

A COMPARISON OF MACHINE LEARNING MODELS FOR SOIL SALINITY ESTIMATION USING MULTI-SPECTRAL EARTH OBSERVATION DATA

Arastou Zarei¹, Mahdi Hasanlou^{1,*}, and Masoud Mahdianpari²

¹ School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

² C-CORE and Department of Electrical and Computer Engineering, Memorial University of Newfoundland, Newfoundland, Canada
arastou.zarei@ut.ac.ir, hasanlou@ut.ac.ir, and m.mahdianpari@mun.ca

KEYWORDS: Soil salinity, Electrical conductivity (EC), Machine learning, Sentinel-2 satellite.

ABSTRACT:

Soil salinity, a significant environmental indicator, is considered one of the leading causes of land degradation, especially in arid and semi-arid regions. In many cases, this major threat leads to loss of arable land, reduces crop productivity, groundwater resources loss, increases economic costs for soil management, and ultimately increases the probability of soil erosion. Monitoring soil salinity distribution and degree of salinity and mapping the electrical conductivity (EC) using remote sensing techniques are crucial for land use management. Salt-affected soil is a predominant phenomenon in the Eshtehard Salt Lake located in Alborz, Iran. In this study, the potential of Sentinel-2 imagery was investigated for mapping and monitoring soil salinity. According to the satellite's pass, different salt properties were measured for 197 soil samples in the field data study. Therefore several spectral features, such as satellite band reflectance, salinity indices, and vegetation indices, were extracted from Sentinel-2 imagery. To build an optimum machine learning regression model for soil salinity estimation, three different regression models, including Gradient Boost Machine (GBM), Extreme Gradient Boost (XGBoost), and Random Forest (RF), were used. The XGBoost method outperformed GBM and RF with the coefficient of determination (R^2) more than 76%, Root Mean Square Error (RMSE) about 0.84 dS m^{-1} , and Normalized Root Mean Square Error (NRMSE) about 0.33 dS m^{-1} . The results demonstrated that the integration of remote sensing data, field data, and using an appropriate machine learning model could provide high-precision salinity maps to monitor soil salinity as an environmental problem.

1. INTRODUCTION

Soil salinization due to natural processes and human factors is a significant environmental hazard in arid and semi-arid regions (Metternicht and Zinck, 2003; Ren et al., 2019). Soil affected by salt reduces the productivity of agricultural products, ecosystem health, and water quality. It also can lead to soil erosion and land degradation (Khan et al., 2005; Wicke et al., 2011). Based on the Food and Agriculture Organization of the United Nations (FAO) estimates, 397 million hectares of land worldwide have been covered by saline soils, and the affected areas are estimated to be expanding at a rate of two million hectares per year (Koochafkan and Stewart, 2012; Peng et al., 2019). In addition to the main salt content in many areas, soil resources are at risk of secondary salinization, mainly due to low precipitation and high evaporation, shallow groundwater level, and irrational activities of farmers (Jiang et al., 2019; Nicolas and Walter, 2006; Zarei et al., 2021). Therefore, careful monitoring, evaluation, and mapping to detect soil salinity can provide sufficient understanding of this threat's temporal and spatial distribution so that it becomes possible to make effective soil restoration for land management (Bannari et al., 2018; Davis et al., 2019).

The EC parameter is commonly used to investigate soil salinity dynamics due to its high correlation with soil salinity (Richards, 1954). Traditional methods to analyze EC are very accurate yet time-consuming, discontinuous, and costly. In the last two decades, remote sensing has been widely used to determine and monitor soil salinity characteristics at different scales (Taghadosi and Hasanlou, 2021; Dale et al., 1986; Dwivedi, 2001; Lizaga et al., 2019; Santra et al., 2015). Multi-spectral data, such as QuickBird, IKONOS, SPOT, Landsat, and Sentinel, are useful in identifying and monitoring soil salinity and environmental

hazards (Farifteh, 2007; Koshal, 2012; Ranjbar et al., 2021; Teggi et al., 2012). Also, the Sentinel-2 satellite that was launched with a multi-spectral instrument (MSI) in 2015 is an essential part of global environmental monitoring, which has been continuously used for the past few years due to its high spatial and spectral resolution to identify areas affected by salinity (Malenovský et al., 2012; Taghadosi et al., 2019; J. Wang et al., 2019).

The spectral reflectance of soil surface salt properties has been widely used in several studies as a direct indicator to detect and monitor soil salinity. In a case study in Malheur County, Landsat TM satellite images were used to map soil salinity, the results of which showed that a large amount of salt in barren soils could be detected in bands 1 to 4 of the Landsat satellite due to the high spectral reflection of salt in this range (Elnaggar and Noller, 2010). In another study, Sentinel-2 satellite data were used to investigate soil salinity with different spectral compositions, the results of which presented salinity indices by combining two or three spectral bands that achieved the highest correlation with ground-truth salinity measurements (J. Wang et al., 2019). In addition to the spectral indices obtained from the combination of satellite image bands, various transformation-based methods were used to extract appropriate properties to assess soil salinity. For example, the principal component analysis (PCA) and spectral indices were used to monitor soil salinity. The results showed that the PCA technique and salinity indices are a suitable method for predicting salinity from satellite images and offers high accuracy (Khan and Abbas, 2007).

In recent years, a wide range of regression methods have been employed to model soil salinity and estimate EC values. These methods' performance varies according to the study area, in-situ data collected, and applied regression methods (Eldeiry and

* Corresponding author

Garcia, 2010; Farifteh et al., 2007; Gorji et al., 2017). For instance, Wang et al. (2007) used correlation analysis, Ordinary Least Square (OLS) method, and spatial regression method to study soil salinity's spatial variation in the yellow river delta. This study showed that the spatial regression model improved the accuracy of soil salinity estimation (Wang et al., 2007). Qu et al. (2008) used the Partial Least Square Regression (PLSR) method to assess soil salinity using hyperspectral data. The results showed that the calibrated PLSR method could predict soil salinity with accurate results (Qu et al., 2008). Recently, soil salinity mapping has been performed accurately using machine learning regression methods, such as Support Vector Regression (SVR) and RF. In 2018, Wu et al. produced the soil salinity map using Landsat satellite data and machine learning regression methods, such as SVR and RF, as well as Multiple Linear Regression (MLR). Comparing different regression methods, the results showed that the RF method with less NRMSE outperformed other models (Wu et al., 2018). Wang et al. (2019) examined different machine learning regression methods for modeling soil salinity at various study sites. The results showed that the Stochastic Gradient Treeboost (SGT) method is the most reliable algorithm for predicting soil salinity in arid regions (F. Wang et al., 2019).

In this study, we investigate soil salinity monitoring and soil EC mapping using Sentinel-2 satellite images. The main objectives of this study are summarized: (i) to understand the spectral reflectance characteristics of saline soil in Eshtehard Salt River, (ii) identifying suitable variables for soil salinity prediction over the study area, and (iii) evaluating and comparing the three machine learning regression algorithms, particularly GBM, XGBoost, and RF, in predicting soil salinity using in-situ data collected in the study area, and (iv) to produce the soil salinity map according to high, moderate and low saline content.

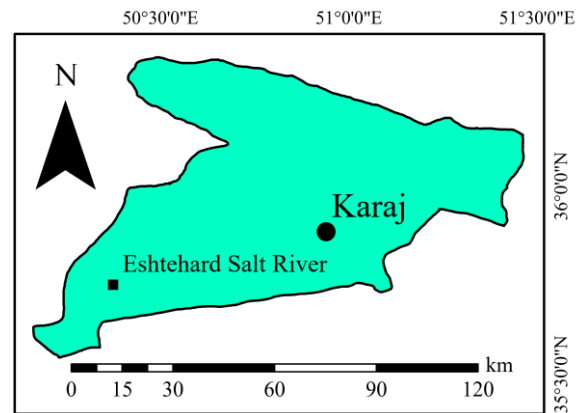
2. CASE STUDY AND DATASET

2.1. Study area

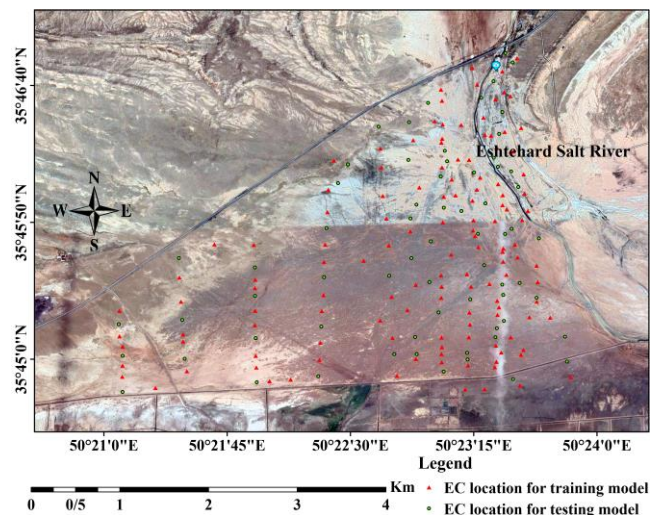
Eshtehard is located in the southwest of Alborz County in the Salt River basin (Figure 1). This region has a diverse structure in terms of hydrology and geohydrology. The study area in this research is located in the Eshtehard Salt River with an area of about 10 Km^2 (latitude= $35^\circ 44' 30'' \text{ N}$ to $35^\circ 46' 30'' \text{ N}$ and longitude= $50^\circ 21' 00'' \text{ W}$ to $50^\circ 24' 00'' \text{ W}$).



(a)



(b)



(c)

Figure 1. (a) Location of Alborz County in Iran, (b) Location of Eshtehard Salt River in the district of Alborz County, and (c) Eshtehard Salt River district image acquired from Google Earth.

2.2. Datasets

2.2.1. In-situ data

To study the soil in the affected areas, measuring salinity, and preparing ground-truth data, *in-situ* data were collected near the Eshtehard Salt River using the TDR-350 device. Soil sampling of this area using the design of ground control points with 197 samples was taken randomly from different parts of the area around the Eshtehard Salt River in August and September of 2020. For each sample, we conducted five different measurements. To do this, a sample is collected in the center of the site, and 4 samples are collected from 4 corners of a square of $10 \text{ m} \times 10 \text{ m}$ (Figure 2), and then the average of these five points represents the salinity in a square of $10 \text{ m} \times 10 \text{ m}$. These steps hold for all samples to optimize the sample values in the pixels of satellite images (Wang et al., 2020). Based on the classes determined by Durand (Durand, 1983), considered five salinity classes (Table 1). Figure 1 shows the *in-situ* operations and the collection and distribution of sample points in the Eshtehard Salt River area. Field data were randomly divided into 70% for the training set and 30% for the testing set according to the two parameters of computational cost and representativeness.

Table 1. Classification of soil based on the electrical conductivity (EC) classification supplied by Duran (1983).

EC (ds/m)	Salinity Classes
$EC < 0.6$	Non-saline soil
$0.6 < EC < 1.0$	Slightly saline soil
$1.0 < EC < 2.0$	Moderately saline soil
$2.0 < EC < 4.0$	Very saline soil
$4 < EC$	Extremely saline soil

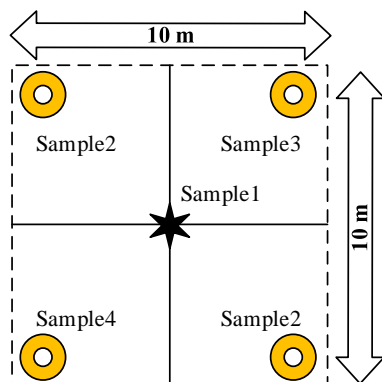


Figure 2. In-situ data collection tool and a sample photo from the landscape of the study area.

2.2.2. Sentinel-2 imagery

Sentinel-2 images were collected for salinity assessment due to data availability at the level-1C (top-of-atmosphere reflectance in cartographic geometry). These should then be converted into a level-2A ortho-image bottom-of-atmosphere corrected reflectance product suitable for regression analysis. Pre-processing of level-1C products was conducted by scene classification and atmospheric correction, which gives level-2A

products (Suhet et al., 2015). It was implemented by Sen2Cor, a processor for the production and molding of the Sentinel-2 L2A product developed by (Müller-Wilm, 2017).

Table 2. The dates of in-situ measurements and Sentinel-2 acquisitions.

Date of in-situ measurements	Date of satellite pass
06-08-2020	05-08-2020
24-08-2020	25-08-2020
30-08-2020	30-08-2020
04-09-2020	04-09-2020

3. METHOD

In this study, the total data set ($n = 197$) was divided into a training set (144 soil samples, 70% of the total soil samples) and a testing set (53 soil samples, 30% of the total soil samples). In the entire data set, according to the sampling order, one sample was selected every four samples as a verification sample. As discussed, the main objective of this study was to evaluate the sustainability of the GBM, XGBoost, and RF algorithms to model the relationship between spectral characteristics of the Sentinel-2 satellite data and the soil salinity parameter over the Eshtehard Salt River. The flowchart of the proposed method is illustrated in Figure 3 and is summarized in the following five steps. Additionally, the main steps of the proposed method are discussed in more detail in the following sections.

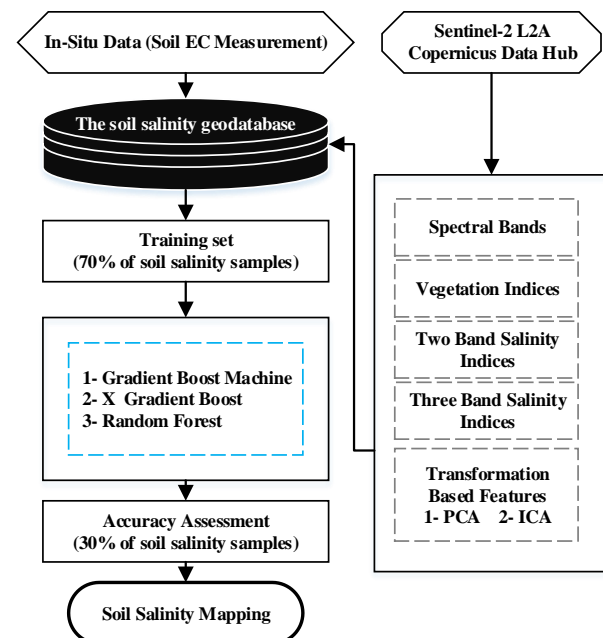


Figure 3. Flowchart of the proposed method for estimating the soil salinity in the Eshtehard Salt River area based on machine learning regression models.

3.1. Feature extraction

To identify the affected areas and determine the relationship between soil samples' EC values and the corresponding pixel values in satellite images, 10 spectral bands in which soil salinity is reflected were used (Taghadosi et al., 2019). Since the study area is mostly barren and the vegetation is very sparse, it is very useful to use salinity indices that highlight salt-affected surfaces' reflection. The use of vegetation indices, which are commonly

used to identify vegetation areas, can effectively analyze soil salinity trends because the negative effects of salinity on plant growth indicate changes in salinity in vegetated areas (Moreira et al., 2015). However, to evaluate the relationship between EC values of soil samples and the corresponding pixel values obtained from salinity and vegetation indices, salinity and vegetation indices were selected as predictor variables in our regression analysis. These indices have shown the best performance for salinity detection in previous studies (Allbed and Kumar, 2013; Scudiero et al., 2014; J. Wang et al., 2019). In image processing and machine learning, feature transformation by converting measured datasets is used to generate datasets that contain useful, informative, and facilitative information (Richards, 2013). This study used PCA and ICA methods as transformation-based features (Lee and Batzoglu, 2003; Ranchordas et al., 2010). A total number of 46 features were extracted from satellite data (Table 3 and 4).

Table 3. Soil salinity features that were extracted from Sentinel-2 satellite data.

Image transformation	Sentinel-2 bands /	Feature	
		Blue	NIR
	Green	Red Edge 4	
	Red	SWIR 1	
	Red Edge 1	SWIR 2	
	Red Edge 2	PC ₁ -PC ₁₀	
	Red Edge 3	ICA ₁ -ICA ₁₀	

Table 4. Vegetation and salinity features were extracted from Sentinel-2 satellite data.

Index	Formula
TBI2	$\frac{\text{RedEdge3}}{(\text{SWIR2} + \text{SWIR2})}$
TBI3	$\frac{(\text{RedEdge2} - \text{SWIR2})}{(\text{RedEdge2} + \text{SWIR2})}$
TBI5	$\frac{(\text{SWIR1} + \text{SWIR2})}{\text{Green}}$
TBI7	$(\text{SWIR1} - 2 \cdot \text{SWIR2} + \text{RedEdge1})$
SI-1	$\frac{(\text{Red} \cdot \text{NIR})}{\text{Green}}$
SI-2	$\frac{(\text{Red} \cdot \text{Blue})}{\text{Green}}$
SI-3	$\frac{(\text{Blue} - \text{Red})}{(\text{Blue} + \text{Red})}$
SI-4	$\sqrt{(\text{Green}^2 + \text{Red}^2)}$
SI-5	$\sqrt{(\text{Green}^2 + \text{Red}^2 + \text{NIR}^2)}$
SI-6	$\sqrt{(\text{Blue} \cdot \text{Red})}$
SI-7	$\frac{\text{Blue}}{\text{Red}}$
BI	$\sqrt{(\text{NIR}^2 + \text{Red}^2)}$
RI	$\frac{\text{Red}}{\text{SWIR2}}$
NDI	$\frac{(\text{SWIR2} - \text{RedEdge3})}{(\text{SWIR2} + \text{RedEdge3})}$
NDSI	$\frac{(\text{Red} - \text{NIR})}{(\text{Red} + \text{NIR})}$
CRSI	$\frac{(\text{Red} \cdot \text{NIR}) - (\text{Blue} \cdot \text{Green})}{(\text{Red} \cdot \text{NIR}) + (\text{Blue} \cdot \text{Green})}$

In this step, the correlation between the collected in-situ samples' electrical conductivity and feature extracted values of satellite

images is carried out to find the relationship between these variables and their efficiency in predicting soil salinity using the linear regression. Therefore, to determine the correlation between saline soil and each feature and among features, correlation matrixes were generated (Figure 4).

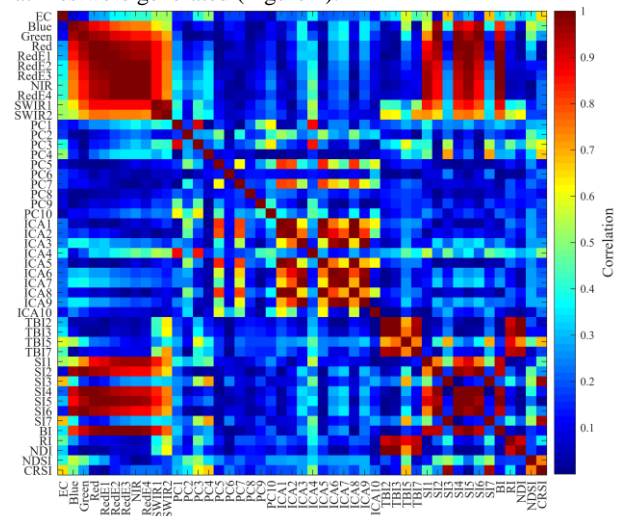


Figure 4. Correlation matrixes document the correlation between electrical conductivity and multi-spectral features, and the correlation between multi-spectral.

3.2. Regression analysis

Regression analysis is generally used to predict the relationship between a dependent and one or more independent variables. In recent years, various regression techniques have been developed in a widespread of applications, which can be used for prediction and model construction (Fan et al., 2015). In this study, GBM, XGBoost, and RF methods were selected for remote sensing inversion of soil salinity.

3.2.1. Gradient Boost Machine

Gradient boost machine algorithm employs ensemble, which simplistically eliminates bias, noise, and variance, which reduce the prediction model's effectiveness. The ensemble uses Boosting methods, which pre-builds many independent models, which are then implemented sequentially to allow the new models to learn from the error of the earlier models. GBM algorithm intuitively follows the concept of running and testing residual models to sustain the new model where the algorithm's cost function was optimized. The user-centric parameters that are optimized for this algorithm using the grid search method include the following: 1) n estimators: the number of sequential trees to be modeled, though GBM is fairly robust at a higher number of trees, it can still overfit at a point, and 2) max features: the number of features to consider while searching for the best split. As a thumb rule, the square root of the total number of features works great. (Friedman, 2001; Ying and Sayed, 2017).

3.2.2. XGBoost

XGBoost is one of the quickest implementations of gradient boosted trees. It does this by tackling gradient boosted trees' significant incapability: considering the potential loss for all possible splits to create a new member. XGBoost address this incapability by looking at the distribution of features across all data points in a leaf and using this information to decrease the search space of possible feature splits. Although XGBoost implements several regularization methods, this fast is the

numerous useful feature of the method, providing many hyperparameter settings to be investigated instantly. The user-centric parameters that are optimized for this algorithm using the grid search method include the following: 1) learning rate: Step size reduction used in the update to prevents overfitting, 2) alpha and lambda: L1 and L2 regularization term on weights. Increasing this value will make the model more conservative, and 3) column sample by tree: This is a family of parameters for the subsampling of columns. (Chen and Guestrin, 2016).

3.2.3. Random Forest

Random forest is a supervised machine learning algorithm effectively used to solve regression and classification problems and determine nonlinear relationships between target and input (Breiman, 2001). This algorithm creates the forest using a set of individual decision trees, each tree with a subset of random features (Belgiu and Drăguț, 2016). Each tree accesses a random subset of training samples and predicts target values. For regression problems, each tree has a vote, and the prediction value is the average estimate of all decision trees. This algorithm can determine the relative importance of each input feature, which is important in understanding each feature's contribution in predicting RF output. The user-centric parameters that are optimized for this algorithm using the grid search method include the following: 1) n estimator: the number of trees in the forest, and 2) max features: the number of features to consider when looking for the best split. (Heung et al., 2016).

3.3. Accuracy assessment

To assess the performance of the regression models, three criteria were chosen: R^2 , RMSE, and NRMSE. The accuracy of the created models was then analyzed based on these criteria (Equations (1–3)):

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

$$RMSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

$$NRMSE = \frac{RMSE}{\bar{y}} \quad (3)$$

where \hat{y}_i is a vector of predicted dependent variables with n data points, y_i is the vector of observed values of the variable being predicted and \bar{y} is the mean of the observed dependent variables.

4. RESULT

4.1. Prediction of soil EC maps

Considering the location of the ground-truth data at the study site, the satellite data's pixel values were extracted for analysis. According to Tables 3 and 4, 46 features were considered for each soil sample, and a matrix containing all *in-situ* data measurements and satellite features was created to be used to train the model. The data matrix was then divided into two sections: 70% for training and 30% for testing, respectively; we developed prediction models using GBM, XGBoost, and RF to estimate soil salinity regarding the independent variables. According to the regression analysis results, the constructed models were

used to plot the EC values for each pixel in the satellite image. Among these models, the best regression method based on R^2 , RMSE, and NRMSE values for mapping soil salinity in the whole image was compared, which will be discussed in the next section. Figure 5 shows the predicted EC map for each regression method.

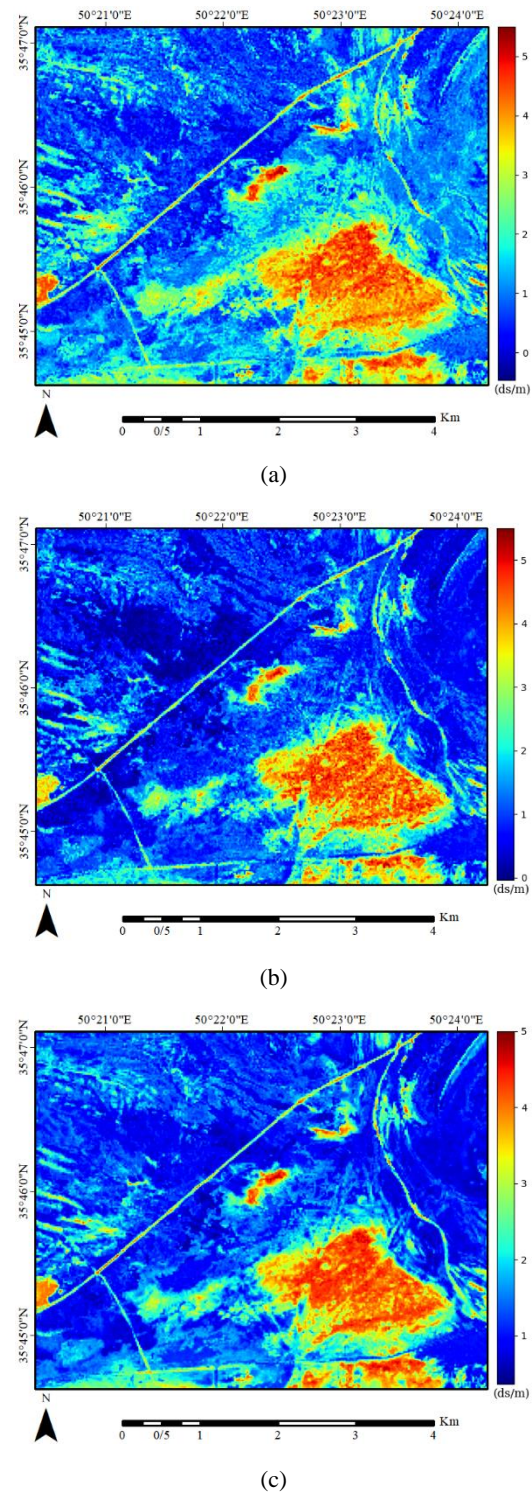


Figure 5. Predicted soil EC maps based on the results of regression analysis, (a) Gradient Boost Machine, (b) XGBoost, and (c) Random Forest.

4.2. Evaluation of the Accuracy of Estimations

The soil salt content estimation models were constructed by GBM, XGBoost, and RF methods, respectively (Table 5). After parameter optimization on each model's training datasets, we found the average estimation to be $R^2 = 0.72$ for the GBM model, $R^2 = 0.76$ for the XGBoost model, and $R^2 = 0.69$ for the RF model. According to the results in Table 5, the XGBoost model shows high precision and accuracy in predicting soil salinity; also, the RF model presents the weakest results. Our proposed method for all algorithms shows the models with higher accuracy compared to the recent study conducted by (Zhang et al., 2020). The two main reasons why the XGBoost model outperformed other models include the: (1) computing second-order gradients, i.e. second partial derivatives of the loss function, which provides more information about the direction of gradients and how to get to the minimum of our loss function. While regular gradient boosting uses our base model's loss function as a substitute for minimizing the overall model's error, XGBoost uses the 2nd order derivative as an approximation. (2) And advanced regularization (L1 and L2), which improves model generalization XGBoost, has further benefits: training is very fast and can be parallelized across batches.

Table 5. Accuracy assessment of three machine learning regression methods.

ML-Regression Method	Train / Test	R ²	RMSE	NRMSE	Number of Feature
GBM	Train	0.98	0.001	0.001	46
	Test	0.72	0.87	0.34	
XGB	Train	0.98	0.019	0.008	46
	Test	0.76	0.84	0.33	
RF	Train	0.94	0.55	0.23	46
	Test	0.69	0.92	0.37	

5. CONCLUSION

In this study, 46 variables, in four main categories (i.e., remote sensing data, terrain characteristics, salinity spectral indices, and vegetation spectral indices) and three models (i.e., GBM, XGBoost, and RF) were selected to estimate soil salinity in the Eshtehard Salt River. The main results are as follows: (1) Overall, the 46 factors considered were significant and contributed to the estimation of soil salinity. (2) The XGBoost model was the best model in this study, with better model performance and accuracy measures ($R^2 = 76\%$, $RMSE = 0.84 \text{ dS m}^{-1}$, $NRMSE = 0.33 \text{ dS m}^{-1}$). (3) According to the EC map obtained from the XGBoost model, the north and the west of the study area showed relatively low soil salinity. Salted soil was mainly found in the southeast to the center of the study area and Salt River.

REFERENCES

Allbed, A., Kumar, L., 2013. Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: a review. *Adv. Remote Sens.* 2013.
 Bannari, A., El-Battay, A., Bannari, R., Rhinane, H., 2018. Sentinel-MSI VNIR and SWIR bands sensitivity analysis for soil salinity discrimination in an arid landscape. *Remote Sens.* 10, 855.
 Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31.
 Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*. pp. 785–794.
 Dale, P.E.R., Hulsman, K., Chandica, A.L., 1986. Classification of reflectance on colour infrared aerial photographs and sub-tropical salt-marsh vegetation types. *Int. J. Remote Sens.* 7, 1783–1788.
 Davis, E., Wang, C., Dow, K., 2019. Comparing Sentinel-2 MSI and Landsat 8 OLI in soil salinity detection: A case study of agricultural lands in coastal North Carolina. *Int. J. Remote Sens.* 40, 6134–6153.
 Durand, J.H., 1983. Les sols irrigables; etude pedologique.
 Dwivedi, R.S., 2001. Soil resources mapping: A remote sensing perspective. *Remote Sens. Rev.* 20, 89–122.
 Eldeiry, A.A., Garcia, L.A., 2010. Comparison of ordinary kriging, regression kriging, and cokriging techniques to estimate soil salinity using LANDSAT images. *J. Irrig. Drain. Eng.* 136, 355–364.
 Elnaggar, A.A., Noller, J.S., 2010. Application of remote-sensing data and decision-tree analysis to mapping salt-affected soils over large areas. *Remote Sens.* 2, 151–165.
 Fan, X., Liu, Y., Tao, J., Weng, Y., 2015. Soil salinity retrieval from advanced multi-spectral sensor with partial least square regression. *Remote Sens.* 7, 488–511. <https://doi.org/10.3390/rs70100488>
 Farifteh, J., 2007. Imaging spectroscopy of salt-affected soils: Model-based integrated method. Utrecht University.
 Farifteh, J., Van der Meer, F., Atzberger, C., Carranza, E.J.M., 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sens. Environ.* 110, 59–78.
 Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 1189–1232.
 Gorji, T., Sertel, E., Tanik, A., 2017. Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey. *Ecol. Indic.* 74, 384–391.
 Heung, B., Ho, H.C., Zhang, J., Knudby, A., Bulmer, C.E., Schmidt, M.G., 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma* 265, 62–77.
 Jiang, Q., Peng, J., Biswas, A., Hu, J., Zhao, R., He, K., Shi, Z., 2019. Characterising dryland salinity in three dimensions. *Sci. Total Environ.* 682, 190–199.
 Khan, N.M., Rastokuev, V. V., Sato, Y., Shiozawa, S., 2005. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agric. Water Manag.* 77, 96–109.
 Khan, S., Abbas, A., 2007. Using remote sensing techniques for appraisal of irrigated soil salinity. *Int. Congr. Model. Simul.(MODSIM), Model. Simul. Soc. Aust. New Zealand, Bright* 2632–2638.
 Koohafkan, P., Stewart, B.A., 2012. Water and Cereals in Drylands, the Food and Agriculture Organization of the United Nations and Earthscan.
 Koshal, A.K., 2012. Spectral characteristics of soil salinity areas in parts of South-West Punjab through remote sensing and GIS. *Int. J. Remote Sens. GIS* 1, 84–89.
 Lee, S.-I., Batzoglou, S., 2003. Application of independent component analysis to microarrays. *Genome Biol.* 4, 1–21.
 Lizaga, I., Quijano, L., Gaspar, L., Ramos, M.C., Navas, A., 2019. Linking land use changes to variation in soil properties in a Mediterranean mountain agroecosystem. *Catena* 172, 516–527.
 Malenovský, Z., Rott, H., Cihlar, J., Schaepman, M.E., García-Santos, G., Fernandes, R., Berger, M., 2012. Sentinels for

- science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land. *Remote Sens. Environ.* 120, 91–101.
- Metternicht, G.I., Zinck, J.A., 2003. Remote sensing of soil salinity: potentials and constraints. *Remote Sens. Environ.* 85, 1–20.
- Moreira, L.C.J., Teixeira, A. dos S., Galvão, L.S., 2015. Potential of multispectral and hyperspectral data to detect saline-exposed soils in Brazil. *GIScience Remote Sens.* 52, 416–436.
- Müller-Wilm, U., 2017. *Sen2Cor Configuration and User Manual*. Telespazio VEGA Deutschl. GmbH 53.
- 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.
- Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., Shi, Z., 2019. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* 337, 1309–1319.
- Qu, Y., Jiao, S., Lin, X., 2008. A partial least square regression method to quantitatively retrieve soil salinity using hyper-spectral reflectance data, in: *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Classification of Remote Sensing Images*. International Society for Optics and Photonics, p. 71471H.
- Ranchordas, A., Pereira, J.M., Araújo, H.J., Tavares, J.M.R.S., 2010. *Computer Vision, Imaging and Computer Graphics: Theory and Applications-International Joint Conference, VISIGRAPP 2009*, Lisboa, Portugal, February 5-8, 2009, Revised Selected Papers. Springer.
- Ranjbar, S., Zarei, A., Hasanlou, M., Akhoondzadeh, M., Amini, J., Amani, M., 2021. Machine learning inversion approach for soil parameters estimation over vegetated agricultural areas using a combination of water cloud model and calibrated integral equation model. *J. Appl. Remote Sens.* 15, 1–17. <https://doi.org/10.1117/1.jrs.15.018503>
- Ren, D., Wei, B., Xu, X., Engel, B., Li, G., Huang, Q., Xiong, Y., Huang, G., 2019. Analyzing spatiotemporal characteristics of soil salinity in arid irrigated agro-ecosystems using integrated approaches. *Geoderma* 356, 113935.
- Richards, J.A., 2013. Supervised classification techniques, in: *Remote Sensing Digital Image Analysis*. Springer, pp. 247–318.
- Richards, L.A., 1954. *Diagnosis and improvement of saline and alkali soils*. LWW.
- Santra, P., Singh, R., Sarathjith, M.C., Panwar, N.R., Varghese, P., Das, B.S., 2015. Reflectance spectroscopic approach for estimation of soil properties in hot arid western Rajasthan, India. *Environ. Earth Sci.* 74, 4233–4245.
- Scudiero, E., Skaggs, T.H., Corwin, D.L., 2014. Regional scale soil salinity evaluation using Landsat 7, western San Joaquin Valley, California, USA. *Geoderma Reg.* 2, 82–90.
- Suhret, B.H., 2015. *Sentinel-2 User Handbook*. (https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook).
- Taghadosi, M.M., Hasanlou, M., 2021. Developing geographic weighted regression (GWR) technique for monitoring soil salinity using Sentinel-2 multispectral imagery. *Environ Earth Sci* 80, 75.
- Taghadosi, M.M., Hasanlou, M., Eftekhari, K., 2019. Retrieval of soil salinity from Sentinel-2 multispectral imagery. *Eur. J. Remote Sens.* 52, 138–154.
- Teggi, S., Costanzini, S., Despini, F., Chiodi, P., Immordino, F., 2012. SPOT5 imagery for soil salinity assessment in Iraq, in: *Earth Resources and Environmental Remote Sensing/GIS Applications III*. International Society for Optics and Photonics, p. 85380V.
- Wang, F., Yang, S., Yang, W., Yang, X., Jianli, D., 2019. Comparison of machine learning algorithms for soil salinity predictions in three dryland oases located in Xinjiang Uyghur Autonomous Region (XJUAR) of China. *Eur. J. Remote Sens.* 52, 256–276.
- Wang, H., Wang, J., Liu, G., 2007. Spatial regression analysis on the variation of soil salinity in the Yellow River Delta, in: *Geoinformatics 2007: Geospatial Information Science*. International Society for Optics and Photonics, p. 67531U.
- Wang, J., Ding, J., Yu, D., Ma, X., Zhang, Z., Ge, X., Teng, D., Li, X., Liang, J., Lizaga, I., 2019. Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China. *Geoderma* 353, 172–187.
- Wang, J., Ding, J., Yu, D., Teng, D., He, B., Chen, X., Ge, X., Zhang, Z., Wang, Y., Yang, X., 2020. Machine learning-based detection of soil salinity in an arid desert region, Northwest China: A comparison between Landsat-8 OLI and Sentinel-2 MSI. *Sci. Total Environ.* 707, 136092.
- Wicke, B., Smeets, E., Dornburg, V., Vashev, B., Gaiser, T., Turkenburg, W., Faaij, A., 2011. The global technical and economic potential of bioenergy from salt-affected soils. *Energy Environ. Sci.* 4, 2669–2681.
- Wu, W., Zucca, C., Muhameed, A.S., Al-Shafie, W.M., Fadhil Al-Quraishi, A.M., Nangia, V., Zhu, M., Liu, G., 2018. Soil salinity prediction and mapping by machine learning regression in C entral M esopotamia, I raq. *L. Degrad. Dev.* 29, 4005–4014.
- Ying, B., Sayed, A.H., 2017. Diffusion gradient boosting for networked learning, in: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 2512–2516.
- Zarei, A., Shah-Hosseini, R., Ranjbar, S., Hasanlou, M., 2021. Validation of non-linear split window algorithm for land surface temperature estimation using Sentinel-3 satellite imagery: Case study; Tehran Province, Iran. *Adv. Sp. Res.* <https://doi.org/10.1016/j.asr.2021.02.019>
- Zhang, Q., Li, L., Sun, R., Zhu, D., Zhang, C., Chen, Q., 2020. Retrieval of the Soil Salinity From Sentinel-1 Dual-Polarized SAR Data Based on Deep Neural Network Regression. *IEEE Geosci. Remote Sens. Lett.* 1–5. <https://doi.org/10.1109/LGRS.2020.3041059>