Nguyen H.D. (2025), Soil salinity monitoring in the Red River Delta using gradient boosting and Sentinel 2 imagery, pp. 18-31. In Gastescu, P., Bretcan, P. (edit., 2025), Water resources and wetlands, *7th International Hybrid Conference Water resources and wetlands, 10-14 September 2025, Tulcea (Romania)*.
Available online at http://www.limnology.ro/wrw2025/proceedings.html

7th International Hybrid Conference Water resources and wetlands, 10-14 September 2025, Tulcea (Romania)

SOIL SALINITY MONITORING IN THE RED RIVER DELTA USING GRADIENT BOOSTING AND SENTINEL 2 IMAGERY

Huu Duy NGUYEN
[1]Faculty of Geography, VNU University of Science, Vietnam National University, Hanoi, Vietnam; nguyenhuuduy@hus.edu.vn

Abstract. Soil salinity is a major environmental problem that causes significant damage to agricultural production, especially in the context of climate change and rising sea levels. The problem is increasingly frequent in the Red River Delta, one of Vietnam's rice granaries. Monitoring soil salinity – in coastal regions in general and the Red River Delta in particular – can support decision-makers or farmers in proposing effective strategies to reduce the impacts of soil salinity and ensure food security. The objective of this study was to construct soil salinity distribution maps for the Red River Delta using machine learning and remote sensing – namely gradient boosting (GB). The statistical indices RMSE, MAE, and R^2 were used to evaluate the accuracy of the GB model. The results showed that this study was successfully able to build the GB model with an accuracy of R^2=0.69. The mapping showed that coastal regions have the highest electrical conductivity value (More than 3 mS/cm). The results of this study may prove important as tools in the selection of appropriate types of agricultural production to ensure food security in the region, especially in the context of climate change.
Keyword: Soil salinity, Red River, Vietnam, machine learning

1.INTRODUCTION

Soil salinity is a pervasive and dangerous ecological challenge, affecting agricultural production and sustainable land development, as well as ecosystem balance, particularly in coastal regions (Duan, Sun et al. 2025, Thangarasu, Mengash et al. 2025). The process of salinization leads to an increase in the accumulation of soluble salts in the soil profile, disrupting the ability of plants to access water and nutrients (Zhang, Fan et al. 2025). This challenge is increasingly serious due to climate change, rising sea levels, and poor irrigation management (Hailegnaw, Awoke et al. 2025, Zhou, Huang et al. 2025). According to the first global estimate in fifty years, 10.7% of land is affected by salinity. This has direct consequences for agriculture and food production, with losses of up to 70% of crops – particularly rice, beans, sugarcane, and potatoes. Therefore, monitoring and assessing soil salinity are considered essential for achieving the United Nations Sustainable Development Goals, as well as for supporting those engaged in agricultural planning to reduce the effects of salinization.

It is now considered important to characterize and monitor the evolution of affected soils. Traditionally, salinity has been measured using in situ methods, particularly measurement of electromagnetic conductivity (Nguyen, Dang et al. 2025, Nguyen, Pham et al. 2025). Although this technique is widely developed in precision agriculture for characterizing the spatial variability of salinity, it is limited when measuring salinity over large areas. Furthermore, this method is very expensive, labor-intensive, and time-consuming. Therefore, more effective methods must be developed for assessing spatial and temporal salinity efficiently and quickly. To limit the shortcomings of the traditional method, in recent years, remote sensing technology has been used to monitor environmental problems in general and soil salinity in particular (Metternicht and Zinck 2003, Wu, Mhaimeed et al. 2014, Gorji, Sertel et al. 2017, Thangarasu, Mengash et al. 2025). Radar and optical techniques have proven effective in providing global information, particularly in monitoring the soil surface (Lhissoui, El Harti et al. 2014, Jiang and Shu 2019, Taghadosi, Hasanlou et al. 2019, Wu, Muhaimeed et al. 2019). However, optical remote sensing products are limited due to cloud cover, and they depend on solar radiation. Radar sensors are considered reliable tools for monitoring the soil surface under any weather and time conditions. For agricultural soil, the radar signal is mainly dependent on surface parameters; for example, soil salinity coupled with humidity influences the dielectric properties of soils and consequently the radar signal. In addition, with the explosion of remote sensing data types, particularly in terms of variety, resolution and size, it is necessary to have robust methods to process these data accurately.

For these reasons, several researchers have integrated data from remote sensing and the Drive model. This model works mainly by learning the relationships between salinity locations and causes. They include statistical models and machine learning models. Although statistical models have been shown to be suitable for monitoring salinity in several regions around the world (Douaik, Van Meirvenne et al. 2007, Fallah Shamsi, Zare et al. 2013), the nonlinear problem is considered a major issue when using this model. This is particularly true in the context of climate change and sea level rise, which make the soil salinity problem increasingly complex and difficult to predict. Therefore, in recent years, some researchers have used machine learning to monitor soil salinity in coastal regions, including methods such as support vector machine (Guan, Wang et al. 2013, Jiang, Rusuli et al. 2019), random forest (Fathizad, Ardakani et al. 2020, Wang, Yang et al. 2021), bagging (Das, Rathore et al. 2022), and XGBoost (Zarei, Hasanlou et al. 2021, Aksoy, Sertel et al. 2024). Machine learning algorithms can analyze large volumes of data from different sources, including satellite images and in situ data, to estimate soil salinity. Studies have shown this method to have higher accuracy than traditional methods. Moreover, machine learning models can improve over time thanks to their continuous learning capacity and the updating of new data. However, a number of previous studies have that as regions have different natural and social conditions, no model can predict problems in all regions. Therefore, it is necessary to select models that adapt to each region of study.

The objective of this study was the monitoring of soil salinity in the Red River Delta using machine learning and Sentinel 2A data, specifically with the Gradient Boosting algorithm. The delta is the country's second most important rice granary, after the Mekong Delta, and so plays a key role in the agricultural development of Vietnam. It accounts for about 18% of the country's agricultural land area and 15-20% of the aquaculture land area. However, in recent years, the delta has been severely affected by saline intrusion, impacting agricultural development and food security in the region and the wider country. Monitoring this intrusion will help inform effective solutions to minimize the impact of this situation.

2.MATERIAL AND METHODOLOGY

2.1. Study area

Nam Dinh is a coastal province located in the Red River Delta, at the coordinates of 19054' to 20040' North latitude and 105055' to 106045' East longitude (Figure 1). The area has relatively flat terrain, mainly composed of lowland plains and coastal plains. The terrain gradually decreases from northwest to southeast, with an average elevation of +2m to +3m. Lowland plains account for most of the natural area of the province. It is also an area with great potential for intensive agricultural development, the processing and mechanical industries, and more traditional industry.

The 72 km-long coastline is fairly flat. The coastal plain has fertile land and great potential for economic development, especially in aquaculture and fisheries. Nam Dinh has a tropical monsoon climate. Average annual precipitation is 1750–1800 mm. The rainy season lasts from May to October, with the dry season from November to February. The study area is located between the downstream reaches of two major rivers in the north of the delta: the Red River and the Day River. The tide in the Nam Dinh region is diurnal. The average tidal range is 1.6 to 1.7 m, the highest being 3.31 m and the lowest being 0.11 m.

Due to its geographical location, the province is affected by storms and tropical depressions, with an average of 4 to 6 storms per year. These phenomena not only cause damage to the region's socio-economic situation but also increase saltwater intrusion in coastal areas such as Hai Hau and Truc Ninh, particularly in the context of climate change. Changes in climatic and hydrological conditions, combined with the degradation of the dike and lock system, have made saline intrusion increasingly serious in Nam Dinh province in general and in coastal areas in particular.

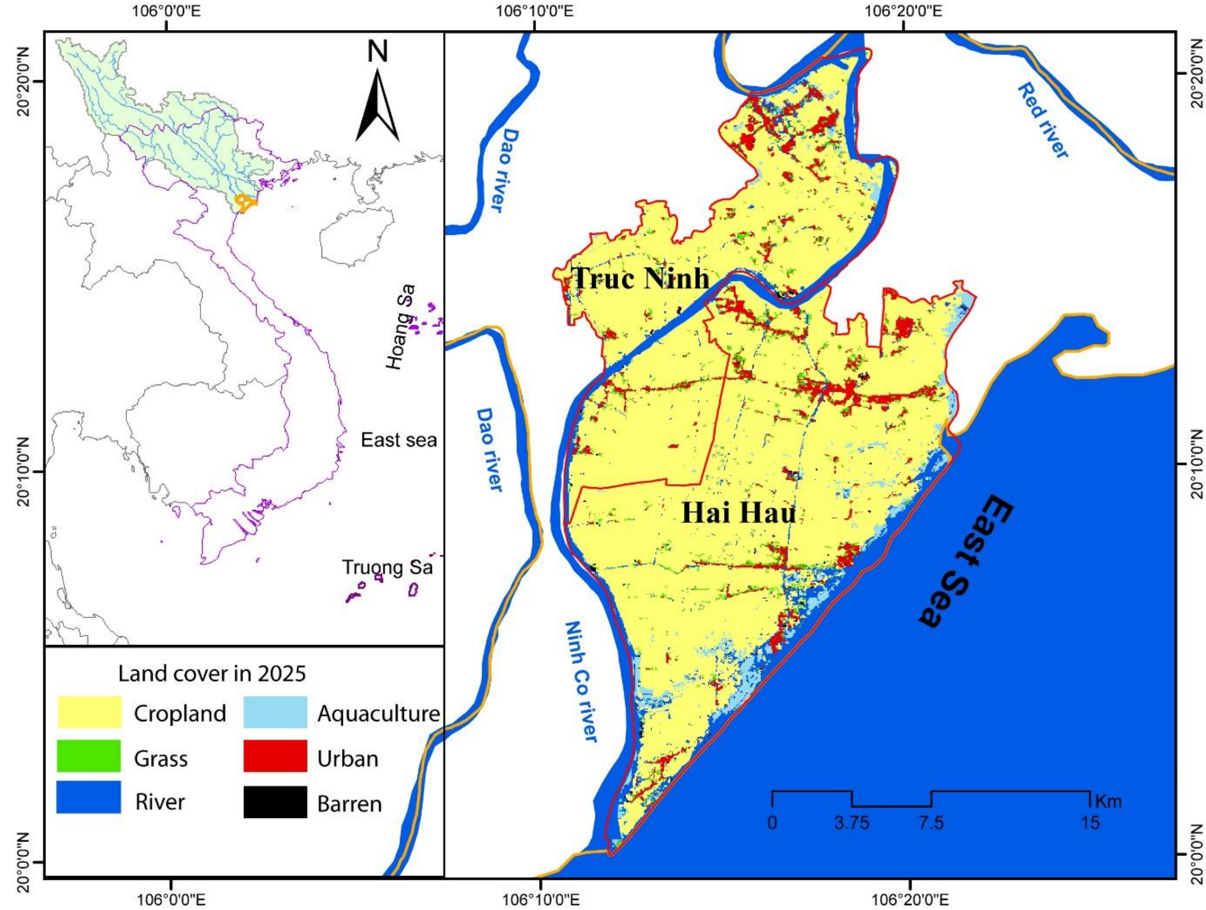

Figure 1. Location of study area

2.2. Soil salinity collection

In this study, 72 soil salinity points were collected in August 2024 in Hai Hau and Truc Ninh districts. We chose the 5-point method for salinity sampling, which means that 4 sub-samples were taken around the main sample and the salinity of each was the average salinity of the 5. After boiling, the samples were sealed in plastic bags and sent to the laboratory for analysis. They were cleaned of impurities before being ground and mixed with distilled water at a ratio of 1:5 (meaning 1 g of soil was mixed with 5 ml of distilled water). The soil and water mixture was then mixed in a thermal oscillator for about 30 minutes to dissolve the soil in the water. The mixture was then allowed to stand, for the liquid to be extracted, and the EC was measured using a water quality analyzer. Finally, 56 soil samples combined with influencing variables were used as input data for the machine learning model. These data were divided into 2 parts: 70% of the data used to build the GB model and 30% to evaluate the model’s accuracy.

2.3.Conditioning factors

Conditioning factors are essential data when using machine learning to monitor soil salinity, because machine learning is based on examining the relationships between salinity locations and causes to estimate salinity. In this study, conditioning factors were collected from different sources, such as remote sensing data and data from the Ministry of Natural Resources and the Environment. 20 conditioning factors were selected to build the machine learning model. These factors were divided into three main groups: topographic factors (altitude, aspect, curvature, slope), hydrometeorology factors (distance to river, rainfall), vegetation groups (NDVI, EVI, ENDVI, RVI, SAVI, GDVI), salinity indices (S1, S2, S3, S5, S6, SI, SI1, SI2, SI3, SI4) and intensity indexes (Int1, Int2). The topographic factors (altitude, aspect, curvature, slope) were extracted by the topographic map with the scale of 1:50,000 (available at the Ministry of Natural Resources and Environment). 5000 / 5000. The groups focusing on vegetation, salinity, intensity, and brightness were extracted from Sentinel 2A imagery from 2024. Annual precipitation was downloaded by https://chrsdata.eng.uci.edu/.

Topographic factors (aspect, altitude, slope, curvature) directly affect the salinity accumulation process and hydrological flow. In particular, altitude affects the salt accumulation capacity in the soil: low-lying areas are susceptible to saltwater intrusion from the sea to the mainland, especially during the dry season when river levels drop and sea levels rise. In addition, low-lying areas have poor drainage, increasing the risk of salt accumulation in the soil (Akramkhanov, Martius et al. 2011, Shahrayini and Noroozi 2022). Aspect is an important topographical factor that greatly affects saltwater intrusion, as it relates to soil moisture, which indirectly affects the salinity intrusion process (Oster and Shainberg 2001). Curvature is an essential element in calculating and monitoring saltwater intrusion, since this factor affects accumulation and surface flow: saltwater often concentrates in areas of low curvature. Slope directly affects drainage capacity (Triki Fourati, Bouaziz et al. 2017); steep slopes often have better drainage, which minimizes saltwater accumulation on the surface. Conversely, areas with low slopes increase the potential for salt accumulation in the soil (Yahiaoui, Douaoui et al. 2015, O'Brien, Almaraz et al. 2019).

Precipitation directly affects salinity intrusion in any region of the world. High rainfall reduces the effects of salt concentrations in the soil through leaching. At the same time, low rainfall reduces river flow, allowing salt water to penetrate deeper inland (Isidoro and Grattan 2011, Dasgupta, Hossain et al. 2015).

Distance to river plays an important role in assessing and monitoring soil salinity, particularly in coastal regions and deltas. Because salt water can enter rivers from estuaries, affecting soil quality, the closer the regions are to the river, the more severe the effects of salinity are (Zhao, Feng et al. 2016, Liu, Wu et al. 2023).

The factors relating to vegetation are important in monitoring saltwater intrusion because they reflect environmental conditions at the surface. In this case, NDVI was used to assess the level of vegetation cover in an area. Areas with a high NDVI index often have a high vegetation density, which indicates that the vegetation is well-developed, helping to maintain soil moisture and reduce salt intrusion (Aldakheel 2011).

Similarly, ENDVI reflects the extent of vegetation cover, especially in areas with low soil moisture or those affected by adverse environmental conditions such as salinity intrusion. They determine the growth status of vegetation under changing soil and climate conditions. Areas with a high ENDVI index represent the vegetation growing in that area. This indicates that these areas are less affected by saltwater intrusion (Wu, Jia et al. 2021).

EVI provides information about the growth status of plants. High EVI values indicate good vegetation growth, suggesting that these areas are less affected by salinity intrusion. A low EVI shows that plants are not growing, which demonstrates the level of influence of salt intrusion (Lobell, Lesch et al. 2010, Ivushkin, Bartholomeus et al. 2017).

RVI and SAVI provide information on the extent of vegetation growth, which helps assess the extent of salinity intrusion affecting an area. Generally, saltwater intrusion reduces the ability of plants to absorb nutrients and water. This results in poor plant growth. Therefore, areas with high

RVI and SAVI values show that the vegetation in that area is well developed, indicating that this area is less affected by salinity intrusion (Teshaev, Mamadaliyev et al. 2020, Zhu, Sun et al. 2021).
GDVI reflects plant health, which is indispensable for assessing and monitoring salinity intrusion. GDVI is inversely proportional to salinity intrusion, meaning that areas with high salinity have low GDVI values, resulting in poor plant health (Wu, Mhaimeed et al. 2014).

Salinity-related factors play an important role in assessing and monitoring saltwater intrusion in any area because they reflect changes in surface conditions and vegetation in an area using near-infrared, short-infrared, red, and blue wavelengths. These indices assess the extent to which salinity intrusion affects plant growth. When the values of the S1, S2, S3, S5, and S6 indices are low, it means that the vegetation in that area is affected by salinity intrusion (Wang, Chen et al. 2019, Wang, Peng et al. 2021). Similarly, the SI, SI1, SI2, SI3, and SI4 indices are often used to assess plant health and soil moisture. When these indices are high, it indicates that the soil is affected by severe salinity intrusion, causing poor plant growth and reduced soil moisture. Assessing salinity intrusion through plant health and soil moisture provides an overview of the status of salinity intrusion (Wang, Chen et al. 2019, Naimi, Ayoubi et al. 2021).

The indices Int1 and Int2 reflect the intensity of environmental factors affecting saltwater intrusion. Int1 provides comprehensive information on the light reflected off the Earth's surface, which is influenced by factors such as soil moisture and texture. With this information, it is possible to indirectly assess the level of impact of saline intrusion on an area. Int2 reflects the reflectance properties of the soil in an area through the intensity of the visible spectrum. This index assesses the condition and surface characteristics of an area, as well as the vegetation growth index. High Int 2 values reflect good vegetation growth and the area is less affected by salinity intrusion (Sidike, Zhao et al. 2014, Zhang, Fan et al. 2022).

The brightness index (BI) reflects the brightness of an area calculated by combining reflectance values from the visible and near-infrared bands. Areas with high brightness have strong surface reflections, indicating that the area has little vegetation or low humidity. This shows that they may be significantly affected by saltwater intrusion (Yahiaoui, Douaoui et al. 2015).

2.4.Methodology

The methodology used to construct the soil salinity map in this study was divided into three main tasks: i) data collection, ii) GB model construction, iii) evaluation of the proposed model accuracy, and iv) construction and analysis of the soil salinity map (Figure 2).

i) Data collection: Databases were divided into two types: soil salinity points and conditioning factors. Salinity points were collected from the field mission in January 2025, during the dry season. Conditioning factors included five groups of factors: topographic, vegetation, salinity, intensity, and brightness. The databases were divided into two groups: one was used to construct the GB model and the other to evaluate the model.

ii) Machine learning model construction: the machine learning model was built on the TensorFlow platform. The accuracy of the GB model depends on the adjustment of parameters such as max depth, learning rate, max feature, and random state. In this study, the trial-and-error method was used to optimize the model parameters. After several trials, the model performed better with the following parameters: n_estimators = 100, max_depth = 3, learning_rate = 0.15, loss = 'squared_error', criterion = 'friedman_mse', max_features = 'sqrt', random_state = None.

iii) Performance evaluation of the proposed model: The statistical indices RMSE, MAE, and R^2 were used to evaluate the performance of the GB model. These indices have been used in several previous studies.

iv) Construction and analysis of the soil salinity map: After training and evaluating the machine learning models, the models were used to estimate salinity intrusion for the entire study area, including Hai Hau District and Truc Ninh District. This process was carried out by assigning influence factor values to all pixels in the study area, from which the model can calculate the salinity value at each pixel.

After calculating the salinity values at each pixel, these values were used to construct a soil salinity distribution map for the study area.

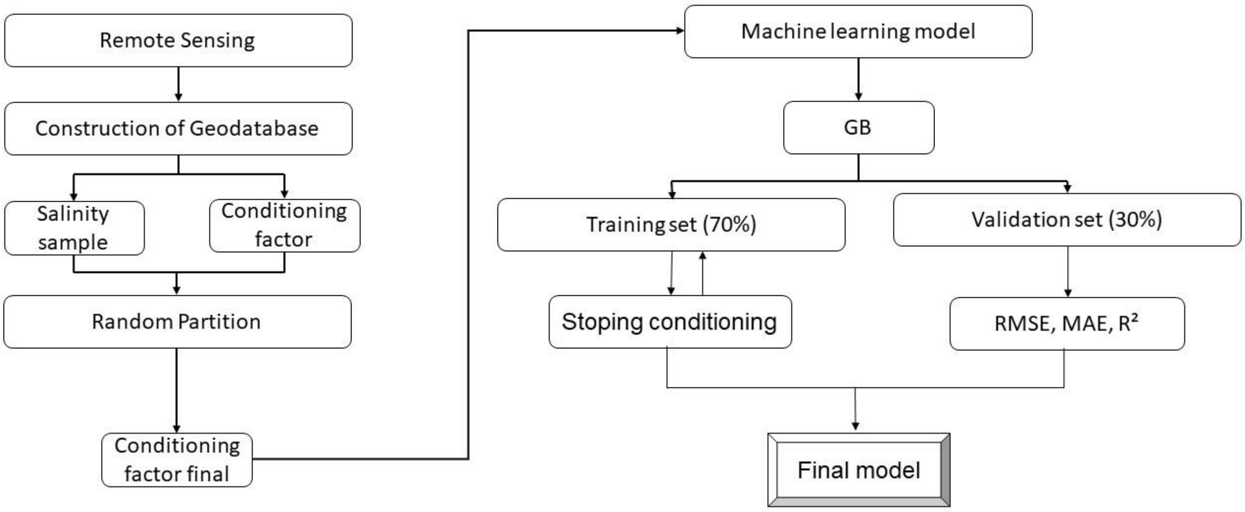

Figure 2. Methodology used for soil salinity model

2.5.Gradient boosting

GB is a machine learning model based on the principle of boosting. This technique was developed based on the idea of progressively combining weak models, most often decision trees, to form a powerful predictive model by training a new model that successfully predicts the errors made by the original model (Natekin and Knoll 2013, Bentéjac, Csörgő et al. 2021). With this idea, for any predictive model, we can improve its accuracy by training a new predictor to predict its current errors. This process is repeated an arbitrary number of times to continuously improve the model's accuracy (Zhang and Haghani 2015). Gradient boosting machines consist of three main components: i) The loss function measures the difference between predicted values and actual values. ii) Base learners built sequentially, each focusing on correcting the errors made by the previous tree. iii) The additive model combines the predictions of all base learners to produce the final prediction (Ayyadevara 2018). The GB model training process includes the following steps (Taieb and Hyndman 2014, Lusa 2017):

i) Model initialization: this process fits the base learner to the dataset to make an initial prediction.

ii) Iteratively fitting weak learners: the model concentrates errors made by previous learners.

iii) Calculating residuals: the difference is measured between the actual value and the current model's prediction.

iv) Fitting a weak learner to the residuals: the learner attempts to correct the errors of the existing model by predicting the errors made by the previous model.

v) Updating the model: this combines the existing model using a "weighted voting" system with the new learner. This is typically done by adding the new learner's predictions to the existing model, scaled by a learning rate.

vi) Loss function optimization: a loss function measures the model's performance in predicting previous errors.

Model assessment

We used three statistical indices, namely RMSE, MAE, and R^2, to evaluate the accuracy of the GB model. Of the three, RMSE and MAE presented the differences between the model values and the realization value. The closer the value was to 0, the more perfect the model was. Meanwhile, the R^2 value measures the quality of the correlation between soil salinity and conditioning factors. This gives, among other things, an indication of the reliability of the salinity estimation based on these factors. An R^2 equal to 75% would mean that 75% of the variations in soil salinity could be explained by the conditioning factors.

3.RESULTS

Conditioning factor analysis

The selection of appropriate conditioning factors plays an important role in accurately estimating soil salinity because the accuracy of machine learning models depends on analyzing the correlation between conditioning factors and soil salinity. Previous studies used various methods to evaluate the importance of conditioning factors. For example, Zhang, Fu et al. (2023) used Pearson's correlation coefficients to evaluate the importance of conditioning factors. The results showed that B6, B7, B10, B11, B4, B5, and SI, NDSI, ENDVI, GDVI are more important for soil salinity monitoring in Kenli district, located in Dongying City of Shandong Province in China. Habibi, Ahmadi et al. (2020) applied sensitivity analysis of the wrapper model to evaluate the importance of conditioning variables. The results showed that SI3, band 5, NDVI, TWI and LS were more important in defining the soil salinity in Farmlands of Saveh Plain in the north and northeast of Qom. Han, Ge et al. (2023) applied Pearson correlation analysis to select appropriate conditioning factors for a soil salinity model. The results showed that BI, VSSI, CRSI, EVI, RVI, SAVI were most impactful on soil salinity in the Songnen Plain of Da'an City.

The importance of individual conditioning factors is clearly different in each region. In this study, we used random forest to select conditioning factors. The results showed that distance to river, ENDVI, RVI, and rainfall were the most important factors for the soil salinity in the Edu region, with RF values of 0.5, 0.23, 0.22, 0.19, respectively.

The Red, Thai Binh, and Ninh Co rivers are the main sources of saline intrusion, especially during the dry season when the tide rises and the river water level decreases. Therefore, areas near rivers will be affected by saltwater intrusion, especially in low-lying areas. ENDVI and RVI are second in importance after distance to the river, as the study area is primarily devoted to rice cultivation. Thus, when affected by salinity intrusion, rice either cannot grow, or it grows poorly. This explains why the RF value is high after distance to the river. The importance of precipitation comes fourth; precipitation plays an important role in regulating soil salinity because it helps remove salts from the soil, especially in estuarine areas. Therefore, precipitation directly affects soil salinity. Aspect, curvature, S6, and SI2 did not influence soil salinity. It should be noted that the model used in this study was the statistical model; therefore, the relationships between salinity locations and conditioning factors are more important. The soil salinity in the Red River Delta in general, and in Hai Hau and Truc Ninh in particular, is influenced by distance to river, rainfall, RVI, and ENDVI (Figure 3).

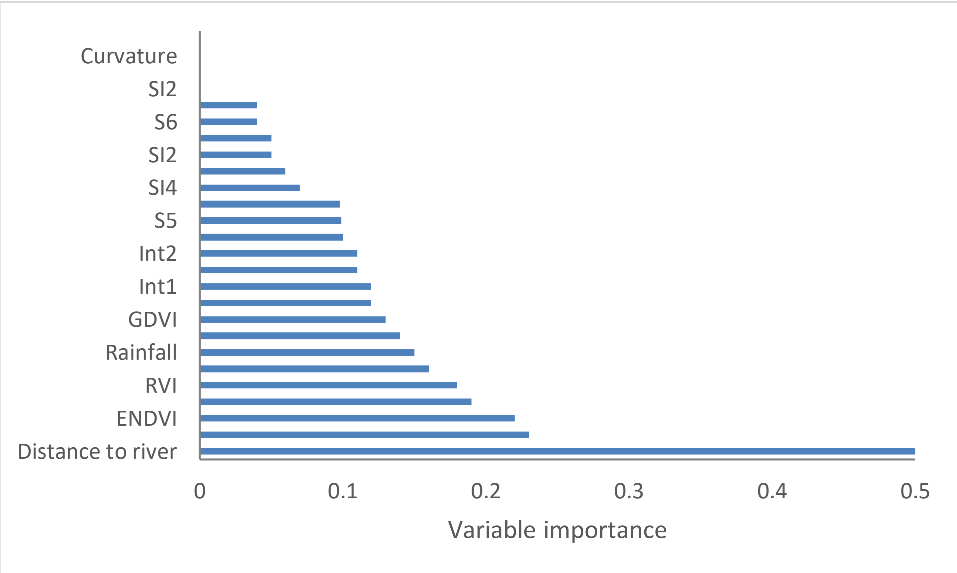

Figure 3. Importance of conditioning factors on the soil salinity model

Modelling accuracy verification

Figure 4 and Table 1 present the performance of the GB model in estimating soil salinity in the study area. The results showed that the R^2 value of the model was 0.69. This can be explained by the fact that 69% of the measured point values could be accurately predicted by the proposed model. In addition, this study used two other indices – RMSE and MAE – to evaluate the accuracy of the model. The results showed that, for the training data, the RMSE and MAE values reached 0.02 and 0.04, while for the validation data, these values increased to 0.3 and 0.2. Overall, even though the model still needs improvement compared to some previous studies, the model in our study remains effective (Wang, Xue et al. 2020, Shi, Hellwich et al. 2021).

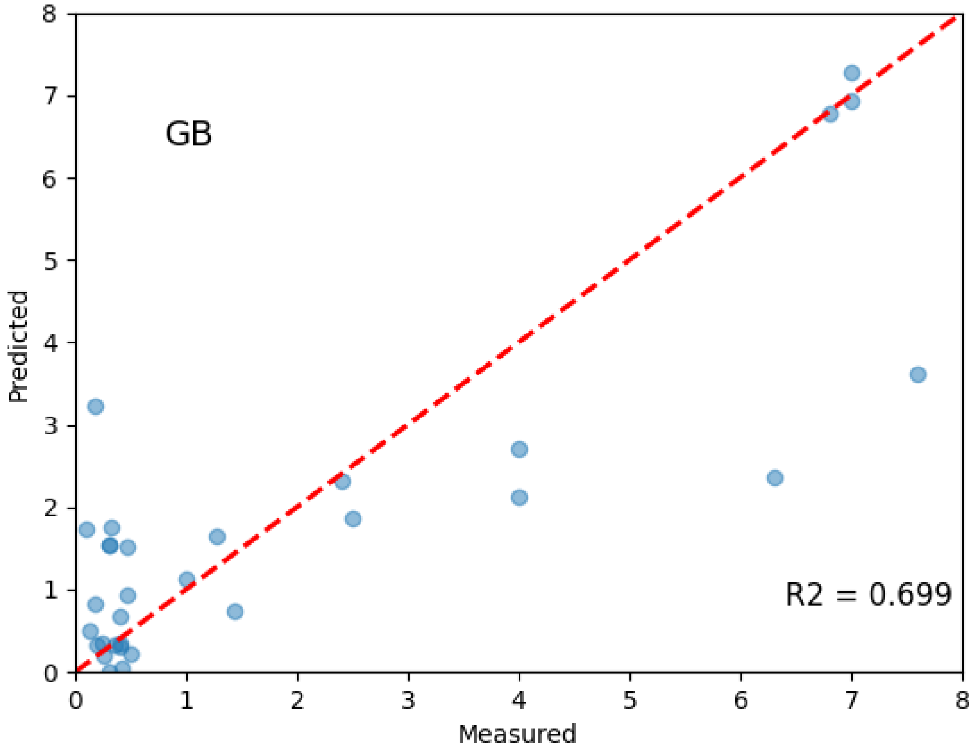

Figure 4: R^2 value for the GB model

Table 1: Model performance for soil salinity model

Model	Training dataset			Validation dataset		
	RMSE	MAE	R2	RMSE	MAE	R2
GB	0.02	0.04	0.98	0.3	0.2	0.69

Soil salinity mapping

Figure 5 shows the distribution of EC values in the study area, which is considered a very important value in the assessment of salinity intrusion. It can be seen that coastal areas are strongly affected by saline intrusion with EC values greater than 3 mS/cm. Meanwhile, further inland, EC values decrease to less than 3.

In deltaic areas in general and in the study area in particular, salinity intrusion is a natural process, but it is also affected by many other factors, such as agricultural production activities and unreasonable groundwater exploitation. This process affects the relationship between river water and seawater. The map shows the study area clearly divided into three zones: red, yellow, and green; salinity decreases from red to green. In red zones, EC values exceed 3; these are concentrated in coastal areas. This area is heavily affected by storms, sea level rise, tides, and water exploitation in the upper Red River basin. Therefore, in some areas, instead of growing rice, people work in aquaculture.

As you move further inland, EC values decrease and become yellow. Based on field surveys conducted in 2024 and 2025, it can be seen that while the yellow areas can grow rice, rice yields have declined, and many areas have switched from growing rice to growing more salt-tolerant crops. The further inland you go, the more significantly salinity decreases. These areas focus on growing rice.

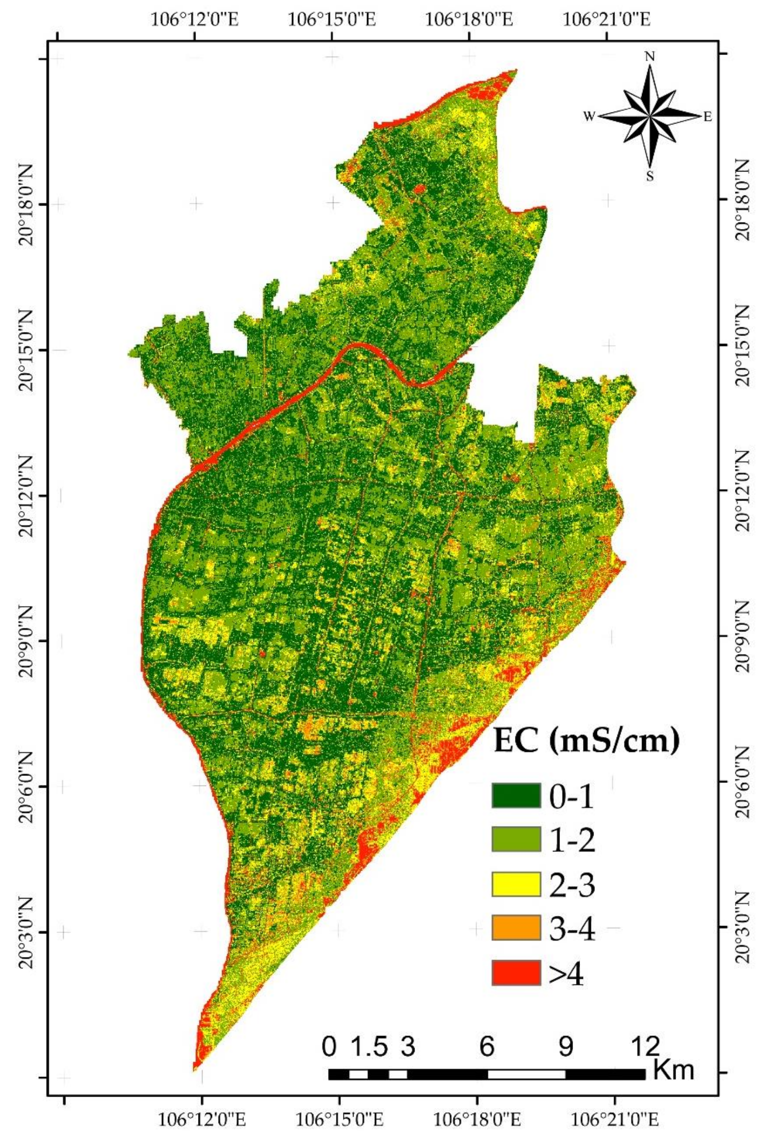

Figure 5. Soil salinity maping product by GB in the study area

4.DISCUSSION

The Red River Delta plays an important role in the socio-economic development of Vietnam, with its population of more than 20 million, of whom 66% of the workforce relies on agriculture and aquaculture. However, in recent years, the region has been strongly affected by drought and saltwater intrusion, especially in the context of climate change. This has seriously affected the food security of the whole country (Phan and Kamoshita 2020, Truc, Mihova et al. 2020). According to a report by the Ministry of Agriculture and Environment, the dry season flow on the Red River-Thai Binh River system has decreased sharply. Since 2001, the water level on the Red River in Hanoi from December to May is often 0.5 to 1.1 m lower than the average level of many years. The water flow downstream has decreased. Also, the water level of rivers in the coastal plains has dropped, combining with rising sea levels and high tides to cause saltwater intrusion to spread further into the area downstream. Therefore, research and monitoring of saline intrusion in the Red River Delta is necessary to support planners and farmers in choosing appropriate agricultural production types (Nguyen, Renaud et al. 2019, Phan and Kamoshita 2020).

In the delta, solutions to mitigate saline intrusion mainly focus on upgrading irrigation infrastructure systems such as sea dikes, river dikes and sluice systems. In addition, adaptation measures also include changes in agricultural production structures such as changing crop calendars and effectively managing irrigation systems (Nguyen, Kamoshita et al. 2017). Other activities include planting mangroves or converting coastal rice fields affected by saline intrusion into aquaculture systems, which has been achieved by many households (Nguyen, Renaud et al. 2019, Phan and Kamoshita 2020). As the results show, in areas far from the sea, saline intrusion is not a problem; only a small number of some areas along the Red River are affected by salinity. Many households in this area have raised their fields with sand from the Red River to minimize saline intrusion. Many households have cultivated rice-vegetable systems, in which rice is mainly used for family consumption, while vegetable prices may determine whether to grow vegetables or other crops. When vegetable prices decrease, they switch to other crops with greater economic benefits (Linh, Linh et al. 2012, Nguyen, Kamoshita et al. 2017).

In the middle areas, affected by moderate saline intrusion, many households often switch to fish or soft-shell turtle farming. Some households switch from rice cultivation to garden-pond-cage systems when salinity increases. Along coastal areas, households may switch from fish farming to shrimp farming. Regardless of the area, farmers in the Red River Delta in general and the study area in particular continue their farming methods, even after many years of crop failure due to salinity intrusion.

Many studies have also emphasized that saline intrusion is becoming more and more serious, in addition to the impact of climate change and sea level rise. The increasing human control over the delta region for the purpose of rice intensification is also one of the causes leading to the worsening situation (Nguyen, Kamoshita et al. 2017, Phan and Kamoshita 2020, Yuen, Hanh et al. 2021). The strategy of shifting from large-scale interventions into nature by means of structures to agricultural production forms suitable for ecological conditions will be the appropriate choice to maintain resources. Although the Mekong Delta and the Red River Delta propose flexible land use strategies to limit the impact of drought and saline intrusion, the application of structural measures can affect the ecosystem (Yuen, Hanh et al. 2021, Hien, Yen et al. 2023). Therefore, to minimize the impact of saline intrusion, it is necessary to identify areas affected by saline intrusion in order to propose appropriate adaptation measures.

This study was successful in constructing a map of saline intrusion distribution in a small study area. It has limitations related to the data used to build the machine learning model. Firstly, we collected 56 saline intrusion points to build the machine learning model. It is in fact necessary to collect more saline intrusion points in different seasons of the year. In addition, saline intrusion is strongly affected by climate change and sea level rise: in the future, we will try to assess the impact of climate change scenarios on saline intrusion, thereby supporting planners and local authorities to come up with appropriate adaptation solutions to minimize the impact of saline intrusion.

5.CONCLUSION

Soil salinity is considered a major environmental problem, causing significant impacts on agricultural production, especially in the context of climate change and rising sea levels. In recent years, the Mekong and Red River deltas have been seriously affected by the problem, leading to significant impacts on food supply in the region. Monitoring soil salinity is an essential task to support decision-makers or farmers in selecting appropriate crop types. Therefore, the objective of this study was to construct a soil salinity distribution map in the Red River Delta using machine learning and remote sensing. The results were as follows.

- This study successfully built a GB model with an R^2 value of 0.69 to construct a soil salinity distribution map for the Red River Delta in general, and the Truc Ninh and Hai Hau districts in particular. This model can be adjusted to construct a map of soil salinity distribution in other regions of the world.
- Coastal regions in the study area are more affected by the soil salinity problem, with an EC value greater than 3 mS/cm. This may be directly linked to climate change and sea level rise.

The results of this study provide a scientific basis for monitoring soil salinity in coastal regions, which can support decision-makers or farmers in sustainable land use planning to reduce the effects of soil salinity.

Funding
This research has been done under the research project QG.24.76 of Vietnam National University, Hanoi.

REFERENCE

Akramkhanov, A., C. Martius, S. J. Park and J. M. Hendrickx (2011). Environmental factors of spatial distribution of soil salinity on flat irrigated terrain. *Geoderma* **163**(1-2): 55-62.

Aksoy, S., E. Sertel, R. Roscher, A. Tanik and N. Hamzehpour (2024). Assessment of soil salinity using explainable machine learning methods and Landsat 8 images. *International Journal of Applied Earth Observation and Geoinformation* **130**: 103879.

Aldakheel, Y. Y. (2011). Assessing NDVI spatial pattern as related to irrigation and soil salinity management in Al-Hassa Oasis, Saudi Arabia. *Journal of the indian Society of Remote Sensing* **39**: 171-180.

Ayyadevara, V. K. (2018). Gradient boosting machine. *Pro machine learning algorithms: A hands-on approach to implementing algorithms in python and R,* Springer**:** 117-134.

Bentéjac, C., A. Csörgő and G. Martínez-Muñoz (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review* **54**: 1937-1967.

Das, B., P. Rathore, D. Roy, D. Chakraborty, R. S. Jatav, D. Sethi and P. Kumar (2022). Comparison of bagging, boosting and stacking algorithms for surface soil moisture mapping using optical-thermal-microwave remote sensing synergies. *Catena* **217**: 106485.

Dasgupta, S., M. M. Hossain, M. Huq and D. Wheeler (2015). Climate change and soil salinity: The case of coastal Bangladesh. *Ambio* **44**: 815-826.

Douaik, A., M. Van Meirvenne and T. Tóth (2007). Statistical methods for evaluating soil salinity spatial and temporal variability. *Soil Science Society of America Journal* **71**(5): 1629-1635.

Duan, R., Y. Sun, J. Gao and B. Zhu (2025). Predicting soil salinity in wastewater land application systems under field conditions with tree-based machine learning approaches. *Irrigation Science*: 1-14.

Fallah Shamsi, S. R., S. Zare and S. A. Abtahi (2013). Soil salinity characteristics using moderate resolution imaging spectroradiometer (MODIS) images and statistical analysis. *Archives of Agronomy and Soil Science* **59**(4): 471-489.

Fathizad, H., M. A. H. Ardakani, H. Sodaiezadeh, R. Kerry and R. Taghizadeh-Mehrjardi (2020). "Investigation of the spatial and temporal variation of soil salinity using random forests in the central desert of Iran." Geoderma **365**: 114233.

Gorji, T., E. Sertel and A. Tanik (2017). Monitoring soil salinity via remote sensing technology under data scarce conditions: A case study from Turkey. *Ecological indicators* **74**: 384-391.

Guan, X., S. Wang, Z. Gao and Y. Lv (2013). Dynamic prediction of soil salinization in an irrigation district based on the support vector machine. *Mathematical and Computer Modelling* **58**(3-4): 719-724.

Habibi, V., H. Ahmadi, M. Jafari and A. Moeini (2020). Machine learning and multispectral data-based detection of soil salinity in an arid region, Central Iran. *Environmental Monitoring and Assessment* **192**: 1-13.

Hailegnaw, N. S., G. W. Awoke, A. d. C. Santos, B. Schaffer, A. I. Vargas, E. R. d. Souza and H. K. Bayabil (2025). Assessing salinity-induced impacts on plant transpiration through machine learning: from model development to deployment. *Modeling Earth Systems and Environment* **11**(3): 1-14.

Han, Y., H. Ge, Y. Xu, L. Zhuang, F. Wang, Q. Gu and X. Li (2023). Estimating soil salinity using multiple spectral indexes and machine learning algorithm in songnen plain, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **16**: 7041-7050.

Hien, N. T., N. H. Yen, M. Balistrocchi, M. Peli, V. M. Cat and R. Ranzi (2023). Salinity dynamics under different water management plans coupled with sea level rise scenarios in the Red River Delta, Vietnam. *Journal of Hydro-environment Research* **51**: 67-81.

Isidoro, D. and S. Grattan (2011). Predicting soil salinity in response to different irrigation practices, soil types and rainfall scenarios. *Irrigation Science* **29**: 197-211.

Ivushkin, K., H. Bartholomeus, A. K. Bregt and A. Pulatov (2017). Satellite thermography for soil salinity assessment of cropped areas in Uzbekistan. *Land degradation & development* **28**(3): 870-877.

Jiang, H., Y. Rusuli, T. Amuti and Q. He (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.

Jiang, H. and H. Shu (2019). Optical remote-sensing data based research on detecting soil salinity at different depth in an arid-area oasis, Xinjiang, China. *Earth Science Informatics* **12**: 43-56.

Lhissoui, R., A. El Harti and K. Chokmani (2014). Mapping soil salinity in irrigated land using optical remote sensing data. *Eurasian Journal of Soil Science* **3**(2): 82-88.

Linh, L. H., T. H. Linh, T. D. Xuan, L. H. Ham, A. M. Ismail and T. D. Khanh (2012). Molecular breeding to improve salt tolerance of rice (Oryza sativa L.) in the Red River Delta of Vietnam. *International journal of plant genomics* **2012**(1): 949038.

Liu, L., Y. Wu, M. Yin, X. Ma, X. Yu, X. Guo, N. Du, F. Eller and W. Guo (2023). Soil salinity, not plant genotype or geographical distance, shapes soil microbial community of a reed wetland at a fine scale in the Yellow River Delta. *Science of the Total Environment* **856**: 159136.

Lobell, D., S. Lesch, D. Corwin, M. Ulmer, K. Anderson, D. Potts, J. Doolittle, M. Matos and M. Baltes (2010). Regional-scale assessment of soil salinity in the Red River Valley using multi-year MODIS EVI and NDVI. *Journal of environmental quality* **39**(1): 35-41.

Lusa, L. (2017). Gradient boosting for high-dimensional prediction of rare events. *Computational Statistics & Data Analysis* **113**: 19-37.

Metternicht, G. I. and J. Zinck (2003). Remote sensing of soil salinity: potentials and constraints. *Remote sensing of environment* **85**(1): 1-20.

Naimi, S., S. Ayoubi, M. Zeraatpisheh and J. A. M. Dematte (2021). Ground observations and environmental covariates integration for mapping of soil salinity: a machine learning-based approach. *Remote Sensing* **13**(23): 4825.

Natekin, A. and A. Knoll (2013). Gradient boosting machines, a tutorial. *Frontiers in neurorobotics* **7**: 21.

Nguyen, H. D., D. K. Dang, T. A. T. Lai, D. D. Tran, H. Shahabi and Q.-T. Bui (2025). Farmers' adaptive capacity towards soil salinity effects using hybrid machine learning in the Red River Delta. *EGUsphere* **2025**: 1-34.

Nguyen, H. D., V. T. Pham, Q.-H. Nguyen and Q.-T. Bui (2025). Soil Salinity Prediction Using Satellite-Based Variables and Machine Learning: Case study in Tra Vinh province, Mekong Delta, Vietnam. *Vietnam Journal of Earth Sciences*: 216-234.

Nguyen, M. T., F. G. Renaud and Z. Sebesvari (2019). Drivers of change and adaptation pathways of agricultural systems facing increased salinity intrusion in coastal areas of the Mekong and Red River deltas in Vietnam. *Environmental science & policy* **92**: 331-348.

Nguyen, Y. T. B., A. Kamoshita, V. T. H. Dinh, H. Matsuda and H. Kurokura (2017). Salinity intrusion and rice production in Red River Delta under changing climate conditions. *Paddy and Water Environment* **15**: 37-48.

O'Brien, F. J., M. Almaraz, M. A. Foster, A. F. Hill, D. P. Huber, E. K. King, H. Langford, M.-A. Lowe, B. S. Mickan and V. S. Miller (2019). Soil salinity and pH drive soil bacterial community composition and diversity along a lateritic slope in the Avon River critical zone observatory, Western Australia. *Frontiers in microbiology* **10**: 1486.

Oster, J. and I. Shainberg (2001). Soil responses to sodicity and salinity: challenges and opportunities. *Soil Research* **39**(6): 1219-1224.
Phan, L. T. and A. Kamoshita (2020). Salinity intrusion reduces grain yield in coastal paddy fields: case study in two estuaries in the Red River Delta, Vietnam. *Paddy and Water Environment* **18**: 399-416.
Shahrayini, E. and A. A. Noroozi (2022). Modeling and mapping of soil salinity and alkalinity using remote sensing data and topographic factors: a case study in Iran. *Environmental Modeling & Assessment* **27**(5): 901-913.
Shi, H., O. Hellwich, G. Luo, C. Chen, H. He, F. U. Ochege, T. Van de Voorde, A. Kurban and P. De Maeyer (2021). A global meta-analysis of soil salinity prediction integrating satellite remote sensing, soil sampling, and machine learning. *IEEE Transactions on Geoscience and Remote Sensing* **60**: 1-15.
Sidike, A., S. Zhao and Y. Wen (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.
Taghadosi, M. M., M. Hasanlou and K. Eftekhari (2019). Soil salinity mapping using dual-polarized SAR Sentinel-1 imagery. *International journal of remote sensing* **40**(1): 237-252.
Taieb, S. B. and R. J. Hyndman (2014). A gradient boosting approach to the Kaggle load forecasting competition. *International journal of forecasting* **30**(2): 382-394.
Teshaev, N., B. Mamadaliyev, A. Ibragimov and S. Khasanov (2020). "The soil-adjusted vegetation index for soil salinity assessment in Uzbekistan. *ICIGIS* **26**: 324-333.
Thangarasu, T., H. A. Mengash, R. Allafi and H. Mahgoub (2025). Spatial prediction of soil salinity: Remote sensing and machine learning approach. *Journal of South American Earth Sciences* **156**: 105440.
Triki Fourati, H., M. Bouaziz, M. Benzina and S. Bouaziz (2017). Detection of terrain indices related to soil salinity and mapping salt-affected soils using remote sensing and geostatistical techniques. *Environmental monitoring and assessment* **189**: 1-11.
Truc, N. N., L. Mihova, T. Mukunoki and D. M. Do (2020). Effect of saline intrusion on the properties of cohesive soils in the Red River Delta, Vietnam. *Marine Georesources & Geotechnology* **38**(1): 23-39.
Wang, F., S. Yang, Y. Wei, Q. Shi and J. Ding (2021). Characterizing soil salinity at multiple depth using electromagnetic induction and remote sensing data with random forests: A case study in Tarim River Basin of southern Xinjiang, China. *Science of the Total Environment* **754**: 142030.
Wang, J., J. Peng, H. Li, C. Yin, W. Liu, T. Wang and H. Zhang (2021). Soil salinity mapping using machine learning algorithms with the Sentinel-2 MSI in arid areas, China. *Remote Sensing* **13**(2): 305.
Wang, N., J. Xue, J. Peng, A. Biswas, Y. He and Z. Shi (2020). Integrating remote sensing and landscape characteristics to estimate soil salinity using machine learning methods: A case study from Southern Xinjiang, China. *Remote Sensing* **12**(24): 4118.
Wang, S., Y. Chen, M. Wang and J. Li (2019). Performance comparison of machine learning algorithms for estimating the soil salinity of salt-affected soil using field spectral data. *Remote sensing* **11**(22): 2605.
Wu, D., K. Jia, X. Zhang, J. Zhang and H. T. Abd El-Hamid (2021). Remote sensing inversion for simulation of soil salinization based on hyperspectral data and ground analysis in Yinchuan, China. *Natural Resources Research* **30**(6): 4641-4656.
Wu, W., A. S. Mhaimeed, W. M. Al-Shafie, F. Ziadat, B. Dhehibi, V. Nangia and E. De Pauw (2014). Mapping soil salinity changes using remote sensing in Central Iraq. *Geoderma Regional* **2**: 21-31.
Wu, W., A. S. Muhaimeed, W. M. Al-Shafie and A. M. Al-Quraishi (2019). Using L-band radar data for soil salinity mapping—a case study in Central Iraq. *Environmental Research Communications* **1**(8): 081004.

Yahiaoui, I., A. Douaoui, Q. Zhang and A. Ziane (2015). Soil salinity prediction in the Lower Cheliff plain (Algeria) based on remote sensing and topographic feature analysis. *Journal of Arid Land* **7**: 794-805.
Yuen, K. W., T. T. Hanh, V. D. Quynh, A. D. Switzer, P. Teng and J. S. H. Lee (2021). Interacting effects of land-use change and natural hazards on rice agriculture in the Mekong and Red River deltas in Vietnam. *Natural Hazards and Earth System Sciences* **21**(5): 1473-1493.
Zarei, A., M. Hasanlou and M. Mahdianpari (2021). A comparison of machine learning models for soil salinity estimation using multi-spectral earth observation data. *ISPRS annals of the photogrammetry, remote sensing and spatial information sciences* **3**: 257-263.
Zhang, M., X. Fan, P. Gao, L. Guo, X. Huang, X. Gao, J. Pang and F. Tan (2025). Monitoring Soil Salinity in Arid Areas of Northern Xinjiang Using Multi-Source Satellite Data: A Trusted Deep Learning Framework. *Land* **14**(1): 110.
Zhang, Y. and A. Haghani (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies* **58**: 308-324.
Zhang, Z., Y. Fan, A. Zhang and Z. Jiao (2022). Baseline-based soil salinity index (BSSI): A novel remote sensing monitoring method of soil salinization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **16**: 202-214.
Zhao, Y., Q. Feng and H. Yang (2016). Soil salinity distribution and its relationship with soil particle size in the lower reaches of Heihe River, Northwestern China. *Environmental Earth Sciences* **75**: 1-18.
Zhou, J., T. Huang, H. Wang, W. Du, Y. Zhan, A. Duan and G. Fu (2025). Using physical method, machine learning and hybrid method to model soil water movement. *Journal of Hydrology* **652**: 132639.
Zhu, K., Z. Sun, F. Zhao, T. Yang, Z. Tian, J. Lai, W. Zhu and B. Long (2021). Relating hyperspectral vegetation indices with soil salinity at different depths for the diagnosis of winter wheat salt stress. *Remote Sensing* **13**(2): 250.