https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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[3,4,5], Himan Shahabi[6], 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

[6] 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 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. The geographical distribution of soil salinity was evaluated by applying machine learning Sentinel 1 and 2A. The 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 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 a key theoretical framework that supplements the previous studies and can support policy-markers and farmers in land resource management, for example 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 significant problems to agricultural progress and global food security (Jia et al., 2024; He et al., 2024; Xiao et al., 2024). 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 (Jia et al., 2024; Aksoy et al., 2024).

The soil salinity problem will occur at the local, regional, and global levels (Liu et al., 2024; Bandak et al., 2024). In Vietnam, many littoral regions affected by soil salinity problems. According to the 2021 Ministry of Agriculture and Rural Development Report on the Current Situation and Planning of Agricultural Development, in 2020, about 200,000 hectares of cropland in Vietnam were touched by soil salinity. The Red River and Mekong Deltas are home to more than 40 million people and represent an extremely key role in the country's agricultural and aquaculture activities. 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). These low-lying coastal areas (Hung and Larson, 2014) are experiencing subsidence (Le Dang et al., 2014), and river water levels are decreasing. As such, they are very susceptible to the effects of climate variability (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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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, the most accurate data would be obtained by measuring soil salinity directly in the field (Rhoades and Ingvalson, 1971; Eldeiry et al., 2008). 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 (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. 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 years, with improvements in computing power, machine learning, and deep learning have greatly provided the growth of 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 showed that allmodels' accuracy was well in assessing soil salinity, with the IBK model proving the most effective. (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.

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 (Vermeulen and Van Niekerk, 2017; Shi et al., 2021). Currently, there are only four studies that have assessed soil salinity in the Mekong Delta (Hoa et al., 2019; Nguyen et al., 2021; Nguyen et al., 2023; Nguyen et al., 2018), 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; Wang et al., 2007; Su et al., 2020). 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.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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, several studies have been conducted in different regions of the world, focusing on the adaptive ability of farmers (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 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).

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 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 for 70%-90% of annual rainfall (Quang et al., 2024). Daily rainfall

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 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 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, 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).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

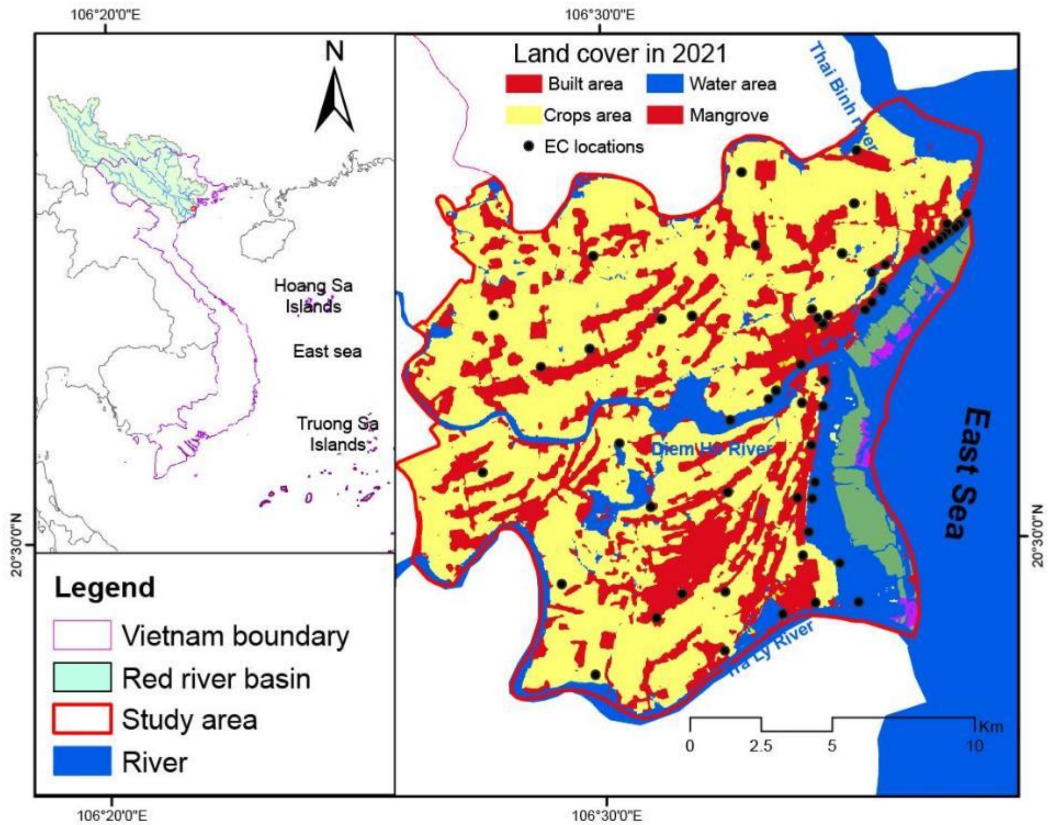


Figure 1: Geographical location of study area

3. Methodology

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

<u>Preparation of soil salinity samples and factors</u>

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

EC Measurements

Soil salinity samples were collected using soil drills using both zigzag and grid techniques. This technique is utilized frequently in small-scale sampling (Jia et al., 2024; Elshewy et al., 2024). The sampling depth depends on the soil

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 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. 62 samples were collected to ensure coverage of the entire field. 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. 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.). 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 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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 (Nguyen et al., 2021; Wang et al., 2021; Hoa et al., 2019).

The indices due to the vegetation reflect the health and growth of vegetation, and 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

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. The XGB model was developed using Python and the Sklearn library. Its accuracy depended on parameter adjustments made using the trial-and-error method.

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 interview area identification

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.

The second strand of our methodology, interviews with populations, was used to assess the adaptive capacity of farms in the study area. A Tân commune was selected for this assessment. A total of 87 households were interviewed. These households were randomly selected from A 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. 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.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

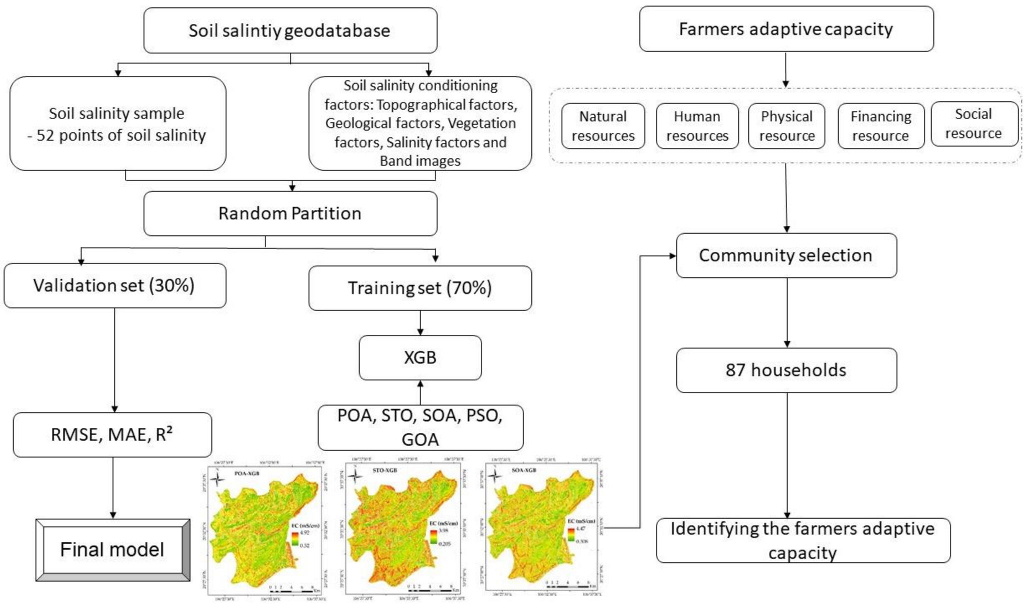


Figure 2: Methodology used for the farmer's adaptive capacity and soil salinity in this study

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). The result of the prediction is, therefore, 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 is called "boosting". 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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 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. 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 (Trojovský and Dehghani, 2022; Alamir et al., 2023).

3.3. Serval Optimization Algorithm (SOA)

SOA is one of the population-based optimization algorithms, and they exploit the searching power of agents in a population. This makes this algorithm powerful in solving optimization problems (Dehghani and Trojovský, 2022).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 to detailed coverage of the search space. The aim of this stage 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 (Sindi et al., 2024; Dehghani and Trojovský, 2022).

When attacking prey, wild cats jump during the chase to prevent the prey from escaping. 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 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 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).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 on the policy of the Siberian tiger when fighting with bears. 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, the STO algorithm is designed to explore the search space extensively and refine 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 optimization function 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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 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 make the model more complex and lead to poor performance.

RF was utilized to measure the suitable factors in this study. It assigns a value to each factor based on the relation within soil salinity samples and 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).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

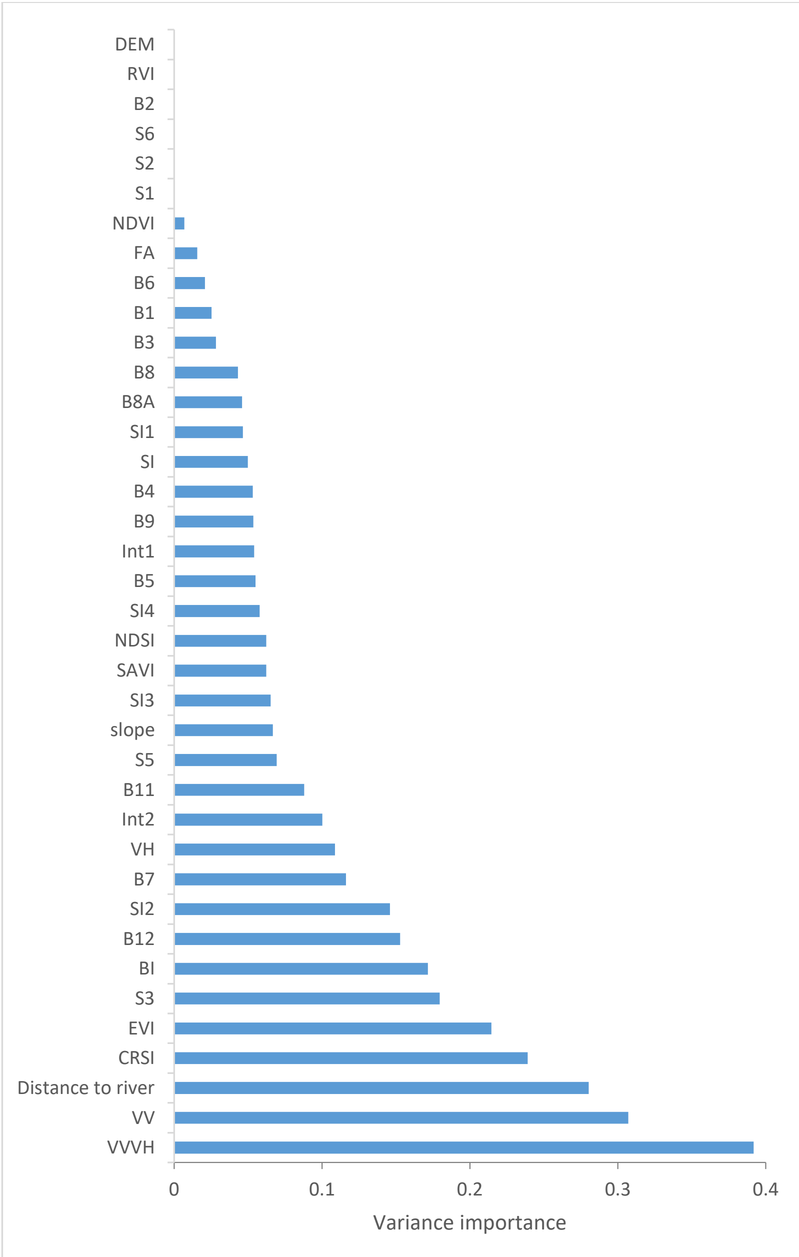
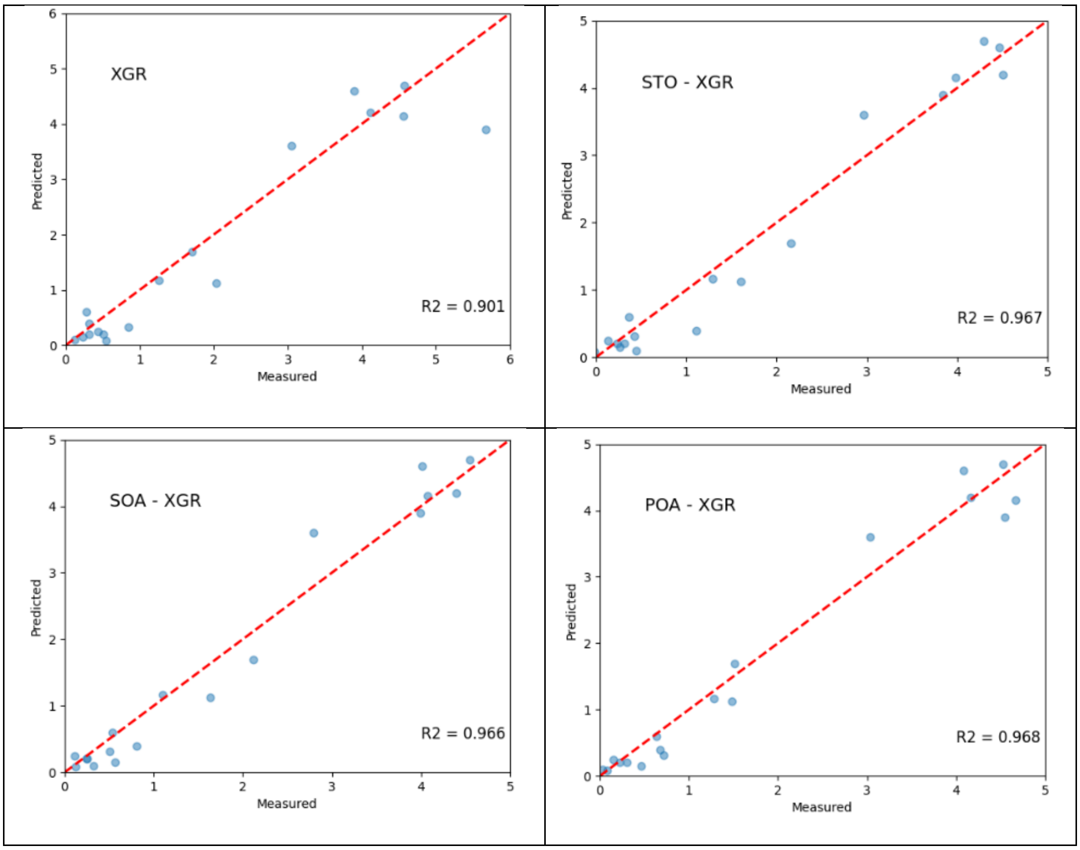


Figure 3: Variables important for soil salinity model using RF.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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).



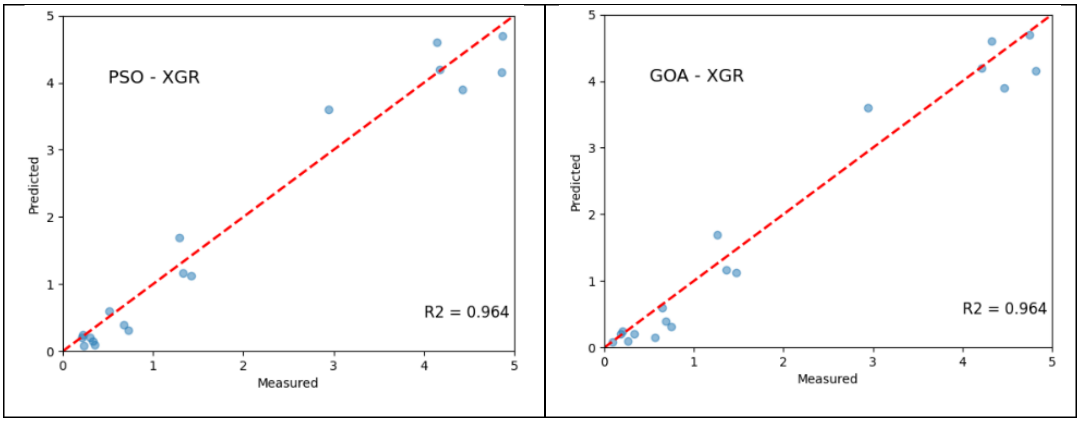


Figure 4: R^2 value for the testing dataset

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).

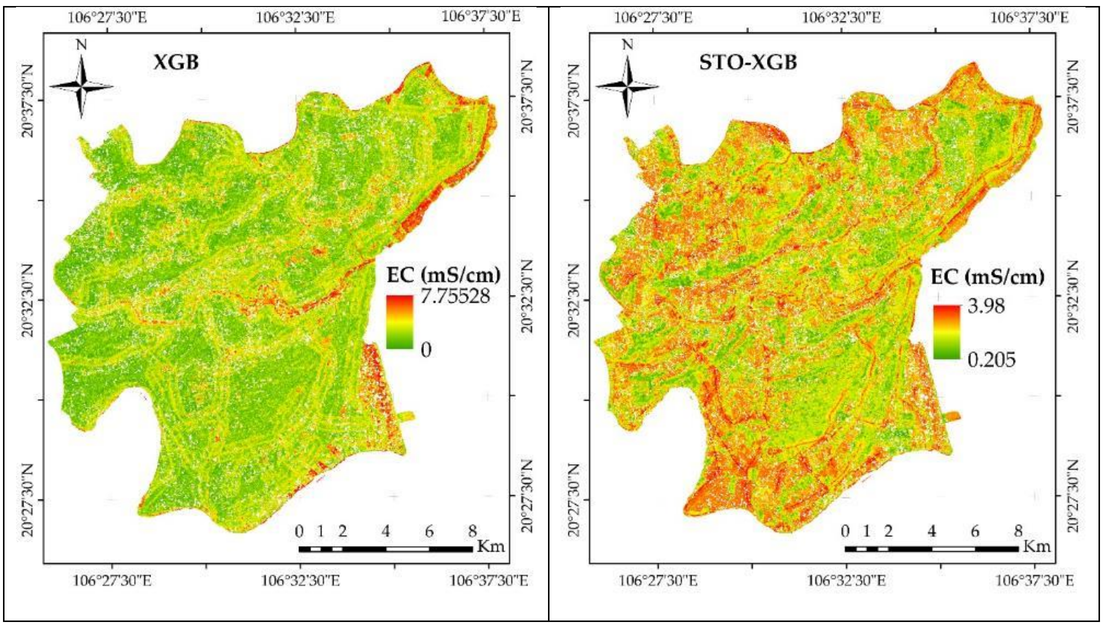
Table 1. Model performance and comparison.

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
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	0.9

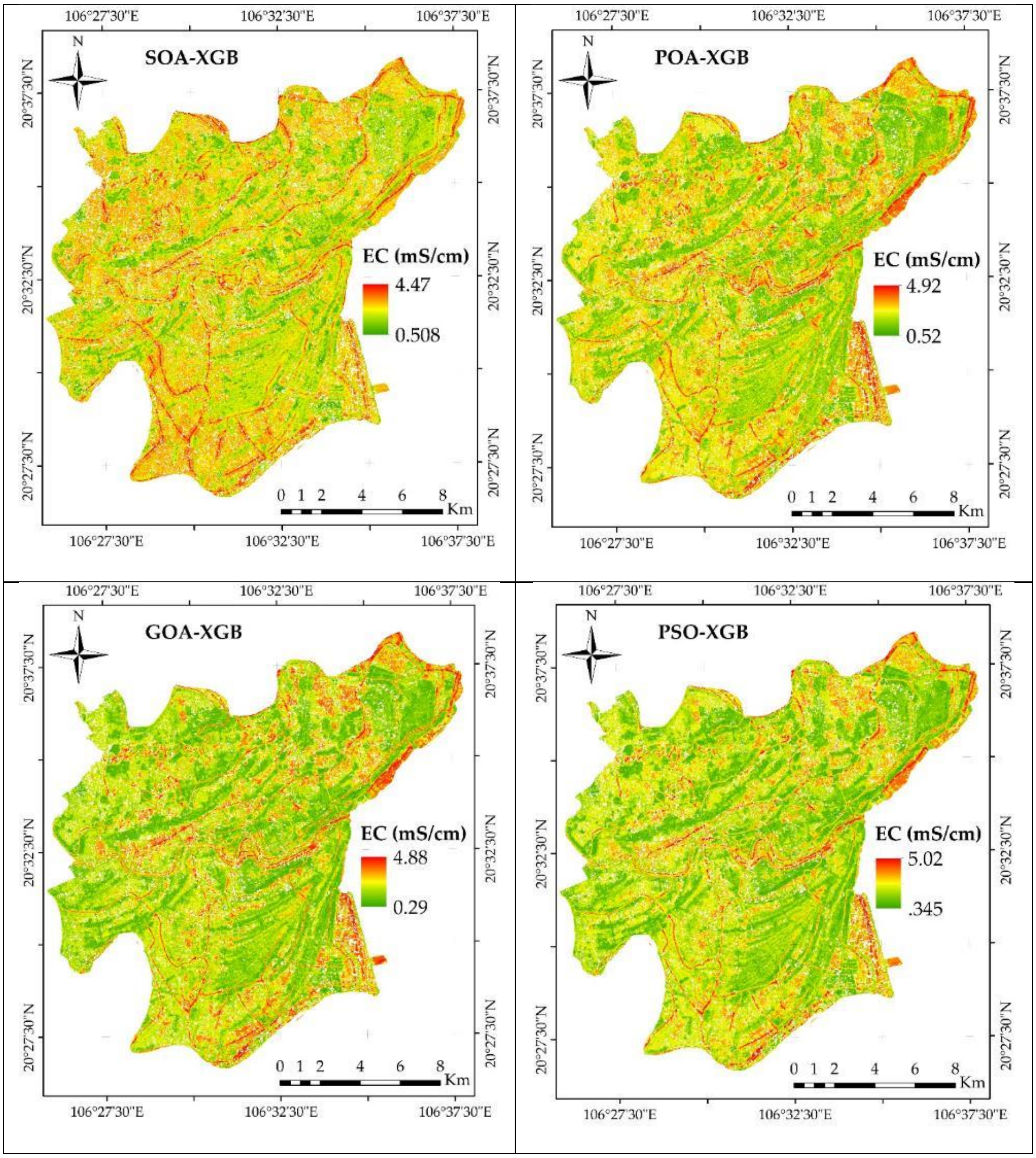
https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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

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).



https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025



375 Figure 5: Soil salinity mapping

376 **4.4. Farmers' Adaptive Capacity Assessment**

377 Soil salinity is a key challenge for deltas worldwide, particularly in deltas where population density is high and
378 socio-economic conditions are poor (Hoque et al., 2016). The Red River Delta is the most densely populated area in

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

Vietnam and one of the most densely populated deltas in Southeast Asia; therefore 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 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, 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 many measures to mitigate the increasingly serious saline intrusion, such as washing the salt after each crop (as instructed by local authorities) and changing the crop structure. 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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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

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

4.4.3. Physical Resources

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

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

4.4.4. Financial Resources

The interviews with 87 households demonstrated that 9% of the households interviewed were poor and near-poor. It can be seen that economic status greatly affects people's ability to adapt and recover from soil salinity. Poor and near-poor households frequently have more difficulty evaluating solutions to mitigate the impact of soil salinity on 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 (Nguyen et al., 2024; Hoang and Hai, 2024). 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.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 (Mo et al., 2019; Liu et al., 2022). 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 (Li et al., 2023; Alamir et al., 2023).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 (Trojovský et al., 2022; Al-Sarray et al., 2024). 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 (Rini et al., 2011; Juneja and Nagar, 2016). 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 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

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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.

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. 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).

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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 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,

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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., and Hamzehpour, N.: Assessment of soil salinity using explainable machine learning methods and Landsat 8 images, International Journal of Applied Earth Observation and Geoinformation, 130, 103879, https://doi.org/10.1016/j.jag.2024.103879, 2024.

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

Alamir, N., Kamel, S., Megahed, T. F., Hori, M., and Abdelkader, S. M.: Developing hybrid demand response technique for energy management in microgrid based on pelican optimization algorithm, Electric Power Systems Research, 214, 108905, 2023.

Asfaw, E., Suryabhagavan, K., and Argaw, M.: Soil salinity modeling and mapping using remote sensing and GIS: The case of Wonji sugar cane irrigation farm, Ethiopia, Journal of the Saudi Society of Agricultural Sciences, 17, 250-258, 2018.

Bandak, S., Movahedi-Naeini, S. A., Mehri, S., and Lotfata, A.: A longitudinal analysis of soil salinity changes using remotely sensed imageries, Scientific Reports, 14, 10383, 2024.

Bhuyan, M. I., Supit, I., Kumar, U., Mia, S., and Ludwig, F.: The significance of farmers' climate change and salinity perceptions for on-farm adaptation strategies in the south-central coast of Bangladesh, Journal of Agriculture and Food Research, 16, 101097, https://doi.org/10.1016/j.jafr.2024.101097, 2024.

Bui, D. T., Pradhan, B., Nampak, H., Bui, Q.-T., Tran, Q.-A., and Nguyen, Q.-P.: Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibilitgy modeling in a high-frequency tropical cyclone area using GIS, Journal of Hydrology, 540, 317-330, 2016.

Castelletti, A., Pianosi, F., Quach, X., and Soncini-Sessa, R.: Assessing water reservoirs management and development in Northern Vietnam, Hydrology and Earth System Sciences, 16, 189-199, 2012.

Cullu, M. A.: Estimation of the effect of soil salinity on crop yield using remote sensing and geographic information system, Turkish Journal of Agriculture and Forestry, 27, 23-28, 2003.

Dasgupta, S., Laplante, B., Murray, S., and Wheeler, D.: Climate change and the future impacts of storm-surge disasters in developing countries, Center for Global Development Working Paper, 2009.

Dehghani, M. and Trojovský, P.: Serval optimization algorithm: a new bio-inspired approach for solving optimization problems, Biomimetics, 7, 204, 2022.

Du, L., Tian, S., Zhao, N., Zhang, B., Mu, X., Tang, L., Zheng, X., and Li, Y.: Climate and topography regulate the spatial pattern of soil salinization and its effects on shrub community structure in Northwest China, Journal of Arid Land, 1-18, 2024.

Eldeiry, A. A., Garcia, L. A., and Reich, R. M.: Soil salinity sampling strategy using spatial modeling techniques, remote sensing, and field data, Journal of irrigation and drainage engineering, 134, 768-777, 2008.

Elshewy, M. A., Mohamed, M. H., and Refaat, M.: Developing a Soil Salinity Model from Landsat 8 Satellite Bands Based on Advanced Machine Learning Algorithms, Journal of the Indian Society of Remote Sensing, 52, 617-632, 2024.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

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

Fu, C., Gan, S., Yuan, X., Xiong, H., and Tian, A.: Determination of soil salt content using a probability neural network model based on particle swarm optimization in areas affected and non-affected by human activities, Remote Sensing, 10, 1387, 2018.

Gong, Z., He, H., Fan, D., Sheng, J., and Zhang, X.: Inversion of sea surface salinity in Bay of Bengal based on Catboost algorithm, Second International Conference on Geographic Information and Remote Sensing Technology (GIRST 2023), 297-304,

Hardie, M. and Doyle, R.: Measuring soil salinity, Plant salt tolerance: methods and protocols, 415-425, 2012.

He, Y., Yin, H., Chen, Y., Xiang, R., Zhang, Z., and Chen, H.: Soil Salinity Estimation based on Sentinel-1/2 Texture Features and Machine Learning, IEEE Sensors Journal, 2024.

Hoa, P. V., Giang, N. V., Binh, N. A., Hai, L. V. H., Pham, T.-D., Hasanlou, M., and Tien Bui, D.: 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, 128, 2019.

Hoang, L. P., Pot, M., Tran, D. D., Ho, L. H., and Park, E.: Adaptive capacity of high- and low dyke farmers to hydrological changes in the Vietnamese Mekong delta, Environmental Research, 224, 115423, https://doi.org/10.1016/j.envres.2023.115423, 2023.

Hoang, N. K. and Hai, D. T.: Assessing surface water salinity intrusion in the Mekong River Delta: a case study in Rach Gia, Vietnam, GeoJournal, 89, 171, 2024.

Hung, N. M. and Larson, M.: Coastline and river mouth evolution in the central part of the Red River Delta, in: Coastal disasters and climate change in Vietnam, Elsevier, 43-79, 2014.

Ingle, K. K. and Jatoth, R. K.: Non-linear Channel Equalization using Modified Grasshopper Optimization Algorithm, Applied Soft Computing, 153, 110091, https://doi.org/10.1016/j.asoc.2023.110091, 2024.

Jia, P., Zhang, J., Liang, Y., Zhang, S., Jia, K., and Zhao, X.: The inversion of arid-coastal cultivated soil salinity using explainable machine learning and Sentinel-2, Ecological Indicators, 166, 112364, 2024.

Jiang, H., Rusuli, Y., Amuti, T., and He, Q.: 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, 284-306, 2019.

Juneja, M. and Nagar, S.: Particle swarm optimization algorithm and its parameters: A review, 2016 International Conference on Control, Computing, Communication and Materials (ICCCCM), 1-5,

Kaplan, G., Gašparović, M., Alqasemi, A. S., Aldhaheri, A., Abuelgasim, A., and Ibrahim, M.: 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, https://doi.org/10.1016/j.pce.2023.103400, 2023.

Kennedy, J. and Eberhart, R.: Particle swarm optimization, Proceedings of ICNN'95-international conference on neural networks, 1942-1948,

Le Dang, H., Li, E., Nuberg, I., and Bruwer, J.: Understanding farmers' adaptation intention to climate change: A structural equation modelling study in the Mekong Delta, Vietnam, Environmental science & policy, 41, 11-22, 2014.

Li, S., Zhang, T., and Yu, J.: Photovoltaic array dynamic reconfiguration based on an improved pelican optimization algorithm, Electronics, 12, 3317, 2023.

Liu, W., Chen, Z., and Hu, Y.: XGBoost algorithm-based prediction of safety assessment for pipelines, International Journal of Pressure Vessels and Piping, 197, 104655, 2022.

Liu, Y., Han, X., Zhu, Y., Li, H., Qian, Y., Wang, K., and Ye, M.: Spatial mapping and driving factor Identification for salt-affected soils at continental scale using Machine learning methods, Journal of Hydrology, 639, 131589, 2024.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

Ma, G., Ding, J., Han, L., Zhang, Z., and Ran, S.: Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms, Regional Sustainability, 2, 177-188, 2021.

Mazumder, M. S. U. and Kabir, M. H.: Farmers' adaptations strategies towards soil salinity effects in agriculture: the interior coast of Bangladesh, Climate Policy, 22, 464-479, 2022.

Mirjalili, S. Z., Mirjalili, S., Saremi, S., Faris, H., and Aljarah, I.: Grasshopper optimization algorithm for multi-objective optimization problems, Applied Intelligence, 48, 805-820, 2018.

Mo, H., Sun, H., Liu, J., and Wei, S.: Developing window behavior models for residential buildings using XGBoost algorithm, Energy and Buildings, 205, 109564, 2019.

Moayedi, H., Nguyen, H., and Kok Foong, L.: Nonlinear evolutionary swarm intelligence of grasshopper optimization algorithm and gray wolf optimization for weight adjustment of neural network, Engineering with Computers, 37, 1265-1275, 2021.

Moayedi, H., Gör, M., Lyu, Z., and Bui, D. T.: Herding Behaviors of grasshopper and Harris hawk for hybridizing the neural network in predicting the soil compression coefficient, Measurement, 152, 107389, https://doi.org/10.1016/j.measurement.2019.107389, 2020.

Mukhamediev, R. I., Merembayev, T., Kuchin, Y., Malakhov, D., Zaitseva, E., Levashenko, V., Popova, Y., Symagulov, A., Sagatdinova, G., and Amirgaliyev, Y.: Soil Salinity Estimation for South Kazakhstan Based on SAR Sentinel-1 and Landsat-8, 9 OLI Data with Machine Learning Models, Remote Sensing, 15, 4269, 2023.

Nguyen, H. D.: Hybrid models based on deep learning neural network and optimization algorithms for the spatial prediction of tropical forest fire susceptibility in Nghe An province, Vietnam, Geocarto International, 37, 11281-11305, 2022.

Nguyen, H. D., Van, C. P., Nguyen, T. G., Dang, D. K., Pham, T. T. N., Nguyen, Q.-H., and Bui, Q.-T.: Soil salinity prediction using hybrid machine learning and remote sensing in Ben Tre province on Vietnam's Mekong River Delta, Environmental Science and Pollution Research, 1-18, 2023.

Nguyen, P. T., Koedsin, W., McNeil, D., and Van, T. P.: Remote sensing techniques to predict salinity intrusion: application for a data-poor area of the coastal Mekong Delta, Vietnam, International journal of remote sensing, 39, 6676-6691, 2018.

Nguyen, T. G., Tran, N. A., Vu, P. L., Nguyen, Q.-H., Nguyen, H. D., and Bui, Q.-T.: Salinity intrusion prediction using remote sensing and machine learning in data-limited regions: A case study in Vietnam's Mekong Delta, Geoderma Regional, 27, e00424, 2021.

Nguyen, V. H., Germer, J., and Asch, F.: Evaluating topsoil salinity via geophysical methods in rice production systems in the Vietnam Mekong Delta, Journal of Agronomy and Crop Science, 210, e12676, 2024.

Nhuan, M. T., Van Ngoi, C., Nghi, T., Tien, D. M., van Weering, T. C., and van den Bergh, G.: Sediment distribution and transport at the nearshore zone of the Red River delta, Northern Vietnam, Journal of Asian Earth Sciences, 29, 558-565, 2007.

Premkumar, N. and Santhosh, R.: Pelican optimization algorithm with blockchain for secure load balancing in fog computing, Multimedia Tools and Applications, 83, 53417-53439, 2024.

Quang, N. H., Viet, T. Q., Thang, H. N., and Hieu, N. T. D.: Long-term water level dynamics in the Red River basin in response to anthropogenic activities and climate change, Science of The Total Environment, 912, 168985, 2024.

Ramraj, S., Uzir, N., Sunil, R., and Banerjee, S.: Experimenting XGBoost algorithm for prediction and classification of different datasets, International Journal of Control Theory and Applications, 9, 651-662, 2016.

Rhoades, J. and Ingvalson, R.: Determining salinity in field soils with soil resistance measurements, Soil Science Society of America Journal, 35, 54-60, 1971.

https://doi.org/10.5194/egusphere-2025-1051
Preprint. Discussion started: 21 March 2025

Rini, D. P., Shamsuddin, S. M., and Yuhaniz, S. S.: Particle swarm optimization: technique, system and challenges, International journal of computer applications, 14, 19-26, 2011.

Ruidas, D., Chakrabortty, R., Islam, A. R. M. T., Saha, A., and Pal, S. C.: A novel hybrid of meta-optimization approach for flash flood-susceptibility assessment in a monsoon-dominated watershed, Eastern India, Environmental earth sciences, 81, 145, 2022.

Shi, H., Hellwich, O., Luo, G., Chen, C., He, H., Ochege, F. U., Van de Voorde, T., Kurban, A., and De Maeyer, P.: A global meta-analysis of soil salinity prediction integrating satellite remote sensing, soil sampling, and machine learning, IEEE Transactions on Geoscience and Remote Sensing, 60, 1-15, 2021.

Sindi, A. O. N., Si, P., and Li, Q.: Secure Task Offloading and Resource Allocation Strategies in Mobile Applications Using Probit Mish-Gated Recurrent Unit and an Enhanced-Searching-Based Serval Optimization Algorithm, Electronics, 13, 2462, 2024.

Song, Y., Gao, M., and Wang, J.: Inversion of salinization in multilayer soils and prediction of water demand for salt regulation in coastal region, Agricultural Water Management, 301, 108970, https://doi.org/10.1016/j.agwat.2024.108970, 2024.

Su, Y., Li, T., Cheng, S., and Wang, X.: Spatial distribution exploration and driving factor identification for soil salinisation based on geodetector models in coastal area, Ecological Engineering, 156, 105961, 2020.

Tan, J., Ding, J., Han, L., Ge, X., Wang, X., Wang, J., Wang, R., Qin, S., Zhang, Z., and Li, Y.: Exploring planetScope satellite capabilities for soil salinity estimation and mapping in arid regions oases, Remote Sensing, 15, 1066, 2023.

Thiam, H. I., Owusu, V., Villamor, G. B., Schuler, J., and Hathie, I.: Farmers' intention to adapt to soil salinity expansion in Fimela, Sine-Saloum area in Senegal: A structural equation modelling approach, Land Use Policy, 137, 106990, 2024.

Tran, V. N. and Kim, J.: Robust and efficient uncertainty quantification for extreme events that deviate significantly from the training dataset using polynomial chaos-kriging, Journal of Hydrology, 609, 127716, 2022.

Trojovský, P. and Dehghani, M.: Pelican optimization algorithm: A novel nature-inspired algorithm for engineering applications, Sensors, 22, 855, 2022.

Trojovský, P., Dehghani, M., and Hanuš, P.: Siberian tiger optimization: A new bio-inspired metaheuristic algorithm for solving engineering optimization problems, Ieee Access, 10, 132396-132431, 2022.

Vermeulen, D. and Van Niekerk, A.: Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates, Geoderma, 299, 1-12, 2017.

Vinh, V. D., Ouillon, S., Thanh, T. D., and Chu, L.: Impact of the Hoa Binh dam (Vietnam) on water and sediment budgets in the Red River basin and delta, Hydrology and Earth System Sciences, 18, 3987-4005, 2014.

Wang, H., Hsieh, Y. P., Harwell, M. A., and Huang, W.: Modeling soil salinity distribution along topographic gradients in tidal salt marshes in Atlantic and Gulf coastal regions, Ecological modelling, 201, 429-439, 2007.

Wang, H., Zhang, L., Zhao, J., Hu, X., and Ma, X.: Application of hyperspectral technology combined with bat algorithm-AdaBoost model in field soil nutrient prediction, Ieee Access, 10, 100286-100299, 2022.

Wang, J., Peng, J., Li, H., Yin, C., Liu, W., Wang, T., and Zhang, H.: Soil salinity mapping using machine learning algorithms with the Sentinel-2 MSI in arid areas, China, Remote Sensing, 13, 305, 2021.

Wang, J., Wang, X., Zhang, J., Shang, X., Chen, Y., Feng, Y., and Tian, B.: Soil Salinity Inversion in Yellow River Delta by Regularized Extreme Learning Machine Based on ICOA, Remote Sensing, 16, 1565, 2024.

Wu, W., Zucca, C., Muhaimeed, A. S., Al-Shafie, W. M., Fadhil Al-Quraishi, A. M., Nangia, V., Zhu, M., and Liu, G.: Soil salinity prediction and mapping by machine learning regression in C entral M esopotamia, I raq, Land degradation & development, 29, 4005-4014, 2018.

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

Xiao, S., Nurmemet, I., and Zhao, J.: 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, 451-469, 2024.

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

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

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