9	SI9	(B4_F×B5_F) / B3_F	Abbas and Khan (2007)
	Salinity index 10	SI10	(B2_F×B4_F) / B3_F	Abbas and Khan (2007)
	Normalized difference salinity index	NDSI	(B4_F-B5_F)/ (B4_F+B5_F)	Bannari et al. (2008)
	Salinity index ASTER	ASTER-SI	(B6_F-B7_F)/ (B6_F+B7_F)	Bannari et al. (2008)
Spectral intensity	Intensity within the visible range	INT1	(B3_F+B4_F) / 2	Douaoui et al. (2006)
	Intensity within the VIS-NIR range	INT2	(B3_F+B4_F+B5_F) / 2	Douaoui et al. (2006)
	Brightness index	BI	$[B4\_F^2+B5\_F^2]^{0.5}$	Douaoui et al. (2006)
Vegetation index	Normalized difference vegetation index	NDVI	(B5_S-B4_S)/ (B5_S+B4_S)	Bouaziz et al. (2011)
	Soil adjusted vegetation index	SAVI	[(B5_S×B4_S) × (1+L1)] / [B5_S+B4_S+L]	Bouaziz et al. (2011)
	Enhanced vegetation index	EVI	2.5×(B5_S-B4_S) / (B5+c1×B4_S-c2×B2_S+L2)	Bouaziz et al. (2011)

Notes: B1_F–B7_F represents the B1–B7 spectral band reflectivity of the Landsat 8 remote sensed image in February 10th, 2017; B4_S and B5_S represent the red and near-red band spectral reflectivity of the Landsat 8 remote sensed image in September 6th, 2017; L1, L2, C1, and C2 are constant equal to 1, 0.5, 6, 7.5, respectively.

2.4 Models and their accuracy comparison for soil salinity estimation

In this study, MLR, GWR, and RFR were used to estimate EC. MLR is a classical method widely used to predict dependent variable values from independent variables. In this study, MLR was realized in R statistics tool.

The GWR model is an extension of the general linear regression model by introducing geographic positions of data into the regression parameters, which allows GWR to estimate the relationship between the response and the covariates locally rather than globally (Li et al., 2012). In the GWR model, local regression coefficients at each location in the study area are derived based on weighted least squares regression, and the spatial extent used for modeling was determined by a given neighborhood or kernel function (Wu et al., 2016). The weight of a sample is a function of the bandwidth and the distance between the position of the estimated point and the observed point in geographic space. The neighborhood size is also determined by the bandwidth. Therefore, the bandwidth is critical for the estimation of regression parameters. Common methods used for selecting bandwidth include cross-validation and the corrected Akaike information criterion (AIC) (Lu et al., 2018). In this study, a Gauss function was chosen as the spatial weighting function, and AIC was selected to optimize the kernel extent. All these procedures were implemented using the spatial statistics tools of ArcGIS 10.5 software.

RFR is an enhancement of DT that generates each regression tree by using a random vector sampled independently from the input vector, making it an ensemble learning method (Immitzer et al., 2012). Both categorical and continuous predictor variables or response variables are allowed in the model. In this research, the randomForest package in R statistics software was used to develop the RFR model (Breiman, 2001). Two parameters, namely n_{tree} (the number of trees) and m_{try} (the number of input variables per node), have a critical impact on model performance. In order to optimize these two parameters, several combinations of n_{tree} and m_{try} were tested, and the optimal parameters value were selected according to the lowest Root Mean Square Error (*RMSE*). Finally, the selected values of n_{tree} and m_{try} in this study were 2000 and 16, respectively. A more detailed discussion of RFR can be found in (Breiman, 2001).

All 184 samples were sorted randomly into two parts using the Geo-statistics analysis tool in ArcGIS 10.5, which can guarantee the spatial consistence of training sets and validation sets. In the end, 75% samples were used for modeling, and the remaining samples were used for validation. The *RMSE*, Mean Absolute Error (*MAE*), coefficient of determination (R^2) and Ratio of Prediction to Deviation (*RPD*) were selected to evaluate the accuracy of the constructed models. In general, lower *RMSE* and *MAE* signify better prediction results (Huang et al., 2019). The values of R^2 and *RPD* higher than 0.91 and 2.5, respectively signify an accurate prediction. A R^2 value between 0.82 and 0.91, and *RPD* higher than 2 indicate a good prediction. If R^2 is between 0.66–0.82 and *RPD* is higher than 1.5, the prediction seems to be false. A R^2 value between 0.5 and 0.66 indicates a poor relationship (Farifteh et al., 2007).

2.5 Soil salinization classification

Due to the strong correlation between EC and SSC, a linear regression model was established with an R^2 of 0.99 (Fig. 2), which indicated that the model could be used to infer SSC from EC. Therefore, EC was used to classify the degree of soil salinization based on the Chinese soil salinization standard (Zhang et al., 2009) and the relationship between EC and SSC. The grading standard is shown in Table 2, which is close to the classification criterion in the Hetao irrigation district near our study area (Yu et al., 2010).

Table 2 Soil salinity classes and their effect on crops in the north of Yinchuan Plain irrigation district, 2017

Salinity class	Soil EC (dS/m)	SSC (g/kg)	Salinity impacts on crops
Non-saline	<0.3568	<1	Salinity impacts are negligible
Slightly saline	0.3568–0.5963	1–2	Affect production slightly for very sensitive crops
Moderately saline	0.5963–1.0753	2–4	Many crops are affected and the yield is restricted
Highly saline	1.0753–2.5123	4–10	Only tolerant crops bear this condition
Extremely saline	≥2.5123	≥10	Only a few very tolerant crops resist

Notes: EC, electrical conductivity; SSC, soil salt content

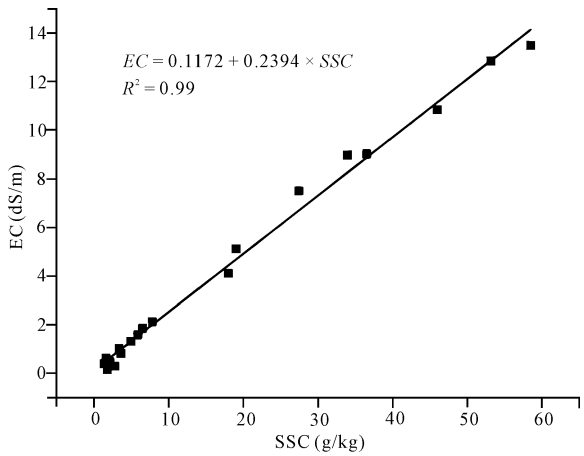

Fig. 2 The linear relationship between soil electrical conductivity (EC) and soil salt content (SSC) of the 19 samples with different salinity level in the study area in 2017

3 Results

3.1 Descriptive statistics

Table 3 shows the descriptive statistics for soil EC of the training sites, validation sites and the whole sample set. According to variance analysis, there was no significant difference between the three datasets. All datasets presented medium spatial variation according to the standard of variability (Huang et al., 2019). Additionally, skewed distribution was observed for all datasets, while the Log-transformed soil EC showed approximately normal distribution. Therefore, the model was constructed based on the Log-transformed data, and then the estimated EC was obtained using the antilogarithmic function.

3.2 Correlation between environmental variables and measured EC

The correlation coefficients between environmental variables and measured EC value are shown in Table 4. Among all the environmental variables, CNBL had the largest absolute *r* value of 0.47, which was followed by DEM. Therefore, the terrain attributes have great potential for soil salinity estimation. From the perspective of imagery texture, there were no significant correlations between EC and image texture indices derived from the remote sensed imagery scanned in September except for the Mean_S index, while there were strong correlations between EC and imagery texture indices obtained from the imagery scanned in February. Therefore, imagery texture index derived from the imagery scanned in dry season was suggested for estimating soil salinity rather than the imagery scanned in wet season. In addition, the reflectivity of seven spectral bands from the two images had significant difference in correlation with soil EC. On one hand, the reflectivity of visible spectral bands from the imagery scanned in February was positively correlated with EC, while there was no significant correlation between soil EC and the reflectivity of visible spectral bands from imagery scanned in September. On the other hand, for infrared wavelength bands, the relationship was contrary to the correlation in visible band. The results show that it would be better to use the bands in visible wavelength to reflect soil salinity in dry season while the infrared band would be more suitable in wet season. Notably, longitude and latitude were also significantly and positively correlated with soil EC, indicating that EC increased from west to east and from south to north, which was consistent with the Yellow River flow direction and the local situation. Additionally, consistent with previous studies, the content of silt and clay was significantly and positively correlated with soil EC. Interestingly, the ET was insignificantly and negatively correlated with soil EC, which is contrary with the known theory. This is likely due to scale conversion or because single-phase images are not representative. Similarly, the three vegetation indexes

Table 3 Statistics of soil EC in training, validation and total datasets

Data sets	Min	Max	Mean	SD	CV	Raw data			Log-transformed data		
						Skew	*Kurt*	*P-value*	*Skew*	*Kurt*	*P-value*
Training sites	0.09	13.49	1.07[a]	2.34	0.46	4.41	20.54	0.00	0.56	–0.15	0.08
Validation sites	0.05	12.38	1.23[a]	2.12	0.58	3.39	12.29	0.00	1.18	1.57	0.41
Total sample	0.05	13.49	1.19[a]	2.17	0.55	3.65	14.28	0.00	0.683	0.097	0.063

Notes: Min, minimum; Max, maximum; SD, standard deviation; CV, coefficient of variation; Skew, Skewness; Kurt, Kurtosis; the letter 'a' above each Mean indicates insignificant difference among the treatments at the level of $P < 0.05$.

Table 4 Correlation coefficients between environmental variables and measured EC

Factor	*r*	Factor	*r*	Factor	*r*	Factor	*r*
DEM	–0.43**	SI1	0.15*	Mean_F	0.079	Mean_S	0.16*
AS	0.10	SI2	0.13	Contras_F	0.24**	Contrast_S	0.06
AH	0.01	SI3	0.15*	Correlation_F	0.09	Correlation_S	–0.04
CNBL	–0.47**	SI4	0.17*	Dissimilarity_F	0.24**	Dissimilarity_S	0.07
CND	–0.24**	SI5	0.16*	Entropy_F	0.26**	Entropy_S	0.05
CD	0.12	SI6	–0.13	Homogeneity_F	–0.24**	Homogeneity_S	–0.06
CI	–0.06	SI7	0.12	Second Moment_F	–0.26**	Second Moment_S	–0.03
LSF	–0.13	SI8	0.16*	Variance_F	0.28**	Variance_S	0.08
RSP	–0.22**	SI9	0.03	B1_F	0.17*	B1_S	0.00
TWI	0.13	SI10	0.17*	B2_F	0.18*	B2_S	0.02
VD	–0.27**	NDSI	0.22**	B3_F	0.17*	B3_S	0.04
DC	–0.32**	ASTR_SI	–0.13	B4_F	0.15*	B4_S	0.03
EVI	–0.05	BI	0.11	B5_F	0.074	B5_S	–0.07
NDVI	–0.04	INT1	0.15*	B6_F	–0.03	B6_S	–0.23*
SAVI	–0.04	INT2	0.14	B7_F	0.01	B7_S	–0.15*
LO	0.20**	LA	0.25**	SC	0.31**	ET	–0.07

Notes: ** Significant at the 0.01 probability level. * Significant at the 0.05 probability level. Meanings of all factors see Table 1

calculated based on the remote sensed images scanned in September 2017 were not significantly correlated with soil EC. In short, out of 64 environmental variables, the correlations between 30 covariates and soil EC were significant with $P < 0.05$. Consequently, the 30 environmental variables were used to construct models.

3.3 Model construction and evaluation

The multicollinity among variables affected the MLR and the GWR models. Therefore, variables with Variance Inflating Factor (VIF) greater than 5 were excluded. Finally, only SC, Variance_F, NDSI, VD, DC and LO were finally fitted into the MLR and GWR models. The established MLR and GWR models are shown in Equations (1) and (2). The coefficient distribution map of GWR model is shown in Fig. 3. All significantly correlated variables were used to establish the RFR model, and the importance of environmental variables is shown in Fig. 4.

$$\lg EC = -40.33 + 33.33X_{\mathrm{SC}} + 85.34X_{\mathrm{NDSI}} - 1589X_{\mathrm{VD}} - 94400X_{\mathrm{DC}} + 2297000X_{\mathrm{LO}} + 686.4X_{\mathrm{Variance\_F}} \quad (1)$$

$$\lg EC = -\beta_0 + \beta_{\mathrm{SC}}X_{\mathrm{SC}} + \beta_{\mathrm{NDSI}}X_{\mathrm{NDSI}} + \beta_{\mathrm{VD}}X_{\mathrm{VD}} + \beta_{\mathrm{DC}}X_{\mathrm{DC}} + \beta_{\mathrm{Variance\_F}}X_{\mathrm{Variance\_F}} \quad (2)$$

where β_m is the local regression coefficient and m is the corresponding independent variable. In this study, the independent variables were SC, NDSI, VD, DC and Variance_F, respectively.

The *RMSE*, *MAE*, R^2 and *RDP* (Ratio of Prediction to Deviation) values of the three models for training datasets and validation datasets are shown in Table 5. In general, the prediction accuracy of RFR model was higher than that of GWR model and the prediction accuracy of the latter was higher than that of MLR model. However, regarding the R^2 and *RDP* values of the validation samples, the prediction results were unreliable irrespective of the model, although the fitting effect of the predicted value and the measured value of the RFR model was slightly better than that of the other two models. In addition, by comparing the R^2 and *RDP* of the RFR model in the training datasets and the validation datasets, it was found that the RFR model had overfitting effect.

3.4 Soil salinity maps derived from constructed model

The soil salinity maps obtained by MLR, GWR, and RFR models are shown in Fig. 5. The spatial distribution pattern of different levels of saline soil predicted by MLR and GWR was consistent, with a gradually increasing trend from south to north. The highly saline soil and extremely saline soil predicted by the three models were all distributed in the west-central part of

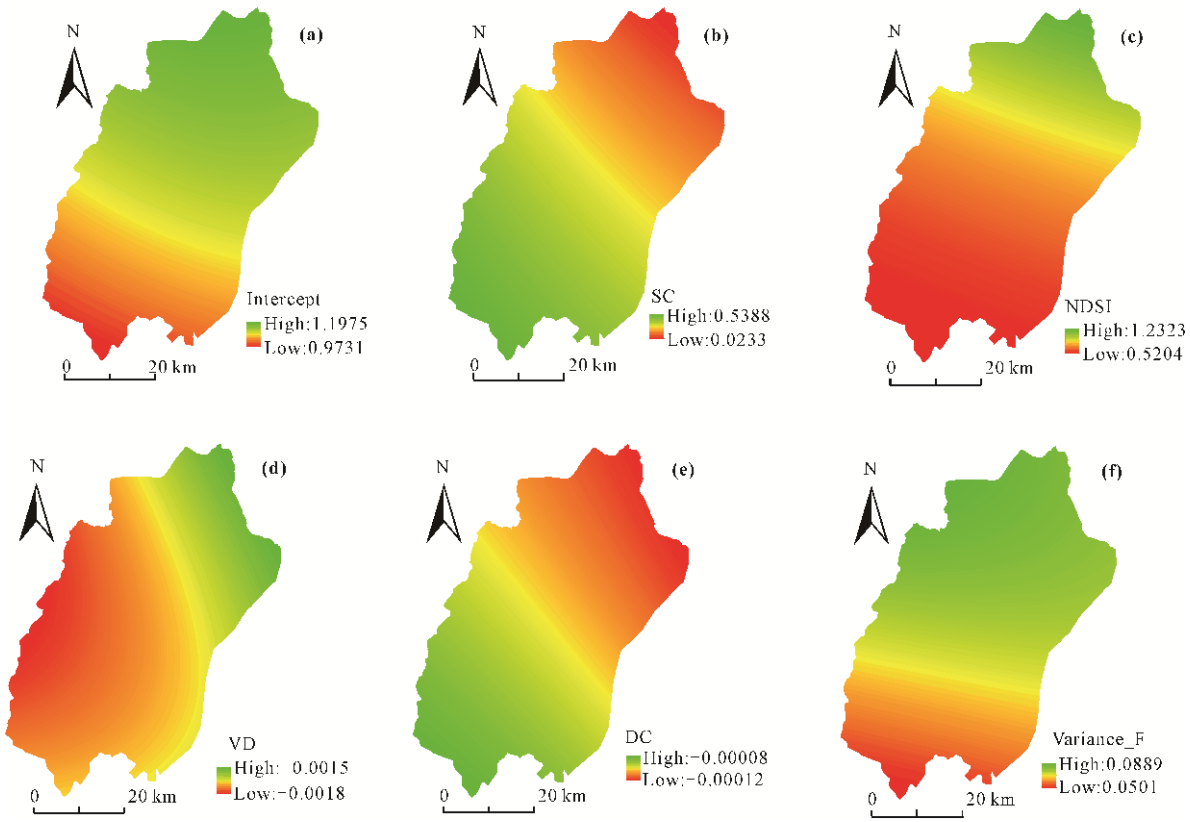

Fig. 3 Map of local regression coefficient and intercept of the geographical weighted regression model. (a) the intercept; (b) the local regression coefficient of the content of sand and clay (SC); (c) the local regression coefficient of the normalized difference salinity index (NDSI); (d) the local regression coefficient of the Valley depth (VD); (e) the local regression coefficient of the Distance to drainage (DC); (f) the local regression coefficient of the Variance_F

the study area, but the predicted areas by different models were different. According to the field investigation, the low terrain and high clay content in this local area restricted soil salt drainage.

Table 6 reflects the area proportion of different grades saline soil predicted by the above models and the proportion of soil samples with different levels of salinity. As can be seen from Table 6 and Fig. 5, the smoothing effect of the RFR model was obvious, which was mainly manifested as the underestimation of highly saline soil and extremely saline soil as well as the overestimation of non-saline soil. The area ratio of highly and extremely saline soil predicted by the GWR model was consistent with that of the highly saline soil and extremely saline soil sample points. This may be related to the concentrated distribution of highly or extremely saline soils, which were easily accumulated in low-lying and poorly drained areas. Thus, the local regression model used by the GWR can extract effective information in this local area. MLR model also had certain smoothing effect, but it was better than RFR model.

4 Discussion

4.1 Effects of band spectral reflectivity on the correlation between soil EC and salinity indices

In this study, NDSI had the highest value of correlation coefficient with soil EC among the 12 salinity indices, indicating that it played an important role in soil salinity mapping. Similarly, the results of Khan et al. (2005) also indicated that NDSI had great potential to estimate EC in dry season. Moreover, among the salinity indices calculated by the visible spectral bands reflectance (SI1, SI2, SI3, SI4, SI5, SI7, SI8, SI9, and SI10), SI1 and SI3 were significantly correlated with EC, which confirmed that SI1–SI3 indices proposed by the Indo Dutch Network Project were suitable to assess salinity

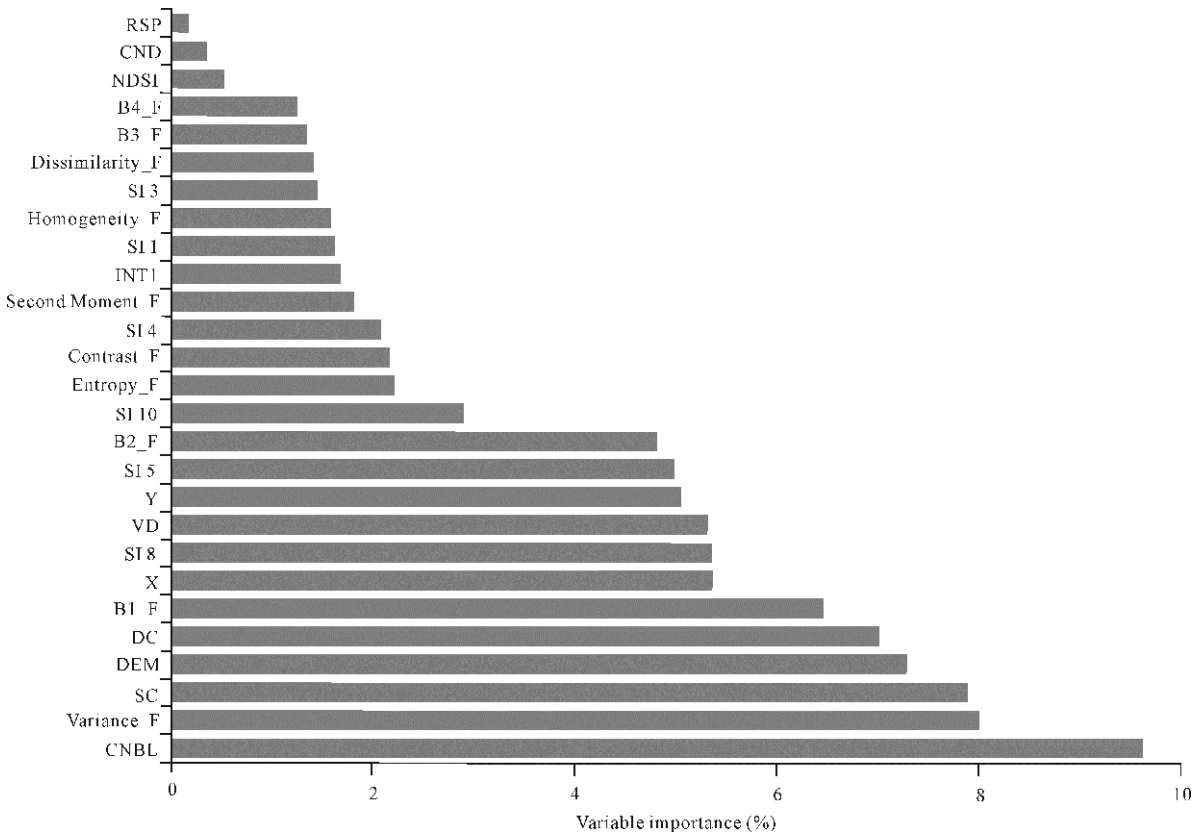

Fig. 4 Importance of environmental variables used in the random forest regression (RFR) model. Each variable was scaled to sum to 100%

Table 5 Prediction accuracy comparison of three models

Data sets	MLR				GWR				RFR			
	RMSE	*MAE*	R^2	*RDP*	*RMSE*	*MAE*	R^2	*RDP*	*RMSE*	*MAE*	R^2	*RDP*
Training sites	0.43	0.35	0.34	0.95	0.41	0.34	0.37	0.99	0.17	0.13	0.93	2.83
Validation sites	0.42	0.33	0.24	1.10	0.41	0.32	0.26	1.11	0.35	0.28	0.45	1.25

conditions based on Landsat-TM data (Bouaziz et al., 2011). Furthermore, the significant correlation coefficient of SI4, SI5, and SI8 indicated that blue and red bands had good potential to estimate soil salinity. Similar to the findings of this study, Fan et al. (2015) found that blue and NIR bands from the Advanced Land Imagery were positively correlated with soil salinity. Unfortunately, SI7 and SI9, which were critical factors for predicting EC in other studies, were not significantly correlated with EC in our study area (Abbas and Khan, 2007; Peng et al., 2019). In addition, except for NDSI, the correlation coefficients were all less than 0.2 (Table 2). This may be related to the large scope of the study area and complex ground conditions. In practice, some soil samples were taken from farmland including dry land and paddy fields, while other samples were taken from fallow land and grassland in this study. In addition, some studies indicated that the salinity indices calculated by the SWIR1 and SWIR2 band reflectance, such as ASTER-SI, had a strong correlation with soil EC (Bannari et al., 2008). However, ASTER-SI were not significantly correlated with soil EC in this study. This was likely because these salinity indices were calculated based on the imagery scanned in February. In the remote sensed imagery scanned in February, soil EC was positively and significantly correlated with the spectral band reflectivity in visible wavelength while it was not significantly correlated with bands reflectance in infrared wavelength (Table 2). In contrast, considering the imagery scanned in September, the reflectivity of the

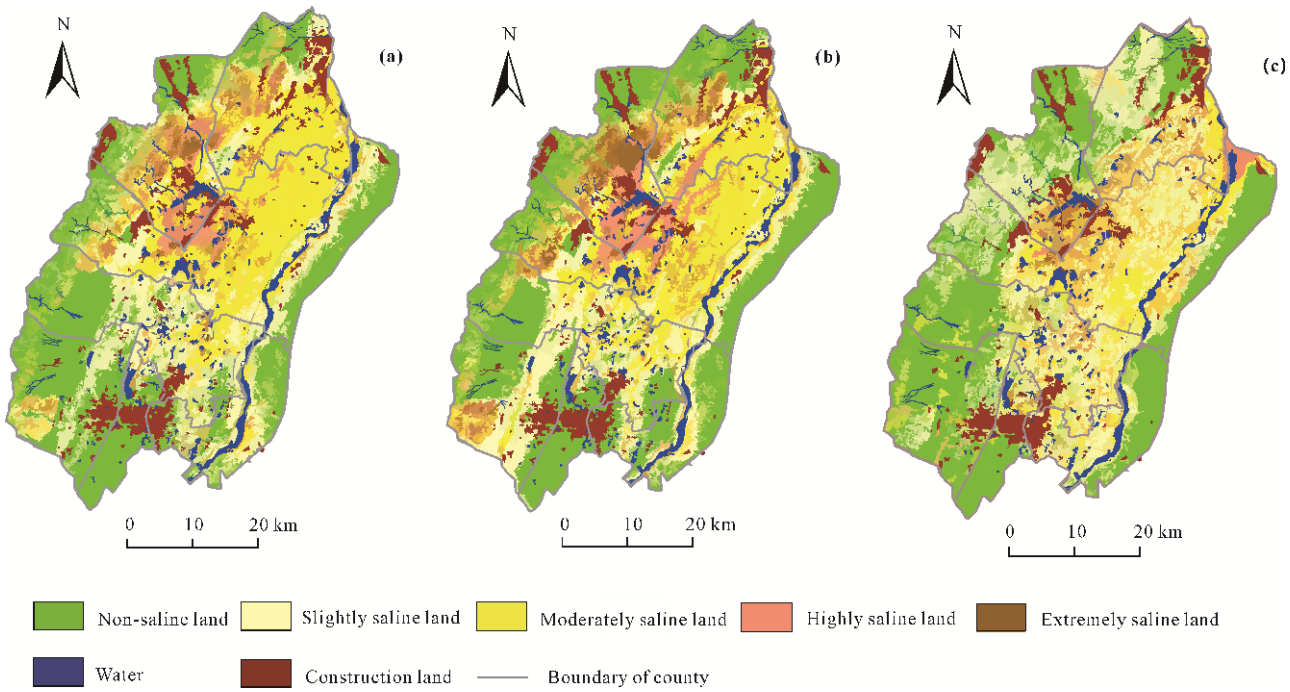

Fig. 5 Soil salinity maps predicted by (a) multiple linear regression model, (b) geographical weighted regression model, and (c) random forest regression model in the north of Yinchuan Plain irrigation district

Table 6 Proportion of estimated area and sampled points for different level of saline soil in the north of Yinchuan Plain irrigation district (%)

Method	Non-saline	Slightly saline	Moderately saline	Highly saline	Extremely saline
MLR	31.27	25.76	27.30	10.75	4.92
GWR	28.85	26.99	22.96	11.95	9.85
RFR	34.51	36.72	19.32	7.31	2.14
Sampled point	42.93	17.93	15.76	14.67	8.70

infrared band was significantly correlated with soil EC, while the reflectivity of the visible band was not significantly correlated. Peng et al. (2019) predicted soil EC using Landsat 8 imagery that was scanned in July 2016, and found that ASTER-SI and SI6 were significantly correlated with soil EC. Therefore, considering the significant differences in the correlation coefficient between spectral band reflectance and soil EC in dry season and wet season, this study concluded that the infrared band was applicable for the prediction of soil salinity in plant growing season, while the visible band reflectance was more applicable in dry season.

4.2 Performance of imagery texture, terrain attributes, and spatial location in soil EC prediction

Among the imagery texture indices, Variance_F had the largest correlation coefficient with soil EC and it was also an important independent variable in the three models. In general, the cells with high value of Variance_F are mainly distributed in the junction of different landscape types, like farmland and village land. Moreover, Variance_F gradually decreases as the distance between farmland and other landscape increases. In fact, the soil EC value also decreased as the distance to village or ditch or some land without irrigation increased. This phenomenon was consistent with the positive correlation coefficient between the Variance_F index and soil EC (Table 4). In essence, this may be because the water flowed from the irrigated area (farmland) to the non-irrigated area and brought salt at the same time, which made the non-irrigated area more salinized. (Konukcu et al., 2006; Gill et al., 2016). Therefore, the direct reason for the significant correlations between soil salinity and imagery texture attributes may be the

landscape variety, and the indirect reason may be the difference in salt accumulation in different landscapes.

Among the terrain attributes, CNBL was the most important environmental variable for the RFR model and both VD and DC were critical independent variables in the three models, indicating that terrain characteristics played a vital role in soil salinity estimation. In addition, DEM, CNBL, CND, RSP, and VD were significantly and negatively correlated with EC in this study, which was consistent with the research results of Peng et al. (2019). In essence, these indicators reflect the height and slope of the terrain and they affect the direction and rate of water flow. Generally, low terrain and small slope result in salt accumulation. Furthermore, consistent with previous conclusions, DC was negatively correlated with EC. In other words, soil salinization was more severe with less distance to the drainage ditch (Zhou et al., 2012). Therefore, topography attributes should be considered to predict soil salinity.

4.3 Model performance on soil salinity estimation

Similar to a precious study, the performance of GWR model was slightly better than that of MLR model in this study (Wu et al., 2016). However, both models had small values of R^2. This may be related to the large scope of the study area and the complex surface landscape type. For instance, Shrestha et al. (2006) used the covariates extracted from Landsat 7 to predict soil EC and found that the R^2 of MLR was 0.23. In another study, Allbed et al. (2014) used OLI images to predict EC in an area of about 20,000 ha and found that the R^2 of MLR was greater than 0.65. Another possible reason may be the scale difference between the ground sampling point and the Landsat 8 image pixel. For example, the reason for high R^2 (more than 0.7) values in some studies was that they built models between apparent conductivity, obtained by electromagnetic induction, and variables derived from remote sensing imagery (Whitney, et al., 2018). Therefore, for soil salinity mapping in large scale, the complexity of surface landscape types should be taken into account, followed by the scale difference between sampling point and the spatial resolution of remote sensed imagery.

Unexpectedly, it was contrary to the conclusion that the prediction accuracy of RFR model is much higher than that of linear regression model (Lu et al., 2018; Wang et al., 2018a). In this study, the prediction accuracy of RFR model for validation datasets was only slightly higher than that of the GWR and MLR models. In addition, there was severe over-fit effect considering the model performance in training dataset and validation dataset. This may be due to the uneven distribution of samples in different salinization levels in this study. In practice, the proportion of non-salinized sample was 42.93%, while that of extremely salinized samples was 8.7%. On the other hand, RFR model adopts the voting scheme to determine the prediction value based on the results of different decision trees (Wang et al., 2018b). Therefore, the prediction results could be easily made smaller when the non-saline soil samples accounted for a large proportion. In conclusion, the sampling scheme can have a great impact on the model performance.

5 Conclusions

In this study, 64 environmental variables related to the formation of saline soil (e.g., terrain, soil matrix and climate) and the apparent spectral characteristics of saline land were summarized to estimate the soil EC. Then the applicability of the 64 indices were evaluated through the correlation coefficient and the prediction accuracies of RFR, MLR and GWR models were compared. The main conclusions of this study are as follows. 1) the remote sensed imagery scanned in dry season was more suitable for soil EC estimation. Among the 15 imagery texture and spectral band reflectivity factors extracted from Landsat 8 image in dry season, 10 indices were significantly correlated with soil EC. however, only 3 of these indices extracted from Landsat 8 image in wet season have significant correlation with soil EC. 2) Channel network base level, one of the terrain attributes, had the largest absolute correlation value of 0.47 and all spatial location factors had significant correlation with soil EC. However, ET extracted from single-phase MODIS images was negatively correlated with EC and vegetation indices were not significantly correlated with soil EC. 3) The comparative results showed that the prediction accuracy of the RFR model was slightly higher than that of the GWR model, and prediction accuracy of the latter was superior to that of the MLR model. However, based on the low *RDP* values of the validated dataset, the predictions of all three models appeared to be unreliable. This study provides reference for the environmental variable selection and

model construction of soil salinization prediction in an arid or semi-arid oasis. In future studies, prediction accuracy could be improved by taking the complex landscape into consideration.

References

Abbas A, Khan S, 2007. Using remote sensing techniques for appraisal of irrigated soil salinity. MODSIM 2007: International Congress on Modelling and simulation: Land, Water and Environmental Management: Integrated Systems for Sustainability.

Abou Samra Rasha M, Ali R R, 2018. The development of an overlay model to predict soil salinity risks by using remote sensing and GIS techniques: a case study in soils around Idku Lake, Egypt. *Environmental Monitoring and Assessment*, 190(12): 706–722. doi: 10.1007/s10661-018-7079-3

Aldabaa A A A, Weindorf D C, Chakraborty S et al., 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. *Geoderma*, 239: 34–46. doi: 10.1016/j.geoderma.2014.09.011

Allbed A, Kumar L, 2013. Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: a review. *Advances in Remote Sensing*, 2: 373–385. doi: 10.4236/ars. 2013.24040

Allbed A, Kumar L, Sinha P, 2014. Mapping and modelling spatial variation in soil salinity in the Al Hassa oasis based on remote sensing indicators and regression techniques. *Remote Sensing*, 6: 1137–1157. doi: 10.3390/rs6021137

Bannari A, Guedon A M, El-Harti A et al., 2008. Characterization of slightly and moderately saline and sodic soils in irrigated agricultural land using simulated data of advanced land imaging (EO-1) sensor. *Communications in Soil Science and Plant Analysis*, 39: 2795–2811. doi: 10.1080/ 0010362080243271

Bao Shidan, 2000. *Soil and Agricultural Chemistry Analysis*. Beijing: Chinese Agricultural press. (in Chinese)

Bouaziz M, Matschullat J, Gloaguen R, 2011. Improved remote sensing detection of soil salinity from a semi-arid climate in Northeast Brazil. *Comptes Rendus Geoscience*, 343: 795–803. doi: 10.1016/j.crte.2011.09.003

Breiman L, 2001. Classification and regression by randomForest. *Machine Learning*, 45(1): 5–32. doi: 10.1023/a: 1010933404324

Cai S M, Zhang R Q, Liu L M et al., 2010. A method of salt-affected soil information extraction based on a support vector machine with texture features. *Mathematical and Computer Modelling*, 51: 1319–1325. doi: 10.1016/j.mcm. 2009.10.037

Conrad O, Bechtel B, Bock M et al., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development Discussions*. 8(2): 2271–2312. doi: 10.5194/gmdd-8-2271-2015

Dou C Y, Kang Y H, Wan S Q et al., 2011. Soil salinity changes under cropping with lycium barbarum l. and irrigation with saline-sodic water. *Pedosphere*, 21: 539–548. doi: 10.1016/S1002-0160(11)60156-2

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: 217–230. doi: 10.1016/j.geod erma.2005.10.009

El Harti A, Lhissou R, Chokmani K et al., 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. doi: 10.1016/j.jag. 2016.03.008

Elnaggar Abdelhamid A, Noller Jay S, 2010. Application of remote-sensing data and decision tree analysis to mapping salt-affected soils over large areas. *Remote Sensing*, 2: 151–165. doi: 10.3390/rs2010151

Fan X W, Liu Y B, Tao J M et al., 2015. Soil salinity retrieval from advanced multi-spectral sensor with partial least square regression. *Remote Sensing*, 7: 488–511. doi: 10.3390/rs70100488

Farifteh J, Van der Meer F, Atzberger C et al., 2007. Quantitative analysis of salt-affected soil reflectance spectra: a comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, 110: 59–78. doi: 10.1016/j.rse.2007.02.005

Gill Bruce C, Terry Alister D, 2016. Keeping salt on the farm-Evaluation of an on-farm salinity management system in the Shepparton irrigation region of South-East Australia. *Agricultural Water Management*, 164: 291–303. doi: 10.1016/j.agwat.2015.10.014

Huang Yajie, Li Zhen, Ye Huichun et al., 2019. Mapping soil electrical conductivity using Ordinary Kriging combined with Back-propagation network. *Chinese Geographical Science*, 29(2): 270–282. doi: 10.1007/s11769-019-1027-1

Immitzer M, Atzberger C, Koukal T, 2012. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sensing*, 4: 2661–2693. doi: 10.3390/rs4092661

Jiang H, Rusuli Y, Amuti T, et al., 2019. Quantitative assessment of soil salinity using multi-source remote sensing data based on the support vector machine and artificial neural network. *International Journal of Remote Sensing*, 40(1): 284–306, doi: 10.1080/01431161.2018.1513180

Khan N M, Rastoskuev V V, Sato Y et al., 2005. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agricultural Water Management*, 77: 96–109. doi: 10.1016/j. agwat.2004.09.038

Konukcu F, Gowing J W, Rose D A, 2006. Dry drainage: A sustainable solution to waterlogging and salinity problems in irrigation areas? *Agricultural Water Management*, 83: 1–12. doi: 10.1016/j.agwat.2005.09.003

Koohafkan P, Stewart B A, 2008. Water and Cereals in Drylands. The Food and Agriculture Organization of the United Nations and Earth scan.

Wang K, Zhang C R, Li W D, 2012. Comparison of geographically weighted regression and regression kriging for estimating the spatial distribution of soil organic matter. *GIScience and Remote Sensing*, 49: 915–932. doi: 10.2747/1548-1603. 49.6.915

Li Zhen, Zhang Shiwen, Cao Meng et al., 2018. Spatial interpolation of soil mechanical composition based on the spherical coordinate transform method. *Transactions of the Chinese society for Agricultural Machinery*, 49(03): 295–302. (in Chinese)

Liu M L, Liu X N, Jiang J L et al., 2013. Artificial Neural Network and Random Forest Approaches for Modeling of Sea Surface Salinity. *International Journal of Remote Sensing Applications*, 3(4): 229–235. doi: 10.14355/ijrsa.2013.0304.08

Lu D S, Li G Y, Moran E et al., 2014. The roles of textural images in improving land-cover classification in the Brazilian Amazon. *International Journal of Remote Sensing*, 35: 8188–8207. doi: 10.1080/01431161.2014.980920

Lu W, Lu D S, Wang G X G et al., 2018. Examining soil organic carbon distribution and dynamic change in a hickory plantation region with Landsat and ancillary data. *Catena*, 165: 576–589. doi: 10.1016/j.catena.2018.03.007

Ma L G, Yang S T, Simayi Z et al., 2018. Modeling variations in soil salinity in the oasis of Junggar Basin, China. *Land Degradation and Development*, 29: 551–562. doi: 10.1002/ldr.2890

Nanni M R, Demattê J A M, 2006. Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Science Society of America Journal*, 70: 393–407. doi: 10.2136/sssaj2003.0285

Peng J, Biswas A, Jiang Q S et al., 2019. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma*, 337: 1309–1349. doi: 10.1016/j.geoderma. 2018.08.006

Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M et al., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71: 804–818. doi: 10.1016/j.oregeorev.2015.01.001

Shrestha R P, 2006. Relating soil electrical conductivity to remote sensing and other soil properties for assessing soil salinity in northeast Thailand. *Land Degradation and Development*, 17: 677–689. doi: 10.1002/ldr.752

Sidike A, Zhao S H, Wen Y M, 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. doi: 10.1016/j.jag. 2013.06.002

Taghizadeh-Mehrjardi R, Ayoubi S, Namazi Z et al., 2016. Prediction of soil surface salinity in arid region of central Iran using auxiliary variables and genetic programming. *Arid Land Research and Management*, 30(1): 49–64. doi: 10.1080/15324982. 2015.1046092

Vermeeulen D, Van Niekert A, 2017. Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates. *Geoderma*, 299: 1–12. doi: 10.1016/j. geoderma.2017.03.013

Wang B, Waters C, Orgill S et al., 2018a. Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. *Ecological Indicators*, 88: 425 – 438. doi: 10.1016/j. ecolind.2018.01.049

Wang B, Waters C, Orgill S et al., 2018b. High resolution mapping of soil organic carbon stocks using remote sensing variables in the semi-arid rangelands of eastern Australia. *Science of Total Environment*, 630: 367–378. doi: 10.1016/j.scitotenv.2018.02.204

Whitney K, Scudiero E, El-Askary H M et al., 2018. Validating the use of MODIS time series for salinity assessment over agricultural soils in California, USA. *Ecological Indicators*, 93: 889 – 898. doi: 10.1016/j.ecolind.2018.05.069

Wu C S, Liu G H, Huang C, 2016. Prediction of soil salinity in the Yellow River Delta using geographically weighted regression. *Archives of Agronomy and Soil Science*, 63: 928–941. doi: 10.1080/03650340.2016.1249475

Yu R H, Liu T X, Xu Y P et al., 2010. Analysis of salinization dynamics by remote sensing in Hetao Irrigation District of North China. *Agriculture Water Management*. 97: 1952–1960. Doi: 10.1016/j.agwat.2010.03.009

Zhang T T, Qi J G, Gao Y et al., 2015. Detecting soil salinity with MODIS time series VI data. *Ecological Indicators*, 52: 480–489. doi: 10.1016/j.ecolind.2015.01.004

Zhang Y P, Hu K L, Li B G et al., 2009. Spatial distribution pattern of soil salinity and saline soil in Yinchuan plain of China. *Transactions of the CSAE*, 25(7): 19–24. (in Chinese)

Zhou D, Lin Z L, Liu L M, 2012. Regional land salinization assessment and simulation through cellular Automaton-Markov modeling and spatial pattern analysis. *Science of Total Environment*, 439: 260 – 274. doi: 10.1016/j.scitotenv.2012.09.013