Predicting Soil Salinity Using Machine Learning and Assessing Farmers' Adaptive Capacity: A Study in the Red River Delta

~~Farmers' adaptive capacity towards soil salinity effects using hybrid machine learning in the Red River Delta~~

Huu Duy Nguyen[1], Dinh Kha Dang[2], Thi Anh Tam Lai[1], Duc Dung Tran[3], Himan Shahabi[4], Quang-Thanh Bui[1]

[1] Faculty of Geography, VNU University of Science, Vietnam National University, Ha Noi, 334 Nguyen Trai, Thanh Xuan district, Hanoi City, Vietnam; nguyenhuuduy@hus.edu.vn; laithianhtam_t65@hus.edu.vn; thanhbq@vnu.edu.vn

[2]Faculty of Hydrology, Meteorology, and Oceanography, VNU University of Science, Vietnam National University, Ha Noi, 334 Nguyen Trai, Thanh Xuan district, Hanoi, Vietnam; dangdinhkha@hus.edu.vn

[3] National Institute of Education, Nanyang Technological University, Singapore, Singapore; [4] Earth Observatory of Singapore and Asian School of the Environment, Nanyang Technological University, Singapore, Singapore; [5]Center of Water Management and Climate Change, Institute for Environment and Resources, Vietnam National University, Ho Chi Minh City, Viet Nam; dungtranducvn@yahoo.com

[4] Departments of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj City, Kurdistan Province, Iran; h.shahabi@uok.ac.ir

Corresponding: Huu Duy Nguyen (nguyenhuuduy@hus.edu.vn)

Abstract

Soil salinity is a grave environmental threat to agricultural development and food security in large parts of the world, especially in the situation of global warming and sea level rise. Reliable information on the adaptive capacity of farms plays a key role in reducing the socioeconomic effects of soil salinization and helps policymakers and farmers propose more appropriate measures to combat the phenomenon. The ~~aim of the research~~research aims ~~is~~ to design a theoretical framework to assess soil salinity and farmers' adaptive capacity~~,~~ based on machine learning, optimization algorithms (namely, Xgboost (XGB), XGB-Pelican Optimization Algorithm (POA), XGB-Siberian Tiger Optimization (STO), XGB-Serval Optimization Algorithm (SOA), XGB-Particle Swarm Optimization (PSO), and XGB-Grasshopper Optimization Algorithm (GOA)), remote sensing, and interviews with local people. We evaluated the geographical distribution of soil salinity by applying machine learning to Sentinel 1 and 2A.~~The geographical distribution of soil salinity was evaluated by applying machine learning Sentinel 1 and 2A.~~ The

Formatted: English (United States)

Formatted: English (United States)

adaptive capacity of farmers was evaluated through interviews with 87 households. The statistical indices, namely the mean absolute error (MAE), the root mean square error (RMSE), and the correlation coefficient (R^2), were used to assess the machine learning models. The outcome of this study demonstrated that all optimization algorithms were successful in improving the accuracy of the XGB model. The XGB-POA ~~was~~ had the most performance, with an R^2 value of 0.968, followed by XGB-STO (R^2=0.967), XGB-SOA (R^2=0.966), XGB-PSO (R^2 = 0.964), and XGB-GOA (R^2=0.964), respectively. The soil salinity map produced by the proposed models also indicated that the coastal and riverside regions were the most affected by soil salinity. The results also showed human and financial resources to be the two most important factors influencing the adaptive capacity of farmers. This study ~~offers~~ provides a key theoretical framework that ~~supplements the~~enhances previous previous ~~studies~~ and can ~~support~~ assist policy-markers and farmers in managing land resource ~~management~~, ~~for example~~such as accurately identifying areas affected by soil salinity for agricultural development in the context of climate change. In addition, this research highlights the importance of integrating machine learning, remote sensing, and socio-economic surveys in soil salinity management, which can support farmers for sustainable agricultural development.

Keywords: Red river, soil salinity, machine learning, adaptive capacity

1. Introduction

Soil salinity is among the greatest threats to land management, ~~poses~~ posing significant problems to agricultural progress and global food security (He et al., 2024; Jia et al., 2024; Xiao et al., 2024). According to FAO, soil salinity affects about 424 million hectares of land surface (with a depth of 0-30 cm)~~According to FAO, about 424 million hectares of land surface (with a depth of 0-30 cm)~~ and more than 833 million hectares of subsoil (30-100 cm) ~~are touched by soil salinity~~. This area is increasing by about 2 million hectares each year and influences more than 100 countries worldwide, causing damage between 12 and 27.3 billion USD (Aksoy et al., 2024; Jia et al., 2024).

The soil salinity problem will occur at the local, regional, and global levels (Bandak et al., 2024; Liu et al., 2024). In Vietnam, many littoral regions are affected by soil salinity problems. According to the 2021 Ministry of Agriculture and ~~Rural Development~~Environmental Report on the Current Situation and Planning of Agricultural Development, in 2020, about 200,000 hectares of cropland in Vietnam ~~were touched~~were already affected by soil salinity. ~~The Red River and Mekong Deltas are~~This problem is increasingly serious in Mekong Delta and Red River Delta - home to ~~more than~~over 40 million people and ~~represent an extremely key role~~playing a key role in Vietnam's agricultural and

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

aquaculture sectors~~the country's agricultural and aquaculture activities~~. ~~-~~ - where ~~They~~ they account for 71% of paddy cultivation, 86% of aquatic farming, and 65% of fruit production (General Statistics Office, 2024; Ministry of Aquaculture, Agriculture and Rural Development, 2013). Because ~~T~~these low-lying coastal areas (Hung and Larson, 2014) are experiencing subsidence (Le Dang et al., 2014), and declining river water levels,~~and river water levels are decreasing. As such~~, they ~~are very~~have become highly susceptible to the effects of climate variability and sea-level rise (Dasgupta et al., 2009). ~~Soil salinization in the littoral regions of these two deltas - partly due to the advance of the sea - is becoming a major threat to crop production while also putting pressure on Vietnam's food abundance.~~ Therefore, monitoring soil salinity is essential to inform agricultural management strategies to ensure food security at local and regional levels.

In order to address the problem, it is important to have the most precise and current data on soil salinity. Traditionally, direct field measurements of soil salinity yielded the most accurate data~~the most accurate data would be obtained by measuring soil salinity directly in the field~~ (Eldeiry et al., 2008; Rhoades and Ingvalson, 1971). This method collects point samples in the areas of interest one by one, which is time-consuming and requires significant manual work. Although this method can accurately identify soil salinity, ~~there is no way for these data to be updated over time without conducting further field missions. To obtain a continuous math function, i.e., raster data suitable for GIS analysis and remote sensing for the areas of interest, freely accessible remote sensing data, such as those obtained via Landsat and Sentinel imagery, can monitor the environment with different spectral bands, at high spatial (10 m) and temporal (3 to 5 days) resolutions~~ it requires ~~so~~ many field missions to collect the data over time, which can be time-consuming and resource-intensive. To reduce these limitations and obtain continuous spatial data (such as raster data) suitable for GIS analysis and environmental monitoring, several studies have used freely available remote sensing data, such as Landsat and Sentinel images. These data provide spatial (10m) and temporal (3 to 5 days) resolution and capture multiple spectral bands (Asfaw et al., 2018; Cullu, 2003). ~~Remote sensing data has been justified by several studies by the ability to monitor soil salinity with high accuracy and faster. This process can be achieved by analyzing the relationships between remote sensing data and in situ data, provided these data are spatially and temporally consistent~~ Several studies have demonstrated the effectiveness of remote sensing to monitor soil salinity accurately and rapidly. By constructing correlations between remote sensing-derived indices and soil salinity points in the field, such as NDVI, VSSI, and NDSI, we can achieve this. Although remote sensing can monitor soil salinity using different spectral responses, slightly or moderately saline soils cannot be distinguished

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

easily because soil minerals and their components modify the spectral capacity of the soil surface. To emphasize the qualities of the land surface, for example, water, vegetation, or saline soil, several studies have highlighted indices such as NDVI, VSSI, and NDSI to monitor soil salinity. Using an index with different wavebands increases the number of variables in the soil salinity modelling process. Although remote sensing can monitor soil salinity using different spectral responses, slightly or moderately saline soils cannot be distinguished easily because soil minerals and their components modify the spectral capacity of the soil surface.

In recent yearsRecently, with improvements in computing power, machine learning, and deep learning, there has been have greatly provided thesubstantial growth of in techniques to construct soil salinity maps with higher accuracy. Algorithms such as random forest (Fathizad et al., 2020), XGBoost (Jia et al., 2024), support vector machines (Jiang et al., 2019), CatBoost (Gong et al., 2023; Wang et al., 2022), and AdaBoost (Wang et al., 2022) are the most popular algorithms to construct soil salinity maps by integrating satellite images and in situ measurements. Some research has used deep learning models to construct soil salinity maps, such as deep neural networks, recurrent neural networks, and Deep Boltzmann machines. (Kaplan et al., (2023) used four machine learning algorithms, namely M5P, RF, Linear, and IBK, integrated with Sentinel 2A data and 393 soil samples collected in situ to construct a soil salinity map for the United Arab Emirates. The study's outcome results indicatedshowed that all models' accuracy wasperformed well in assessing soil salinity, with the IBK model proving demonstrating the most effectivehighest effectiveness. (Aksoy et al., (2024) used XGBoost and random forest with 26 environmental covariates from Landsat 8 OLI to evaluate soil salinity in Iran's Lake Urmia. The study's outcome showed that machine learning integrated with Landsat 8 OLI data successfully monitored soil salinity, with XGB yielding more accurate results than random forest. (Jia et al., (2024) applied nine models, namely PLSR, Lasso, CART, RF, ERT, GBDT, LightGBM, XGBoost, and AdaBoost, integrated with Sentinel 2A imagery, to evaluate soil salinity in the Ningxia Yellow River Diversion Irrigation Area. The results showed that the AdaBoost model performed better than the others.

Field Code Changed

Field Code Changed

Field Code Changed

Previous studies show that although machine learning methods have been utilized to assess soil salinity in many regions of the world, their application for this purpose is still limited in the Mekong and Red River Deltas (Shi et al., 2021; Vermeulen and Van Niekerk, 2017). Currently, there are only four studies that have assessed soil salinity in the Mekong Delta (Hoa et al., 2019; Nguyen et al., 2023; Nguyen et al., 2018; Nguyen et al., 2021), and no work has

been done in this field for the Red River Delta. In addition, most previous studies have developed state-of-the-art methods, such as integrating machine learning and remote sensing, to identify the geographical distribution of soil salinity in different regions of the world (Hardie and Doyle, 2012; Su et al., 2020; Wang et al., 2007). While several studies have highlighted the importance of assessing the adaptation capacity of the community to strengthen their resilience to soil salinity and other types of natural hazards, however, very few studies integrate this aspect into the identification of spatial soil salinity. ~~Few studies have integrated the identification of the geographical distribution of soil salinity with the adaptive capacity of farms. Several studies have, however, highlighted the importance of measuring such adaptive capacity to improve the resilience of farms against soil salinization in particular and natural hazards in general. (~~Hoang et al.~~, (~~2023) reported that assessing the ability of farms to adapt to soil salinization is the key to reducing vulnerability and contributes significantly to the development of sustainable livelihoods.

The adaptive capacity is defined as the capability of the community to cope, adjust, and adapt to the impacts of growing soil salinity. It measures the ability to predict, respond, and recover from the phenomenon. It is assessed on different scales, using different approaches, according to the region in question (Mazumder and Kabir, 2022; Thiam et al., 2024). ~~The IPCC in 2014 indicated that farm adaptive capacity depends on five main factors: natural capital, human capital, material resources, financial resources, and social capital.~~ Furthermore, understanding the adaptive capacity of communities plays an important role in reducing the negative effects of salinity intrusion in coastal regions in general and the Red River Delta in particular. By assessing adaptation at multiple scales with site-specific methods, researchers and local governments can identify interventions (such as crop variety changes, crop calendars, irrigation systems) that are effective. The IPCC in 2014 indicated that farm adaptive capacity depends on five main factors: natural capital, human capital, material resources, financial resources, and social capital. Therefore, integrating the adaptive capacity of populations with the soil salinity map improves the accuracy of predictions and proposes adaptation strategies that strengthen the overall resilience of communities.

The research aims to improve a theoretical framework to assess soil salinity and farmers' adaptive capacity based on machine learning, optimization algorithms (namely XGB, XGB- POA, XGB- STO, XGB- SOA, XGB- PSO, and XGB- GOA), remote sensing, and interviews with local people. ~~This study used machine learning to construct the soil salinity map. From the literature reviews, s~~Several studies have ~~been conducted in different regions of the world, focusing on the adaptive ability of farmers~~examined farmer's adaptive capacity to environmental stressors in

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

different regions (Bhuyan et al., 2024; Thiam et al., 2024). However, no studies comprehensively analyze farmers' adaptive ability to combat soil salinity in a given region based on machine learning, remote sensing, and interviews with local people. In addition, several studies combine machine learning with Sentinel 1 or Sentinel 2 to assess soil salinity (Wang et al., 2021; Xiao et al., 2023); however, there are rarely studies that combine machine learning with Sentinel 1 and Sentinel 2 to monitor soil salinity in the Red River Delta. ~~Several studies have pointed out that~~ By the combination of Sentinel 1 and Sentinel 2, advanced machine learning, and the information from farmers themselves, this study filled a critical gap and provided a novel, comprehensive framework for monitoring and responding to soil salinity in the Red River Delta~~the combination of Sentinel 1 and Sentinel 2 can reduce the limitations of each image, which can improve the accuracy of soil salinity monitoring in the context of climate change~~ (Ma et al., 2021).

In general, salinity intrusion harms agricultural development and people's livelihoods. Therefore, it is necessary to develop a theoretical framework to address the soil salinity problem in terms of natural and social factors. However, previous studies have mainly assessed the spatial distribution of salinity or the community's adaptive capacity, and hardly any studies have assessed both the spatial distribution of salinity and the community's adaptive capacity. Thus, the global contribution of this study is to fill the knowledge gap about the spatial distribution of soil salinity and the adaptive capacity of communities in the Red River Delta in general and Thai Binh Province in particular by relying on modern methods to improve this important and understudied understanding. The results of this study can play an important role in mitigating the impact of salinity intrusion on agricultural development and can help policymakers and planners develop effective strategies to mitigate this impact, especially in the context of climate change.

Formatted: Vietnamese

~~This map gives us a panoramic picture of the saline intrusion situation in the Thai Thuy district in particular and the Red River Delta region in general. This map will be the basis for identifying areas affected by saline intrusion, thereby assessing people's adaptability to this situation. The results of this study can inform farmers of developing strategies to reduce the impacts of soil salinity on agriculture and ensure food security in the region. To our knowledge, this study is considered the first study to assess the soil salinity and farmer's adaptive capacity of the population based on machine learning, remote sensing and interviews with the local people. The finding of this paper provides important information for policymakers or local authorities based on evidence and, ultimately, supports researchers or decision makers in developing effective strategies for smallholder farmers.~~

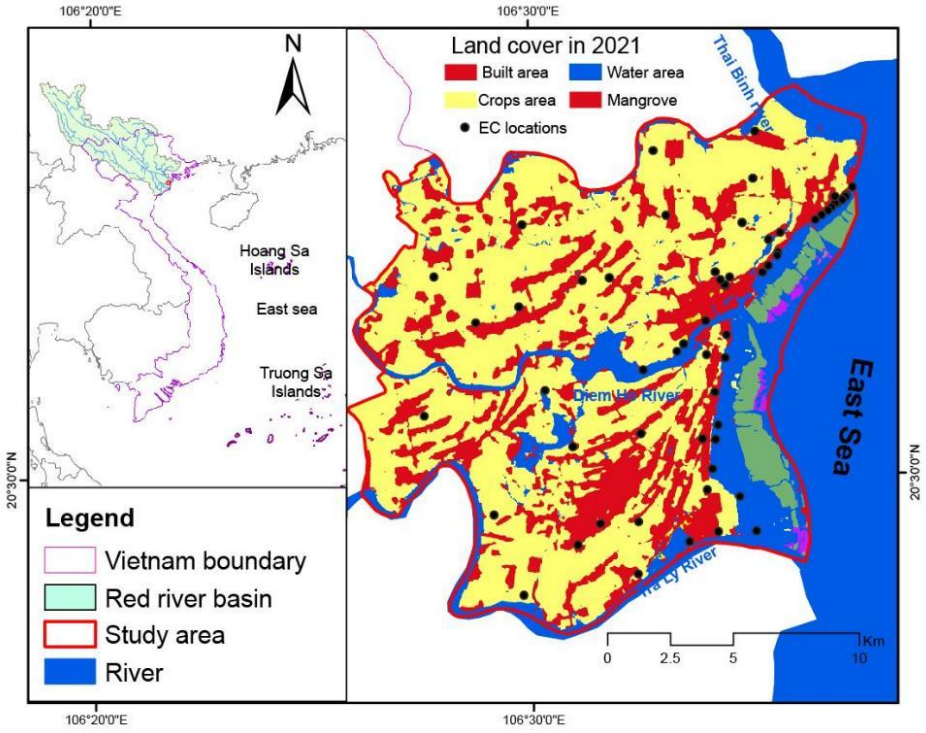
2. Study Area

The Red River Basin covers a total area of 169,000 km² and spans China (48%), Laos (0.7%), and Vietnam (51.3%). The river system has a total length of 1,150 km, with around 500 km in the territory of Vietnam before discharging into the Gulf of Tonkin. The topography is mainly mountainous terrain that comprises about 70% of the total area at elevations above 500 meters. In the lowlands, elevations range from approximately 0.4 to 9 m, ~~with the~~ characterized by a tropical climate with summer monsoons from the south and winter monsoons from the northeast (Vinh et al., 2014), the basin experiences average annual precipitation ranging from 800 to 3000 mm. The rainy season occurs from May to October and ~~registers~~ accounts for 70%-90% of annual rainfall (Quang et al., 2024). Daily rainfall varies from 300-400 mm during this period. The average temperature ranges from 22 to 27 °C, with winter temperatures potentially below 10 °C and summer temperatures above up to 40 °C.

The basin flows into the Gulf of Tonkin through nine river mouths, of which the Tra Ly, Van Uc, and Ba Lat are the main channels for water conveyance. These channels transport a substantial sediment load of approximately 120×10^6 tons annually to the Red River Delta region (Vinh et al., 2014). The littoral region has a semi-diurnal tidal regime, with tidal ranges ranging from 2 to 4 m. Saline intrusion significantly influences the littoral region during the dry season with average and maximum wave heights of about 0.7-1.3 m and 3.5-4.5 m, respectively. However, during major storms, wave heights can ~~obtained at~~reach 5 m (Nhuan et al., 2007).

The Red River Delta is influenced by several natural hazards, such as flooding, soil salinity, and sea level rise (Castelletti et al., 2012). Several studies have highlighted that rising sea levels are having an increasingly severe impact on inland regions, leading to soil salinity (Nguyễn Văn Đào, 2023). ~~In recent years~~Recently, the Red River Delta in general and the Thai Thuy district in particular have been affected by soil salinity, causing significant damage to agricultural development and negatively impacting residents' livelihoods (Figure 1). According to the FAO report, if the sea level rises by 50 cm, about 11.8% of coastal land is at risk of being flooded by salt water; this figure increases to about 31.4% if the sea level rises by 100 cm in Thai Binh province. Recently, saltwater intrusion has clearly affected agricultural production in Thai Binh province and Thai Thuy district. Typically, in the 2015-2016 winter-spring crop, the salinity in the main rivers exceeded the threshold of 1‰, causing significant damage to crops. In 2020, the salinity in the main rivers, 28 km from the sea, exceeded the threshold of 3.75‰, far exceeding the allowable threshold of 2.75‰. This phenomenon greatly affects agricultural production in the area.

Formatted: Vietnamese



193

Formatted: Centered

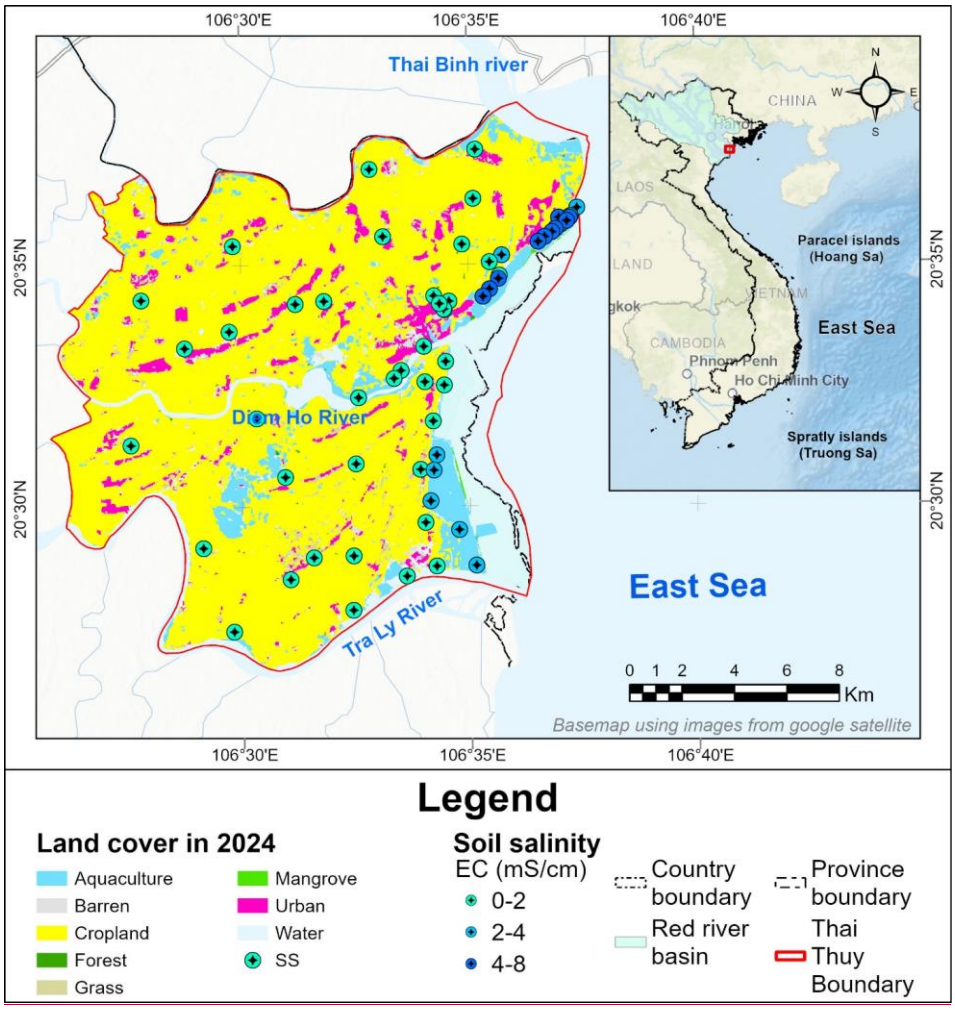


Figure 1: Geographical location of study area: The red boundary on the map represents the Thai Thuy district, located in the Thai Binh province in the Red River Delta of Vietnam. The green points are the soil salinity samples collected in April 2024. The land use in the Thai Thuy district is divided into eight types: aquaculture, barren land, cropland, forest, grass, mangrove, urban, and water body. While the aquaculture area is located in the coastal zone, cropland takes up a large part of the study area, and Urban is located in the center and along the road.

3. Methodology

The first strand of the methodology was the identification of the soil salinity mapping. We divided this process into four main steps ~~This process was divided into four main steps~~ (Figure 2):

Preparation of soil salinity samples and factors

The data for constructing the soil salinity map were divided into two main types: EC and conditioning factors.

EC Measurements

We collected soil salinity samples using soil drills, employing both zigzag and grid techniques. Small-scale sampling frequently uses this technique ~~Soil salinity samples were collected using soil drills using both zigzag and grid techniques. This technique is utilized frequently in small-scale sampling~~ (Elshewy et al., 2024; Jia et al., 2024). The sampling depth depends on the soil salinity assessment for each specific crop. This study monitors soil salinity with the objective of agricultural development; therefore, soil samples were obtained from a ~~profoundness~~ depth of 0 to 30 cm. The sampling process occurred in the dry season, between March and April 2024. In addition, when sampling in the field, it is necessary to consider the homogeneity of the soil. We collected 62 samples to cover the entire field~~62 samples were collected to ensure coverage of the entire field~~. We collected samples along the road to identify different types of soil, and the farmers labeled the samples accordingly.~~Samples were collected along the road to get the different types of soil, and the farmers indicated samples.~~ The samples were locked in bags until analysis in the laboratory. We noted the positions of the samples, including longitude and latitude, during the sampling process. When the samples were sent to labs, they were stored in enameled jars, and impurities like stones, wood, and branches were removed~~The information on the positions of the samples, such as longitude and latitude, was noted in the sampling process. It should be noted that when the samples were transferred to laboratories, they were stored in enameled jars, and the impurities present in the samples, such as stones, wood, and branches, were removed~~. These soil samples were then finely ground. The electrical conductivity (EC) was then calculated from a 1:5 soil/deionised water suspension. A soil/water suspension was created by adding 7 g of soil to 35 ml of distilled water and then mixed with a mechanical stirrer for 60 minutes to dissolve the salt. The EC value was measured using a conductivity probe. Finally, the samples were split into two parts: the first part (70%) was used to build the proposed models, while the other part (30%) was applied to confirm the model.

Remote Sensing Data

Due to the effects of the earth's cycle, the salt accumulated in the soil is closely linked to climatic conditions, hydrology, soil characteristics, and surface vegetation density, for example, topographic characteristics (Wang et al., 2024; Xie et al.). We calculated these factors using optical Sentinel 2A images and microwave Sentinel 1 images to determine the soil salinity value.~~These factors were calculated using the optical Sentinel 2A images and microwave Sentinel 1 images to identify the soil salinity value.~~ The Sentinel-2A images were calculated by running Sen2Cor for atmospheric correction to ensure the transition between apparent atmospheric reflectance and surface reflectance, and these images were obtained using Google Earth Engine. To reduce the influence of clouds, Sentinel 2A images for March-April 2024 were selected with a cloud cover rate of less than 10%. To enhance the quality of these images, the median value of each pixel was calculated at a resolution of 10 m. As for the Sentinel 1 images, they were acquired in dual cross-polarization mode and broadband interferometric mode. The median value of the Sentinel 1 image obtained on March 31, 2024, was computed to acquire microwave remote sensing data at a scale of 10 m. As well as 12 bands of the Sentinel 2A image (from band 1 to band 12) and 3 indices of the Sentinel 1 image (VV, VH, and VVVH). In addition, 20 spectral indices extracted from Sentinel 2A image~~extractés à partir de l'image Sentinel 2A~~ were selected to build the soil salinity model, namely Brightness index (BI), Canopy Response Salinity Index (CRSI), Enhanced Vegetation Index (EVI), Intensity index 1 (Int1), Indensity index 2 (Int2), Normalized Difference Salinity Index (NDSI), Normalized Difference Vegetation Index (NDVI), Ratio Vegetation Index (RVI), Salinity index (S1), Salinity index (S2), Salinity index (S3), Salinity index (S5), Salinity index (S6), Soil Adjusted Vegetation Index (SAVI), Salinity Index 1 (SI), Salinity index 2 (SI1), Salinity index 3 (SI2), Salinity index 4 (SI3) and Salinity index 5 (SI4). These factors were divided into three main groups: vegetation indices (NDVI, CRSI, RVI, SAVI, and EVI), water indices (flow direction and distance to river), salinity indices (SI, SI1, SI2, SI3, SI4, S1, S2, S3, S5, S6, and NDSI), topography indices (elevation and slope), and chlorophyll spectral indices (BI, Int1, Int2). These indices have been used frequently in previous studies (Hoa et al., 2019; Nguyen et al., 2021; Wang et al., 2021).

The indices due to the vegetation reflect the health and growth of vegetation, and the present study~~This~~ indirectly reflects the level of soil salinity in any region. The increase in soil salinity has a negative effect on the development of vegetation due to the difference in the absorption of water and nutrients; therefore, it leads to a decrease in the

values of NDVI, RVI, SAVI, and EVI and an increase in the value of CRSI (Jia et al., 2024; Wang et al., 2024). Water indices play an important role because regions near rivers or along the flow path are more affected by the salinity risk. River flow is often affected by tides or seawater intrusion; therefore, when the distance to the river decreases, the salinity risk increases due to the infiltration of salty river water into the soil. Furthermore, the flow direction influences the propagation and infiltration of water in the soil. Salty water can penetrate further inland if the flow is from the sea to the river, especially in the dry season (Nguyen et al., 2021).

Topography indices are key in constructing soil salinity models, as salty water penetrates low-lying regions more easily. In the Red River Delta, the low-lying regions are located along the coastline, and as such, these regions are more affected by the risk of soil salinity. Salinity indices highlight the value of spectral reflectance in regions affected by saltwater intrusion. The higher the salinity, the higher the spectral reflectance value (Du et al., 2024).

Construction of hybrid machine learning models

We built six machine-learning models to identify the spatial distribution of soil salinity~~Six machine-learning models were built to identify the spatial distribution of soil salinity~~. This involved two main steps: constructing an individual XGB model and then creating hybrid models by integrating each of five optimization algorithms with the XGB model. We developed the XGB model using Python and the Sklearn library~~The XGB model was developed using Python and the Sklearn library~~. ~~Its accuracy depended on parameter adjustments made using the trial-and-error method.~~

The machine learning model-building process was divided into two main steps: the first was the XGB model building, and the second was the hybrid model building (the integration of XGB with optimization algorithms). The accuracy of the machine learning model depends on the parameter adjustments of the XGB model. In this study, the XGB model parameters were selected using the trial-and-error method. Finally, the XGB parameters were n_estimators=100, max_depth=4, subsample=0.5, and colsample_bytree=0.5. While the hybrid model was built by integrating the XGB model and optimization algorithms, namely GOA, POA, SOA, STO, and PSO. To integrate the XGB model with optimization algorithms, we first need to construct an objective function $F(\theta)$ that returns the error value of XGB on the validation set when using the parameter sets θ. That is, each parameter set has a different error value. Next, determine the search space of the hyperparameters (n_estimators, max_depth, subsample, colsample_bytree) as discrete value intervals. Then, the optimization algorithms will initialize the population of

individuals with the size and parameters characteristic of each optimization algorithm. This study was tested with 500 iterations: at each iteration, each individual will generate a combination of θi, and the optimization algorithms will update the position or velocity of the individuals according to their own rules. This process is repeated until a stopping threshold is set. Finally, the results are the optimal parameters. The parameters of the model are as follows: problem_size = 3, batch_size = 25, epoch = 500, pop_size = 50, "fit_func": fun_avr2, "lb": [0] *problem_size*, *"ub": [1]* problem_size, c_min = 0.00004, c_max = 2.0 for **XGB-GOA**; problem_size = 3, batch_size = 25, epoch = 500, pop_size = 50, "fit_func": fun_avr2, "lb": [0] *problem_size*, *"ub": [1]* problem_size, c1=2.05, c2=2.05, w_min=0.4 for **XGB-PSO** ; problem_size = 3, batch_size = 25, epoch = 500, pop_size = 50, "fit_func": fun_avr2, "lb": [0] *problem_size*, *"ub": [1]* problem_size for **XGB-POA**; problem_size = 3, batch_size = 25, epoch = 500, pop_size = 50, "fit_func": fun_avr2, "lb": [0] *problem_size*, *"ub": [1]* problem_size for **XGB-SOA**; problem_size = 3, batch_size = 25, epoch = 500, pop_size = 50, "fit_func": fun_avr2, "lb": [0] *problem_size*, *"ub": [1]* problem_size for **XGB-STO**.

Formatted: Vietnamese

Evaluation of model accuracy

The statistical indices, the root mean square error (RMSE), the mean absolute error (MAE), and the correlation coefficient (r) were used to assess the accuracy of the proposed models.

Analysis of spatial distribution and identifying the farmers adaptive capacity~~interview area identification~~

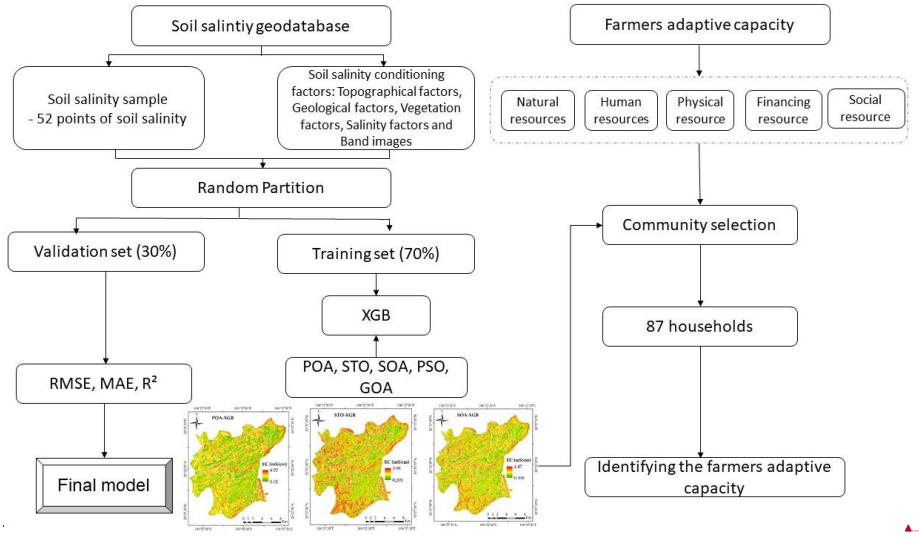
Formatted: English (United States)

We validated the models and then used them to assess soil salinity in the study area at a pixel scale with a resolution of 10x10 m~~After validating the models, they were utilized to assess soil salinity in the study area at a pixel scale with a resolution of 10x10 m~~. Approximately 30 million pixels were assessed, and a soil salinity map was constructed using the Point to Raster tool in ArcGIS 10.8.

We used the second strand of our methodology, interviews with populations, to evaluate the adaptive capacity of farms in the study area~~The second strand of our methodology, interviews with populations, was used to assess the adaptive capacity of farms in the study area~~. ~~A Tân~~We selected An Tan commune in the Thai Thuy district to participate in structured interviews~~was selected for this assessment~~. A total of 87 households were interviewed. These households were randomly selected from An Tan commune in the Thai Thuy district to participate in structured interviews. The commune is located in the coastal region, which has the lowest altitude and ~~is,~~ often

affected by soil salinity. Residents mainly worked in rice and fish farming. We analyzed all 87 responses to assess farmers' ability to adapt to soil salinity~~Residents mainly worked in rice and fish farming. All 87 responses were used to analyze farmers' adaptive capacity to soil salinity~~.

The structured interviews focused on five main elements: natural resources, human resources, physical resources, financial resources, and social resources. There was a particular focus on soil salinity in 2023, allows to evaluate the stress of environment on the livelihood of population~~2023.~~



Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

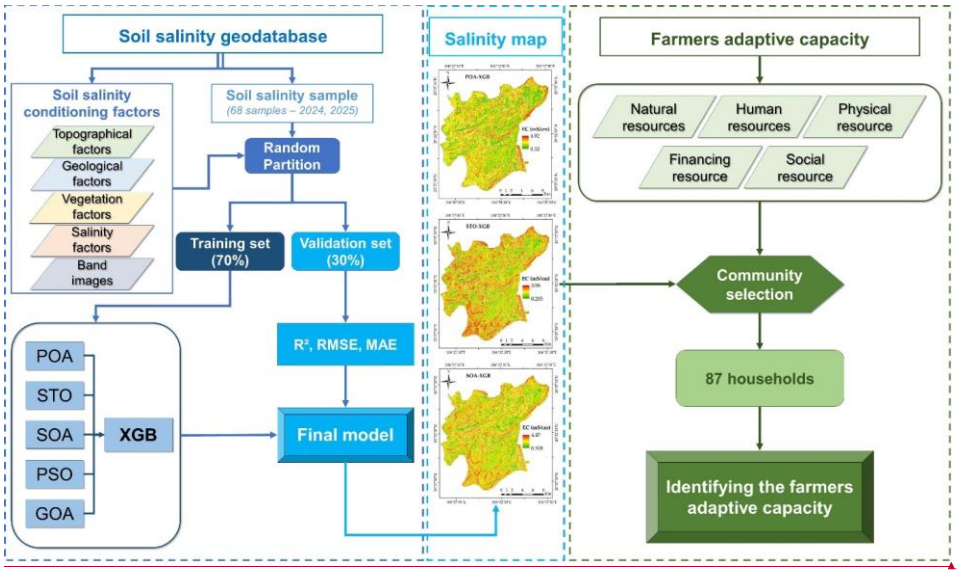


Figure 2: Methodology used for the farmer's adaptive capacity and soil salinity in this study: The methodology used in this study is to design a theoretical framework to assess soil salinity and farmers' adaptive capacity based on machine learning, optimization algorithms, remote sensing, and interviews with local people. We divided this process into four main steps: i) data preparation; ii) machine learning model construction; iii) machine learning model evaluation; and iv) analysis of spatial distribution and identifying the farmers' adaptive capacity.

3.1. XGBoost (XGB)

XGB is a popular gradient-boosted tree algorithm that can solve classification and regression problems. The main idea of learning with XGBoost is to train several models sequentially and combine them successively by correcting errors iteration after iteration to obtain the most potent ensemble model possible (Zhang et al., 2022). Therefore, th~~The~~ prediction result ~~of the prediction is, therefore,~~consists of a set of chained decision trees ~~made up of the prediction of the set of chained decision trees~~. This method increases the performance and stability of the model while minimizing its variance (Zhang et al., 2022). The XGB model functions in three main steps: i) it constructs an individual model (tree) by taking predictions on the training data, ii) it computes the mistakes of these predictions for the real values, and iii) it constructs another tree to predict and correct these mistakes. The process is repeated, and each new tree is built to correct the mistakes of the previous one. This phase is called "boosting". We then sum

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

up the predictions of all trees to determine the final predictions~~The predictions of all trees are then summed to take the final predictions~~ (Mukhamediev et al., 2023).

To optimize the accuracy of the XGB model, three main parameters need to be adjusted: learning rate (reducing the value of this parameter can avoid the overfitting problem), alpha, and lambda (increasing the value of these parameters makes the model more conservative), and column sample by tree (adjusting this parameter has the objective of obtaining the subsample of columns~~;~~) (Tan et al., 2023).

3.2. Pelican Optimization Algorithm (POA)

Agents searching for prey in nature have a mechanism similar to that of agents searching for optimal solutions. Therefore, based on this perspective, the search agents that comprise a population seek to achieve the optimal solution more quickly. Each agent is an optimal solution whose position is determined in the search space. From a mathematical perspective, agents are vectors, and the population of agents forms matrices (Dehghani and Trojovský, 2022). Among the values used to calculate the aim function, the top value of the agents is determined as the top solution of the agents.

One such optimization algorithm is POA proposed by ~~(~~Trojovský and Dehghani~~,~~ (2022). This algorithm is designed based on Pelicans' foraging and communication processes. This is a large bird with a long beak and a large pouch in the throat to hold prey during hunting. Hundreds of individuals may flock together. They can weigh up to 15 kg, grow to a height of 1.8 m, and have a wingspan of up to 3 m. These characteristics greatly assist them in finding food, such as fish, frogs, and turtles.

POA is based on the simulation of the ~~comportment~~ behavior and plan of pelicans when attacking prey. Pelican hunting strategies are divided into two stages. First, the bird moves towards its prey, then it spreads its wings and glides along the water surface to attack. In the first stage, the pelicans determine the situation of the prey and move toward the identified prey area. Identifying prey areas represents the determination of the search space in the POA model. The positions of the prey are randomly produced in the search space, which increases the exploration power of POA in the process of searching and solving optimization problems. After locating the prey area in the second stage, pelicans spread their wings and move on the water surface to attack the prey and store it in their throat pouch.

Formatted: English (United States)

This strategy allows them to capture more prey. Modelling this behavior of pelicans makes the POA model easier to converge and improves local search ability (Alamir et al., 2023; Trojovský and Dehghani, 2022).

3.3. Serval Optimization Algorithm (SOA)

SOA is one of the population-based optimization algorithms, and ~~they~~ it exploit the searching power of agents in a population. This property makes this algorithm powerful in solving optimization problems (Dehghani and Trojovský, 2022). Each iteration of SOA continuously determines and updates the agents' situations~~In SOA, the agents' situation are continuously determined and up-to-date in each iteration~~. The updating process simulates the behavior of serval cats in the wild, divided into two stages: i) exploration of the search space and ii) local exploitation in the search space (Sindi et al., 2024).

Wild cats are some of the most efficient predators, using hearing to identify and attack prey. In the first stage of SOA, the situation of the servals is repeatedly up-to-date after each move: the continuous updating of positions ~~guides~~ leads to detailed coverage of the search space. The ~~aim of~~ this stage aims is to raise the power to search and explore in the search space. The situation of the best member in the SOA is considered the situation of the prey and, therefore, the optimal solution (Dehghani and Trojovský, 2022; Sindi et al., 2024).

Formatted: English (United States)

When attacking prey, wild cats jump during the chase to prevent the prey from escaping. These strategies also serve~~In SOA, these strategies are also used~~ to update the position of the SOA population. The simulation of the chase process can cause small changes in agents' positions in the SOA. However, this phase aims to increase the search space mining capability of SOA, which helps to improve the local search capability in the search space (Dehghani and Trojovský, 2022).

3.4. Siberian Tiger Optimization (STO)

The STO algorithm is a new biologically inspired hyper-heuristic algorithm modeled after Siberian tigers' hunting behavior (Trojovský et al., 2022). STO replicates the Siberian tiger tracking and capture strategy, using a population-based approach to explore the search space efficiently and quickly (Trojovský et al., 2022).

STO works by simulating the way Siberian tigers move and communicate with each other while hunting their prey. Each agent in the STO algorithm represents a Siberian tiger, ~~each~~ exploring a different region in the search space. The tigers communicate and share information about their locations with each other to find the optimal location. The

location update process of Siberian tigers in STO is carried out in two main phases: the hunting and bear-fighting phases (Trojovský et al., 2022).

In the hunting phase, since the Siberian tiger is a powerful predator, it hunts by attacking ~~assaulting~~ different prey, so the agents in the STO are up-to-date based on the simulation of this hunting strategy. After choosing the prey, the Siberian tiger will chase, attack, and kill its prey in this phase. The population continuously updates the members' positions based on the selection and attack of the prey~~The members' positions of the population are continuously updated on the choice and attack of the prey~~. This phase causes sudden changes in the members' positions and increases the search ability in the search space (Al-Sarray et al., 2024).

During hunting, the Siberian tiger has to fight with brown and black bears. Therefore, in the second phase, the members of the STO ~~are up-to-date~~ stay update on the ~~policy~~ strategies used by the~~of the~~ Siberian tiger when ~~fighting with bears~~bear-fighting. When fighting with bears, the Siberian tiger ambushes and then assaults the bear until it kills it (Al-Sarray et al., 2024; Trojovský et al., 2022).

One of the key features of STO is its ability to balance exploration and exploitation. In ~~other words~~put it another way, the design of the STO algorithm ~~is designed to explore~~involves extensive exploration of the search space ~~extensively~~ and ~~refine~~ refinement of promising solutions in the most promising areas. Thus, STO avoids local optimization problems and increases the likelihood of global optimization (Trojovský et al., 2022).

3.5. Particle Swarm Optimization (PSO)

PSO was proposed by ~~(~~Kennedy and Eberhart~~,~~ (1995). It is founded on the principles of self-organization that allow one or more groups of living organisms to move together in a complex way (Fu et al., 2018). PSO simulates the movement process of some animals, such as flocks of birds. In this model, birds move randomly by following three rules: i) they track the same path as their friends, ii) they are enticed to the average situation of the group, and iii) they maintain a certain space between each other to avoid collisions (Ruidas et al., 2022).

PSO explores the search space through the birds' position and flight paths. The position of each bird in the search space is considered a potential solution. More precisely, the position and speed of the birds are represented by vectors with D dimensions, and the initial speed is determined randomly. In the PSO algorithm, the position and velocity of each bird are updated continuously after each iteration until an optimal solution is reached. The

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

Formatted: English (United States)

optimization function assesses the position and velocity quality ~~is used to evaluate the quality of the position and velocity~~(Bui et al., 2016). In this study, PSO was used to optimize the XGB model.

3.6. Grasshopper Optimization Algorithm (GOA)

GOA was first proposed by ~~(~~Mirjalili et al. (~~,~~2018). This algorithm is based on the swarm behavior of locusts during foraging to solve optimization problems. Grasshoppers move quickly to explore spaces, and then they move locally to exploit resources in the foraging space. GOA models the behavior of a virtual swarm of grasshoppers, where each position represents an optimization solution to the problem (Moayedi et al., 2021; Nguyen, 2022). Movement is influenced by several factors: social interaction, gravity, and wind advection. Social interaction plays an important role in finding the optimal position because grasshoppers interact with each other to exchange information about precise positions. This social communication allows grasshoppers to find the right solutions. Then, gravity allows grasshoppers to explore the foraging spaces in a balanced manner, hence avoiding the local optimization problem (Ingle and Jatoth, 2024). Finally, wind advection represents the external effects that can influence the movement of grasshoppers, leading them to some areas of the search space. In the optimization process, an equilibrium between exploration and exploimportants is important to accurately approach the true global optimum (Moayedi et al., 2020).

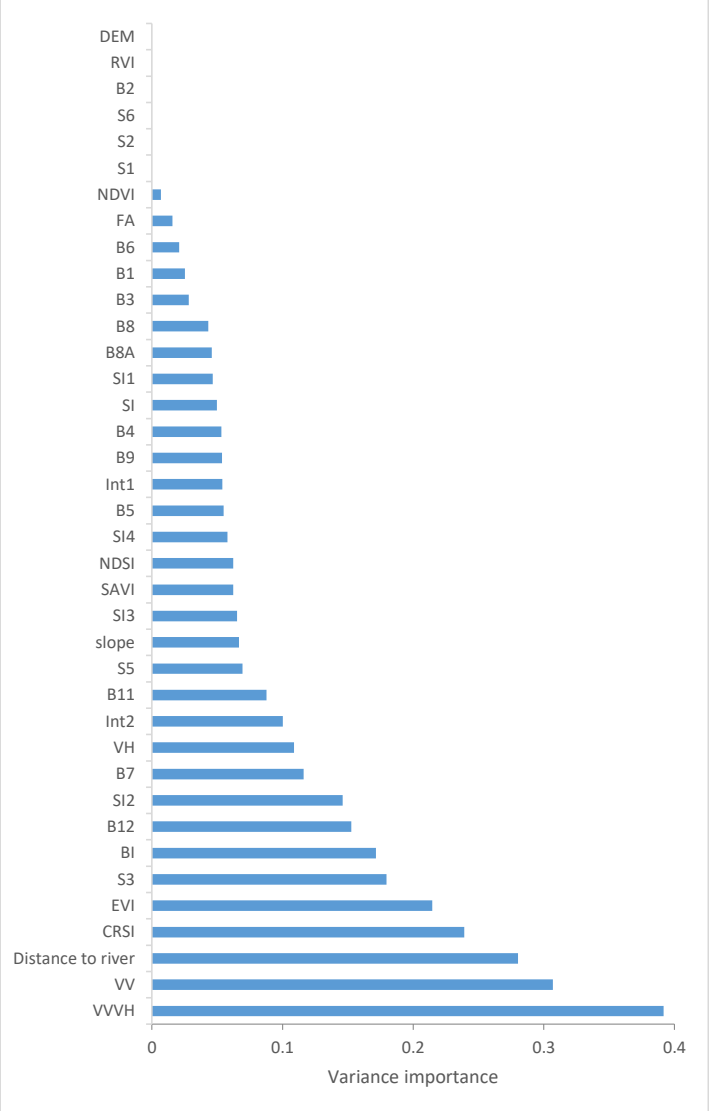
4. Results

4.1. Soil Salinity Predictors

The choice of suitable factors plays a key role when using machine learning to determine the geographical distribution of soil salinity in any region. Conditioning factors represent the causes of soil salinity, so improper selection of these factors can result in inaccurate prediction. Data redundancy may complicate the model and lead to poor performance~~Data redundancy may make the model more complex and lead to poor performance~~.

In this study, we used RF to measure the appropriate factors~~RF was utilized to measure the suitable factors in this study~~. It assigns a value to each factor based on ~~the relation~~its relationship with the soil salinity ~~within soil salinity~~ samples and the conditioning factors. The most important factor is the one with the greatest importance in determining soil salinity zones. In addition, after using RF to determine the importance of factors, we used trial and error to continue eliminating factors that affected the precision of the model.

The outcome showed that six factors (DEM, RVI, B2, S6, S2, and S1) had an RF value of zero, so these factors did not affect the determination of the spatial distribution of saline areas. In addition, two factors (NDVI and flow accumulation) were eliminated using the trial-and-error method. The other 30 factors were used to build the model. VVVH (0.39), VV (0.3), distance from the river (0.28), CRSI (0.24), and EVI (0.21) had a strong influence on the soil salinity in the study area. S3 (0.17), BI (0.17), B12 (0.15), SI2 (0.14), B7 (0.11), VH (0.1), and Int2 (0.1) have moderate relationships with the soil salinity. B11 (0.08), S5 (0.06), slope (0.06), SI (0.06), SAVI (0.06), NDSI (0.06), SI4 0.05), B5 (0.05), Int1 (0.05), B9 (0.05), B4 (0.05), SI (0.04), SI1 (0.04), B8 (0.04), B3 (0.02), B1 (0.02), and B6 (0.02) had only a weak relationship on soil salinity (Figure 3).



441

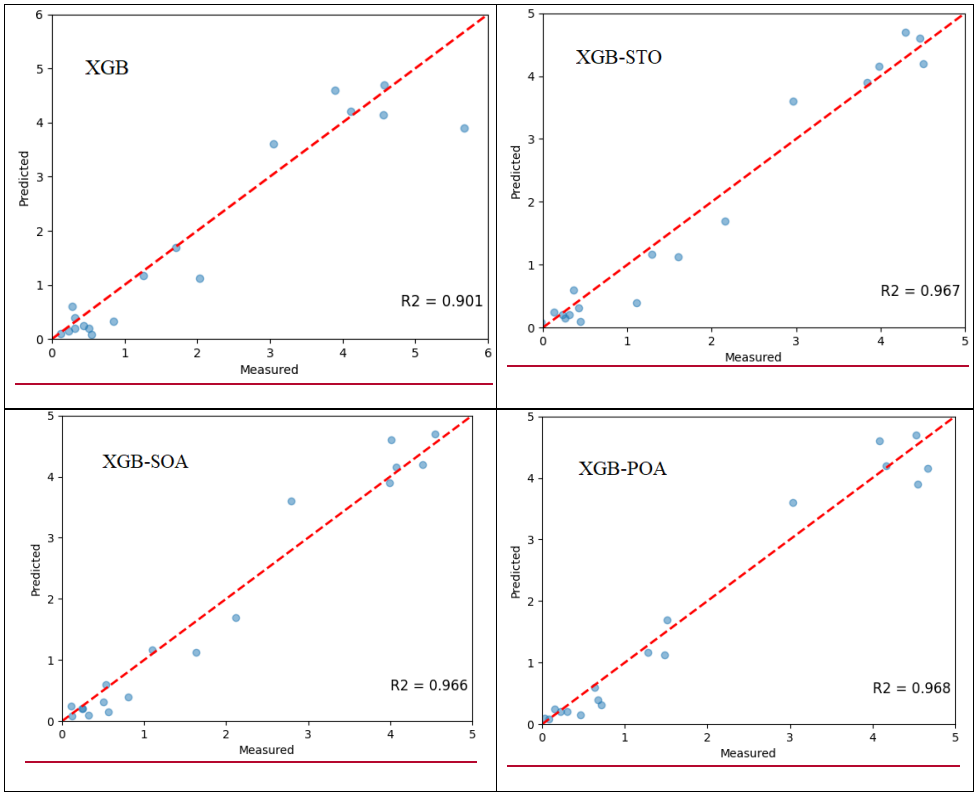
442 **Figure 3:** Variables important for soil salinity model using RF.

443

444

4.2. Model Accuracy Validation

R^2 was used to assess the performance of the machine learning models. The outcome of this study demonstrated that all optimization algorithms enhanced the performance of the XGB model. The XGB-POA model was the greatest, with an R^2 value of 0.968, followed by XGB-STO (R^2=0.967), XGB-SOA (R^2=0.966), XGB-PSO (R2 = 0.964), and XGB-GOA (R^2=0.964; Figure 4).



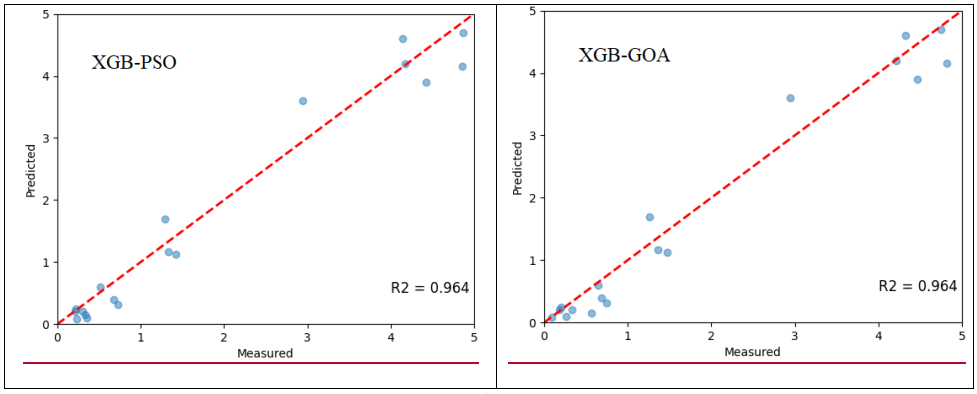


Figure 4: R^2 value for the testing dataset

The RMSE and MAE were also used to evaluate the accuracy of the machine learning models. The XGB-POA model performed better on training and validation data (RMSE=0.28, MAE=0.18 for learning data and RMSE=0.31 and MAE=0.242 for verification data). The XGB-STO was ranked second with an RMSE value of 0.3 and MAE of 0.22 for learning data, and an RMSE value of 0.32 and MAE of 0.244 for verification data. The XGB-SOA model was ranked third, with RMSE=0.31 and MAE=0.23 for learning data and RMSE=0.33 and MAE=0.25 for verification data. XGB-GOA model came fourth, with RMSE=0.33 and MAE=0.25 for learning data and RMSE=0.34 and MAE=0.26 for verification data. The XGB-PSO model performed less well than the other models, with RMSE=0.335 and MAE=0.26 for learning data and RMSE=0.341 and MAE=0.27 for verification data (Table 1).

Table 1. Model performance and comparison.

Models	**Training dataset**			**Validation dataset**		
	RMSE	**MAE**	**R^2**	**RMSE**	**MAE**	**R^2**
XGB-POA	0.28	0.18	0.99	0.31	0.242	0.968
XGB-STO	0.3	0.22	0.987	0.32	0.244	0.967
XGB-SOA	0.31	0.23	0.98	0.33	0.25	0.~~965~~966
XGB-GOA	0.33	0.25	0.97	0.34	0.26	0.964
XGB-PSO	0.335	0.26	0.97	0.341	0.27	0.964

XGB	0.35	0.32	0.91	0.37	0.34	0.901

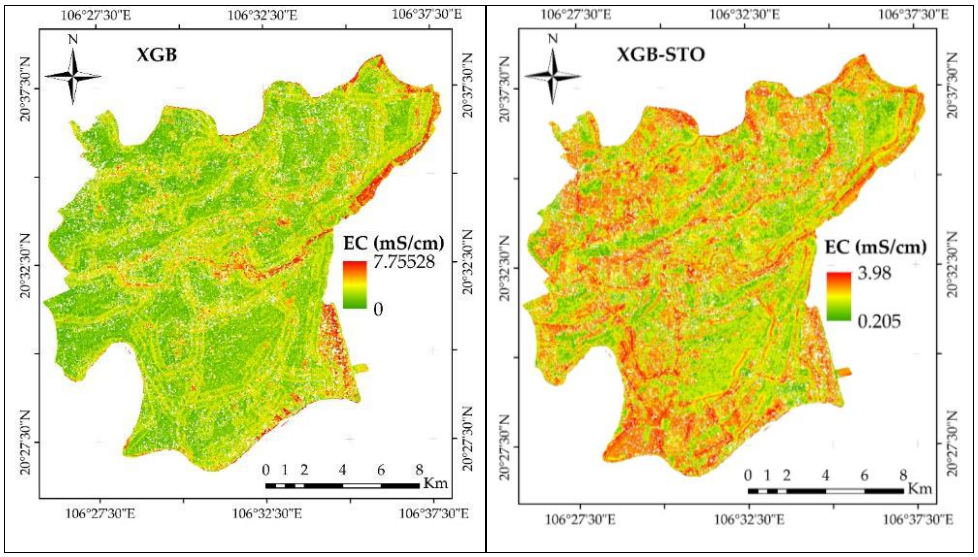
4.3. Spatial distribution of soil salinity in the Thai Thuy district of the Red River Delta

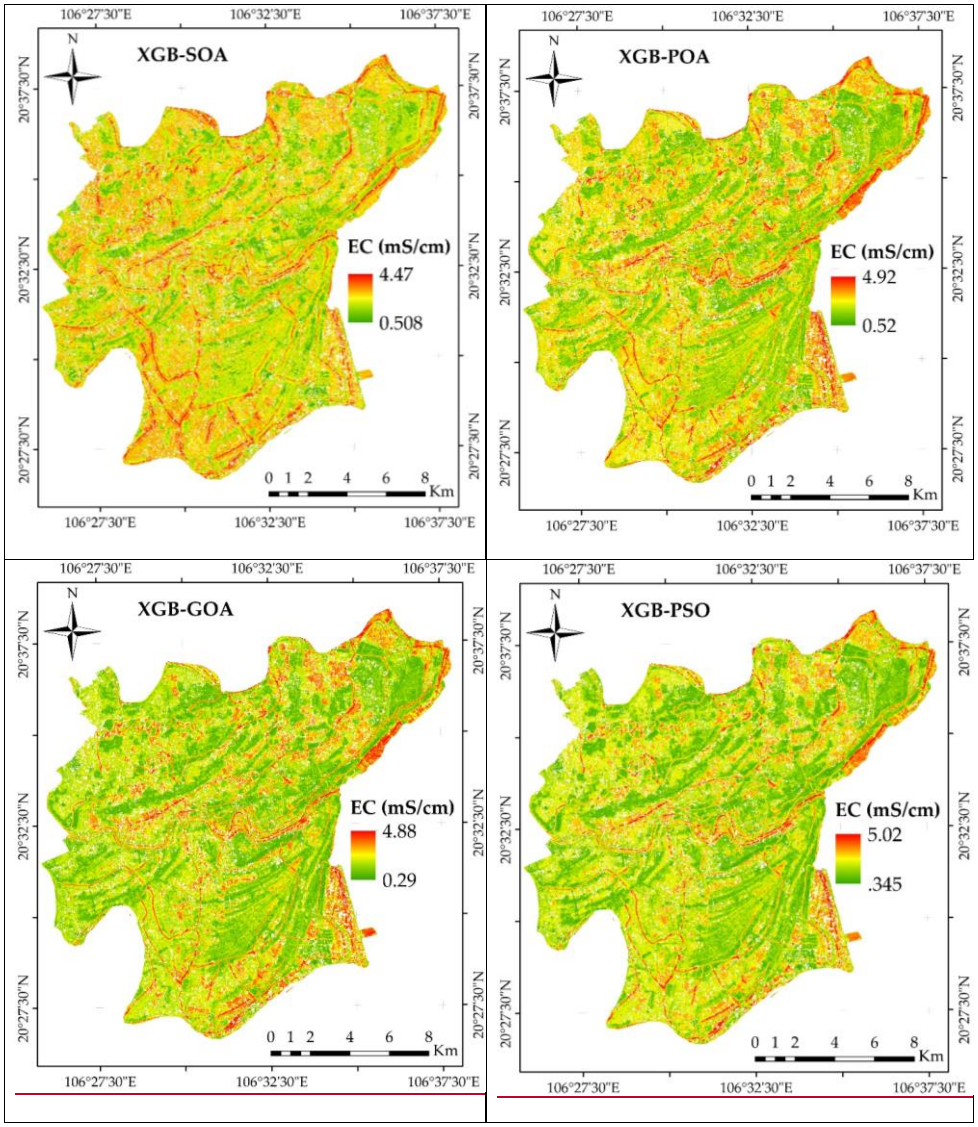
Following validation, we constructed a geographical distribution map of soil salinity using the proposed models. We carried out the process by assigning conditioning factors to the 30 million pixels across the entire study area. Depending on each model, the EC value varies from 0.29 to 7.7 mS/cm. On the map, the color varies from green to red, representing different EC values. The areas with green colors are located far from the continent (EC = 0.29), while the areas with red colors are located on the coast, with EC values superior to 7.7 mS/cm. This graph shows that these areas are directly affected by saltwater intrusion from the sea (Figure 5).

According to the FAO, salinity can be divided into 5 levels: non-saline, slightly saline, moderately saline, heavily saline, and very heavily saline. In which the EC value is below 2 mS/cm, the soil is considered not saline, and the plants grow completely normally. The EC value ranges from 2- 4 mS/cm; the soil is slightly saline and has very little effect on the plants. Specifically, flower crops may grow slowly, while rice and fruit trees may have reduced height. The EC value ranges from 4-8 mS/cm; the soil is moderately saline and reduces crop yields. Specifically, rice can have a 10-20% reduction in yield. The EC value ranges from 8 - 16 mS/cm; the soil is heavily saline and affects the growth of plants. The soil becomes extremely saline and uncultivated if the EC value surpasses the threshold of 16 mS/cm. This study utilizes the XGB-POA model, which boasts the highest accuracy, to analyze the areas affected by saline intrusion. Specifically, in the study area, about 65 km² of land area is not affected by saline intrusion, 165 km² is slightly affected, and 1.8 km² is greatly affected by saline intrusion. Compared to the land cover/land use map, it can be seen that the land area greatly affected by saline intrusion is mainly aquaculture; therefore, this area will not be significantly affected in terms of productivity. Meanwhile, 165 km2 of land affected by saline intrusion is rice land, which can slow down rice growth and reduce productivity.

In the study area of Thai Thuy District, Thai Binh Province, about 70-75% of the agricultural land is irrigated by a canal system that draws water from the Red River and the Day River. Thanks to the water source from the Red River and the Day River, agricultural production areas are regularly washed away with salt, so the EC value in these areas is often below 4 mS/cm. In coastal areas, salinity intrusion at river mouths forces people to use shallow groundwater for irrigation, which leads to salt accumulation and prompts a shift from rice cultivation to

aquaculture.After validation, the proposed models were used to construct a geographical distribution map of soil salinity. The process was done by assigning conditioning factors to the 30 million pixels for the entire study area. It can be seen that the EC value varies from 0.29 to 7.7 mS/cm, depending on each model. On the map, the color varies from green to red, representing different EC values. The areas with green color are located far from the continent (EC=0.29), while the areas with red color are located on the coast, with EC value superior 7.7 mS/cm. This shows that these areas are directly affected by saltwater intrusion from the sea (Figure 5).





493 Figure 5: Soil salinity mapping in the Thai Thuy district

494 **4.4. Farmers' Adaptive Capacity Assessment**

Soil salinity is a key challenge for deltas worldwide, ~~-~~particularly in deltas where population density is high and socio-economic conditions are poor (Hoque et al., 2016). The Red River Delta is the most densely populated area in Vietnam and one of the most densely populated deltas in Southeast Asia; therefore it is crucial~~it is key~~ to consider the impact of soil salinity in this delta on the farmer's life and their adaptive capacity. Spatial distribution maps of soil salinity show that this phenomenon occurs in many areas of the Thai Thuy district, especially in coastal areas. This phenomenon certainly has a significant impact on people's living conditions and production areas, posing challenges to their livelihoods. In this section, we address the adaptive capacity of farmers in An Tan commune, a coastal area, through five elements: i) natural resources including land use, awareness of saline intrusion, perceived impacts on agricultural activities, and adaptation measures taken, ii) human resources such as household demographics, education, and farming experience, iii) physical resources, including the availability of farming equipment and infrastructure, iv) financial resources focusing on household income, income structure, credit access, and changes over time, and v) social resources addressing support from government and social networks, community cooperation, and participation in collective adaptation activities. We interviewed 87 households in the An Tan commune to analyze the community's ability to adapt to saline intrusion.~~This certainly significantly impacts people's living and production areas and poses challenges to their livelihoods. In this section, we address the adaptive capacity of farmers in A Tan commune, a coastal area, through five elements: natural resources, human resources, physical resources, financial resources, and social resources. To analyze the community's ability to adapt to saline intrusion, 87 households in the A Tan commune were interviewed.~~

4.4.1. Natural Resources

Due to the process of ~~salinity~~salinization, we are facing a major threat to agriculture and sustaining arable land. Excess salinity adversely influences soil structure and fertility, plant growth, crop yield, and microorganisms (Tarolli et al., 2024). Soil salinity is frequently associated with water salinity. Groundwater in littoral regions of the Red River Delta is characterized by high salinity (Hoque et al., 2016). The scarcity of freshwater poses significant challenges for crop irrigation. Irrigating with saline water exacerbates soil salinity. In addition, soil salinity also creates a scarcity of grazing land and fodder cropland of coastal areas. Coastal livestock is harshly suffering from food inaccessibility and poultry farming in the coastal districts. All the mentioned factors impose considerable risks

to the coastal inhabitant's livelihood and food security, who rely mainly on agricultural activities such as growing rice and crops such as onions, garlic, watermelon, and tobacco, according to our interviews.

Indeed, the results showed that 59% of interviewed households said that salinization ~~had~~ had a medium to high impact on agricultural production in the area in recent years, especially during the 2023 saline intrusion. Meanwhile, 38% of households stated that saline intrusion has little or very little impact on agricultural production, mainly households in areas far from the coast and so less affected. Households in the study area have ~~taken~~ implemented ~~many~~ various measures to mitigate the increasingly serious saline intrusion, such as washing the salt from the fields after each crop (as instructed by local authorities) and ~~changing the crop structure~~adjusting their cropping systems by adopting salt-tolerant crop varieties or transitioning to aquaculture practices. 66% of interviewed households said they had to change the crop structure to suit the saline intrusion or switch to non-agricultural occupations to earn more income. Of the 87 households interviewed, 28% had their main income from non-agricultural activities. A large number of people switching to non-agricultural activities can meet their livelihood needs, but in the long-term, farmers abandoning agricultural activities to seek jobs in factories or migrate to cities also poses many negative environmental and social consequences, such as the decline of agrobiodiversity or labor shortages in agriculture, etc. (Subedi et al., 2022).

4.4.2. *Human Resources*

In agricultural production and the adaptability of the community to salinity intrusion, demographics are considered one of the most important factors contributing to the creation of labour resources, directly affecting crop productivity. The results of interviews with 87 households showed that each household has an average of 3.5 members, comprising 2.5 workers and 1 dependent member. Using available family resources reduces labor costs, thereby increasing production profits. However, the quality of human resources poses a concern when adapting to saline intrusion. One of the criteria for evaluation is education level. Most workers in households had a junior high school degree (66%), 5% of interviewees had a high school degree, and less than 3% had a university degree. According to previous studies, education is an important factor in determining workers' income. In the context of climate change and sea level rise, agriculture is negatively affected by these phenomena: low levels of education mean a lesser ability to absorb new knowledge and methods in organizing production to reduce the negative impacts of saline intrusion. Although 83% of the interviewed households had more than 20 years of experience in

548 agricultural production, their knowledge of saline intrusion and climate change was still limited. Specifically, people
549 lack the adequate knowledge and skills to adapt to changes in environmental conditions, leading to difficulties in
550 choosing appropriate livelihood models.

551 Furthermore, saline intrusion has led to several health insecurity. Coastal residents in saline areas are at risk of
552 consuming high salt above the recommended levels. It is evaluated that over 7 million coastal populations in
553 Bangladesh, India, and Vietnam suffer from hypertension and cardiovascular diseases as a result of long-term
554 ingestion of saline groundwater (Hoque et al., 2016). This has profound consequences for the development and
555 quality of human resources in these areas.

556 *4.4.3. Physical Resources*

557 Material resources include essential items that serve people's daily life and livelihoods. The majority of the
558 households interviewed were engaged in agriculture, so the means of production were mainly related to agricultural
559 activities. Of the 87 households interviewed, 90% were equipped with agricultural production equipment such as
560 pumps, sprayers, and tractors, while 100% had access to tractors and harvesters for farming. The interviewed
561 households used equipment to exploit water sources for agricultural production; however, because the town of An
562 Tan is located in a coastal area, groundwater and surface water are often affected by saline intrusion. Thus, they still
563 faced challenges related to water resources, especially in the context of saline intrusion

564 Moreover, in response to saline intrusion, accessing freshwater for daily use and irrigation often leads to the
565 spontaneous extraction of groundwater through tubewells, a common practice in coastal areas of Vietnam and
566 Southeast Asia (Hoque et al., 2016). However, excessive groundwater extraction and improper irrigation practices
567 also pose many potential risks of increasing water resource depletion and accelerating salinization processes (Tarolli
568 et al., 2024). This will likely undermine the long-term adaptive capacity of coastal communities.

569 *4.4.4. Financial Resources*

570 The interviews with 87 households demonstrated that 9% of the households interviewed were poor and near-poor. It
571 can be seen that economic status greatly affects people's ability to adapt and recover from soil salinity. Poor and
572 near-poor households frequently have more difficulty evaluating solutions to mitigate the impact of soil salinity on
573 agricultural production. The primary source of income for a large part of the population mainly comes from

agricultural activities: 72% of the households interviewed confirmed that their main livelihood was agriculture. However, their agricultural income is frequently unstable. This stresses the vulnerability of people's livelihoods due to the strong effects of soil salinity on agricultural activities. Meanwhile, 56% had a stable source of income from factory work. This emphasizes the need to diversify income sources for inhabitants in soil salinity areas. In addition, although the income from agricultural production was enough to cover farmers' daily life, most of the households interviewed could not save. Therefore, with increasing saline intrusion in the context of climate change, it is very difficult for these households to have an effective response or adaptation solutions. Borrowing capital to overcome the negative impacts of saline intrusion is one of the adaptation strategies reported by the interviewed households. 36% of households borrowed capital from relatives, 11% borrowed from credit funds or banks, 14% from local organisations, and 33% from distribution agents. Meanwhile, 8% of the interviewed households could not borrow any capital to overcome the consequences of saline intrusion.

Diversifying external sources of capital can help households overcome the consequences of saline intrusion and support agricultural production more generally. However, there are still several people who cannot access capital sources, and training to the adaptability and resilience of the people. This increases the impact of soil salinity on the community in the study area. Furthermore, a capital utilization strategy must be carefully considered to ensure efficient use of resources to improve livelihoods and enhance adaptability to saline intrusion.

4.4.5. *Social Resources*

Social resources play a key role in mitigating the impacts of soil salinity on the adaptability and resilience of the people. In the study area, households received support from various sources, such as local communities, volunteer organizations, non-governmental organizations, and mutual assistance among households. This support includes exemption from land use tax, support for production equipment, crops, and food supply for people. Regarding people's awareness of climate change and its impact on salinity intrusion, about 82% of households said they learned about this issue through local authorities and media propaganda.. Furthermore, 100% of the households surveyed reported that local authorities also had organised training sessions and drills to respond to saline intrusion and sea level rise. However, as mentioned above, the knowledge and skills of inhabitants are still limited. This raises questions about the effectiveness of training sessions. In addition, these training activities occur infrequently. For

example, during the 2soil salinity, people did not receive timely support and assistance from these organisations. This led to a reduction in the community's ability to adapt tsoil salinity.

5. Discussion

Soil salinity is a global environmental threat, a key cause of food insecurity worldwide (Song et al., 2024). Therefore, it is essential to monitor it with high precision, as is identifying the adaptive capacity of those in vulnerable regions.

In Vietnam, two large deltas ensure food security not only, but also in other countries. Although several previous studies have been conducted to assess soil salinity in Vietnam, most have focused on assessing soil salinity and farmers' adaptive capacity in the Mekong Delta (Hoang and Hai, 2024; Nguyen et al., 2024). Research on the Red River Delta is scarce~~Few studies have been conducted in the Red River Delta~~. The Red River Delta is one of the key agricultural regions in Southeast Asia. Therefore, assessing soil salinity and farms' adaptive capacity in this area is necessary. In this study, remote sensing, machine learning, and community interviews were used to evaluate soil salinity and the adaptive capacity of farms in the delta.

Remote sensing plays a key role in analyzing soil salinity because the salt in the soil has a significant effect on the spectral reflectance of the soil. Soils with different salinity levels will have different spectral characteristics; for example, areas covered with white salt often have higher spectral reflectance levels and salinity (Hoa et al., 2019; Wu et al., 2018; Xiao et al., 2023). This is the basis for using remote sensors to monitor saltwater intrusion. However, one of the challenges in using Sentinel 2 satellite images in soil salinity monitoring is that sometimes, the spectral reflectance level is not consistent with soil salinity. Many studies have integrated vegetation indices in soil salinity monitoring to minimize this limitation because different areas will have different soil salinity. It should also be noted that the difference depends on the vegetation type in each area. Therefore, this study has integrated Sentinel 1 images in soil salinity monitoring. Sentinel 1 images use radar signals to monitor moisture and dielectric properties providing accurate information on soil salinity. This is particularly important in coastal areas, where surface moisture is high, reducing the accuracy of optical imagery. This approach identifies areas severely affected by salinity intrusion while supporting the assessment of the adaptive capacity of communities in the area (Hoa et al., 2019). However, with the increase in the volume, type, and speed of remote sensing data collection, bottlenecks in

the data analysis process may occur (due to the inadequacy of the structure of current models for processing large datasets).

XGB is one of the most powerful algorithms for identifying the spatial distribution of natural hazards, such as floods, landslides, and soil salinity. Its advantages include the ability to avoid the overfitting problem and fast convergence. Additionally, XGB effectively handles missing values (Liu et al., 2022; Mo et al., 2019). However, both the configuration and interpretation of XGB are more complex, and the parameters of this model are also complex to tune. Incorrect parameter selection can reduce performance (Ramraj et al., 2016). Therefore, it is necessary to use optimization algorithms to select the parameters of this model. In this study, five optimization algorithms, namely POA, STO, SOA, PSO, and GOA, were utilized to optimize the parameters of XGB. The XGB-POA model outperformed the other models as it is easy to carry out, has few parameters to adjust, has faster convergence capability, and can avoid local minima - which enables it to find the best global solution (Premkumar and Santhosh, 2024). Previous studies have indicated that POA also can solve complex problems with a large number of variables and non-linear properties (Alamir et al., 2023; Li et al., 2023). XGB-STO model ranked second. The STO algorithm maintains a good equilibrium with exploration and exploitation processes. This allows it to avoid local minima problems, which improves model performance (Al-Sarray et al., 2024; Trojovský et al., 2022). The XGB-SOA model came third in terms of accuracy. SOA can solve complex problems with a large number of variables or continuous, discrete, or multi-objective problems, so it is a versatile tool for several different applications. In addition, inspired by the serval's precise jumps and fast movements, SOA can converge quickly with high accuracy (Dehghani and Trojovský, 2022; Sindi et al., 2024). The XGB-PSO model was ranked fourth. In addition to ease of use, PSO has the advantage of equilibrium of the exploration and exploitation processes. This can avoid the local optimization problem (Juneja and Nagar, 2016; Rini et al., 2011). The XGB-GOA model was less accurate than other models because it tends to concentrate exploration at the beginning of the process to avoid the local optimization problem. This may lead to slow convergence (Mirjalili et al., 2018; Zhao et al., 2019). When comparing the models proposed in this study on the ability to predict natural hazards such as soil salinity, each model has different characteristics that influence the real-time prediction ability. Three models (XGB-POA, XGB-STO, and XGB-SOA) can converge quickly because of the faster learning speed. Therefore, these models best suit adaptation for real-time applications because fast updates are necessary to support those tasked with developing mitigation strategies.

The results of this study not only confirmed the effectiveness of XGBoost models in soil salinity prediction but also showed the potential for improving accuracy by combining them with optimization algorithms. Compared with previous studies, the models in this study outperform traditional models. (Wang and Sun, (2024) used three machine learning models, namely random forest (RF), support vector machine (SVM), and artificial neural network (ANN), to predict the soil salinity in Huludao City, China. The results indicated that the RF model performed better with an R^2 value of 0.84. The model accuracy in Wang and Sun's 2024 study was lower than the models in our study. (Aksoy et al., (2024) applied two models, namely XGB and RF to predict the soil salinity in western and southeastern Lake Urmia Playas (LUP) in the Northwest of Iran. The results showed that the XGB model was more efficient with the R^2 value of 0.83. It was less accurate than in our study. (Elshewy et al., (2024) evaluated the soil salinity in Sharkia Governorate, Egypt, using four machine learning models, namely support vector machines (SVM), regression trees, Gaussian linear regression, and tree-based ensemble. The results indicated that the SVM model performed better with an R^2 value of 0.86. Comparison with previous studies showed the potential of the machine learning model in this study to predict soil salinity.

Saline intrusion in the Red River Delta and the study area reflects the interaction between natural factors and human activities. One cause of the increasingly serious saline intrusion in the study area is the reduced flow in the delta caused by the construction of dams and reservoirs upstream in China. This reduction reduces the ability of the river system to repel salt water, creating conditions for seawater to penetrate deep into the inland. (Hien et al., (2023) have emphasized that by 2050, saltwater intrusion is likely to extend about 20 km inland from the river mouth, related to sea level rise and reduced discharge from the upper river. (Nguyen et al., (2017) reported that the increasing trend of saline intrusion is the result of sea level rise, combined with the decline of the Red River water level, especially in the dry season. Specifically, the sea level increased by 0.19 m in the period 1901-2010, with an average rate of 3.2 m from 1993-2010. In addition, the phenomenon of saline intrusion is increasingly severe due to subsidence related to groundwater exploitation. In many areas of the Red River Delta in general and the study area in particular, uncontrolled groundwater exploitation for agricultural production and aquaculture contributes to subsidence, increasing the impact of tides. (Nguyen and Takewaka, (2020) have emphasized that the subsidence phenomenon in the delta can reach -12.3 mm/year, which is one of the causes that aggravate the problem of saltwater intrusion, especially in the context of rising sea levels.

Formatted: English (United States)
Field Code Changed
Formatted: English (United States)
Formatted: English (United States)
Field Code Changed
Formatted: English (United States)
Field Code Changed
Formatted: English (United States)
Formatted: English (United States)
Formatted: English (United States)
Formatted: English (United States)

The results of this study explored the adaptive capacity of farms in the Thai Thuy district of Thai Binh province. Riverine farmers in areas affected by saltwater intrusion are prepared. They rely on their local communities and expect support from local authorities and voluntary organisations. Our results are similar to those of previous studies investigating the adaptive capacity of residential communities to natural hazards, including saltwater intrusion. The key to adaptation is education, knowledge, and resources to cope with saltwater intrusion. These resources can be natural, physical, financial, social, and human resources.

The community's adaptive capacity in the study area faces many challenges, especially in the context of global warming and growing saltwater intrusion. Although most households surveyed have more than 20 years of experience in agricultural production and benefit from available labor resources, their adaptive capacity to saltwater intrusion remains limited. This is in part because these households lack the knowledge to change their livelihood patterns in addressing varying environmental situations. In addition, the main sources of agricultural income are often unstable, and the ability to accumulate finances is low, leading to difficulties in adapting to and recovering from saltwater intrusion. People's adaptation strategies, such as uncontrolled groundwater extraction and conversion to non-agricultural activities, also present long-term environmental and social risks. Furthermore, policies and support programs for residents, such as training sessions and lending programs provided by stakeholders, also raise concerns regarding their effectiveness and inclusiveness. Although people in the study area have access to capital from many different sources, some households still cannot access these sources of capital to overcome the consequences of saltwater intrusion. All of these factors impact agriculture and human life, leading to increased household vulnerability. To enhance people’s adaptive capacity, it is important to emphasize the role and effectiveness of policies of local governments, policymakers and stakeholders in supporting people to understand better and respond to saline intrusion. Information and knowledge sharing can be done through direct outreach to people to raise awareness of saline intrusion among communities. Lending policies of local governments and stakeholders need to cover all households while improving the efficiency of capital use. Effective management of natural and physical resources and enhancing social capital through the development of cooperative community models are important factors contributing to people’s adaptive capacity to saline intrusion. This study has successfully built a theoretical framework using machine learning with optimization algorithms, remote sensing, and farmer interviews to determine the spatial distribution of soil salinity and farmers' adaptation capacity. However, to apply this theoretical framework in different regions, it is necessary to use factors specifically pertinent to each

region. Machine learning models must be provided with the local characteristics of the region in question. However, data collection in any region is difficult, often due to restrictive data-sharing policies or limited financing resources to maintain and distribute the data.

From field surveys, it can be seen that in the Red River Delta, adaptation options to soil salinity mainly rely on upgrading the sea dike system, river dikes, and saline prevention sluice systems. In addition, other adaptation options mainly include increasing the resilience of the current agricultural system, such as changing the crop calendar, changing crop varieties, using fertilizers, and planting mangroves. Many households have transitioned from rice cultivation to aquaculture in coastal areas, where soil salinity has a significant impact. In addition, some fish farming households have also switched to shrimp farming or fish farming due to increased saline intrusion. Some households do not have the capital to convert their agricultural systems, and while agricultural productivity decreases due to saline intrusion, they consider finding non-agricultural jobs or migrating to the city to find jobs with more stable incomes. Households located further inland, less affected by saline intrusion, still maintain traditional agriculture. Some households practice intercropping by growing rice alongside vegetables to increase their income. Thus, it can be seen that the adaptability of households in the Red River Delta is not only based on strengthening the system of sea dikes, river dikes, and salinity prevention sluices, but also on transforming the traditional agricultural system to minimize the impact of salinity intrusion. However, capital barriers force many households to abandon agriculture, seriously affecting the food security situation in the region.

Formatted: Vietnamese

A significant problem when using machine learning is that of extrapolation. Each model built is adapted only to one set of data. Therefore, evaluating the soil salinity in other regions is challenging. ~~Machine learning models require stationarity due to the abrupt and non-stationary nature of the system~~General, there is only one model that fits each training dataset. In theory, this would not be a problem if enough training data were collected and all extreme events were included. However, in practice, it is very difficult to collect data for all these events, especially in the context of climate change and sea level rise.~~.~~. To solve this problem, several studies have pointed out that integrating machine learning with conventional models for example, remote sensing or hydrodynamic models can be effective, as such traditional models can provide the training data to use as the input file of the machine learning model. Another solution is to combine machine learning with optimization algorithms, as in this study, to enhance the prediction capability of the machine learning model (Tran and Kim, 2022).

Formatted: English (United States)

This study emphasizes the significance of combining machine learning methods to analyze the spatial distribution of salinity intrusion with the community's adaptive capacity to soil salinity. The salinity intrusion map from the machine learning model shows a clear difference in the level of salinity intrusion between coastal, riverside, and inland areas.s. Coastal and estuarine areas often have high levels of salinity intrusion, with EC values exceeding 7 mS/cm. These are also areas where communities must apply appropriate adaptation strategies, including crop restructuring, selecting more salinity-tolerant plant varieties, combining agriculture and fisheries, or livelihood conversion. In contrast, inland areas, where the level of salinity intrusion is lower, have less variation in agricultural production models, and communities in these areas still mainly maintain traditional agricultural practices. The findings may indicate that the coping strategies and adaptive capacity of the communities depend on the level of salinity intrusion in the areas. In addition, it can be seen that in areas with high salinity intrusion, people have difficulty in accessing fresh water for agricultural production; therefore, the communities in this area tend to depend more on non-agricultural sources of income. Previous studies– (Nguyen et al., 2019; Yuen et al., 2021), have demonstrated this trend.

The results of the study emphasize that the integration of advanced machine learning models and sociological surveys not only improves the comprehensive research ability from natural factors to socio-economic factors but can also support policymakers and planners to develop appropriate adaptation solutions. Identifying areas affected by saline intrusion by using machine learning models and qualitative analysis of the adaptive capacity of the community is a solid scientific basis for developing policies to minimize the impact of saline intrusion, especially in the context of climate change, to ensure agricultural development and food security.

This study was successful in building machine learning models integrated with optimization algorithms to identify the spatial of soil salinity, as well as evaluating farmers' adaptive capacity in the study area. However, in terms of data, this study collected 62 soil salinity samples to build the machine learning model; therefore, the soil salinity map constructed by the proposed models cannot present the trend and drive of soil salinity in time series. Furthermore, soil salinity is significantly affected by climate change and rising sea levels, so it is necessary to assess the effects of this change on soil salinity in the future.

As hydrological conditions change, those living in deltas are confronting increased risk. The Red River Delta is one of the largest deltas in the world and, thanks to its fertile floodplains, is home to about 21 million inhabitants. In

Formatted: English (United States)

Formatted: English (United States)

Formatted: Vietnamese

Formatted: Vietnamese

Formatted: Vietnamese

recent years, in the context of global warming and rising sea levels, these deltas are confronting growing flooding and soil salinity problems, which affect food security in the region and the country. Policies must be implemented to improve the agricultural system and the adaptive capacity of farmers. A proactive approach is required, envisaging multiple scenarios to provide appropriate support for agriculture. These scenarios may include activities and programs adaptive to the different influences of global warming on soil salinity.

6. Conclusion

Soil salinity is a key environmental threat, which will have a growing effect on the development of agriculture and food security globally. A lack of assessment of local adaptive capacities exacerbates the problem. Therefore, this research's objective was toconstruct a theoretical framework to assess soil salinity and farmers' adaptive capacity based on machine learning, optimization algorithms, remote sensing, and interviews with local people. The results in this study represent a novel contribution to the literature for researchers worldwide and can support policy-makers and farmers to establish suitable strategies to limit damage related to soil salinity. The outcome of this research is as follows.

- This study justified the effectiveness of machine learning and remote sensing in soil salinity monitoring in the Red River Delta. The results of this study can be opened to realize in different regions.

- Five optimization algorithms, namely POA, STO, SOA, PSO, and GOA, were successful in optimizing the accuracy of the XGB model. All these algorithms were successful in improving the accuracy of XGB. Of these, the XGB-POA model showed the greatest performance, with an R^2 value of 0.968. This was followed by XGB-STO (R^2=0.967), XGB-SOA (R^2=0.966), XGB-PSO (R^2=0.964), and XGB-GOA (R^2=0.964).

- The models in this research were utilized to construct soil salinity maps. The maps demonstrated that littoral areas and those along the rivers were the most influenced by the soil salinity problem because these regions are influenced by seawater. In addition, when the river levels are lower during the dry season, it creates the conditions for seawater to penetrate the land.

- Five factors were analyzed to consider farmers' adaptive capacity: natural capital, human capital, material resources, financial resources, and social capital. The results show that people have awareness and actions in improving their adaptive capacity to increasingly severe saline intrusion; however, there are still many limitations

related to lack of awareness and finance. As a recommendation, the participation of multiple stakeholders is required, with a particular emphasis on the role of policies in sustainably and effectively enhancing people's adaptive capacity.

The outcome of this research provides key knowledge on the spatial distribution of soil salinity and farmers' adaptive capacity to growing salinization, to support local authorities or farmers in proposing appropriate measures to reduce soil salinity damage. This can complement a theoretical framework in the existing literature on soil salinity management and adaptive capacity.

Statements and Declarations

Funding

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 105.08-2023.13

Competing Interests

The authors declare that they have no conflict of interest.

Ethics and consent to participate

Not applicable

Consent for publication

Not applicable

Authors Contributions

Huu Duy Nguyen: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Quang-Thanh Bui:** Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Thi Anh Tam Lai:** Data curation**. Duc Dung Tran:** Data curation, Investigation, Writing – original draft, Writing – review & editing. **Dinh Kha Dang:** Data curation, Investigation. **Himan Shahabi:** Writing – original draft, Writing – review & editing.

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request

Reference

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

Al-Sarray, N.H.S., Rahebi, J., Demirhan, A., 2024. Detection of DDoS attacks in SDN with Siberian Tiger Optimization algorithm and deep learning.

822 Alamir, N., Kamel, S., Megahed, T.F., Hori, M., Abdelkader, S.M., 2023. Developing hybrid demand
823 response technique for energy management in microgrid based on pelican optimization
824 algorithm. Electric Power Systems Research 214, 108905.
825 Asfaw, E., Suryabhagavan, K., Argaw, M., 2018. Soil salinity modeling and mapping using remote sensing
826 and GIS: The case of Wonji sugar cane irrigation farm, Ethiopia. Journal of the Saudi Society of
827 Agricultural Sciences 17(3), 250-258.
828 Bandak, S., Movahedi-Naeini, S.A., Mehri, S., Lotfata, A., 2024. A longitudinal analysis of soil salinity
829 changes using remotely sensed imageries. Scientific Reports 14(1), 10383.
830 Bhuyan, M.I., Supit, I., Kumar, U., Mia, S., Ludwig, F., 2024. The significance of farmers' climate change
831 and salinity perceptions for on-farm adaptation strategies in the south-central coast of
832 Bangladesh. Journal of Agriculture and Food Research 16, 101097.
833 Bui, D.T., Pradhan, B., Nampak, H., Bui, Q.-T., Tran, Q.-A., Nguyen, Q.-P., 2016. Hybrid artificial
834 intelligence approach based on neural fuzzy inference model and metaheuristic optimization for
835 flood susceptibilitgy modeling in a high-frequency tropical cyclone area using GIS. Journal of
836 Hydrology 540, 317-330.
837 Castelletti, A., Pianosi, F., Quach, X., Soncini-Sessa, R., 2012. Assessing water reservoirs management
838 and development in Northern Vietnam. Hydrology and Earth System Sciences 16(1), 189-199.
839 Cullu, M.A., 2003. Estimation of the effect of soil salinity on crop yield using remote sensing and
840 geographic information system. Turkish Journal of Agriculture and Forestry 27(1), 23-28.
841 Dasgupta, S., Laplante, B., Murray, S., Wheeler, D., 2009. Climate change and the future impacts of
842 storm-surge disasters in developing countries. Center for Global Development Working Paper
843 (182).
844 Dehghani, M., Trojovský, P., 2022. Serval optimization algorithm: a new bio-inspired approach for
845 solving optimization problems. Biomimetics 7(4), 204.
846 Du, L., Tian, S., Zhao, N., Zhang, B., Mu, X., Tang, L., Zheng, X., Li, Y., 2024. Climate and topography
847 regulate the spatial pattern of soil salinization and its effects on shrub community structure in
848 Northwest China. Journal of Arid Land, 1-18.
849 Eldeiry, A.A., Garcia, L.A., Reich, R.M., 2008. Soil salinity sampling strategy using spatial modeling
850 techniques, remote sensing, and field data. Journal of irrigation and drainage engineering
851 134(6), 768-777.
852 Elshewy, M.A., Mohamed, M.H., Refaat, M., 2024. Developing a soil salinity model from landsat 8
853 satellite bands based on advanced machine learning algorithms. Journal of the Indian Society of
854 Remote Sensing 52(3), 617-632.
855 Fathizad, H., Ardakani, M.A.H., Sodaiezadeh, H., Kerry, R., Taghizadeh-Mehrjardi, R., 2020. Investigation
856 of the spatial and temporal variation of soil salinity using random forests in the central desert of
857 Iran. Geoderma 365, 114233.
858 Fu, C., Gan, S., Yuan, X., Xiong, H., Tian, A., 2018. Determination of soil salt content using a probability
859 neural network model based on particle swarm optimization in areas affected and non-affected
860 by human activities. Remote Sensing 10(9), 1387.
861 Gong, Z., He, H., Fan, D., Sheng, J., Zhang, X., 2023. Inversion of sea surface salinity in Bay of Bengal
862 based on Catboost algorithm, Second International Conference on Geographic Information and
863 Remote Sensing Technology (GIRST 2023). SPIE, pp. 297-304.
864 Hardie, M., Doyle, R., 2012. Measuring soil salinity. Plant salt tolerance: methods and protocols, 415-
865 425.
866 He, Y., Yin, H., Chen, Y., Xiang, R., Zhang, Z., Chen, H., 2024. Soil Salinity Estimation based on Sentinel-1/2
867 Texture Features and Machine Learning. IEEE Sensors Journal.

Hien, N.T., Yen, N.H., Balistrocchi, M., Peli, M., Cat, V.M., Ranzi, R., 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.
Hoa, P.V., Giang, N.V., Binh, N.A., Hai, L.V.H., Pham, T.-D., Hasanlou, M., Tien Bui, D., 2019. Soil salinity mapping using SAR sentinel-1 data and advanced machine learning algorithms: A case study at Ben Tre Province of the Mekong River Delta (Vietnam). Remote Sensing 11(2), 128.
Hoang, L.P., Pot, M., Tran, D.D., Ho, L.H., Park, E., 2023. Adaptive capacity of high- and low dyke farmers to hydrological changes in the Vietnamese Mekong delta. Environmental Research 224, 115423.
Hoang, N.K., Hai, D.T., 2024. Assessing surface water salinity intrusion in the Mekong River Delta: a case study in Rach Gia, Vietnam. GeoJournal 89(4), 171.
Hung, N.M., Larson, M., 2014. Coastline and river mouth evolution in the central part of the Red River Delta, Coastal disasters and climate change in Vietnam. Elsevier, pp. 43-79.
Ingle, K.K., Jatoth, R.K., 2024. Non-linear Channel Equalization using Modified Grasshopper Optimization Algorithm. Applied Soft Computing 153, 110091.
Jia, P., Zhang, J., Liang, Y., Zhang, S., Jia, K., Zhao, X., 2024. The inversion of arid-coastal cultivated soil salinity using explainable machine learning and Sentinel-2. Ecological Indicators 166, 112364.
Jiang, H., Rusuli, Y., Amuti, T., He, Q., 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.
Juneja, M., Nagar, S., 2016. Particle swarm optimization algorithm and its parameters: A review, 2016 International Conference on Control, Computing, Communication and Materials (ICCCCM). IEEE, pp. 1-5.
Kaplan, G., Gašparović, M., Alqasemi, A.S., Aldhaheri, A., Abuelgasim, A., Ibrahim, M., 2023. Soil salinity prediction using Machine Learning and Sentinel – 2 Remote Sensing Data in Hyper – Arid areas. Physics and Chemistry of the Earth, Parts A/B/C 130, 103400.
Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, Proceedings of ICNN'95-international conference on neural networks. ieee, pp. 1942-1948.
Le Dang, H., Li, E., Nuberg, I., Bruwer, J., 2014. Understanding farmers' adaptation intention to climate change: A structural equation modelling study in the Mekong Delta, Vietnam. Environmental science & policy 41, 11-22.
Li, S., Zhang, T., Yu, J., 2023. Photovoltaic array dynamic reconfiguration based on an improved pelican optimization algorithm. Electronics 12(15), 3317.
Liu, W., Chen, Z., Hu, Y., 2022. XGBoost algorithm-based prediction of safety assessment for pipelines. International Journal of Pressure Vessels and Piping 197, 104655.
Liu, Y., Han, X., Zhu, Y., Li, H., Qian, Y., Wang, K., Ye, M., 2024. Spatial mapping and driving factor Identification for salt-affected soils at continental scale using Machine learning methods. Journal of Hydrology 639, 131589.
Ma, G., Ding, J., Han, L., Zhang, Z., Ran, S., 2021. Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms. Regional Sustainability 2(2), 177-188.
Mazumder, M.S.U., Kabir, M.H., 2022. Farmers' adaptations strategies towards soil salinity effects in agriculture: the interior coast of Bangladesh. Climate Policy 22(4), 464-479.
Mirjalili, S.Z., Mirjalili, S., Saremi, S., Faris, H., Aljarah, I., 2018. Grasshopper optimization algorithm for multi-objective optimization problems. Applied Intelligence 48, 805-820.
Mo, H., Sun, H., Liu, J., Wei, S., 2019. Developing window behavior models for residential buildings using XGBoost algorithm. Energy and Buildings 205, 109564.

Formatted: French (France)

Formatted: French (France)

914 Moayedi, H., Gör, M., Lyu, Z., Bui, D.T., 2020. Herding Behaviors of grasshopper and Harris hawk for
915 hybridizing the neural network in predicting the soil compression coefficient. Measurement 152,
916 107389.
917 Moayedi, H., Nguyen, H., Kok Foong, L., 2021. Nonlinear evolutionary swarm intelligence of grasshopper
918 optimization algorithm and gray wolf optimization for weight adjustment of neural network.
919 Engineering with Computers 37(2), 1265-1275.
920 Mukhamediev, R.I., Merembayev, T., Kuchin, Y., Malakhov, D., Zaitseva, E., Levashenko, V., Popova, Y.,
921 Symagulov, A., Sagatdinova, G., Amirgaliyev, Y., 2023. Soil Salinity Estimation for South
922 Kazakhstan Based on SAR Sentinel-1 and Landsat-8, 9 OLI Data with Machine Learning Models.
923 Remote Sensing 15(17), 4269.
924 Nguyen, H.D., 2022. Hybrid models based on deep learning neural network and optimization algorithms
925 for the spatial prediction of tropical forest fire susceptibility in Nghe An province, Vietnam.
926 Geocarto International 37(26), 11281-11305.
927 Nguyen, H.D., Van, C.P., Nguyen, T.G., Dang, D.K., Pham, T.T.N., Nguyen, Q.-H., Bui, Q.-T., 2023. Soil
928 salinity prediction using hybrid machine learning and remote sensing in Ben Tre province on
929 Vietnam's Mekong River Delta. Environmental Science and Pollution Research, 1-18.
930 Nguyen, M.T., Renaud, F.G., Sebesvari, Z., 2019. Drivers of change and adaptation pathways of
931 agricultural systems facing increased salinity intrusion in coastal areas of the Mekong and Red
932 River deltas in Vietnam. Environmental science & policy 92, 331-348.
933 Nguyen, P.T., Koedsin, W., McNeil, D., Van, T.P., 2018. Remote sensing techniques to predict salinity
934 intrusion: application for a data-poor area of the coastal Mekong Delta, Vietnam. International
935 journal of remote sensing 39(20), 6676-6691.
936 Nguyen, Q.H., Takewaka, S., 2020. Land subsidence and its effects on coastal erosion in the Nam Dinh
937 Coast (Vietnam). Continental Shelf Research 207, 104227.
938 Nguyen, T.G., Tran, N.A., Vu, P.L., Nguyen, Q.-H., Nguyen, H.D., Bui, Q.-T., 2021. Salinity intrusion
939 prediction using remote sensing and machine learning in data-limited regions: A case study in
940 Vietnam's Mekong Delta. Geoderma Regional 27, e00424.
941 Nguyen, V.H., Germer, J., Asch, F., 2024. Evaluating topsoil salinity via geophysical methods in rice
942 production systems in the Vietnam Mekong Delta. Journal of Agronomy and Crop Science
943 210(1), e12676.
944 Nguyen, Y.T.B., Kamoshita, A., Dinh, V.T.H., Matsuda, H., Kurokura, H., 2017. Salinity intrusion and rice
945 production in Red River Delta under changing climate conditions. Paddy and Water Environment
946 15, 37-48.
947 Nhuan, M.T., Van Ngoi, C., Nghi, T., Tien, D.M., van Weering, T.C., van den Bergh, G., 2007. Sediment
948 distribution and transport at the nearshore zone of the Red River delta, Northern Vietnam.
949 Journal of Asian Earth Sciences 29(4), 558-565.
950 Premkumar, N., Santhosh, R., 2024. Pelican optimization algorithm with blockchain for secure load
951 balancing in fog computing. Multimedia Tools and Applications 83(18), 53417-53439.
952 Quang, N.H., Viet, T.Q., Thang, H.N., Hieu, N.T.D., 2024. Long-term water level dynamics in the Red River
953 basin in response to anthropogenic activities and climate change. Science of The Total
954 Environment 912, 168985.
955 Ramraj, S., Uzir, N., Sunil, R., Banerjee, S., 2016. Experimenting XGBoost algorithm for prediction and
956 classification of different datasets. International Journal of Control Theory and Applications
957 9(40), 651-662.
958 Rhoades, J., Ingvalson, R., 1971. Determining salinity in field soils with soil resistance measurements. Soil
959 Science Society of America Journal 35(1), 54-60.
960 Rini, D.P., Shamsuddin, S.M., Yuhaniz, S.S., 2011. Particle swarm optimization: technique, system and
961 challenges. International journal of computer applications 14(1), 19-26.

962 Ruidas, D., Chakrabortty, R., Islam, A.R.M.T., Saha, A., Pal, S.C., 2022. A novel hybrid of meta-
963 optimization approach for flash flood-susceptibility assessment in a monsoon-dominated
964 watershed, Eastern India. Environmental earth sciences 81(5), 145.
965 Shi, H., Hellwich, O., Luo, G., Chen, C., He, H., Ochege, F.U., Van de Voorde, T., Kurban, A., De Maeyer, P.,
966 2021. A global meta-analysis of soil salinity prediction integrating satellite remote sensing, soil
967 sampling, and machine learning. IEEE Transactions on Geoscience and Remote Sensing 60, 1-15.
968 Sindi, A.O.N., Si, P., Li, Q., 2024. Secure Task Offloading and Resource Allocation Strategies in Mobile
969 Applications Using Probit Mish-Gated Recurrent Unit and an Enhanced-Searching-Based Serval
970 Optimization Algorithm. Electronics 13(13), 2462.
971 Song, Y., Gao, M., Wang, J., 2024. Inversion of salinization in multilayer soils and prediction of water
972 demand for salt regulation in coastal region. Agricultural Water Management 301, 108970.
973 Su, Y., Li, T., Cheng, S., Wang, X., 2020. Spatial distribution exploration and driving factor identification
974 for soil salinisation based on geodetector models in coastal area. Ecological Engineering 156,
975 105961.
976 Tan, J., Ding, J., Han, L., Ge, X., Wang, X., Wang, J., Wang, R., Qin, S., Zhang, Z., Li, Y., 2023. Exploring
977 planetScope satellite capabilities for soil salinity estimation and mapping in arid regions oases.
978 Remote Sensing 15(4), 1066.
979 Thiam, H.I., Owusu, V., Villamor, G.B., Schuler, J., Hathie, I., 2024. Farmers' intention to adapt to soil
980 salinity expansion in Fimela, Sine-Saloum area in Senegal: A structural equation modelling
981 approach. Land Use Policy 137, 106990.
982 Tran, V.N., Kim, J., 2022. Robust and efficient uncertainty quantification for extreme events that deviate
983 significantly from the training dataset using polynomial chaos-kriging. Journal of Hydrology 609,
984 127716.
985 Trojovský, P., Dehghani, M., 2022. Pelican optimization algorithm: A novel nature-inspired algorithm for
986 engineering applications. Sensors 22(3), 855.
987 Trojovský, P., Dehghani, M., Hanuš, P., 2022. Siberian tiger optimization: A new bio-inspired
988 metaheuristic algorithm for solving engineering optimization problems. Ieee Access 10, 132396-
989 132431.
990 Vermeulen, D., Van Niekerk, A., 2017. Machine learning performance for predicting soil salinity using
991 different combinations of geomorphometric covariates. Geoderma 299, 1-12.
992 Vinh, V.D., Ouillon, S., Thanh, T.D., Chu, L., 2014. Impact of the Hoa Binh dam (Vietnam) on water and
993 sediment budgets in the Red River basin and delta. Hydrology and Earth System Sciences 18(10),
994 3987-4005.
995 Wang, H., Hsieh, Y.P., Harwell, M.A., Huang, W., 2007. Modeling soil salinity distribution along
996 topographic gradients in tidal salt marshes in Atlantic and Gulf coastal regions. Ecological
997 modelling 201(3-4), 429-439.
998 Wang, H., Zhang, L., Zhao, J., Hu, X., Ma, X., 2022. Application of hyperspectral technology combined
999 with bat algorithm-AdaBoost model in field soil nutrient prediction. Ieee Access 10, 100286-
1000 100299.
1001 Wang, J., Peng, J., Li, H., Yin, C., Liu, W., Wang, T., Zhang, H., 2021. Soil salinity mapping using machine
1002 learning algorithms with the Sentinel-2 MSI in arid areas, China. Remote Sensing 13(2), 305.
1003 Wang, J., Wang, X., Zhang, J., Shang, X., Chen, Y., Feng, Y., Tian, B., 2024. Soil Salinity Inversion in Yellow
1004 River Delta by Regularized Extreme Learning Machine Based on ICOA. Remote Sensing 16(9),
1005 1565.
1006 Wang, W., Sun, J., 2024. Estimation of soil salinity using satellite-based variables and machine learning
1007 methods. Earth Science Informatics, 1-13.

Wu, W., Zucca, C., Muhaimeed, 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. Land degradation & development 29(11), 4005-4014.

Xiao, C., Ji, Q., Chen, J., Zhang, F., Li, Y., Fan, J., Hou, X., Yan, F., Wang, H., 2023. Prediction of soil salinity parameters using machine learning models in an arid region of northwest China. Computers and Electronics in Agriculture 204, 107512.

Xiao, S., Nurmemet, I., Zhao, J., 2024. Soil salinity estimation based on machine learning using the GF-3 radar and Landsat-8 data in the Keriya Oasis, Southern Xinjiang, China. Plant and Soil 498(1), 451-469.

Xie, J., Shi, C., Liu, Y., Wang, Q., Zhong, Z., Wang, X., He, S., Feature Variable Selection Methods for Inversion of Soil Salinity at the Irrigation District Scale Based on Machine Learning. Available at SSRN 4865442.

Yuen, K.W., Hanh, T.T., Quynh, V.D., Switzer, A.D., Teng, P., Lee, J.S.H., 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.

Zhang, W., Zhang, W., Liu, Y., Zhang, J., Yang, L., Wang, Z., Mao, Z., Qi, S., Zhang, C., Yin, Z., 2022. The role of soil salinization in shaping the spatio-temporal patterns of soil organic carbon stock. Remote Sensing 14(13), 3204.

Zhao, R., Ni, H., Feng, H., Song, Y., Zhu, X., 2019. An improved grasshopper optimization algorithm for task scheduling problems. Int. J. Innov. Comput., Inf. Control 15, 1967-1987.