Modeling And Mapping of Soil Salinity And Alkalinity Using Remote Sensing Data And Topographic Factors

Elham Shahrayini (✉ e.shahrayini@yahoo.com)
  Tarbiat Modares University

Ali Akbar Noroozi
  Soil Conservation and Watershed Management Research Institute (SCWMRI)

Research Article

Abstract

Soil salinity and alkalinity seriously threaten crop production, soil productivity and sustainable agriculture, especially in arid and semi-arid areas, leading to land degradation, therefore, spatial distribution of these parameters are really important for successful management of such areas. The surface soil salinity and sodium adsorption ratio (SAR) have been modeled in this article. Auxiliary data were terrain attributes derived from digital elevation model (DEM), remote sensing spectral bands, and indices of vegetation and salinity derived from Landsat 8 OLI satellite. In total, 118 soil samples were collected from depth of 0-15 cm in homogenous units at Doviraj plain in the southern part of Ilam province, western Iran. Saturated electrical conductivity (ECe), SAR and other soil properties were analyzed and calculated. To model ECe and SAR parameters with the auxiliary data, stepwise multi linear regression (SMLR) and random forest (RF) regression were applied. The highest accuracy were obtained through RF model with validation coefficient of determination (R^2val) =0.82 and 0.83 and validation root mean square error (RMSEval)=7.40 dS/m and 11.20 for ECe and SAR respectively. Furthermore, results indicated that strongest influence on the prediction of soil salinity followed by Band10, principal component analysis (PC3), Vertical Distance to Channel Network (VDCN) and Analytical Hill Shading (AH). Also, Band10, Band11, Flow Accumulation (FA) and Topographic Wetness Index (TWI) were the important covariate in alkalinity prediction through RF model. Finally, it is suggested that similar techniques can be used to map and monitor soil salinity and alkalinity in other parts of arid regions.

1. Introduction

Soil salinity and alkalinity are serious environmental concern in the world, which have negative impact on our limited soil resources and lead to ecological health degradation [1]. Soil salinity also contributes to desertification, land degradation, and a significant decrease in soil productivity, crop yield, and plant resistance to various stresses [2, 3].

Salt-affected lands are estimated to be about one billion hectares of land, while around 20 percent of the 300 million hectares of irrigated farmland have been found to be impacted by soil salinity on a global scale. In addition, salinity has affected approximately 21% of Iran's land area, which equates to 30 million hectares [4]. Traditional soil salinity measurement techniques are limited in time and space, and can only prepare point-wise data for continuous monitoring of saline soils. Satellite methods generate cost-effective, quick, qualitative and quantitative spatial data on saline soils [5]. Multispectral and microwave remote sensing (RS) spatiotemporal resolutions data on different (e.g., Landsat, Spot, Ikonos, Aster, Modis, IRS, and Radar, etc.) have been used in recent years to track, evaluate and map soil properties with reasonable accuracy as a cost-effective approach, particularly for large scale applications [6, 7].

RS technology can directly provide information about the presence of salt on barren soils, and indirectly on vegetated areas depending on vegetation characteristics.

Different indices of salinity and vegetation were created by combining spectral bands to identify salt-affected regions and they have been assessed with different results, as salinity level and extent of the vegetation cover differs for each case study, choosing and using a similar index may not attain the best outcome in all situations [8]. In most cases, multiple bands have been combined into one index, which is more sensitive to soil salinity than a single band [9, 10]. The spectral absorption characteristics are related to the amount of salinization. Even so, the presence of vegetation and soil crust changed the roughness of the soil's surface, producing a strong reflective effect in visible and near infrared electromagnetic waves and greater reflection than non-saline or moderately saline soil [11].Visual interpretation of false color composite (FCC) can identify various levels of alkalinity and salinity in soils [12]. The analysis and interpretation are based on human experience, which makes the results incomparable in different seasons. For identification of soil salinity and alkalinity utilizing multi-spectral imagery, digital image processing approaches are also applied. Six TM bands were used by Metternich and Zinck [13] to categorize types of salt and sodium-affected soils. Bannari et al. [14], determined that short-wave infrared (SWIR) has advantages in soil salinity and alkalinity detection. Mapping of soil salinity at different scales is critical for assessing the status and trends of agricultural soil salinity as well as management of land degradation [15]. The rapid spread of soil salinization is mainly affected by altitude, because topography controls the speed of salt transport through different soil layers [16]. Several studies have highlighted the effect of micro-topography on the spatial soil salinity distribution in the dry and rainy seasons [17]. The Terrian environment plays an important role in the distribution and redistribution of salt. Although high and steep terrain is conducive to salt migration, low-lying terrain is lead to salt accumulation [10]. In Digital Soil Mapping (DSM), collecting auxiliary variables to predict soil salinity parameters is a prerequisite for improving the prediction accuracy of spatial distribution modeling, which is composed of RS data, DEM and their derivatives (e.g. slope, topographic wetness index, curvatures, etc.), for spatial analysis modeling of soil characteristics [18].

The general idea of selecting stepwise multiple linear regression (SMLR) and random forest (RF) analysis in this research is that SMLR method improves regression models from easily acquired variables to predict data more difficult to obtain, in which statistical tests determine the addition or reduction of predictive variables to the model and produce a final equation. Various researchers have applied a

variety of methods, including SMLR to evaluate large data sets of continuous and categorical variables [19]. RF is another method, which has been increasingly applied for predictions [20]. The main benefits of RF consist of non-parametric nature and high classifying precision, which can be achieved by the use of numerical and categorical variables, and the ability to measure the significance of factors [21]. However, unlike SMLR, this method does not generate the final equation of the model. Therefore, it is often referred to as a form of black box [22]. At the same time, some studies have shown that compared with other methods, this method is reliable and provides better outcomes in both spatial and non-spatial predictions. [23].

In recent years, much progress has been made in RS-based land cover mapping, including the use of machine learning classifiers such as artificial neural networks (ANN) and RF to extract saline land [24, 25]. Few studies have focused on using topographical factors and RS data to map the spatial distribution of soil salinity and alkalinity. Doviraj plain is severely affected by salinity, which poses a great threat to agriculture in this area. Hence, the objectives of this research in modeling and mapping of soil salinity and alkalinity were to i) define the most important covariates among terrain and RS data also, vegetation and salinity spectral indices derived from Landsat 8 OLI image ii) evaluate the performance of RF and SMLR model in this regard.

2. Materials And Methods

2.1. Study area

The field of study is situated on Doviraj plain in the southern part of Ilam province, western Iran. It is located between 47°, 25' – 47°, 35' East longitudes and 32°, 24' – 32°, 28' N latitude with an area of 2600 ha (Fig. 1).The climate of the study area is arid with mean annual precipitation, mean annual temperature and annual potential evaporation of 274.8 mm, 26.4°C, and 2054 mm, respectively. The soil moisture and temperature regimes are aridic and hyperthermic, respectively. The elevation ranges from 87 to 127 m, with a rising gradient from the south to the north. There are different types of land use including irrigated land, dry land, pasture, and the most common land use of the region is bare land. Furthermore, Doviraj plain has different types of physiography such as flood plain, piedmont alluvial plain, plateau, river alluvial plain, river terrace and valley bottom. Dominant soil texture within the area is Loam. The soils in the study area have been classified into Aridisols order and five subgreat groups including Typic Haplocalcids, Typic Haplocambids, Typic Calcigypsids, Typic Haplogypsids and Gypsic Haplosalids [26].

2.2. Data collection and soil sample analysis

In March 2018, by considering to elevation, slope, aspect, soil cover, land use, physiographic units and interpreting satellite images, homogenous unit map prepared and 118 soil samples were taken (Fig. 1).

The samples were air-dried at room temperature then crushed to pass through a 2-mm sieve prior to the analysis.Sodium adsorption ratio (SAR) is an indicator of sodic or non-sodic characteristics and soil salinity is measured by saturated electrical conductivity (ECe).

ECe, pH, soluble calcium, magnesium and sodium (in meq/l for calculation of SAR) [27] and particle size distribution [28] were measured. Also, sodium absorption ratio (SAR) was calculated according to the Eq. 1 [29].

Equation 1
$$SAR = \frac{Na^+}{\sqrt{(Ca^{2+}+Mg^{2+})/2}}$$

2.3. Terrain attributes

Sheng et al.[30] used exclusively terrain attributes to track changes in soil salinity in China.

Topographic wetness index (TWI) can represent water content of soil stationary and can display the areas with potential of salic horizons [31]. However, terrain attributes such as Multi-resolution index of Valley Bottom Flatness (MrVBF) is an important index for flat areas and used to explain relief in flat areas also, identification for the of flat valley bottoms and thus suggested possible sediment transport zones [32]. Analytical hill-shading (AH) technique is defined as topographical images, which produced from digital elevation model (DEM) and can be lit from a specific direction, this helps in shaded-relief images [33] and could recognize lineaments or defects that are difficult to see by

the normal airphoto interpretation methods [34]. 17 terrain maps from a DEM (30m) were determined using the open source System for Automated Geoscientific Analyses (SAGA) software, which are listed in Table 1.

Table 1
terrain attributes used to predict soil EC

Covariate	Description	Reference/source
Elevation (DEM)	Height above sea level (m)	[35]
Analytical Hill Shading (AH)	shaded-relief images	[33]
Aspect (AS)	The direction the plane faces of slope is the aspect Aspect is expressed in positive degrees from 0 to 360, measured clockwise from the north	[36]
Catchment Area (CA)	The catchment area of a cell indicates the area upslope that cell whose flow will eventually reach it	SAGA
Channel Network Base Level (CNBL)	For cells outside the channel network a cell elevation can be interpolated using the elevation values of channel cells. Doing this, a grid with channel network base level elevations can be calculated	SAGA
Closed Depressions (CD)	Closed depression may be complex, comprising flat areas and other smaller nested depressions	[37]
Convergence Index (CI)	Convergence index is used to determine whether water flow from neighboring cells diverges or converges	[38]
Cross-Sectional Curvature (CSC)	Cross-sectional curvature measures the curvature perpendicular to the down slope direction	SAGA
Flow Accumulation (FA)	Flow accumulation is directly proportional to the total amount of water which would pass through a cell as it flowed downhill from higher elevations	SAGA
LS Factor (LS)	Another erosion related parameter is the LS factor. the original equation for the LS factor which used slope and slope length as main parameters	SAGA
Relative Slope Position (RSP)	represents slope position of cell and its relative position between valley floor and ridgetop	[36]
Slope (S)	Slope measures the rate of change of elevation at a surface location. Slope may be expressed as percent slope or degree slope	SAGA
Topographic Wetness Index (TWI)	A measure of water accumulation or soil saturation	[38]
Valley Depth (VD)	The vertical height below summit accumulation	[38]
Vertical Distance to Channel Network (VDCN)	The vertical height above the channel network. Also known as Altitude Above Channel Network	[38]
Multi-resolution Ridge-Top Flatness Index (MrRTF)	Measure of flatness and lowness	[39]
Multi-resolution Valley Bottom Flatness Index (MrVBF)	Measure of flatness and upness	[39]

2.4. Remote sensing data

Landsat 8 OLI image obtained on March 17th, 2018 from USGS (https://glovis.usgs.gov) and utilized in this research. The resolution of spectral bands in the coastal, visible, near infrared, two short waves infrared and Cirrus was 30 meters, for panchromatic band was 15 meter and resolution of two thermal infrared bands was 100 meters. The atmospheric correction was carried out in order to take away or diminish

the atmospheric effect. To lessen the whole number of information layers and enhance the discrimination between saline and non-saline soil, principal component analysis (PCA) was performed. Since most of the study area was bare (Fig. 1), Landsat spectral data

would clearly detect the presence of salts on the surface and one of the important variables in this field is Bare Soil Index (BSI). Normalized Difference Vegetation Index (NDVI) is a critical factor for defining vegetation cover, higher values of NDVI suggest greater vegetation coverage [40]. Therefore, in this paper NDVI and other vegetation spectral indices such as Enhanced Vegetation Index(EVI), Ratio Vegetation Index (RVI), Generalized Difference Vegetation Index (GDVI) and Canopy Response Salinity Index (CRSI) was used to illustrate the effect of vegetation coverage on salinity prediction accuracy. In addition, salinity and vegetation spectral indices were applied as auxiliary variables at the study area, as described in Table 2.

Table 2
Remote sensing data, vegetation spectral indices and salinity spectral indices (derived from Landsat 8 OLI satellite) used to predict soil EC

Auxiliary data	Covariate	Abbreviated	Formulations	References
Remote sensing data	Band1(Coastal / Aerosol)	B1	0.43−0.45 µm	
	Band2(Visible blue)	B2	0.45−0.51 µm	
	Band3(Visible green)	B3	0.53−0.59 µm	
	Band4(Visible red)	B4	0.64−0.67 µm	
	Band5(Near-infrared)	B5	0.85−0.88 µm	
	Band6(Short wavelength infrared)	B6	1.57−1.65 µm	
	Band7(Short wavelength infrared)	B7	2.11−2.29 µm	
	Band8(Panchromatic)	B8	0.50−0.68 µm	
	Band9(Cirrus)	B9	1.36−1.38 µm	
	Band10 (Thermal infrared)	B10	10.60−11.19 µm	
	Band11(Thermal infrared)	B11	11.50−12.51 µm	
	Principal component analysis	PC1, PC2, PC3		
	Brightness index	BI	$\left[\left(B4^2\right)+\left(B5^2\right)\right]^{0.5}$	[41]
	Bare soil index	BSI	$[(B6+B4)-(B5+B2)]\,/\,[(B6+B4)+(B5+B2)]$	[42]
	Clay index	CI	$B6/B8$	[43]
Vegetation spectral Indices	Normalized difference vegetation index	NDVI	$(B5-B4)\,/\,(B5+B4)$	[43]
	Enhanced vegetation index	EVI	$g*(B5-B4)\,/\,(B5+C1*B4-C2*B2+L)$	[44]
	Ratio Vegetation Index	RVI	$B5/B4$	[45]
	Generalized difference vegetation index	GDVI	$\left(B5^2-B4^2\right)/\left(B5^2+B4^2\right)$	[46]
	Canopy response salinity index	CRSI	$\{[(B5*B4)-(B3*B2)]\,/\,[(B5*B4)+(B3*B2)]\}^{0.5}$	[47]
	Soil-adjusted vegetation index	SAVI	$[(B5-B4)*(1+L)]\,/\,(B5+B4+L)$	[48]
Salinity spectral indices	Salinity index	SI	$(B4*B2)^{0.5}$	[41]
	Salinity index 1	SI1	$(B4*B3)^{0.5}$	[49]
	Salinity index 2	SI2	$\left[\left(B5^2\right)+\left(B4^2\right)*\left(B3^2\right)\right]^{0.5}$	[49]

Auxiliary data	Covariate	Abbreviated	Formulations	References
	Salinity index 3	SI3	$\left[\left(\text{B4}^2\right) + \left(\text{B3}^2\right)\right]^{0.5}$	[49]
	Salinity index I	S1	$\text{B2}/\text{B4}$	[50]
	Salinity index II	S2	$(\text{B2} - \text{B4}) / (\text{B2} + \text{B4})$	[50]
	Salinity index III	S3	$\text{B3*B4}/\text{B2}$	[50]
	Salinity index IV	S4	$\text{B2*B4}/\text{B3}$	[50]
	Salinity index V	S5	$\text{B4*B5}/\text{B3}$	[50]
	Salinity index VI	S6	$\text{B6}/\text{B7}$	[14]
	Salinity index VII	S7	$(\text{B6} - \text{B7}) / (\text{B6} + \text{B7})$	[14]
	Salinity index VIII	S8	$(\text{B3} + \text{B4}) /2$	[49]
	Salinity index IX	S9	$(\text{B3} + \text{B4} + \text{B5}) /2$	[49]

Note: the aerosol and soil correcting parameters for EVI: g, C1, C2 and L are set to 2.5, 6, − 7.5 and 1, respectively and soil brightness correction factor (L) for SAVI defined as 0.5 to accommodate most land cover types.

2.5 Analysis of data and modeling

The average, maximum and minimum values, standard deviation and skewness coefficient were calculated for the laboratory-measured data (Table 3). Also, to model and map the soil salinity and alkalinity, two strategies including SMLR and RF have been explored.

Table 3
Descriptive statistics of soil salinity and alkalinity measurements

Depth/Parameter	Min	Max	Average	Std.deviation	Skewness
ECe (dS/m)	0.49	74	9.30	15.1	0.834
SAR	0.47	100.39	7.13	19.9	0.910

2.5.1 Random forest regression

RF is generated by a set of growing decision trees, it relies on the accumulation of random variables, and starts from a few bootstrap samples, which can be randomly extracted from the main training data set [20]. A key procedure in RF is to use Bagging (Bootstrap Aggregating) in conjunction with random feature selection, as Bagging can significantly decrease the variance of unstable procedures such as increasing tree and improve predicting accuracy [20]. RF were calculated and mapped with R software (version 3.1.2). In this model the number of trees (ntree) was set to 800 and both the size of the variables subset (mtry) and the minimum number of nodes (nodesize) were set to 5. All information was randomly and automatically divided to calibration data (80% of the total data) to fit the model and validation data sets (20% of the total data). The significance of the variables additionally decided with the aid of this algorithm. So, the most important variables were used in modeling.

2.5.2 Stepwise multiple linear regressions

These regressions evaluate the coefficients of the linear equation using two or more independent variables to predict the dependent variable. Data the same as subset, which was used for RF, selected to calibrate and validate the model. The stepwise procedure was applied. R^2 and adjusted coefficient of determination (R^2adj) were calculated for assessing of the model performance.

2.6. Accuracy of models

An independent subset of data (not used in the modeling, including 20% of the total data) was used to determine whether the models'predictions were valid for other observed data. For this purpose, the estimated values for each sample of the independent subset were determined and the accuracy of the models was evaluated using R^2, R^2adj, and RMSE. Finally, the models with the highest values of R^2 and R^2adj also, the lowest RMSE values in relation to observed and estimated data were considered as the best models for predicting soil salinity.

3. Results And Discussion

3.1. Descriptive statistics of ECe and SAR data

Summary statistics of ECe and SAR were described in Table 3. In general, the average soil salinity and alkalinity values were above 9.30 dS/m and 7.13 respectively (Since ECe is more than 4 dS/m, the average of soils are salty). As shown in the statistical results, the minimum and maximum ECe and SAR values were 0.49 dS/m, 0.47 and 74 dS/m, 100.39 respectively, which ranged from severely saline (EC > 32 dS/m) to non-saline (0 < EC < 2 dS/m) and from non sodic (SAR < 13) to sodic soils(SAR ≥ 13), FAO [51]. According to skewness values, ECe and SAR conformed to a normal distribution.

3.2. Remote sensing and terrain data processing

Variance Inflation Factor (VIF) was used to assess collinearity (relationship between at least two predictor variables) for Landsat 8 OLI spectral bands and derived indices as well as terrain data. Collinearity leads to uncertainty in the regression method. VIF demonstrates the grade that each independent variable is illustrated by the other independent variable, VIF is used for linear and generalized linear models [52].

About the multivariate collinearity issue does not need to concern for RF. Since RF method in training sample is not sensitive to noise. So, the accuracy of models could be greater than the other machine learning and the conventional statistical regressions [53]. Therefore, only variables in Table 4 were used for SMLR model as input parameters, which were selected by VIF method from all parameters in Table 1 and Table 2.

3.3 Stepwise multi linear regression

At first, the correlation between several RS data and terrain attributes (which were chosen by VIF method) with measured ECe was investigated by Pearson correlation coefficients method (Table 4). This table indicated that the covariates which had**, were significantly correlated with ECe at P < 0.01, covariates that had*, were also significantly related to ECe at P < 0.05 (but this correlation was weaker than **). According to Table 4 the strongest positive correlation was between B10 and EC, with a correlation coefficient of 0.514 as well, B11 and SAR with a correlation coefficient of 0.299, there were significant at P < 0.01.

Different regression models have been constructed between the measured ECe and SAR values and specific RS data as well as terrain attributes (parameters which were chosen according to their significantly, Table 4). Table 5 shows the results of best equations for ECe, SAR and auxiliary data. As it can be seen in equations 1 to 6 and Table 5, the most important covariates for ECe were band10, band11, relative slope position (RSP), Generalized difference vegetation index (GDVI) and Enhanced vegetation index (EVI) with R^2val = 0.62, 0.52, 0.54 and R^2adjval = 0.58, 0.47 and 0.52 respectively. The best equations for SAR were also obtained using of band10, band11, Ratio Vegetation Index (RVI) and GDVI covariates with R^2val = 0.49, 0.46, 0.45 and R^2adjval = 0.46, 0.45, 0.43. The best regressions for auxiliary data EC and SAR were presented as follows:

$$EC = 80.79 + 2.60\,(b10) - 2.31\,(b11) - 0.259\,(RSP)$$

1

$$EC = 53.94 + 2.78\,(b10) - 2.42b11 + 0.5\,(GDVI)$$

2

$$EC = 102.08 + 2.90\,(\text{b10}) - 2.60\,(\text{b11}) + 0.034\,(\text{EVI})$$

3

$$SAR = 231.08 + 2.97\,(\text{b10}) - 2.67\,(\text{b11}) - 0.043\,(\text{PC3})$$

4

$$SAR = 27.64 + 2.81\,(\text{b10}) - 2.39\,(\text{b11}) + 0.168\,(\text{RVI})$$

5

$$SAR = 27.64 + 2.81\,(\text{b10}) - 2.39\,(\text{b11}) + 0.168\,(\text{GDVI})\,(6)$$

Table 4
Pearson correlation coefficients between auxiliary data and measured ECe and SAR

Covariate	EC(dS/m) 0−15 cm	Covariate	SAR
RSP	-0.237**	RSP	-0.067
S	-0.064	S	-0.085
TWI	0.070	TWI	0.052
VD	0.093	VD	0.027
CNBL	0.207*	CNBL	0.157
VDCN	-0.152*	VDCN	-0.106
AH	-0.19	AH	0.101
AS	-0.097	AS	-0.209^{*}
CA	0.067	CA	0.026
LSF	-0.127	CI	0.107
MRRTF	0.204*	LSF	0.038
FA	0.105	MRRTF	0.167
S5	0.210*	FA	0.059
S7	-0.190*	S5	0.005
S9	0.249**	S7	-0.202^{*}
SI	0.285**	S9	0.088
SI1	0.290**	SI	0.141
CRSI	0.093	SI1	0.141
GDVI	-0.327**	CRSI	-0.242^{*}
B2	0.267**	GDVI	-0.293^{**}
B9	0.107	b2	0.138
B10	0.514**	b5	-0.135
B11	0.483**	b9	0.059
PC3	0.161	b10	0.297^{**}
EVI	0.330**	b11	0.299^{**}
RVI	-0.327**	PC3	0.280^{**}
NDVI	-0.327**	EVI	-0.145
BSI	0.283**	RVI	-0.293^{**}
		NDVI	-0.256^{*}
		BSI	0.106

Table 5

Evaluation data of the models generated by stepwise multiple linear regression

Soil Depth (cm)	Equation	Calibration		Validation	
		R^2cal	$R^2_{adj}cal$	R^2val	$R^2_{adj}val$
EC (dS/m)	1	0.69	0.66	0.62	0.58
	2	0.59	0.55	0.52	0.47
	3	0.57	0.54	0.54	0.52
SAR	4	0.51	0.48	0.49	0.46
	5	0.48	0.46	0.46	0.45
	6	0.46	0.45	0.45	0.43

According to the results, evaluation of SMLR equations shows moderate to weak predictive ability using terrain and RS data as well spectral indices derived from Landsat 8 OLI data. An et al. [54] used SMLR in a study in combination with soil spectra measured in field condition and satellite-based remote-sensing images, along with laboratory measurements of soil sample salinity. Their best model for the prediction of soil salinity using the RS data indicated R^2 of 0.896, verification R^2 of 0.867 and RMSE of 0.264. Rahmati and Hamzehpour [55] reported that the constructed regression relations could show a robust prediction of the soil salinity with the R_{adj}^2 up to 0.875 and the best equation was related to the data set with NDVI values above 0.35. In a research was conducted by Hihi et al. [56] results demonstrated that applying linear regression model with combining the Sentinel_2 SWIR bands and the salinity index could illustrate 48% of the spatial variation of soil salinity in the study area.

3.4 Random forest regression

RF was used to model the relationship between ECe, SAR and total auxiliary data. Results demonstrated that the RF model could provide a good relationship between auxiliary data and soil salinity and alkalinity (Table 6). The highest accuracy were obtained with R^2val = 0.82 and RMSEval = 7.35 dS/m for EC and R^2val of 0.76 also, RMSEval of 11.20 were achieved for SAR. Figure 2 shows the scatter plots of the measured versus predicted ECe and SAR for calibration and validation data set. Spatial distribution of soil salinity and alkalinity maps as well the important parameters in each depth are shown in Fig. 3. Two variable importance (VI) indicators were calculated using the RF model, included the percent increase of the mean squared error (%IncMSE) and the cumulative increase in node purity (IncNodePurity) [57].

Table 6

Evaluation data of the models generated by random forest regression

parameters	Calibration		Validation	
	R^2cal	RMSEcal (dS/m)	R^2val	RMSEval (dS/m)
EC	0.80	7.73	0.82	7.40
SAR	0.88	4.63	0.83	3.37

Results showed that the use of Landsat 8 OLI images and terrain data can lead to an acceptable accuracy in soil salinity and alkalinity estimation. Evaluation of those maps in this paper based on Fig. 2 showed strong predictive ability of RF model. In addition, band 10 values (Thermal Infrared Band, 10.60-11.19 μm) were found to be highly correlated with ECe and SAR (Fig. 3).

The temperature of soil surface is affected by internal as well as external factors. Thermal conductivity and heat capacity are considered as internal factors. The rate at which heat passes through a substance is measured with thermal conductivity. The soil's thermal conductivity relies on physical characteristics of the soil including soil particles, air, moisture and porosity. The external variables that affect the surface temperature are meteorological conditions such as, solar radiation, air temperature, relative humidity, wind speed and cloudiness. Thermal infrared bands specially B10 was broadly utilized in the investigation of soil salinity and soil water [58]. The B10 value describes the surface temperature and the high value associated with the high surface temperature. Land surface temperature is mainly affected by soil moisture. The zones with low soil moisture content, where the areas with high salinity in soil surface [59]. The capacity of thermal band Landsat TM for monitoring of soil salinity was evaluated by Alavipanah and Goossens [60]. The results of this research revealed that the addition of the thermal band information contained some helpful information that could play an important role in soil salinity and alkalinity studies.

The flow accumulation (FA) map includes values of cumulative hydrologic flow, which show the quantity of information pixels that contribute any water to outlet. The activity was utilized to understand the drainage pattern of the terrain. According to results of Elmahdy and Mohamed [61], there is a good relationship between FA, groundwater salinity, topographic features and salt-affected soil under irrigated agriculture in arid regions. By comparing Fig. 3 with Fig. 4, it can be found that in places with the highest salinity and alkalinity B10 and B11 maps had the maximum value. Furthermore, B10 and PC3 had the lowest values in irrigated agriculture lands (Fig. 1 and Fig. 4). According to Fig. 3, PC3 was another important parameter that playeda key role in salinity modeling. The PCA is computing based on the eigen vectors and eigen values. Csillag et al [62] stated that principle component analysis is used to separate saline from non saline soils by the stable brightness of PC1 and the stable greenness of PC2, while the differential brightness in PC3 and the differential greenness in PC4 are used to understand the changes that occurring in salinity. According to RS images, large amounts of PC2 were distributed mostly in salt spots. In addition, farmland or wetland had mainly low PC2 values [63]. Figure 4 shows the spatial distribution of TWI, the small values are generally associated with plateau, the intermediate values are related to parts of the piedmont alluvial plain, and the larger values of TWI is corresponded to river alluvial plain and flood plain, which area showed high potential of accumulation of soluble salts such as sodium, calcium and magnesium and caused to higher alkalinity in this area. Moore et al. [64] showed strong relationship between soil salinity and TWI. Also, they have previously been used to classify areas with saline soils by TWI, which displaying landscape degree of wetness and hydrology. BSI is composed of blue, red, near infrared and short wave infrared spectral bands. In order to evaluate the soil mineral compound short wave infrared and the red band are utilized, whereas to increase the attendance of vegetation, the blue and the near infrared spectral bands are used [65]. Some spectral indices including the BSI, Normalized Difference Salinity Index (NDSI), and Salinity Index (SI) have been suggested in order to recognize and map salt-affected soils [66]. In a research performed by Noroozi et al. [67] on 288 soil samples, result revealed that mid-infrared band (TM Band-7), visible band (TM Band-1), Tasseled cap3, Wetness index and PCA2 had strong correlation with the observed EC values in soil surface. RS has been shown satisfactory results in predicting soil salinity. Meanwhile, the spatial distribution of soil salinity seems to be correlated with one more variables based on the properties of region under research therefore, there is no universal spectral index which can use with the best outputs in any environmental conditions [8].Results showed that the most important terrain data, which used in RF modeling, wereVDCN, AH, FA, TWI (Fig. 3 and Fig. 4).Since, the most common land use of study area was bare land, AH covariate was identified as an important terrain data, which could distinguish the land without vegetation cover. Analytical hill shading images are helpful not only display landforms but also to recognize lineaments, to because the shaded relief images show bare land surfaces that are not covered by vegetation [68]. Allbed et al. [9] and Taghizadeh-Mehrjardi et al. [31] used Landsat and terrain data in order to predict the soil surface salinity, they reported R^2 values around 0.65 and 0.87 by genetic programming. According to Pal [69] and Wu et al. [25], both SVM and RF could attain equally well land cover mapping with very high preciseness about 95.7– 96.8% of local sites despite taking much longer processing time than the maximum likelihood. There are several studies have been carried out on soil surface, however in this research, RS datawere used along with terrain data for different soil depths in addition to the topsoil, and the results showed good accuracy in modeling of EC with these covariates up to 60 cm. As can be seen in Fig. 1 and Fig. 3, soils with the highest salinity and alkalinity were located on the sides of the river in bare land as well as in flood plain and river alluvial plain according to the land use and physiographic maps, respectively. However, the soil with the lowest level of salinity and alkalinity can be found in the irrigated agriculture and piedmont alluvial plain based on the land use and physiographic maps. Since, the mean amount of EC and sodium adsorption ratio (SAR) of Doviraj river were 4220 μS/cm and 3.8 respectively, therefore, water quality for this river was classified as C4S1, which indicates very high salinity and slight sodicity [70]. The other reasons of salinity in the study area were the strong evaporation and low precipitation, which caused difficulty for leaching the salts, the presence of gypsum and carbonate calcium materials in soils, and the water table which was between two and three meters from the surface of the soil around the river. In addition, the soils in this area have been classified into four suborders of cambids, calcids, gypsids and salids [26].

4. Conclusions

Accelerating food demand from the rapidly growing population has made the existing land hard to meet the needs of local inhabitants. Reclamation of soil salinity and alkalinity could be an effective method to handle the conflict between man and land. Using RS data and topographic factors to model and map soil salinity and alkalinity also, define the most important covariates in modeling are the objectives of this research. Soil salinity and alkalinity variations at 0-15 cm were investigated at Doviraj plain in Iran. Different models were constructed using the measured ECe, SAR and auxiliary data, including RS spectral bands, vegetation and salinity indices as well terrain data by applying SMLR and RF models. The results of this study confirmed that the RF model was a reliable approach to create continuous soil salinity and alkalinity maps. Furthermore, B10, PC3, VDCN, and AH covariates could construct the most robust model for soil salinity with R^2val=0.82 and RMSE val=7.40 dS/m. B10, FA, TWI, B11 covariates caused to develop the most powerful model for soil alkalinity with R^2val=0.83 and RMSE val=3.37. Also, it is found that thermal bands (B10 and B11) are highly correlated with ECe and SAR. It is necessary to providing proper soil salinity and alkalinity maps with acceptable accuracy in order to achieve sustainable agriculture and economic development. Owing to the fact that RS information and terrain data revealed acceptable accuracy with potentially quick and inexpensive

method in modeling and mapping of soil salinity and alkalinity spatial variations using RF regression in this research, it is recommended to test this technique for the prediction of soil salinity and alkalinity in other regions with different amount of ECe & SAR, variety in climate, vegetation and relief conditions.

Declarations

References

1. Li, H. Y., Webster, R., & Shi, Z. (2015). Mapping soil salinity in the Yangtze delta: REML and universal kriging (E-BLUP) revisited. *Geoderma*, 237–238. doi:10.1016/j.geoderma.2014.08.008

2. Gorji, T., Sertel, E., & Tanik, A. (2017). Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey. *Ecological Indicators*, 74, 384–391. https://doi

3. Shahrayini, E., Fallah, M., Shabanpour, M., Ebrahimi, E., & Saadat, S. (2018). Investigation of soil compaction on yield and agronomic traits of wheat under saline and non-saline soils. *Archives of Agronomy and Soil Science*, 64, 1329–1340. doi:10.1080/03650340.2018.1431832

4. FAO. The State of Food Insecurity in the World. (2008). Food and Agriculture Organization of the United Nations, Rome, Italy ISBN 978-92-5-106049-0.

5. Gorji, T., Alganci, U., Sertel, E., & Tanik, A. (2018). Comparing two different spatial interpolation approaches to characterize spatial variability of soil properties in Tuz Lake Basin – Turkey, Conference: Environmental and health inequities-socio-economic determinants of exposure. https://www.researchgate.net/publication/320755488.

6. Abuelgasima, A., & Ammad, R. (2019). Mapping soil salinity in arid and semi-arid regions using Landsat 8 OLI satellite data. *Remote Sensing Applications: Society and Environment*, 13, 415–442. https://doi.org/10.1016/j.rsase.2018.12.010

7. Periasamy, S., & Shanmugam, R. S. (2017). Multispectral and microwave remote sensing models to survey soil moisture and salinity. *Land Degrad&Development*, 28, 1412–1425. doi:10.1002/ldr.2661

8. 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 https://doi:10.4236/ars.2013.24040

9. Allbed, A., Kumar, L., & Aldakheel, Y. Y. (2014). Assessing soil salinity using soil salinity and vegetation indices derived from IKONS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma*, 230–231, 1–8. doi:10.1016/j.geoderma.2014.03.025

10. 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

11. Bouaziz, M., Matschullat, J., & Gloaguen, R. (2011). Improved remote sensing detection of soil salinity from semi-arid climate in Northeast Brazil. *Comptes Rendus Geoscience*, 343, 795–803101016201109003

12. Wu, H. S., & Liu, Z. L. (2007). Remote sensing and mapping of saline sodic land based on spectral characteristics for Da'an city. *Syst. Sci. Compr. Stud. Agric*, 23, 178–182

13. Metternicht, G., & Zinck, J. A. (1997). Spatial discrimination of salt- and sodium-affected soil surfaces. Int. J. Remote Sens, 1997, 18, 2571–2586.

14. Bannari, A., Guedon, A. M., El-Harti, A., Cherkaoui, F. Z., & El-Ghmari, A. (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 Journal*, 39(19–20), 2795–2811. https://doi.org/10.1080/00103620802432717

15. Gorji, T., Yıldırım, A., Sertel, E., & Tanık, A. (2019). Remote sensing approaches and mapping methods for monitoring soil salinity under different climate regimes. *International Journal of Environment and Geoinformatics*, 6(1), 33–49. https:// doi

16. Mulder, V. L., Bruin, S. D., Schaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping A review.Geoderma, 162,1−19. doi:10.1016/j.geoderma.2010.12.018

17. Wang, J., Liu, Y., Wang, S., Liu, H., & Fu, G. (2019). Spatial distribution of soil salinity and potential implications for soil management in the Manas River watershed, China. *Soil Use and Managment*, 36, 93−163

18. Metelka, V., Baratoux, L., Jessell, M. W., Barth, A., Ježek, J., & Naba, S. (2018). Automated regolith landform mapping using airborne geophysics and remote sensing data, Burkina Faso, West Africa. *Remote Sensing of Environment*, 204, 964−978. https://doi.org/10.1016/j.rse.2017.08.004

19. Juhos, K., Szabo, S., & Ladanyi, M. (2015). Influence of soil properties on crop yield: A multivariate statistical approach. *International Agrophysics*, 29(4), 433−440. https://doi

20. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5−32 https://doi.org/10.1023/A:1010933404324.

21. Rodriguez-Galiano, V. F., Chica-Olmo, M., Abarca-Hernandez, F., Atkinson, P. M., & Jeganathan, C. (2012). Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sensing of Environment*, 121, 93−107

22. Grimm, R., et al. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island - Digital soil mapping using random forests analysis. *Geoderma*, 146(1−2), 102−113 https://doi.org/10.1016/j.geoderma.2008.05.008

23. Souza, E., et al. (2016). Pedotransfer functions to estimate bulk density from soil properties and environmental covariates: Rio Doce basin. *Scientia Agricola*, 73(6), 525−534. https://doi:10.1590/0103-9016-2015-0485

24. Belgiu, M., & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24−31101016201601011

25. Wu, W., Zucca, C., Karam, F., & Liu, G. (2016). Enhancing the performance of regional land cover mapping. *International Journal of Earth Observation and Geoinformation*, 52, 422−432101016201607014

26. Soil Survey Staff. (2014). *Keys to Soil Taxonomy, United states. Department of Agriculture* (12nd ed.). -Natural resources, conservation service

27. Sparks, D. L., Page, A. L., Helmke, P. A., Leoppert, R. H., Soltanpour, P. N., Tabatabai, M. A. … Summer, M. E. (1996). *Methods of Soil Analysis, Soil Science Society of American Journal*. Madison, Wisconsin: Book Series No. 5. ASA and SSSA

28. Gee, G. W., & Or, D. (2002). Particle-size analysis, In J.H. Dane and G.C. Topp, Eds. Methods of Soil Analysis, Part 4, Physical Methods. Soil Science Society of America, Book Series No. 5, Madison, 255−293.

29. Suarez, D. L. (1981). Relation between pHc and sodium adsorption ratio (SAR) and an alternative method of estimating SAR of soil or drainage waters. *Soil Science Society of American Journal*, 45, 469−475

30. Sheng, et al. (2010). Used exclusively terrain attributes to monitor soil salinity variations in China

31. Taghizadeh-Mehrjardi, R., et al. (2016). Prediction of soil surface salinity in arid region of central Iran using auxiliary variables and genetic programming. *Taylor & Francis*, 30(1), 49−64. http://dx.doi.org/10.1080/15324982.2015.1046092

32. Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., & Malone, B. P. (2014). Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*, 213, 15−28101016201307020

33. Thelin, G. P., & Pike, R. J. (1991). Landforms of the conterminous United States-a digital shaded-relief portrayal. Manual of US Geological Survey Map I-2206. https://doi.org/10.3133/i2206

34. Oguchia, T., Aokib, T., & Matsutac, N. (2003). Identification of an active fault in the Japanese Alps from DEM-based hill shading. *Computers & Geosciences*, 29, 885−891. https://doi:10.1016/S0098-3004(03)00083-9

35. National Cartographic Center. Research Institute of NCC: Tehran, Iran. www.ncc.org.ir

36. Evans, S. (1972). General Geomorphometry, Derivatives of Altitude, and Descriptive Statistics. In Chorley, R. J. (Ed.), *Spatial Analysis in Geomorphology* (pp. 17−90). London: Methuen & Co. Ltd.

37. Wang, L., & Liu, H. (2006). An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modeling. *International Journal of Geographical Information Science*, 1−21. https://doi

38. McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1), 3−52. doi: 10.1016/S0016-7061(03)00223-4

39. Gallant, J. C., & Dowling, T. I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39, 1347−1360

40. Zhang, C., Lu, D., Chen, X., Zhang, Y., Maisupova, B., & Tao, Y. (2016). The spatiotemporal patterns of vegetation coverage and biomass of the temperate deserts in Central Asia and their relationships with climate controls. *Remote Sensing of Environment*, 175, 271−281. https://doi

41. Khan, N. M., Rastoskuev, V. V., Sato, Y., & Shiozawa, S. (2005). Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agriculture Water Managment*, 77(1–3), 96–109. https://doi

42. Li, S., & Chen, X. (2014). A new bare-soil index for rapid mapping developing areas using LANDSAT 8 data. The International Archives of Photo grammetry. *Remote Sensing and Spatial Information Sciences*, 40(4), 139–144105194

43. Boettinger, J. L., Ramsey, R. D., Bodily, J. M., Cole, N. J., Kienast-Brown, S., Nield, S. J. … Stum, A. K. (2008). Landsat spectral data for digital soil mapping. In Hartemink, A. E., McBratney, A. B., & Mendonca-Santos, M. L. (Eds.), *Digital Soil Mapping with Limited Data* (pp. 193–203). Australia: Springer Science

44. Raymond Hunt, E., Daughtry, C. S. T., Eitel, J. U. H., & Long, D. S. (2011). Remote sensing leaf chlorophyll content using a visible band index. *Agronomy Journal*, 103, 1090–1099. https://doi:10.2134/agronj2010.0395.

45. Shen, L., & Li, C. (2010). Water body extraction from Landsat ETM + imagery using adaboost algorithm. In Geo informatics, 2010 18th International Conference on 2010 Jun 18 (pp. 1–4). IEEE. https://doi:10.1109/GEOINFORMATICS.2010.5567762

46. Wu, W., Al-Shafie, W., Mhaimeed, A., Ziadat, F., Nangia, V., & Payne, W. (2014). Soil salinity mapping by multi scale remote sensing in mesopotamia, Iraq. *IEEE J.Sel.Top.Appl*, 7, 4442–4452. https://doi:10.1109/jstars.2014.2360411

47. Scudiero, E., Corwin, D., Anderson, R., & Skaggs, T. (2016). Moving Forward on Remote Sensing of Soil Salinity at Regional Scale.Environmental Science. Volum4,Article65. https://dio: 10.3389/fenvs.2016.00065

48. Alhammadi, M. S., & Glenn, E. P. (2008). Detecting date palm trees health and vegetation greenness change on the eastern coast of the United Arab Emirates using SAVI. *Int. J. Remote Sens*, 29, 1745–1765 https://doi.org/10.1080/01431160701395195

49. 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 https://doi:10.1016/j.geoderma.2005.10.009

50. Abbas, A., & Khan, S. (2007). Using remote sensing techniques for appraisal of irrigated soil salinity. In: MODSIM 2007: International Congress on Modeling and Simulation: Land, Water and Environmental Management: Integrated Systems for Sustainability, 2632–2638.

51. Food and Agriculture Organization (FAO) of the United Nations. The salinity and alkalinity status of arid and semi-arid lands. Bull 39, Rome. 1988: 131.

52. Dormann, C. F., et al. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

53. Chen, L., Wang, Y., Ren, C., Zhang, B., & Wang, Z. (2019). Assessment of multi-wavelength SAR and multispectral instrument data for forest aboveground biomass mapping using random forest kriging. *Forest ecology and management*, 447, 12–25. https://doi:10.3390/ijgi8110511

54. An, D., Zhao, G., Chang, C., Wang, Z., Li, P., Zhang, T., & Jia, J. (2016). Hyperspectral Field Estimation and Remote Sensing Inversion of Salt Content in Coastal Saline Soils of the Yellow River Delta. *International Journal of Remote Sensing*, 37(2), 455–470. https://doi:10.1080/01431161.2015.1129562.

55. Rahmati, M., & Hamzehpour, N. (2016). Quantitative remote sensing of soil electrical conductivity using ETM + and ground measured data. *INTERNATIONAL JOURNAL OF REMOTE SENSING*, 38(1), 123–140. http://dx.doi.org/10.1080/01431161.2016.1259681

56. Hihi, S., Rabah, Z. B., Bouaziz, M., Chtourou, M. Y., & Bouaziz, S. (2019). Prediction of Soil Salinity Using Remote Sensing Tools and Linear Regression Model. *Advances in Remote Sensing*, 8, 77–88. https://doi.org/10.4236/ars.2019.83005

57. Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *Am Stat*, 63, 308–319. https://doi.org/10.1198/tast.2009.08199

58. Goossens, R., & Van Ranst, E. (1998). The use of remote sensing to map gypsiferous soils in the Ismailia Province (Egypt). *Geoderma*, 87(1–2), 47–56

59. Metternicht, G. I., & Zinck, J. A. (2003). Remote sensing of soil salinity: potentials and constraints. *Remote Sensing of the Environment*, 85(1), 1–20. https://doi.org/10.1016/S0034-4257(02)00188-8

60. Alavipanah, S. K., & Goossens, R. (2001). Relationship between the Landsat TM, MSS data and soil salinity. *Journal of agricultural science and technology (JAST)*, 3, 101–111

61. Elmahdy, S. I., & Mohamed, M. M. (2012). Topographic Attributes Control Groundwater Flow and Groundwater Salinity of Al Ain, UAE: A Prediction Method Using Remote Sensing and GIS. *Journal of Environment and Earth Science*, 2, 1–13. (Paper) ISSN 2225 – 0948 (Online).

62. Csillag, F., Pa´sztor, L., & Biehl, L. (1993). Spectral band selection for the characterization of salinity status of soils. *Remote Sensing of Environment*, 43, 231–242. https://doi.org/10.1016/0034-4257(93)90068-9

63. Gutierrez, M., & Johnson, E. (2010). Temporal variations of natural soil salinity in an arid environment using satellite images. *(J. South Am. Earth Sci.)*, 30, 46–57. https://doi

64. Moore, D., Grayson, R. B., & Ladson, A. R. (1991). Digital terrain modeling: review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5, 3–30

65. ThiLoi, D., Yin, Chou, T., & Min Fang, Y. (2017). Integration of GIS and Remote Sensing for Evaluating Forest Canopy Density Index in Thai Nguyen Province, Vietnam. *International Journal of Environmental Science and Development*, Vol. 8, No. 8. https://doi:10.18178/ijesd.2017.8.8.1012

66. Nematolahi, M. J., Alavipanah, S. K., Zehtabian, G. R., Jafari, M., & Matinfar, J., E. (2012). *Playa. Desert*, 17, 241–248Assessment of ASTER Data for Soils Investigation Using Field Data and GIS in Damghanhttps://doi 10.22059/JDESERT.2013.35188

67. Noroozi, A. A., Homaee, M., & Farshad, A. (2012). Integrated Application of Remote Sensing and Spatial Statistical Models to the Identification of Soil Salinity: A Case Study from Garmsar Plain,Iran. ENVIRONMENTAL SCIENCES, Vol.9, No.1, Autumn 2012.

68. Kawabata, D., Okada, A., & Takemura, K. (2000). Identification of active tectonic base systems for civil consultants. *Geoinfomatics (Japan Society of Geoinfomatics)*, 11, 183–188. (in Japanese with English abstract).

69. Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, 26(1), 217–222101080014311604123311269698

70. US Salinity Laboratory Staff. (1954). *Diagnosis and improvement of saline and alkali soils* (60Handbook No. vol., pp. 160). US Department of Agriculture

Figures
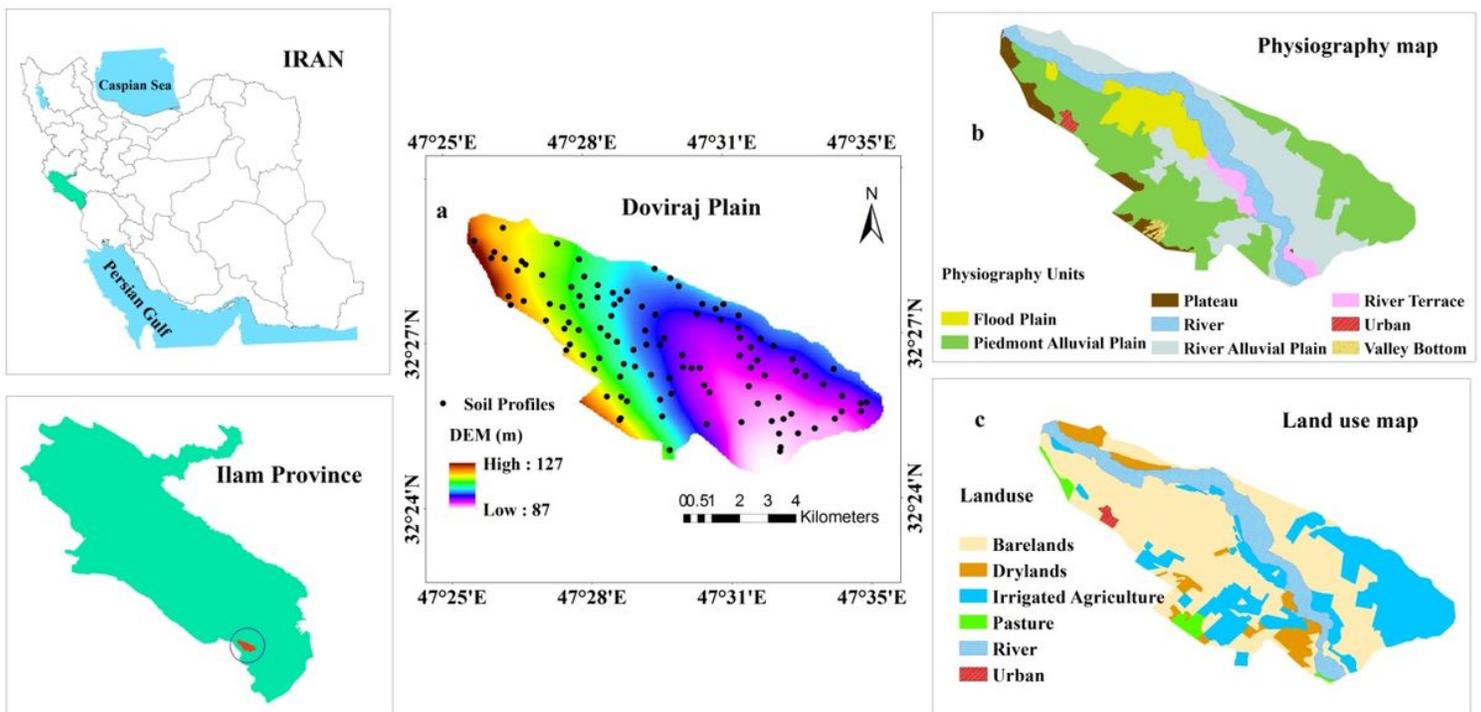
Figure 1

Doviraj Plain study area, including spatial distribution of soil profiles Super imposed on DEM (a), the physiography (b), and land use (c) maps
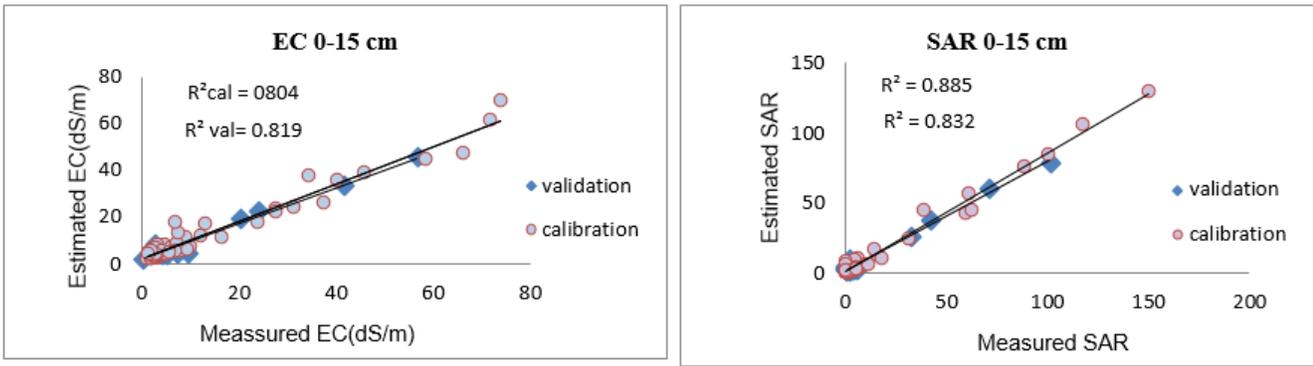
Figure 2

Scatter plots of measured against predicted ECe and SAR using RF model
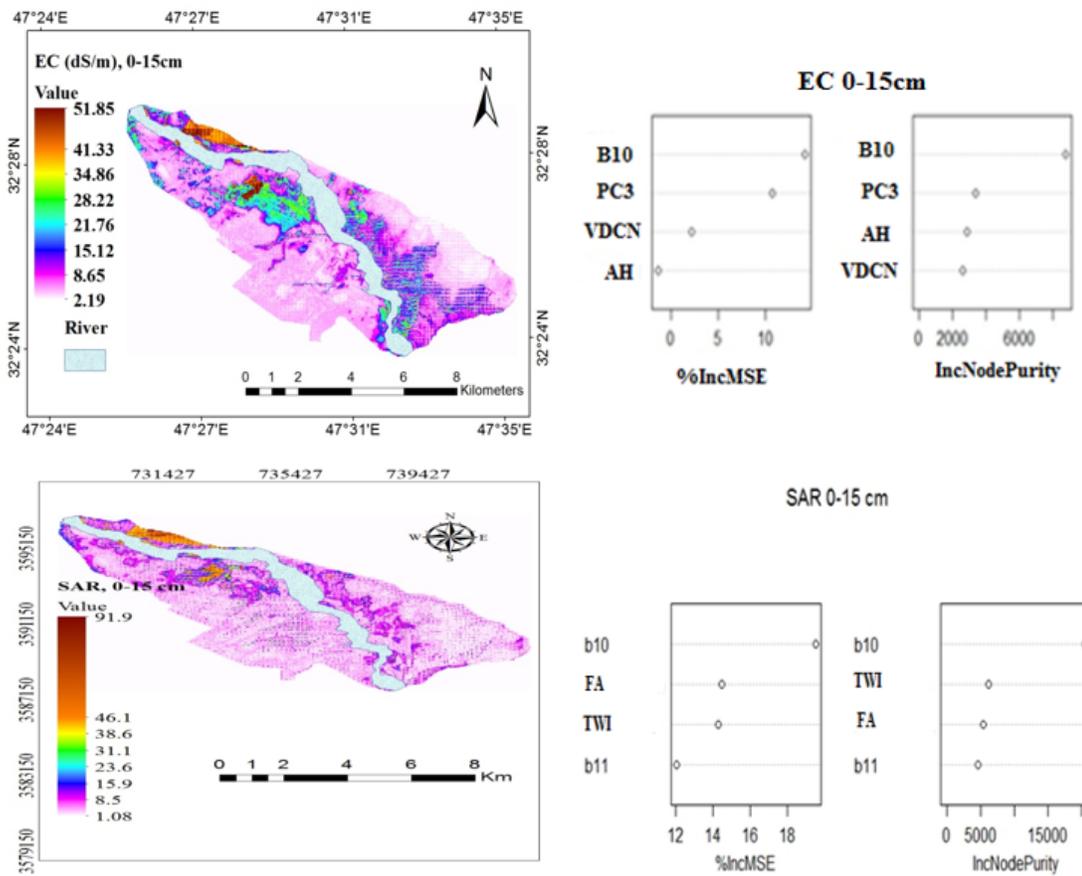
Figure 3

Predicted soil salinity and alkalinity maps with selection of the most important covariates by RF model

Figure 4

The auxiliary data maps, which were used in RF modeling for predicting soil salinity and alkalinity