



Article

Soil Salinity Inversion in Yellow River Delta by Regularized Extreme Learning Machine Based on ICOA

Jiajie Wang¹, Xiaopeng Wang^{1,*}, Jiahua Zhang², Xiaodi Shang¹, Yuyi Chen¹, Yiping Feng¹ and Bingbing Tian¹

¹ College of Computer Science & Technology, Qingdao University, Qingdao 266071, China; wangjiajie@qdu.edu.cn (J.W.); shangxd@qdu.edu.cn (X.S.); chenyyi@qdu.edu.cn (Y.C.); fengyiping@qdu.edu.cn (Y.F.)

² Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; zhangjh@radi.ac.cn

* Correspondence: wxp@qdu.edu.cn; Tel.: +86-0532-8595-0208

Abstract: Soil salinization has seriously affected agricultural production and ecological balance in the Yellow River Delta region. Rapid and accurate monitoring of soil salinity has become an urgent need. Traditional machine learning models tend to fall into local optimal values during the learning process, which reduces their accuracy. This paper introduces Circle map to enhance the crayfish optimization algorithm (COA), which is then integrated with the regularized extreme learning machine (RELM) model, aiming to improve the accuracy of soil salinity content (SSC) inversion in the Yellow River Delta region. We employed Landsat5 TM remote sensing images and measured salinity data to develop spectral indices, such as the band index, salinity index, vegetation index, and comprehensive index, selecting the optimal modeling variable group through Pearson correlation analysis and variable projection importance analysis. The back propagation neural network (BPNN), RELM, and improved crayfish optimization algorithm–regularized extreme learning machine (ICOA-RELM) models were constructed using measured data and selected variable groups for SSC inversion. The results indicate that the ICOA-RELM model enhances the R^2 value by an average of about 0.1 compared to other models, particularly those using groups of variables filtered by variable projection importance analysis as input variables, which showed the best inversion effect (test set R^2 value of 0.75, MAE of 0.198, RMSE of 0.249). The SSC inversion results indicate a higher salinization degree in the coastal regions of the Yellow River Delta and a lower degree in the inland areas, with moderate saline soil and severe saline soil comprising 48.69% of the total area. These results are consistent with the actual sampling results, which verify the practicability of the model. This paper's methods and findings introduce an innovative and practical tool for monitoring and managing salinized soils in the Yellow River Delta, offering significant theoretical and practical benefits.

Keywords: soil salinity inversion; Yellow River Delta; regularized extreme learning machine; improved crayfish optimization algorithm; chaotic mapping



Citation: Wang, J.; Wang, X.; Zhang, J.; Shang, X.; Chen, Y.; Feng, Y.; Tian, B. Soil Salinity Inversion in Yellow River Delta by Regularized Extreme Learning Machine Based on ICOA. *Remote Sens.* **2024**, *16*, 1565. <https://doi.org/10.3390/rs16091565>

Academic Editors: Diego González-Aguilera and Pablo Rodríguez-González

Received: 27 March 2024

Revised: 22 April 2024

Accepted: 25 April 2024

Published: 28 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Salts accumulate in the soil due to groundwater-associated salinity, non-groundwater-associated salinity, and irrigation-induced salinity, resulting in the occurrence of soil salinization. This phenomenon can have significant impacts on agricultural production, environmental health, and regional economies [1]. In China, the total area of saline soils accounts for about 4.88% of the country's total usable land area. Saline soils are primarily found in arid, semi-arid, and semi-humid regions [2]. The Yellow River Delta, situated in Dongying City, Shandong Province, represents China's best-preserved, largest, and most recent wetland ecosystem within the warm-temperature zone. The area experiences low precipitation and high evaporation, leading to the accumulation of salts on the soil surface. Its proximity to the Bohai Sea facilitates seawater intrusion, exacerbated by its low-lying

topography and inadequate drainage, resulting in high soil salinity levels. Over time, due to the combined effects of climate conditions and seawater, as well as China's limited knowledge of saline land management and outdated technology, more than half of the land in the Yellow River Delta has become salinized soil, resulting in the formation of the current semi-humid saline area [3]. Accurate monitoring of salinity in the soil and improvement of soil salinization has become an urgent need.

Before the adoption of remote sensing, soil salinity was monitored through time-consuming and labor-intensive traditional methods like gravimetric analysis and electrical conductivity measurements, limiting the scope of salinity assessment across large areas [4,5]. Remote sensing technology offers the advantage of wide-range monitoring with high spatial and temporal resolution, providing immediate and cost-effective data, which play a crucial role in monitoring soil salinity content [6]. Soil salinity can be effectively characterized by band indices [7]. Subsequently, researchers have incorporated the other spectral indices as modeling variables in salinity inversion studies, yielding favorable outcomes [8–10]. Salt inversion models can be broadly classified into two categories: linear fitting models and nonlinear models. Linear models mainly include multivariate spline autoregressive model (MSA) [11], multivariate linear regression model (MLR) [12], exponential fitting model (EF) [13], partial least squares regression model (PLSR) [14], and so on. Nonlinear models primarily encompass the BPNN model [15], support vector machine model (SVM) [16], random forest model (RF) [17], extreme learning machine model (ELM) [18], and other machine learning models.

In scenarios where machine learning models are employed for regression analysis, researchers often encounter a significant challenge: the models may contain numerous parameters, and the initial settings of these parameters critically influence the model's final performance. Traditionally, these parameters are initialized randomly; however, this method has a clear drawback: it may lead the model to converge to local minima during training, preventing it from reaching the global optimum, thereby limiting its predictive capability [19]. The use of intelligent optimization algorithms has proven beneficial in effectively addressing this challenge. Genetic algorithms, simulated annealing, and particle swarm optimization are capable of performing global searches within the parameter space. During the model's iterative training process, intelligent optimization algorithms continuously adjust its fitness value, thereby evaluating and enhancing selected parameters, reducing prediction errors, and enhancing accuracy [20]. The integration of intelligent optimization algorithms with traditional machine learning models has produced excellent outcomes in regression prediction research. This approach enhances both the predictive accuracy and the generalizability of the models across various domains. The integrated models, such as the particle swarm optimization–extreme learning machine (PSO-ELM) [21], bat optimization algorithm–extreme learning machine (BOA-ELM) [22], the estimation distribution algorithm–extreme learning machine (EDA-ELM) [23], genetic algorithm–support vector machine (GA-SVM) [24], and whale optimization algorithm–random forest (WOA-RF) [25], incorporate intelligent optimization algorithms into machine learning models with the aim of optimizing parameter selection and enhancing model performance. These methods have demonstrated efficacy in various fields, including financial market forecasting, bioinformatics, environmental monitoring, and energy consumption prediction. This integrated model is extensively employed in SSC inversion studies. Zhao Wenju and his team enhanced the BPNN model using PSO, mind evolutionary algorithm (MEA), and GA, selecting the western corridor of China's Taolai River Basin as their study area. It was demonstrated that the accuracies of the PSO-BPNN, MEA-BPNN, and GA-BPNN models surpassed those of the standalone BPNN model, with the GA-BPNN model emerging as the most effective salinity inversion model, achieving an R^2 of 0.6659 [26]. Yang Lianbing and his colleagues employed the GA and the Bayesian optimization algorithm (BOA) to optimize subsets of inversion parameters and RF model parameters, respectively, subsequently constructing GA-RFR and BOA-RFR salt inversion models. Results indicated that the BOA-RFR model achieved the highest predictive accuracy [27]. To enhance the

predictive performance of the SVM, Xiaohong Zhou and his colleagues employed PSO, gray wolf optimization (GWO), and differential evolution algorithms (DE) for SVM parameter optimization, resulting in the development of PSO-SVM, GWO-SVM, and DE-SVM models for the Ebinur Lake Wetland National Nature Reserve (ELWNNR) area's SSC inversion. Ultimately, the DE-SVM model demonstrated superior performance, evidenced by an R^2 value of 0.56. Utilizing this model, the authors mapped soil salinity in the ELWNNR area for August 2018 and May 2019 [28]. The primary strength of this methodology lies in its systematic exploration of the parameter space, eschewing dependence on randomness or unidirectional gradient descent. During the iterative process, the algorithm enables the model to effectively circumvent local optima and progressively advance towards a solution that more closely approximates the global optimum. It has been proven that integrating intelligent optimization algorithms into the training process of machine learning models can significantly improve the performance of the models on regression tasks, both in terms of prediction accuracy and generalization ability of the models [29].

The crayfish optimization algorithm (COA) [30], proposed in 2023, simulates the crayfish's heat avoidance, competition, and foraging behaviors in varying environmental temperatures, and has the ability of fast searching speed, strong searching ability, and the ability to effectively balance the global search and local search. The COA algorithm exhibits sluggish convergence toward the optimal solution during the search phase. In this paper, the COA is improved by using the chaotic population initialization method to improve the search capability of the model. Furthermore, the improved COA (ICOA) is integrated with a machine learning model to analyze soil salinity information in the Yellow River Delta region. This integration aims to mitigate the influence of random initialization parameters on the performance of the salt inversion model.

This paper is centered on the Yellow River Delta region as the study area. Twenty-nine spectral indices across four categories (band index, salinity index, vegetation index, and composite index) were extracted from Landsat5 TM image data. Two optimal sets of input variables were determined using two different variable screening methods. The SSC inversion model was constructed by using BPNN, RELM, and ICOA-RELM. Comparative analysis was performed to evaluate the performance of different combinations of modeling variables and models. The most accurate and stable model was selected to create a spatial distribution map of soil salinity in the Yellow River Delta region.

2. Materials and Methods

The content is organized into two primary sections, namely the Materials and Methods sections. The Materials section mainly includes the basic overview of the study area, the acquisition and analysis of experimental data, and the processing of remote data. The Methods section is a critical component of the paper that describes in detail the intelligent optimization algorithm used and the improvement of the optimization algorithm. Finally, it describes the technical route of the research in this paper.

2.1. Study Area

The Yellow River Delta, situated in Dongying City in Shandong Province, is located between the Bohai Sea and Laizhou Bay, with its coordinates ranging from 118°9' to 119°18'E and 37°16' to 38°9'N, as shown in Figure 1. The Yellow River Delta covers an area of approximately 5400 square kilometers. The region experiences an average annual temperature of 12.3–12.8 °C, with an average of 2590–2830 h of sunshine per year. The average annual precipitation ranges from 542.3–842 mm, with the majority occurring during the summer season. Additionally, the average evapotranspiration in the area is 750–2400 mm. The region falls within the warm-temperate zone and is characterized by a semi-moist continental monsoon climate [31]. The vegetation in the area can be categorized into two main types: artificial forests consisting of acacia, poplar, cotton, and other cultivated plants, and natural vegetation including reeds, tamarisks, winged alkali fluff, swede, and white fescue. The region is characterized by its vastness and topographic

complexity. However, several factors contribute to the accumulation of salts in the soil, including poor runoff drainage, low topography, high evaporation rates compared to rainfall, and recurrent storm surges. These factors have a significant impact on the region's ecology and agricultural productivity [32].

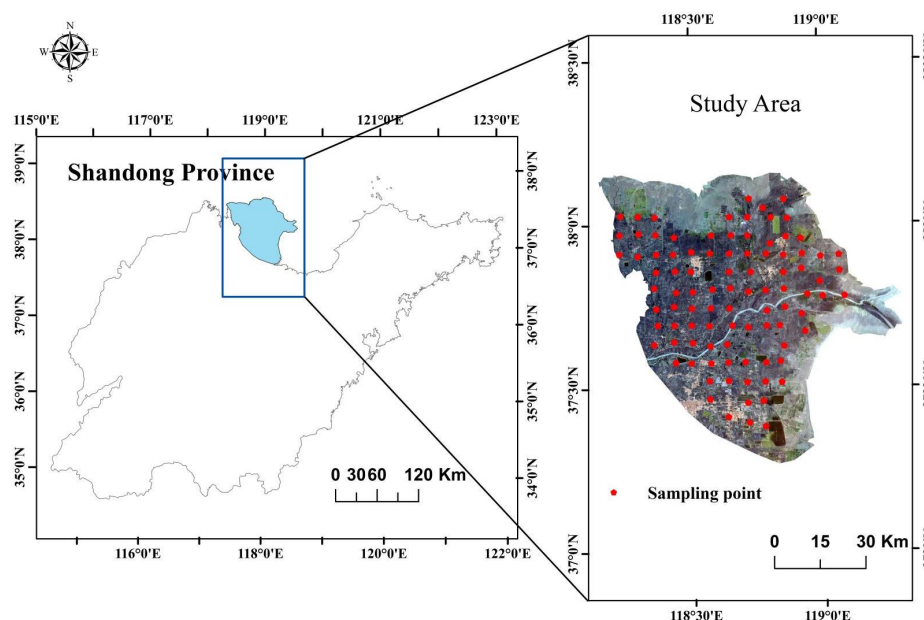


Figure 1. Study area and sampling point location distribution map.

2.2. Test Data Acquisition and Preprocessing

2.2.1. Soil Sample Analysis

The soil salinity data used in this paper were obtained from the National Science and Technology Resources Shared Service Platform–National Earth System Science Data Center (<http://www.geodata.cn>) (accessed on 1 March 2023). Soil samples were collected from the field in October 2003. A total of 94 sampling points were evenly distributed throughout the study area, with an average spacing of 6 km and sampling depths ranging from 30 to 40 cm. The locations of the sampling points can be seen in Figure 1. At the time of sample collection, the sample number, longitude, latitude, feature type, groundwater level, and collection time were recorded. The data were processed and analyzed in the laboratory to obtain the pH, organic matter content, ion concentrations, and SSC. The data characteristics of the sampling points are shown in Table 1. The maximum value of SSC is 2.036%, the minimum value is 0.044%, the average value is 0.547%, the standard deviation is 0.463%, and the coefficient of variation is 0.846. (The coefficient of variation reflects the dispersion of the sampling point values, with a coefficient of variation less than 0.1 indicating weak variability, between 0.1 and 1 indicating moderate variability, and greater than 1 indicating strong variability.) In this paper, the coefficient of variation falls within the range of moderate variability.

Table 1. Mathematical statistics of soil salt content.

Dataset	Sample Size	Min (%)	Max (%)	Avg (%)	Standard Deviation (%)	Coefficient of Variation
SSC	94	0.044	2.036	0.547	0.463	0.846

2.2.2. Remote Sensing Data Acquisition and Preprocessing

In this paper, we selected the Landsat5 TM remote sensing image captured on 26 October 2003, which corresponds to the sampling time of the soil salinity measurement data. The image has a spatial resolution of 30 m and was acquired from the geospatial data cloud (<https://www.gscloud.cn/>) (accessed on 2 March 2023). The image was preprocessed

using Envi 5.3.1.0 software, which included radiometric calibration, atmospheric correction, and cropping. This process resulted in obtaining corrected images of the study area. Finally, we extracted the reflectance data corresponding to the sampling points from the corrected images using ArcGIS 10.8 software in order to construct each spectral index. The Landsat5 TM sensor parameters for each band are shown in Table 2 (Landsat5 was the fifth satellite in the U.S. Landsat series, launched on March 1, 1984, from Vandenberg Air Force Base, California). The preprocessing process of the Landsat5 TM remote sensing data for Yellow River Delta is shown in Figure 2.

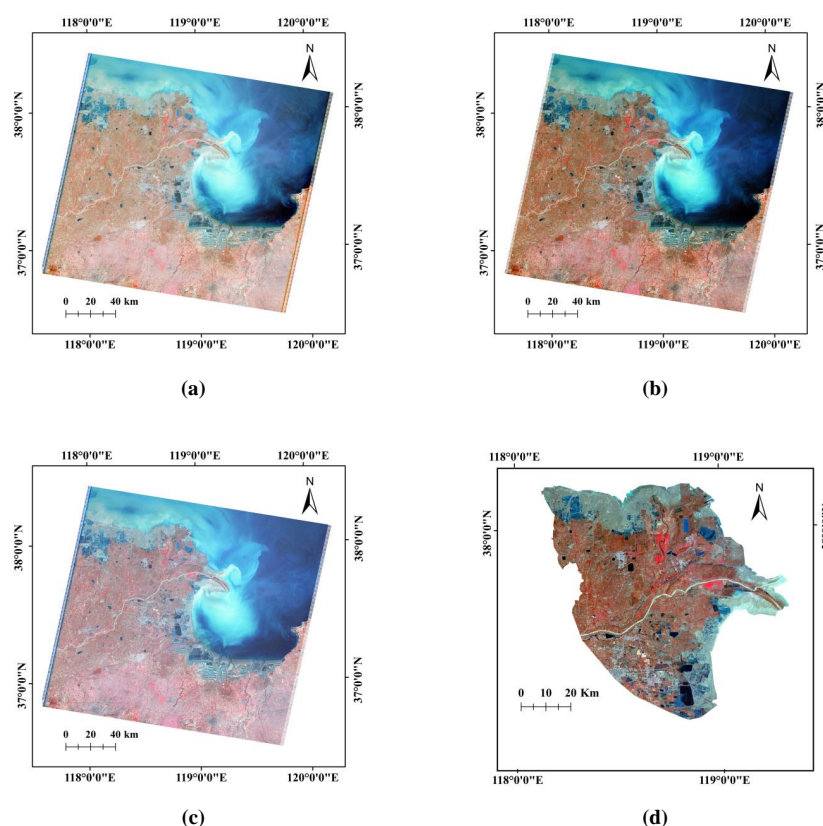


Figure 2. Preprocessing of the Landsat5 TM remote sensing image for Yellow River Delta. All images above are false color composite images. (a) The original remote sensing image; (b) the remote sensing image after radiometric calibration; (c) the remote sensing image after atmospheric correction; (d) the remote sensing image after data cropping.

Table 2. Landsat5 TM sensor parameters.

Band	Band Name	Spectrum Range (μm)	Resolution (m)
Band 1	BLUE	0.45–0.52	30
Band 2	GREEN	0.52–0.60	30
Band 3	RED	0.63–0.69	30
Band 4	NIR	0.76–0.90	30
Band 5	SWIR1	1.55–1.75	30
Band 6	LWIR	10.40–12.50	120
Band 7	SWIR2	2.08–2.35	30

2.2.3. Construction of Spectral Indices

Building on prior research in the field of SSC inversion, we screened spectral indices that are suitable for SSC inversion. The effectiveness of these indices has been validated in multiple studies. Each index and its corresponding literature are detailed in Table 3.

The band reflectance data of the Landsat image are extracted at each sampling point based on their respective locations. A total of twenty-nine spectral indices are calculated for salinity inversion purposes. The indices are classified into four groups as follows. (1) Band index group: BLUE, GREEN, RED, NIR, SWIR1, SWIR2; (2) Salinity index group: SI1 (salinity index 1), SI2 (salinity index 2), SI3 (salinity index 3), SI4 (salinity index 4), SI5 (salinity index 5), SI7 (salinity index 7), SI8 (salinity index 8), SI9 (salinity index 9), SIT (salinity index T), NDSI (normalized differential salinity index); (3) Vegetation index group: MSAVI (modified soil adjusted vegetation index), ALBEDO, NDVI (normalized difference vegetation index), ENDVI (extended normalized difference vegetation index), ERVI (extended ratio vegetation index), EDVI (extended difference vegetation index), NDWI (normalized difference water index), GRVI (green band ratio vegetation index); (4) Composite index group: SDI (salinization detection index), SRSI (remote sensing index of salinization), CORSI (combined spectral response index), EEVI (extended enhanced vegetation index), SIMSAVI (salinity index–modified soil adjusted vegetation index). The formula for each indicator is shown in Table 3.

Table 3. Remote sensing spectral indices.

Category	Abbreviation	Formula	Reference
Band indices	BLUE/GREEN/RED/ NIR/SWIR1/ SWIR2	—	—
Salinity indices	SI1	$\sqrt{\text{Blue} \times \text{Red}}$	[33]
	SI2	$\sqrt{\text{Green}^2 + \text{Red}^2 + \text{NIR}^2}$	[34]
	SI3	$\sqrt{\text{Green}^2 + \text{Red}^2}$	[34]
	SI4	SWIR1 / NIR	[34]
	SI5	$(\text{RED} - \text{SWIR1}) / (\text{RED} + \text{SWIR1})$	[35]
	SI7	RED × NIR / GREEN	[7]
	SI8	SWIR1 – SWIR2	[36]
	SI9	$(\text{SWIR1} \times \text{SWIR2} - \text{SWIR2} \times \text{SWIR2}) / \text{SWIR1}$	[36]
	SIT	RED / NIR × 100	[37]
	NDSI	$(\text{RED} - \text{NIR}) / (\text{RED} + \text{NIR})$	[37]
Vegetation indices	MSAVI	$\frac{2 \times \text{NIR} + 1 - (\sqrt{(2 \times \text{NIR} + 1)^2 - 8 \times (\text{NIR} - \text{Red})})}{2}$	[38]
	ALBEDO	$0.356 \times \text{BLUE} + 0.13 \times \text{RED} + 0.373 \times \text{NIR} + 0.085 \times \text{SWIR1} + 0.072 \times \text{SWIR2} - 0.0018$	[39]
	NDVI	$(\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED})$	[16]
	ENDVI	$(\text{NIR} + \text{SWIR2} - \text{RED}) / (\text{NIR} + \text{SWIR2} + \text{RED})$	[16]
	ERVVI	$(\text{NIR} + \text{SWIR2}) / \text{GREEN}$	[16]
	EDVI	$\text{NIR} + \text{SWIR1} - \text{RED}$	[16]
	NDWI	$(\text{GREEN} - \text{NIR}) / (\text{GREEN} + \text{NIR})$	[40]
	GRVI	NIR / GREEN	[41]
Composite indices	SDI	$\sqrt{(\text{NDVI} - 1)^2 + \text{SI1GYH}^2}$	[38]
	SRSI	$\sqrt{(\text{NDVI} - 1)^2 + \text{SI1}^2}$	[42]
	CORSI	$(\text{GREEN} + \text{BLUE}) / (\text{RED} + \text{NIR}) \times \text{NDVI}$	[43]
	EEVI	$(2.5 \times \text{EDVI}) / (\text{NIR} + \text{SWIR1} + 6 \times \text{RED} - 7.5 \times \text{BLUE} + 1)$	[16]
	SIMSAVI	$\sqrt{(\text{MSAVI} - 1)^2 + \text{BLUE} \times \text{RED}}$	[44]

2.3. Crayfish Optimization Algorithm and Its Improvement

To enhance the performance of the machine learning model for soil salinity inversion, this paper introduces the COA for optimizing the parameters of the RELM model. The COA is affected by different population initialization methods when searching for the optimal solution. By employing chaotic mapping to initialize the crayfish population, we have enhanced the algorithm's convergence speed and mitigated the risk of converging to local optimal solutions.

2.3.1. Crayfish Optimization Algorithm

The COA [30] is a recently proposed swarm intelligence algorithm. The algorithm aims to find the optimal solution to the problem by simulating the heat avoidance, competitive, and foraging behaviors of crayfish. It consists of two main stages: exploration and exploitation. Both stages' behaviors are influenced by the environmental temperature. The algorithm utilizes X to represent the initial population position, where X_{ij} represents the position of crayfish i in dimension j . The value of X_{ij} is calculated using the following equation:

$$x_{i,j} = lb_j + (ub_j - lb_j) \times rand \quad (1)$$

where lb_j denotes the lower bound of the j th dimension, ub_j denotes the upper bound of the j th dimension, and $rand$ is a random number.

Crayfish thrive in environments with temperatures ranging from 15 to 30 °C. When the temperature variable "temp" exceeds 30 °C, crayfish seek refuge in caves. In situations where the number of available caves is limited, cave scrambling events may transpire. To represent the absence of a cave scrambling event, $rand < 0.5$ is employed, and in such cases, the following formula is utilized to indicate the crayfish's entrance into the cave.

$$X_{i,j}^{t+1} = X_{i,j}^t + C_2 \times rand \times (X_{shade} - X_{i,j}^t) \quad (2)$$

where t denotes the current iteration number, C_2 is the decreasing curve, and X_{shade} denotes the location of the cave.

If the temperature (temp) exceeds 30 and the random variable (rand) is greater than or equal to 0.5, crayfish engage in competition for burrows. This behavior is represented by the following equation:

$$X_{i,j}^{t+1} = X_{i,j}^t - X_{z,j}^t + X_{shade} \quad (3)$$

where z is a random individual of the crayfish; $z = \text{round}(\text{rand} \times (N - 1)) + 1$.

Crayfish initiate feeding when the temperature variable "temp" is less than or equal to 30. Due to their limited body size, crayfish exhibit two distinct feeding behaviors. When the food is excessively large, crayfish employ their claws to shred the food into manageable pieces before feeding, utilizing their second and third walking feet in an alternating manner. Conversely, when the food size is suitable, crayfish engage in direct feeding. The foraging behaviors for oversized and normal-sized food are represented by the following equations:

$$X_{i,j}^{t+1} = X_{ij}^t + X_{food} \times p \times (\cos(2 \times \pi \times rand) - \sin(2 \times \pi \times rand)) \quad (4)$$

$$X_{i,j}^{t+1} = (X_{i,j}^t - X_{food}) \times p + p \times rand \times X_{i,j}^t \quad (5)$$

where X_{food} represents the location of the food, and p represents a mathematical model of crayfish intake. Sine and cosine functions are employed to depict the alternating feeding behavior of crayfish.

The crayfish, as it enters the cave and consumes the food, symbolizes the optimal solution during various stages of the algorithm. By continuously updating the position of the crayfish, it remains in close proximity to the target variable, thus achieving the optimization function of the algorithm. The pseudocode for COA is shown in Table 4.

Table 4. Crayfish optimization algorithm pseudo-code.

Crayfish Optimization Algorithm Pseudo-Code
Initialization iterations T, population N, dimension dim
Randomly generate an initial population
Calculate the fitness value of the population to get X_G, X_L
While $t < T$
Defining temperature temp
End
If temp > A30
Define cave X_{shade}
If rand < 0.5
Crayfish conducts the summer resort stage according to Equation (2)
Else
Crayfish compete for caves through Equation (3)
End
Else
Define the food intake p and food size Q
If $Q > 2$
Crayfish shreds food
Crayfish foraging according to Equation (4)
Else
Crayfish foraging according to Equation (5)
End
End
Update fitness values, X_G, X_L
$t = t + 1$
End

2.3.2. The Improved Crayfish Optimization Algorithm

The COA population initialization method limits the speed and directionality with which the optimal solution is found, thereby affecting the overall performance of the algorithm. To enhance its global search capability, chaotic mapping is introduced to improve the population initialization method. Chaotic mapping is utilized for sequences that exhibit characteristics of ergodicity, randomness, and orbital instability. Commonly used chaotic mapping functions include the Logistic map, Circle map, Sin map, Singer map, and Tent map. Among them, the Circle map is known for its stability and wider range of chaotic values. In this paper, the Circle map is employed to initialize the population crayfish, and its formula is as follows:

$$x_{n+1} = \text{mod} \left(x_n + 0.2 - \frac{0.5}{2\pi} \sin(2\pi x_n), 1 \right) \quad (6)$$

where n is the dimension of the solution.

2.4. Model Construction

ELM is a machine learning method for training single-layer feedforward neural networks, used for improving the issue of slow training speed in the BP algorithm and demonstrating good generalization performance. However, ELM, based on the principle of empirical risk minimization does not consider the impact of noise and outliers on model performance, making the model prone to overfitting. RELM has made improvements upon ELM by introducing a regularization parameter to balance between empirical risk and structural risk, thereby enhancing model performance and improving prediction accuracy [45]. A soil salinity prediction model for the research area based on RELM is established in this paper, using MATLAB 2022 software, with three sets of different variables as input data and measured soil salinity data as output data.

The architecture diagram of the ICOA-RELM model used in this paper is shown in Figure 3.

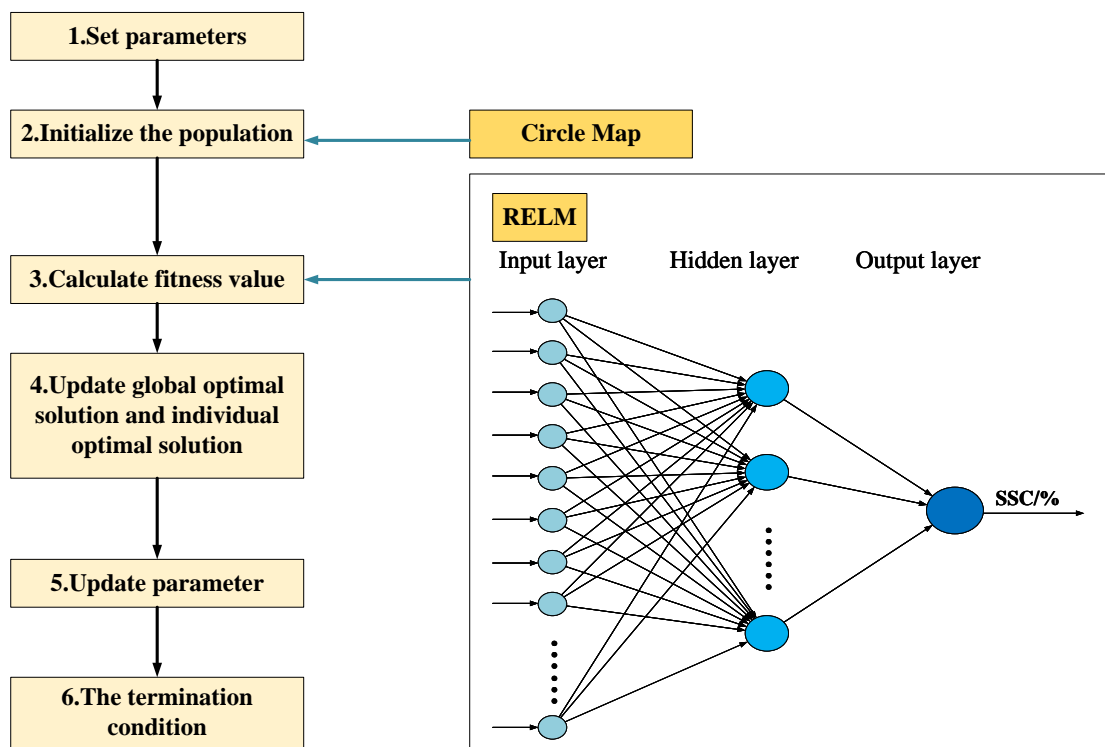


Figure 3. ICOA-RELM model architecture diagram.

The process begins with the initialization of parameters for the COA and the RELM, which include the number of crayfish populations, limits of parameter values, maximum iterations, input and hidden nodes, and the regularization parameters for RELM. Next, crayfish populations are initialized using the Circle map. The fitness value for each population is then calculated by the RELM model. Subsequently, the model updates global and individual optimal solutions according to fitness values, and populations adjust their positions accordingly. Parameters are updated iteratively, and the algorithm cycles through steps 3 to 5 until the predetermined number of iterations is completed. In the end, the optimal solution of the regression algorithm is represented by the parameters of the population that achieves the best fitness.

2.5. Accuracy Evaluation

We input the feature variables into the pre-trained model to generate predictions for SSC in the test set samples. The performance of each model was evaluated using three metrics: the coefficient of determination (R^2), the root mean square error (RMSE), and the mean absolute error (MAE). Models with a high R^2 , low RMSE, and small MAE exhibit better performance. This demonstrates that the model exhibits strong predictive ability and high stability. The formulas for each evaluation indicator are provided below:

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

where n represents the number of samples, y_i denotes the measured value of SSC, \hat{y}_i refers to the predicted value of SSC, and \bar{y}_i refers to the average value of the measured value of SSC.

2.6. Flowchart

Based on remote sensing image data and measured soil salinity data, BP and RELM models were established based on the full variable group and two groups of preferred variables. Circle chaotic mapping was introduced to optimize the initialization mode of crayfish population, and the ICOA was used to optimize the RELM model. The accuracy of all models was compared, and the optimal model was selected to establish the salt inversion image of the study area. The technical methods adopted in this paper are shown in Figure 4.

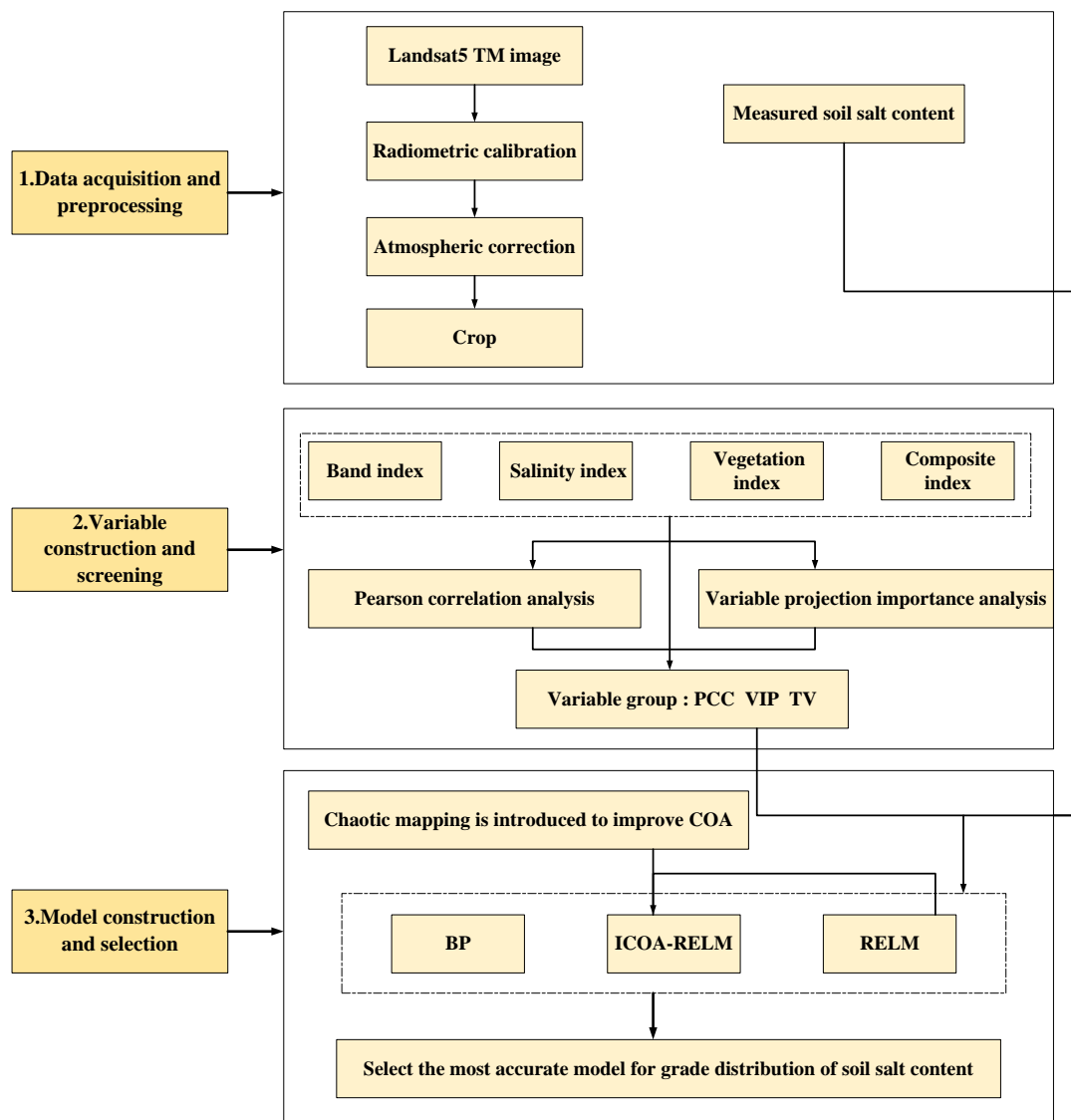


Figure 4. The working flowchart of this paper.

3. Results and Analysis

3.1. Statistical Analysis of Soil Salt Content Characteristics

According to Feng Xueli's study on soil salinization monitoring in the irrigation domain of Jiefangzha, Hetao Irrigation District, Inner Mongolia [46], the level of soil salinization was classified into five classes: non-saline, slight saline, moderate saline, severe saline, and extreme saline. The distribution of various salinity classes across the 94 sets of

measured data is presented in Table 5. Non-saline soil samples accounted for the largest percentage (38.3%) among all samples. Sample points in both slight saline and severe saline soil accounted for 21.28% of all samples. Moderate saline soil samples accounted for 19.14% all samples. No extreme saline soil samples were found in any of the samples in this paper.

Table 5. Descriptive statistics of soil salinity.

SSC (%)	Sample Number	Percent (%)
Non-saline (0–0.3)	36	38.3
Slight saline (0.3–0.5)	20	21.28
Moderate saline (0.5–1)	18	19.14
Severe saline (1–2.2)	20	21.28
Extreme saline (>2.2)	0	0
Total	94	100

3.2. Filtering of Input Variables

As the calculation of various spectral indices occurs within the basic band of the image, these indices are often significantly correlated. Using correlated variables to train a machine learning model often leads to overfitting. Consequently, before model training, constructed variables must be analyzed and screened to mitigate overfitting, simplify the model, and enhance its efficiency. In this section, two analysis methods are employed to screen the constructed variables.

Pearson correlation analysis is a statistical method used to measure the strength and direction of a linear relationship between two continuous variables, and it can help to understand the degree of association between the variables, which is useful for feature selection, variable screening, and understanding patterns and associations in the data. The variable importance score is calculated by considering both the predictive performance of the model and the contribution of the independent variables. Generally, a higher score signifies that the associated independent variable contributes more significantly to the model. This method aids in identifying the independent variables that hold the most importance for the prediction objective. Consequently, it facilitates feature selection and model optimization.

3.2.1. Correlation between Spectral Indices and Soil Salt Content

The original band data at the corresponding locations were extracted from the corrected remote sensing image using ArcMap 10.8.12790 software, based on the latitude and longitude information of the measured soil samples. The values of all characteristic variables corresponding to each measured soil sample were calculated using IBM SPSS Statistics 24.0.0.0 software. Pearson's correlation coefficients were calculated between the four types of characterization variables and the measured values of soil salinity. The correlation heat map showing the relationship between SSC and different categories of characterization variables is presented in Figure 5. The color red indicates a positive correlation between the variables, whereas blue indicates a negative correlation. The intensity of the color darkens as the correlation increases. Based on the graph, we can conclude that: (1) Among the band index group, SWIR1 exhibited the strongest negative correlation with SSC, with a correlation coefficient of -0.6 . The GREEN band showed the weakest correlation with SSC, with a correlation coefficient of 0.21 . Within the salinity indices group, SI5 demonstrated a robust positive correlation with SSC, evidenced by a high correlation coefficient of 0.76 , in contrast to SI2, which exhibited the weakest correlation. In the vegetation indices category, ENDVI presented the strongest correlation with SSC, whereas ALBEDO displayed the weakest. Within the composite indices category, CORSI and SWIR1 each registered a correlation coefficient of -0.6 with SSC. (2) The mean absolute values of the correlation coefficients between each group of variables and the SSC were calculated separately. The mean absolute correlation values for the band indices group, the salinity indices group, the vegetation indices group, and the composite indices group were 0.375 , 0.469 , 0.571 , and

0.42, respectively. The vegetation indices group had the highest mean absolute correlation with SSC, whereas the band indices group had the lowest. (3) SI5 emerged as the variable with the strongest correlation to SSC, succeeded by ENDVI and EDVI from the vegetation indices category, both manifesting negative correlations with SSC. GREEN and SI2 each displayed the weakest correlation with SSC, with their correlation coefficients having an absolute value of 0.21. Following the correlation analysis, four variables—SWIR1, SI5, ENDVI, and COSRI—were chosen to constitute the input variable group PCC.

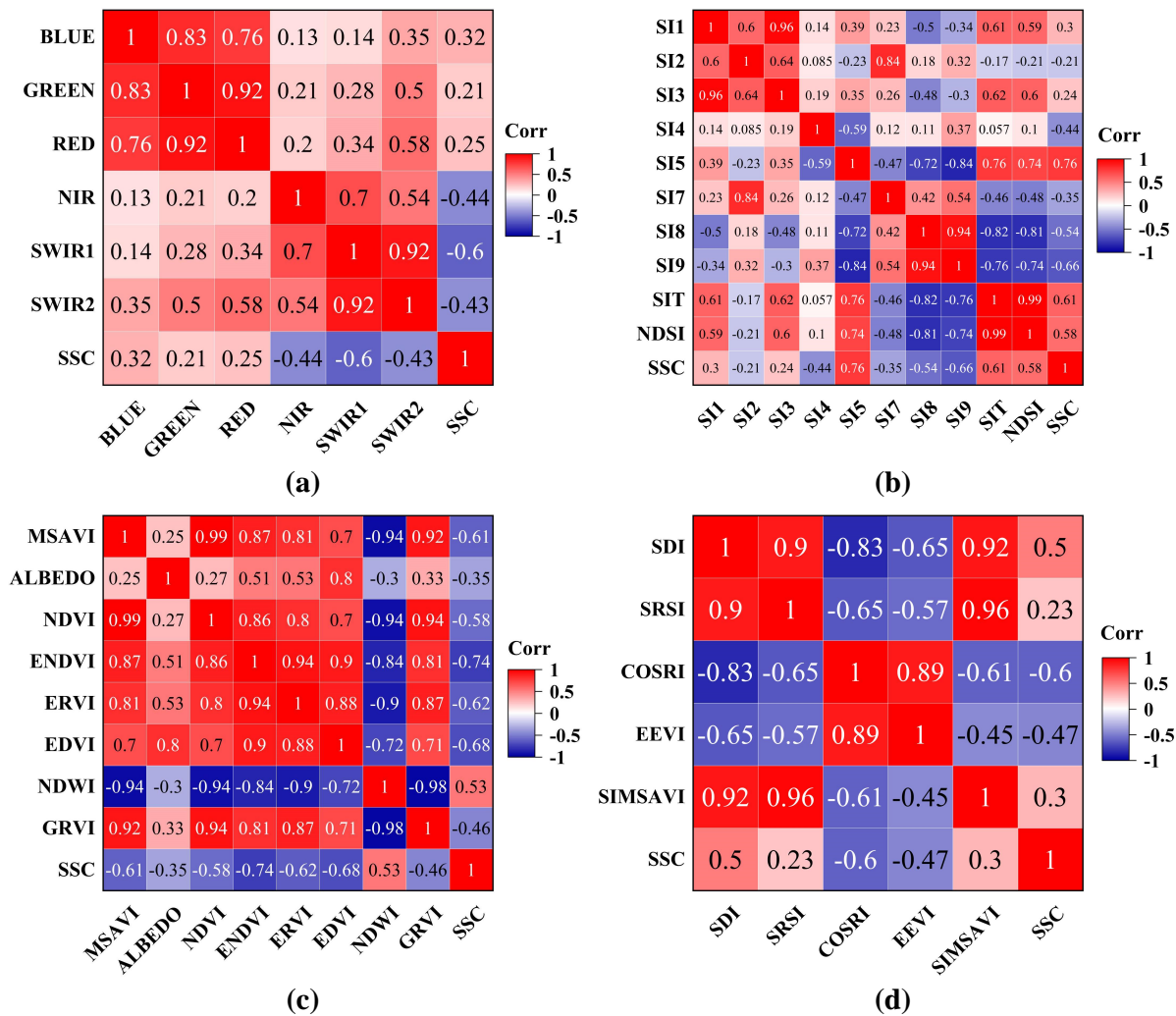


Figure 5. Heat maps of Pearson correlation analysis between four spectral indices and SSC: (a) band indices; (b) salinity indices; (c) vegetation indices; (d) composite indices.

3.2.2. Importance Analysis of Characteristic Variables

Variable importance in the projection analysis was used to screen twenty-nine spectral indices, including six band indices, ten salinity indices, eight vegetation indices, and five composite indices. The results are presented in Figure 6. In the figure, the blue dots represent the projected importance value of the variable, and the red circles represent the positions where the projected importance value is 1. The figure reveals that certain band and composite indices exhibit low VIP value, whereas several vegetation and salinity indices possess VIP value greater than 1. Among them, SI5 has the highest importance, indicated by a VIP value of 1.444, whereas SI3 has the lowest VIP value of 0.578. (A VIP value greater than 1 indicates that the variable is highly important for the dependent variable, a value greater than 0.5 but less than 1 suggests unclear importance, and a value less than 0.5 indicates that the variable is not important for the dependent variable.) Sixteen variables

were selected for the input variable group VIP, comprising SI5, ENDVI, SI4, SWIR1, EDVI, ERVI, SI9, COSRI, SWIR2, SIT, MSAVI, EEVI, GRVI, NDWI, NDSI, and NDVI.

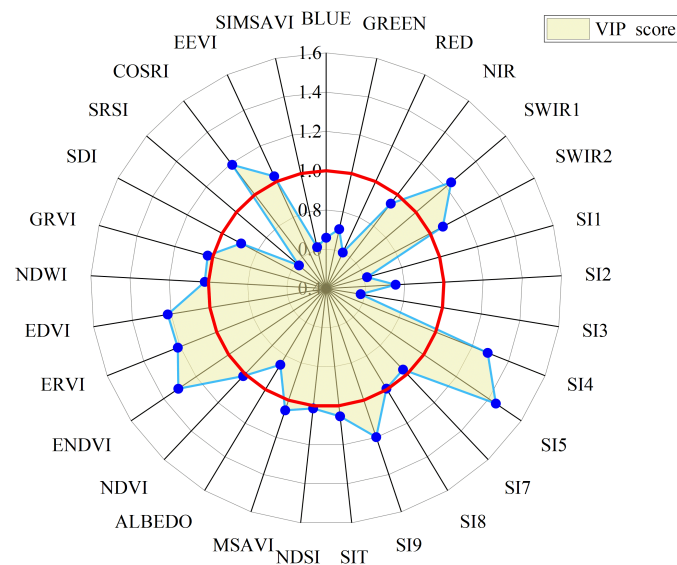


Figure 6. Characteristic importance values for all spectral indices.

After variable analysis and screening, three different groups of input variables were finally created for the experimental part of this paper. The spectral indices of the three input variable groups are shown in Table 6.

Table 6. Different input variable groups.

Name	Variables
PCC	SWIR1, SI5, ENDVI, COSRI
VIP	SI5, ENDVI, SI4, SWIR1, EDVI, ERVI, SI9, COSRI, SWIR2, SIT, MSAVI, EEVI, GRVI, NDWI, NDSI, NDVI
TV	BLUE, GREEN, RED, NIR, SWIR1, SWIR2, SI1, SI2, SI3, SI4, SI5, SI7, SI8, SI9, SI-T, NDSI, MSAVI, ALBEDO, NDVI, ENDVI, ERVI, EDVI, NDWI, GRVI, SDI, SRSI, CORSI, EEVI, SIMSAVI

3.3. Soil Salinity Inversion Model

The variable group PCC, variable group VIP, and full variable group TV are used as modeling variables to build three machine learning models—RELM, BP, and ICOA-RELM, respectively—and the performance of each model is evaluated using R^2 , RMSE, and MAE.

The R^2 values of the nine model test sets and the fitting equations for the measured and predicted values are depicted in Figure 7. The results indicate that optimizing the RELM model using ICOA substantially improved the model’s performance, resulting in enhanced predictions of SSC in the study area. Among all the models, ICOA-RELM-TV achieved the second-best performance with a test set R^2 value of 0.728, followed by BP-TV (R^2 value of 0.676). RELM-TV exhibited a similar R^2 value of 0.676, with higher MAE and RMSE compared to BP-TV. BP-PCC attained an R^2 value of 0.661, whereas RELM-VIP had an R^2 value of 0.607. ICOA-RELM-PCC achieved an R^2 value of 0.6, BP-VIP had an R^2 value of 0.594, and RELM-PCC obtained an R^2 value of 0.589.

The prediction results of SSC based on datasets selected by different feature band selection methods and different models are shown in Table 7. The analysis of the table

reveals that the RELM and BP models exhibit the best performance when utilizing the full set of variables as modeling variables. They achieved an R^2 value above 0.67 for both the training and test sets, along with lower MAE and RMSE. In the ICOA-RELM model, ICOA-RELM-VIP demonstrates the best performance and the highest inversion accuracy, with a test set R^2 of 0.75, MAE of 0.198, and RMSE of 0.249. ICOA-RELM-TV follows closely, with R^2 of 0.748 and 0.728 for the training and test sets, respectively. On the other hand, ICOA-RELM-PCC performs the worst, with R^2 below 0.7 for all the models. Comparatively, the models constructed using ICOA-RELM outperformed the unoptimized RELM model across all three input variables, yielding an average improvement of approximately 0.1 in the test set's R^2 .

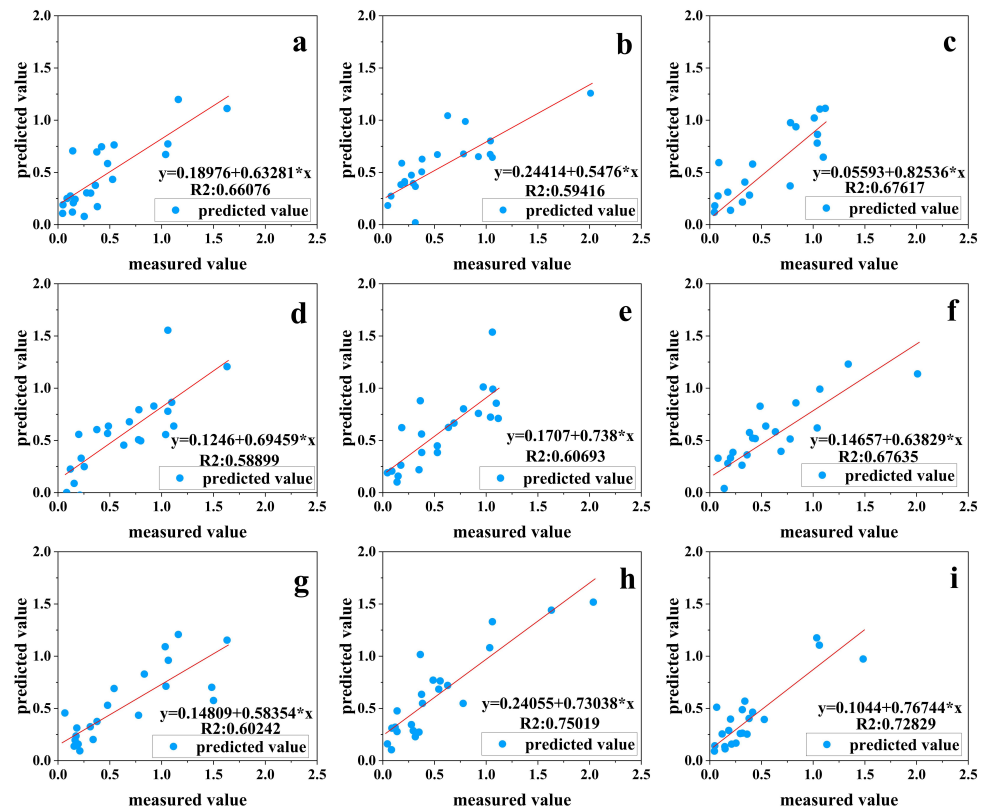


Figure 7. Scatter plots of measured and estimated SSC based on different models for different input variable groups. (a) BP-PCC; (b) BP-VIP; (c) BP-TV; (d) RELM-PCC; (e) RELM-VIP; (f) RELM-TV; (g) ICOA-RELM-PCC; (h) ICOA-RELM-VIP; (i) ICOA-RELM-TV. The red line is the fitting line between the measured and predicted values.

Table 7. Quantitative statistics of the SSC inversion.

Model	Input Variables	Training Set			Test Set		
		R^2	MAE	RMSE	R^2	MAE	RMSE
BP	PCC	0.708	0.183	0.261	0.661	0.18	0.238
	VIP	0.641	0.219	0.282	0.594	0.253	0.293
	TV	0.736	0.195	0.253	0.676	0.183	0.229
RELM	PCC	0.619	0.203	0.299	0.589	0.213	0.263
	VIP	0.656	0.221	0.295	0.607	0.167	0.231
	TV	0.706	0.191	0.254	0.676	0.191	0.266
ICOA-RELM	PCC	0.63	0.19	0.27	0.6	0.2	0.32
	VIP	0.771	0.149	0.217	0.75	0.198	0.249
	TV	0.748	0.182	0.244	0.728	0.135	0.186

3.4. Inversion of Soil Salt Spatial Distribution Based on ICOA-RELM-VIP Model

The Yellow River Delta wetlands include both perennial and seasonal storage wetlands. Perennial wetlands, predominantly characterized by mudflat ecosystems, consist of rivers, lakes, estuaries, and various types of ponds, including those for salt, shrimp, and crab. Conversely, seasonal storage wetlands comprise heavily saline supratidal areas, marshes, wet meadows, and paddy fields. As a result, salinization levels in the Yellow River Delta region vary significantly [47]. Soil salinization in this region arises from both natural factors and human activities. The Yellow River Delta's unique geographic location causes an imbalance in precipitation and distribution, exacerbated by an arid climate and scarce rainfall. This leads to soil moisture evaporation exceeding recharge, resulting in inadequate moisture and subsequent soil salinization. Additionally, a significant decline in the water table accelerates salt migration in groundwater, worsening surface soil salinization. Extensive soil erosion alters the nutrient composition of the land, further contributing to soil salinization. Excessive reclamation and rapid industrialization have disrupted the land's nutrient composition. Prolonged irrigation and improper water management have further exacerbated soil salinization [48].

The 16 spectral indices, selected through variable projection importance analysis, were utilized as model inputs. The ICOA-RELM model which performed best, was then employed for field inversion of the study area to obtain the distribution of soil total salinity classes in October 2003, as depicted in Figure 8. Subsequently, the percentage of soils in each class was tabulated, and the results are presented in Table 8.

In the study area, spatial distribution patterns reveal higher salinity levels along the coastal regions and lower salinity levels inland. The southeastern coastal region, accompanied by segments of the northwestern coast and the northeastern countryside, predominantly features soil with extreme and severe salinization, encompassing approximately 2351.5 square kilometers, which constitutes 43.36% of the entire study area. These areas are prone to repeated saltwater intrusion, exacerbated by drought and high temperatures, which promotes salt accumulation in the soil. Moderate saline soils are predominantly found in the central region, characterized by granite terraces and fluvial uplands at higher elevations, covering approximately 1266.94 square kilometers, accounting for 23.36% of the study area. Slight saline soils are primarily located along both sides of the Yellow River and in the northwestern region, where irrigation is extensively utilized. This area consists of river terraces, flatlands, and lowlands. Despite the influence of shallow groundwater levels and significant capillary action, these areas benefit from freshwater recharge. This category spans approximately 948.87 square kilometers, comprising 17.5% of the study area. Non-saline soils, the least represented category, comprise 15.78% of the entire study area, covering approximately 855.93 square kilometers, and are primarily found in the northeastern region, excluding coastal areas.

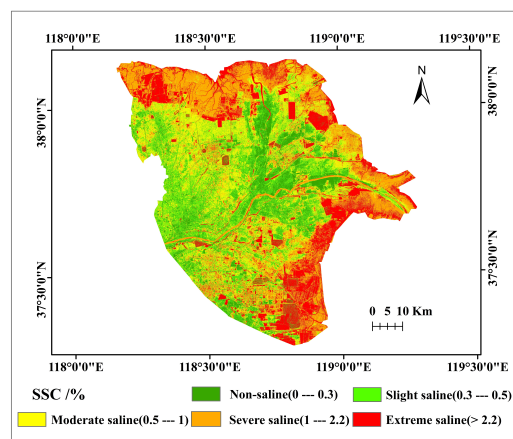


Figure 8. Spatial distribution map of soil salinity.

Table 8. Soil area and proportion of different salinization levels.

Salinization Level Total	Non-Saline	Slight Saline	Moderate Saline	Severe Saline	Extreme Saline
Area/km ²	855.93	948.87	1266.94	1373.66	977.84
Percent	15.78%	17.5%	23.36%	25.33%	18.03%

4. Discussion

In order to explore the effect of combining intelligent optimization algorithms with traditional machine learning models for inversion of SSC, there have been scholars combining the two for inversion of SSC, using intelligent optimization algorithms such as GA [26], seagull optimization algorithm (SOA) [49], sparrow search algorithm (SSA), bird swarm algorithm (BSA), moth search algorithm (MSA), Harris hawk optimization algorithm (HHO), grasshopper optimization algorithm (GOA), particle swarm optimization algorithm (PSO) [50], and so on. In this paper, on the basis of the previous studies, using measured SSC and different combinations of spectral indices as modeling input, we improve the crayfish optimization algorithm based on the one proposed in 2023 and combine the improved optimization algorithm with the RELM model to train the SSC inverse model. Circle chaotic mapping was introduced to improve the initialization of crayfish populations, which improved the convergence ability of the algorithm and the speed of searching for optimal solutions, as well as the accuracy of SSC inversion model. The results show that the use of the model of ICOA-RELM can realize the monitoring of soil salinity conditions in the Yellow River Delta region, which is conducive to the soil management in the region.

Comparative analysis of the final inversion model's accuracy demonstrates that, across all three input variable groups, the ICOA-RELM model introduced in this paper enhances the accuracy of estimating SSC in the study area when compared with the unoptimized model. This enhancement indicates that the optimization algorithm positively impacts the model's inversion capability. Overall, non-saline, slight saline, and moderate saline soils intersect throughout the central part of the study area. Non-saline soils tend to form dendritic patterns following the direction of the water network's runoff. Extreme saline soils are primarily found in tidal flats and tidal ditches, as well as other water bodies. The degree of salinization generally increases toward the seaward direction, closely linked to tidal infiltration and ground elevation. In the northern part of the study area, the former estuary area of the old Yellow River channel is dominated by severe saline soils, and the inner part is wrapped by a small amount of extreme saline soils. The coastal area in the southern part of the study area is dominated by extreme saline soils, and this part of the area is mainly tidal flats. The overall salinization level of the soil in the inversion results is consistent with the measured data from the actual sampling.

The accuracy of the model for inverting the SSC in the study area is influenced to some extent by the resolution and band information of the remote sensing images. The modeling effectiveness is limited by the use of reflectance data extracted from Landsat5 TM imagery, collected in 2003, to construct the spectral indices. The availability of higher quality imagery was not utilized. Currently, there are satellite data with higher resolution and quality, such as Sentinel 1 and 2, Planet, and Landsat 8, that can be utilized for studying the subsequent soil salinization levels using more recent data. Additionally, environmental factors such as soil moisture and soil utilization type can affect the level of salinity. Incorporating these environmental covariates into the input spectral index allows for an investigation of their relationship with SSC, thereby enhancing the accuracy of the inversion model [51]. The sample size used in this paper is limited, which restricts the application of new inversion modeling techniques. Collecting a larger sample size in future studies would be beneficial. Additionally, the utilization of deep learning algorithms in salinity inversion can enhance the accuracy of soil salinity level identification. The applicability of the ICOA-RELM inversion model in this paper in other regions needs to be further verified, and comprehensive testing and evaluation are needed to determine the performance of the model under different environmental conditions.

5. Conclusions

This paper presents the construction of 29 modeling variables, which encompass band indices, salinity indices, vegetation indices, and composite indices. These variables were derived from the band data obtained from a Landsat5 TM remote sensing image. The input variable groups PCC and VIP were obtained by screening the modeling indices using Pearson correlation analysis and variable importance ranking methods, respectively. A total of nine machine learning models, BP, RELM, and ICOA-RELM, were built based on the three modeling variable combinations of PCC, VIP, and TV and the SSC, respectively. The model with the highest accuracy was chosen based on the modeling results to generate a distribution map depicting the levels of soil salinity in the Yellow River Delta. The following conclusions have been drawn:

1. In the Pearson correlation analysis between spectral indices and SSC, SI5 showed the highest correlation with SSC with a correlation coefficient of 0.76; GREEN and SI2 had the least significant relationship with SSC. Among the four groups of spectral indices, the vegetation indices exhibited the highest average correlation, with a mean absolute correlation coefficient value of 0.571. The importance of the 29 spectral indices was ranked based on the VIP score. The four variables with the highest importance were identified as SI5, ENDVI, SI4, and SWIR1, with importance levels of 1.44, 1.31, 1.29, and 1.23, respectively. SI3 had the lowest importance value of 0.58.
2. The ICOA-RELM model was tested using the variable group VIP as the input, resulting in an R^2 value of 0.75, an MAE of 0.198, and an RMSE of 0.249. The model exhibits higher predictive accuracy and stability. The application of this model in the inversion of soil salinization in the Yellow River Delta region carries valuable reference significance.
3. This information is obtained from the distribution map depicting soil salinity levels in the Yellow River Delta region. The dominant soil types in the region are severe saline soils, followed by moderate saline soils. Severe saline soils dominate the northern part of the study area and the eastern portion of the central region, whereas the majority of the extreme saline soils are concentrated in the southeastern part of the region. A smaller proportion of extreme saline soils can also be found in the northern part of the region. Non-saline, slight saline, and moderate saline soils are evenly distributed in the central region of the district.

Author Contributions: Methodology, J.W. and X.W.; software, Y.C. and J.W.; validation, X.S. and Y.C.; formal analysis, Y.F.; investigation, B.T.; resources, J.Z.; data curation, J.W.; writing—original draft preparation, J.W.; writing—review and editing, X.W.; visualization, J.W. and X.S.; project administration, X.S.; funding acquisition, J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Shandong Key Research and Development Project (No. 2018GNC110025), in part by the National Natural Science Foundation of China (No.42301380), in part by “Taishan Scholar” Project of Shandong Province (No. TSXZ201712), in part by Qingdao Natural Science Foundation Grant (No.23-2-1-64-zyyd-jch), and in part by the Science and Technology Support Plan for Youth Innovation of Colleges and Universities of Shandong Province of China (No.2023KJ232).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Rengasamy, P. Soil Salinization. In *Oxford Research Encyclopedia of Environmental Science*; Oxford University Press: Oxford, UK, 2016.
- Li, J.; Pu, L.; Han, M.; Zhu, M.; Zhang, R.; Xiang, Y. Soil salinization research in China: Advances and prospects. *J. Geogr. Sci.* **2014**, *24*, 943–960. [[CrossRef](#)]
- Xiaomei, F.; Gaohuan, L.; Zhipeng, T.; Longcang, S. Analysis on main contributors influencing soil salinization of Yellow River Delta. *J. Soil Water Conserv.* **2010**, *24*, 139–144.
- Rhoades, J. Electrical conductivity methods for measuring and mapping soil salinity. *Adv. Agron.* **1993**, *49*, 201–251.
- Zhuang, Q.; Shao, Z.; Huang, X.; Zhang, Y.; Wu, W.; Feng, X.; Lv, X.; Ding, Q.; Cai, B.; Altan, O. Evolution of soil salinization under the background of landscape patterns in the irrigated northern slopes of Tianshan Mountains, Xinjiang, China. *CATENA* **2021**, *206*, 105561. [[CrossRef](#)]
- Metternicht, G.I.; Zinck, J. Remote sensing of soil salinity: Potentials and constraints. *Remote Sens. Environ.* **2003**, *85*, 1–20. [[CrossRef](#)]
- Abbas, A.; Khan, S.; Hussain, N.; Hanjra, M.A.; Akbar, S. Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Phys. Chem. Earth Parts A/B/C* **2013**, *55*, 43–52. [[CrossRef](#)]
- Ennaji, W.; Barakat, A.; Karaoui, I.; El Baghdadi, M.; Arioua, A. Remote sensing approach to assess salt-affected soils in the north-east part of Tadla plain, Morocco. *Geol. Ecol. Landscapes* **2018**, *2*, 22–28. [[CrossRef](#)]
- Ramos, T.B.; Castanheira, N.; Oliveira, A.R.; Paz, A.M.; Darouich, H.; Simionesei, L.; Farzamian, M.; Gonçalves, M.C. Soil salinity assessment using vegetation indices derived from Sentinel-2 multispectral data. Application to Lezíria Grande, Portugal. *Agric. Water Manag.* **2020**, *241*, 106387. [[CrossRef](#)]
- Yunhao, D.; Yao, G.; Chunyong, F.; Min, J.; Xinghong, H. Extraction and analysis of soil salinization information of Alar reclamation area based on spectral index modeling. *Remote Sens. Nat. Resour.* **2023**, *35*, 205–212.
- Wang, D.; Jia, W. Retrieving coastal soil saline based on landsat image in Chongming Dongtan. *J. Agric. Sci. Technol.* **2018**, *20*, 55–63.
- Zhang, S.M.; Zhao, G.X.; Wang, Z.R.; Xiao, Y.; Lang, K. Remote sensing inversion and dynamic monitoring of soil salt in coastal saline area. *J. Agric. Resour. Environ.* **2018**, *35*, 349–358.
- Huang, Q.; Xu, X.; Lü, L.; Ren, D.; Ke, J.; Xiong, Y.; Huo, Z.; Huang, G. Soil salinity distribution based on remote sensing and its effect on crop growth in Hetao Irrigation District. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 102–109.
- Weng, Y.L.; Qi, H.p.; Fang, H.b.; Zhao, F.; Lu, Y. PLSR-Based hyperspectral remote sensing retrieval of soil salinity of ChaKa-GongHe basin in QingHai province. *Acta Pedofil. Sin* **2010**, *47*, 1255–1263.
- Liu, X.; Yungang, B.; Zhongping, C.; Zhang, J.; Zhu, J.; Bangxin, D.; Zhang, C. Multispectral remote sensing inversion and seasonal difference in soil salinity of cotton field in typical oasis irrigation area. *J. Agric. Resour. Environ.* **2023**, *40*, 598.
- Hongyan, C.; Gengxing, Z.; Jingchun, C.; Ruiyan, W.; Mingxiu, G. Remote sensing inversion of saline soil salinity based on modified vegetation index in estuary area of Yellow River. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 107–114.
- Chengzhi, F.; Ziwen, W.; Xingchao, Y.; Yongkai, L.; Xuexin, X.; Bin, G.; Zhenhai, L. Machine Learning Inversion Model of Soil Salinity in the Yellow River Delta Based on Field Hyperspectral and UAV Multispectral Data. *Smart Agric.* **2022**, *4*, 61–73.
- Cao, X.; Ding, J.; Ge, X.; Wang, J. Estimation of soil electrical conductivity based on spectral index and machine learning algorithm. *Acta. Pedofil. Sin* **2020**, *57*, 867–877.
- Adnan, R.M.; Mostafa, R.R.; Kisi, O.; Yaseen, Z.M.; Shahid, S.; Zounemat-Kermani, M. Improving streamflow prediction using a new hybrid ELM model combined with hybrid particle swarm optimization and grey wolf optimization. *Knowl.-Based Syst.* **2021**, *230*, 107379. [[CrossRef](#)]
- Bai, L.; You, Q.; Zhang, C.; Sun, J.; Liu, L.; Lu, H.; Chen, Q. Advances and applications of machine learning and intelligent optimization algorithms in genome-scale metabolic network models. *Syst. Microbiol. Biomanuf.* **2023**, *3*, 193–206. [[CrossRef](#)]
- Han, F.; Zhao, M.R.; Zhang, J.M.; Ling, Q.H. An improved incremental constructive single-hidden-layer feedforward networks for extreme learning machine based on particle swarm optimization. *Neurocomputing* **2017**, *228*, 133–142. [[CrossRef](#)]
- Tripathi, D.; Edla, D.R.; Kuppili, V.; Bablani, A. Evolutionary Extreme Learning Machine with novel activation function for credit scoring. *Eng. Appl. Artif. Intell.* **2020**, *96*, 103980. [[CrossRef](#)]
- Li, Q.; Du, Y.; Liu, Z.; Zhou, Z.; Lu, G.; Chen, Q. Drought prediction in the Yunnan–Guizhou Plateau of China by coupling the estimation of distribution algorithm and the extreme learning machine. *Nat. Hazards* **2022**, *113*, 1635–1661. [[CrossRef](#)]
- Chi, D.; Zhang, L.; Li, X.; Wang, K.; Wu, X.; Zhang, T. Drought prediction model based on genetic algorithm optimization support vector machine (SVM). *J. Shenyang Agric. Univ.* **2013**, *44*, 190–194.
- Liu, D.; Fan, Z.; Fu, Q.; Li, M.; Faiz, M.A.; Ali, S.; Li, T.; Zhang, L.; Khan, M.I. Random forest regression evaluation model of regional flood disaster resilience based on the whale optimization algorithm. *J. Clean. Prod.* **2020**, *250*, 119468. [[CrossRef](#)]
- Zhao, W.; Ma, H.; Zhou, C.; Zhou, C.; Li, Z. Soil Salinity Inversion Model Based on BPNN Optimization Algorithm for UAV Multispectral Remote Sensing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 6038–6047. [[CrossRef](#)]
- Yang, L.; Chen, C.; Zheng, H.; Luo, G.; Shang, B.; Hellwich, O. Retrieval of soil salinity content based on random forests regression optimized by Bayesian optimization algorithm and genetic algorithm. *J. Geo-Inf. Sci.* **2021**, *23*, 1662–1674.
- Zhou, X.; Zhang, F.; Liu, C.; Kung, H.t.; Johnson, V.C. Soil salinity inversion based on novel spectral index. *Environ. Earth Sci.* **2021**, *80*, 501. [[CrossRef](#)]

29. He, B.; Jia, B.; Zhao, Y.; Wang, X.; Wei, M.; Dietzel, R. Estimate soil moisture of maize by combining support vector machine and chaotic whale optimization algorithm. *Agric. Water Manag.* **2022**, *267*, 107618. [[CrossRef](#)]
30. Jia, H.; Rao, H.; Wen, C.; Mirjalili, S. Crayfish optimization algorithm. *Artif. Intell. Rev.* **2023**, *56*, 1919–1979. [[CrossRef](#)]
31. Dehaan, R.; Taylor, G. Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced soil salinization. *Remote Sens. Environ.* **2002**, *80*, 406–417. [[CrossRef](#)]
32. Dwivedi, R.; Rao, B. The selection of the best possible Landsat TM band combination for delineating salt-affected soils. *Int. J. Remote Sens.* **1992**, *13*, 2051–2058. [[CrossRef](#)]
33. Ding, J.; Yu, D. Monitoring and evaluating spatial variability of soil salinity in dry and wet seasons in the Werigan–Kuqa Oasis, China, using remote sensing and electromagnetic induction instruments. *Geoderma* **2014**, *235*, 316–322. [[CrossRef](#)]
34. Nicolas, H.; Walter, C.; Douaoui, A.E.K. Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* **2006**, *134*, 217–230.
35. Bannari, A.; Guedon, A.; El-Harti, A.; Cherkaoui, F.; El-Ghmari, A. Characterization of slightly and moderately saline and sodic soils in irrigated agricultural land using simulated data of advanced land imaging (EO-1) sensor. *Commun. Soil Sci. Plant Anal.* **2008**, *39*, 2795–2811. [[CrossRef](#)]
36. Cheng, T.; Zhang, J.H.; Zhang, S.; Yun, B.; Wang, J.; Li, S.; Javid, T.; Meng, X.; Pangali Sharma, T.P. Monitoring soil salinization and its spatiotemporal variation at different depths across the Yellow River Delta based on remote sensing data with multi-parameter optimization. *Environ. Sci. Pollut. Res.* **2022**, *29*, 24269–24285. [[CrossRef](#)] [[PubMed](#)]
37. Allbed, A.; Kumar, L.; Aldakheel, Y.Y. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma* **2014**, *230*, 1–8. [[CrossRef](#)]
38. Bian, L.; Wang, J.; Guo, B.; Cheng, K.; Wei, H. Remote sensing extraction of soil salinity in Yellow River Delta Kenli County based on feature space. *Remote Sens. Technol. Appl.* **2020**, *35*, 211–218.
39. Liang, S. Narrowband to broadband conversions of land surface albedo I: Algorithms. *Remote Sens. Environ.* **2001**, *76*, 213–238. [[CrossRef](#)]
40. Farahmand, N.; Sadeghi, V. Estimating soil salinity in the dried lake bed of Urmia Lake using optical Sentinel-2 images and nonlinear regression models. *J. Indian Soc. Remote Sens.* **2020**, *48*, 675–687. [[CrossRef](#)]
41. Xi, X.; Zhao, G.x.; Gao, P.; Cui, K.; Li, T. Inversion of soil salinity in coastal winter wheat growing area based on sentinel satellite and unmanned aerial vehicle multi-spectrum—A case study in Kenli district of the Yellow River delta. *Sci. Agric. Sin.* **2020**, *53*, 5005–5016.
42. Alhammad, M.; Glenn, E. Detecting date palm trees health and vegetation greenness change on the eastern coast of the United Arab Emirates using SAVI. *Int. J. Remote Sens.* **2008**, *29*, 1745–1765. [[CrossRef](#)]
43. Fernandez-Buces, N.; Siebe, C.; Cram, S.; Palacio, J. Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. *J. Arid. Environ.* **2006**, *65*, 644–667. [[CrossRef](#)]
44. Zhang, T.; Wang, L.; Zeng, P.; Wang, T.; Geng, Y.; Wang, H. Soil salinization in the irrigated area of the Manas River basin based on MSAVI-SI feature space. *Arid Zone Res.* **2016**, *33*, 499–505.
45. Martínez-Martínez, J.M.; Escandell-Montero, P.; Soria-Olivas, E.; Martín-Guerrero, J.D.; Magdalena-Benedito, R.; Gómez-Sanchis, J. Regularized extreme learning machine for regression problems. *Neurocomputing* **2011**, *74*, 3716–3721. [[CrossRef](#)]
46. Feng, X.; Liu, Q. Regional soil salinity monitoring based on multi-source collaborative remote sensing data. *Trans. Chin. Soc. Agric. Mach.* **2018**, *49*, 127–133.
47. Guo, B.; Liu, Y.; Fan, J.; Lu, M.; Zang, W.; Liu, C.; Wang, B.; Huang, X.; Lai, J.; Wu, H. The salinization process and its response to the combined processes of climate change–human activity in the Yellow River Delta between 1984 and 2022. *Catena* **2023**, *231*, 107301. [[CrossRef](#)]
48. Ivushkin, K.; Bartholomeus, H.; Bregt, A.K.; Pulatov, A.; Kempen, B.; de Sousa, L. Global mapping of soil salinity change. *Remote Sens. Environ.* **2019**, *231*, 111260. [[CrossRef](#)]
49. Xiao, D.; Wan, L. Remote sensing inversion of saline and alkaline land based on an improved seagull optimization algorithm and the two-hidden-layer extreme learning machine. *Nat. Resour. Res.* **2021**, *30*, 3795–3818. [[CrossRef](#)]
50. Nguyen, H.D.; Van, C.P.; Nguyen, T.G.; Dang, D.K.; Pham, T.T.N.; Nguyen, Q.H.; Bui, Q.T. Soil salinity prediction using hybrid machine learning and remote sensing in Ben Tre province on Vietnam’s Mekong River Delta. *Environ. Sci. Pollut. Res.* **2023**, *30*, 74340–74357. [[CrossRef](#)]
51. Wang, F.; Yang, S.; Ding, J.; Wei, Y.; Ge, X.; Liang, J. Environmental sensitive variable optimization and machine learning algorithm using in soil salt prediction at oasis. *Trans. Chin. Soc. Agric. Eng* **2018**, *34*, 102–110.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.