Check for updates

Spatial heterogeneity response of soil salinization inversion cotton field expansion based on deep learning

Jinming Zhang[1,2], Jianli Ding[2,3*], Jinjie Wang[1,2], Zihan Zhang[1,2], Jiao Tan[1,2] and Xiangyu Ge[1,2]

[1]College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi, China, [2]Xinjiang Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi, China, [3]Xinjiang Institute of Technology, Aksu, China

Soil salinization represents a significant challenge to the ecological environment in arid areas, and digital mapping of soil salinization as well as exploration of its spatial heterogeneity with crop growth have important implications for national food security and salinization management. However, the machine learning models currently used are deficient in mining local information on salinity and do not explore the spatial heterogeneity of salinity impacts on crops. This study developed soil salinization inversion models using CNN (Convolutional Neural Network), LSTM (Long Short-Term Memory Network), and RF (Random Forest) models based on 97 field samples and feature variables extracted from Landsat-8 imagery. By evaluating the accuracy, the best-performing model was selected to map soil salinity at a 30m resolution for the years 2013 and 2022, and to explore the relationship between soil electrical conductivity (EC) values and the expansion of cotton fields as well as their spatial correlation. The results indicate that:(1) The CNN performs best in prediction, with an R^2 of 0.84 for the training set and 0.73 for the test set, capable of capturing more local salinity information. (2) The expansion of cotton fields has reduced the level of soil salinization, with the area of severely salinized and saline soils in newly added cotton fields decreasing from 177.91 km^2 and 381.46 km^2 to 19.49 km^2 and 1.12 km^2, respectively. (3) Regions with long-term cotton cultivation and newly reclaimed cotton fields exhibit high sensitivity and vulnerability to soil salinity. This study explores the excellent performance of deep learning in salinity mapping and visualizes the spatial distribution of cotton fields that are highly sensitive to soil salinity, providing a scientific theoretical basis for accurate salinity management.

1 Introduction

Soil salinization poses a significant challenge to global agriculture, with saline soils now present in over 100 countries and regions worldwide (Abbas et al., 2013). For example, approximately one-quarter of the irrigated land in Pakistan has been salinized, with annual economic costs estimated between $0.26 and $9.4 billion. The yields of wheat and rice grown in saline-alkali soils have decreased by 32% and 48%, respectively (Sheikh et al., 2022). Similarly, in Iraq's Mesopotamian region, crop yields in salinized farmland have decreased by 30-60% (Wu et al., 2018). In China, approximately 37.72% of the total irrigated land in Xinjiang is affected by salinization (Muhetaer et al., 2022).The accumulation of salt ions in the soil leads to physiological drought in plants, inhibits nutrient absorption, ultimately resulting in poor development, decreased yields, and even death (Ma and Tashpolat, 2023). Soil salinization results in a disruption of the water-salt balance in affected areas, which presents a significant threat to the ecological environment and biosphere (Zhang et al., 2022). In order to formulate optimized soil improvement policies to address the persistent degradation of land in typical arid oasis areas, the reversal of soil salinization has emerged as a significant area of investigation within the field of salinization.

Oases are a distinctive geographical phenomenon that emerge under specific natural geographical and climatic conditions. They represent major areas for human habitation and development (Huang et al., 2007). The Wei-Ku Oasis is one of the traditional agricultural areas in Xinjiang. Since the 1950s, the Chinese government has implemented policies to promote agricultural growth, which has led to a notable expansion of cultivated land (Shabiti et al., 2008). Based on long-term field surveys, our research team has found that in the past decade (2013-2022), humans have continuously reclaimed peripheral wasteland. Saline-alkali land has been improved into arable land through drainage, flushing, salt drainage, and appropriate agricultural measures, with cotton as the main crop type. Cotton, as a significant cash crop, has made outstanding contributions to the local economy. In previous studies, research on cotton has mainly focused on identifying planting distribution and predicting growth and yield. Exploring the relationship between cotton and soil salinization can contribute to the improvement of saline-alkali land, the safeguarding of China's red line of 120 million hectares of arable land and the assurance of food security are of paramount importance (Chen et al., 2019).

Over the past decade, there has been a significant advancement in the field of research pertaining to the reversal of soil salinization. The three elements of available remote sensing image data—spatial resolution, temporal resolution, and spectral resolution—have all improved, and inversion methods have evolved from traditional geostatistical analysis to the use of machine learning models. Geostatistical methods permit the analysis of the spatial distribution characteristics of soil salinity, employing semi-variance functions and Kriging interpolation (Gong et al., 2012; Rongjiang and Jingsong, 2009). Nevertheless, these methods do not exhibit significant variability. With the widespread use of machine learning algorithm since the late 20th century. It requires fewer parameters defined by researchers in soil salinity prediction studies, have higher computational speed and efficiency, and can handle advantages such as numerical ordinal and discrete predictor variables (Wang et al., 2019). The model effectively addresses the nonlinearity between soil and environmental factors. Commonly used machine learning models include RF (Ma et al., 2021), GBDT (Chen et al., 2021) and PLSR (Chengzhi et al., 2022). Over the past decade, the computer field has experienced a period of rapid development, deep learning has gradually emerged. It can be used not only for image recognition tasks but also for sequence prediction and other tasks. Representative deep learning models include CNN (Amarasinghe et al., 2017), LSTM (Hochreiter and Schmidhuber, 1997), BPNN (Hao et al., 2021), etc. By comparing the mapping results of deep learning models and machine learning models, it is possible to provide more robust scientific evidence for the regional management of soil salinization.

Currently, the main methods for exploring the relationship between two factors include Pearson correlation analysis (Gogtay and Thatte, 2017) and grey relational analysis (Chakraborty et al., 2023). However, these methods overlook their spatial heterogeneity. Geographical Detector (Yu et al., 2021) addresses this issue by better exploring the spatial heterogeneity between variables and explaining their interactions. However, the Geographical Detector cannot explore the local spatial expression of variables according to their correlations. The Multiscale Geographically Weighted Regression (MGWR) model (Fotheringham et al., 2017), as a local modelling method derived from the GWR model has been employed to analyze spatial relationships in ecological processes. Unlike the GWR model, it can search for the optimal bandwidth (scale) for regression analysis, thereby providing more detailed regression coefficient estimates, which allows for a more accurate capture of spatial heterogeneity. Consequently, this study introduces the MGWR model in an innovative manner to investigate the interrelationship between salinization and cotton fields, thereby providing scientific evidence for the improvement of salinization and the rational expansion of cotton fields. The proposed method will achieve local sustainable development.

In soil salinity prediction, previous studies primarily used machine learning models for regression predictions. Although good prediction results were achieved, they did not integrate the rapidly developing deep learning technologies of recent years. This study explores two established deep learning models, CNN and LSTM, as well as the most widely used machine learning model, RF. The optimal model is selected to create digital maps of soil salinity. Through the MGWR model, this study innovatively investigates the local expression of how soil salinity affects the spatial heterogeneity of cotton field yields, providing new directions for targeted and regional management of cotton field soil salinization. The main research objectives are as follows: (1) To compare the differences in the inversion effects and accuracy of soil salinization between the CNN, LSTM, and RF models to better depict soil salinity maps; (2) To map the distribution of cotton planting and predict yield distribution; (3) To study the improvement effects and spatial heterogeneity of cotton field expansion on saline-alkali land.

2 Materials and methods

Figure 1 presents the study's flowchart, which encompasses three principal sections: a) Construction of CNN, LSTM and RF inversion models for mapping of soil salinity distribution in 2013 and 2022; b) Identifying the distribution of cotton fields in 2013 and 2022, and calculating image-by-image metric yields of the cotton fields; and c) Analyzing the soil salinity status of the newly added cotton fields and exploring the spatial heterogeneous response relationship.

2.1 Study area

The Wei-Ku Oasis is situated in the Aksu region of Xinjiang (81°28′30″~84°05′06″E、39°29′51″~42°38′01″N) (Figure 2). The oasis includes three counties: Xinhe, Shayar, and Kuqa. The salinization problem of soil in the Wei-Ku Oasis is prominent, with a total area of 523.76×10^4 hm^2 for the three counties. Shayar County has an average of 2667 hm^2 of land planted with crops that do not yield profits annually. Additionally, land salinization has led to the degradation of grassland area, reaching 4.56×104 hm^2. The degradation area of grassland in Xinhe and Kuqa counties has also reached 3.72×104 hm^2 (Xueping, 2009). The climate of the Wei-Ku Oasis is characterized by a temperate continental arid climate, with an annual average evaporation of 1991.0 to 2864.3

mm and an annual average precipitation of only 51.3 mm. The multi-year average temperature ranges from 10.6°C to 14.8°C, with the highest and lowest temperatures recorded at 41.3°C and -28.7°C, respectively. High evapotranspiration ratios allow soil salts to accumulate.

2.2 Data sources

2.2.1 Field soil sampling and laboratory analysis

The soil data sampling points were primarily selected based on field surveys, taking into account the topography, vegetation cover, and previous research results in the study area. This study selected 37 and 63 representative sampling points in the study area from August 25 to September 5, 2013, and from June 20 to July 10, 2022, respectively (with land use types recorded for the 2022 samples). Portable Global Positioning System (GPS) devices were used to record the geographical location of sampling sites. Soil samples were collected at each sampling point in a uniform manner within a 1-meter radius at a depth of 0-10 cm using a soil sampler. Five samples were collected and mixed together within this radius, and approximately 500 grams of the mixed sample were placed in labeled waterproof sealed bags. All soil samples were subjected to a series of treatments including natural drying, grinding, and sieving (2.0mm) in the laboratory. Twenty grams of processed soil and 100ml of distilled water were thoroughly shaken and left to stand for 24 hours. The electrical conductivity (EC) values of the leachate after standing were measured at 25°C using a multiple parameter measuring instrument (WTWinoLab® Multi3420 Set B, WTW GmbH, Germany).Following the removal of outliers, a total of 97 sample data points were obtained. Consequently, the EC value of the soil leachate can be employed as a reference index for the evaluation of soil salinity (Hardie and Doyle, 2012).

2.2.2 Remote sensing data and processing

The Landsat-8 imagery used in the study came from Google Earth Engine (GEE).As the soil sampling periods are concentrated in late June, late August, and early September, we utilized cloud masking functions in GEE and selected Landsat-8 images with cloud cover below 40% from May to September to composite mean images for the years 2013 and 2022. The cotton field identification task involved synthesizing median composite images from mid-April to mid-November for the full growing season of cotton to obtain the 2013 and 2022 remote sensing images. Cotton yield prediction utilized median composite remote sensing images synthesized during the bolling period from July to mid-August.

2.3 Spectral bands selection for salinity modeling

Spectral reflectance characterizes the degree of soil salinization by the reflectance values at specific wavelengths; higher reflectance values indicate more severe soil salinization at the land surface (Al-Ali et al., 2021). Therefore, it is essential to include spectral reflectance in the model's input feature variables. We selected B1-
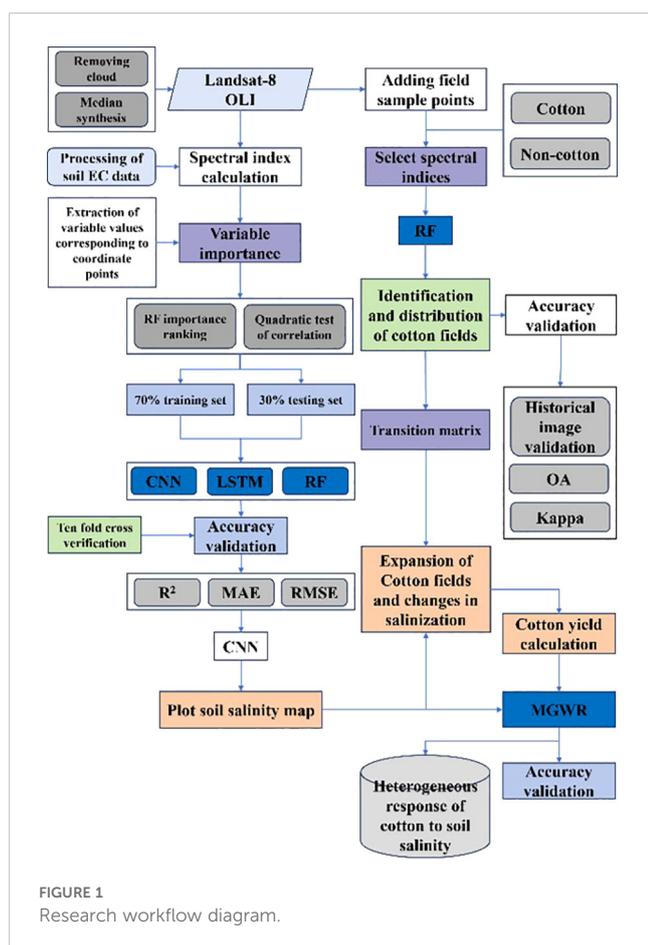
FIGURE 1
Research workflow diagram.

FIGURE 2
Schematic map of study area and distribution of sampling points.
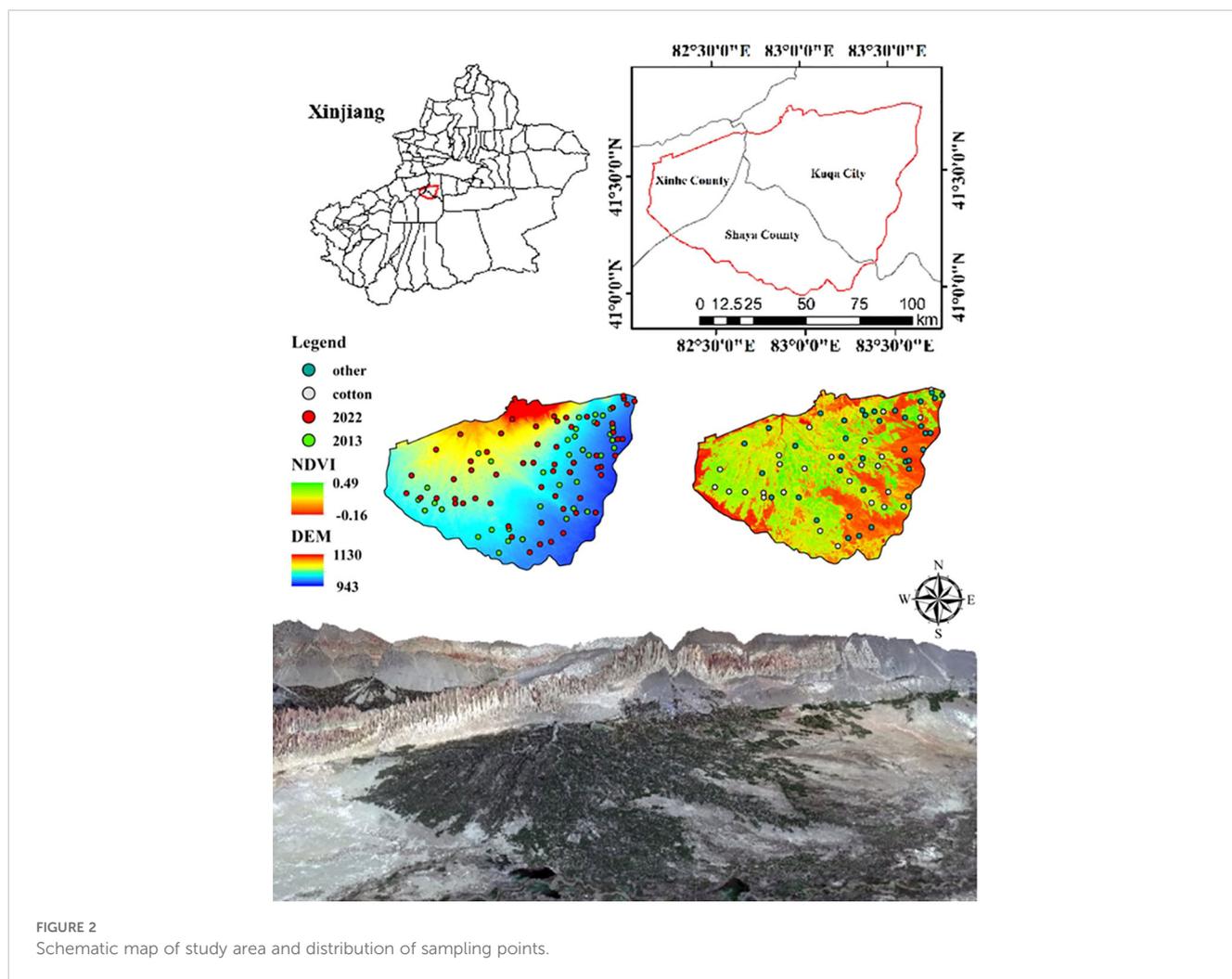
B7 as feature variable input models from the raw spectral bands provided by the OLI sensor of Landsat-8.

2.3.1 Vegetation indices

The spectral reflectance of vegetation under salt stress can serve as an indirect indicator of salt presence, especially in arid regions with low vegetation coverage, where its sensitivity is higher. Previous studies (Du et al., 2021) have indicated that vegetation indices are employed in machine learning methods to effectively elucidate the close relationship between vegetation and soil salinity, thereby enhancing the accuracy of soil salinity information retrieval. The formula for calculating the vegetation index selected for this study is shown in Table 1.

2.3.2 Soil indices

Soil texture is one of the fundamental conditions for salt accumulation. Gypseous desert soil is a type of desert soil with a distinct layer enriched in gypsum, widely distributed in the Xinjiang region. In addition to its elevated gypsum content, the substance in question also exhibits a high salt content. High soluble salt content in the soil is reflected below the gypsum accumulation layer,

typically forming sulfate salts at the gypsum surface. Gypsum index (GYEX) can reflect the salt content status in bare soil areas. Carbonate index (CAEX) and clay index (CLEX) reflect the soil's water retention capacity based on the content of carbonate and clay in the soil, thus indicating the salt migration capability. The formulae used in this study to calculate the soil indices are shown in Table 2.

2.3.3 Salinity indices

Salinity indices can directly reflect the degree of soil salinization either on spectral bands or by influencing canopy reflectance. The interaction among soil salinity, water content, and vegetation maintains the dynamic equilibrium of ecosystems. In this study, a total of 20 salt indices proposed by various researchers were calculated, and the formulas are shown in Table 3.

2.3.4 Other indices

Considering that NDBI and MNDWI can characterize the wetness or dryness of water content, we added NDBI and MNDWI indices as covariates, with the calculation formulas shown in Table 4.

TABLE 1 Vegetation indices.

Vegetation Index	Formula	Reference
NDVI	$(NIR - R)/(NIR + R)$	(Hong et al., 2023)
EVI	$2.5[\frac{(NIR - R)}{(NIR + 6 \times R - 7.5 \times B + 1)}]$	(Li et al., 2023)
DVI	$NIR - R$	(Haq et al., 2023)
GDVI	$(NIR^2 - R^2)/(NIR^2 + R^2)$	(Zhang et al., 2023)
MSAVI	$[2NIR + 1 - ((2NIR + 1)^2 - 8(NIR - R))^{0.5}]/2$	(Guo et al., 2023)
NLI	$(NIR^2 - R)/(NIR^2 + R)$	(Huete et al., 2002)
RVI	NIR/R	(Abulaiti et al., 2022)
OSAVI	$(NIR - R)/(NIR + R + 0.16)$	(Tian and Min, 1998)
ENDVI	$(NIR + SWIR1 - R)/(NIR + SWIR2 + R)$	(Song and Park, 2020)
IPVI	$NIR/(NIR + R)$	(Moradi et al., 2022)
GARI	$\{NIR - [G + \gamma \times (B - R)]\}/\{NIR + [G + \gamma \times (B - R)]\}$	(Gitelson et al., 1996)
GVMI	$[(NIR + 0.1) - (SWIR1 + 0.02)]/[(NIR + 0.1) + (SWIR1 + 0.02)]$	(Xiao et al., 2023)

B represents the blue band; G represents the green band; R represents the red band; NIR represents the near-infrared band; SWIR1 represents the shortwave infrared band 1; SWIR2 represents the shortwave infrared band 2; γ is the parameter controlling atmospheric correction.

2.4 Model construction

2.4.1 Convolutional neural network

CNN (LeCun et al., 1998) is a type of artificial neural network widely used in image recognition (Figure 3). The core features of CNN include their design with local connections and weight sharing, which reduce the number of network parameters and alleviate overfitting issues. Considering the one-dimensional nature of our measured soil salinity data, we constructed a one-dimensional convolutional neural network to address the regression prediction problem of soil salinity. This model can effectively extract local pattern features from time series data and achieve good convergence within a relatively short time. The model consists of two one-dimensional convolutional layers. The first layer uses 16 convolutional kernels with a kernel size of 3, a stride of 1, and padding of 1. The second layer uses 32 convolutional kernels with a kernel size of 3, a stride of 1, and padding of 1. Following the convolutional layers is a one-dimensional max-pooling layer with a pooling window size of 2, which reduces the size of the feature maps. After the convolutional and pooling operations, the feature maps are flattened and passed to two fully connected layers, with the first layer having 120 nodes, the second layer having 84 nodes, and the final output being a regression value. The model employs the ReLU activation function to introduce non-linearity, uses the Adam optimizer with a learning rate of 0.0006, and has a loss function of mean squared error (MSE), with training conducted over 300 epochs.

2.4.2 Long short-term memory

LSTM (Hochreiter and Schmidhuber, 1997) is a specific type of recurrent neural network (RNN) (Figure 3). Unlike standard RNN, LSTM has memory units capable of learning long-term dependencies and can effectively handle sequential data. The core component of the model is the LSTM layer, which has an input feature dimension of 8 and maps to 256 hidden units. The LSTM layer is responsible for capturing both short-term and long-term dependencies in the time series. Next, a fully connected layer maps the output of the LSTM to the regression result. During the forward propagation process, the LSTM's hidden states and memory cells are initialized at each iteration, with the final step involving the generation of the final prediction result through the fully connected layer. The model uses mean squared error (MSE) as the loss function, Adam as the optimizer, with a learning rate set to 0.0009, and is trained over 5000 epochs.

2.4.3 Random forest

RF (Breiman, 2001) is mainly used for classification and regression problems (Figure 3). The core idea of Random Forest includes random sampling, random feature selection, and prediction through majority voting. The parameters of the Random Forest model were set to 100 decision trees, a random seed of 5, a maximum depth of 5, a minimum of 5 samples required

TABLE 2 Soil indices.

Soil index	Formula	Reference
CLEX	$SWIR1/SWIR2$	(Prout et al., 2021)
GYEX	$(SWIR1 - NIR)/(NIR + SWIR2)$	(Taghizadeh-Mehrjardi et al., 2014)
CAEX	G/B	(Taghizadeh-Mehrjardi et al., 2014)
BI	$(G^2 + B^2)^{0.5}$	(Rafik et al., 2022)

Salinity index	Formula	Reference
S1	B/R	(Allbed et al., 2014)
S2	$(B - R)/(B + R)$	(Allbed et al., 2014)
S3	$G \times R/B$	(Allbed et al., 2014)
S5	$B \times R/G$	(Allbed et al., 2014)
S6	$R \times NIR/G$	(Allbed et al., 2014)
S7	$(SWIR1 - SWIR2)/((SWIR1 + SWIR2)$	(Bannari et al., 2008)
S8	$(G + R)/2$	(Nicolas and Walter, 2006)
S9	$(G + R + NIR)/2$	(Nicolas and Walter, 2006)
SI	$(G + R)^{0.5}$	(Khan et al., 2005)
SI1	$(G \times R)^{0.5}$	(Khan et al., 2005)
SI2	$(G^2 + R^2 + NIR^2)^{0.5}$	(Khan et al., 2005)
SI3	$(G^2 + R^2)^{0.5}$	(Khan et al., 2005)
SI4	$SWIR1/NIR$	(Khan et al., 2005)
SIT	$(\frac{R}{NIR}) \times 100$	(Jia et al., 2022)
SSSI1	$R - NIR$	(Bannari et al., 2008)
SSSI2	$(R \times NIR - NIR \times NIR)/R$	(Bannari et al., 2008)
NDSI	$(NIR - SWIR1)/(NIR + SWIR1)$	(Ghazali et al., 2022)
CRSI	$[\frac{NIR \times R - G \times B}{NIR \times R + G \times B}]^{0.5}$	(Ma et al., 2023)

to split a node, and a minimum of 2 samples required to be at a leaf node. The model was then trained.

2.5 Accuracy validation of soil salinization inversion

In this study 70% samples are used as training set and 30% samples are used as test set. The optimal parameters for all models were obtained by tenfold cross-validation. The performance of the model is assessed by the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error ($RMSE$), with the respective formulas as follows:

Other index	Formula	Reference
NDBI	$(SWIR1 - NIR)/(SWIR1 + NIR)$	(Guo et al., 2015)
MNDWI	$(B - SWIR1)/(B + SWIR1)$	(Moisa et al., 2022)

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

$$MAE = \frac{1}{m} \sum_{i=1}^{m} |\hat{y}_i - y_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2} \quad (3)$$

In the formulas: y_i represents the measured soil EC value; $\hat{y}_i$ denotes the predicted soil EC value; $\bar{y}_i$ stands for the mean measured soil EC value; and m represents the number of sampling points.

2.6 Cotton planting distribution and yield prediction

2.6.1 Environmental covariates

When extracting cotton planting areas, we selected Landsat-8 bands B1-B7, as well as NDVI and EVI spectral indices to identify cotton planting regions (Wang et al., 2017). The calculation formulae for NDVI and EVI are shown in Table 1.

2.6.2 Construction of classification models

Selecting the Random Forest model for cotton classification on the GEE platform, as described previously. We set the specific parameters of the Random Forest model in the GEE as follows: the number of trees in the Random Forest is set to 100, with a sampling proportion of 0.7 for the cotton classification task (Figure 3). Similarly, we divided the sample 70:30.

2.6.3 Accuracy validation of cotton field classification

To evaluate the classification performance of cotton fields, the confusion matrix is used in this study. Currently, the confusion matrix represents the most widely utilized methodology for evaluating the accuracy of classification algorithms. This method includes four evaluation metrics: producer's accuracy, user's accuracy, overall accuracy, and Kappa coefficient. The calculation formulas are as follows:

$$OA = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Kappa = \frac{PA-PE}{1-PE} \quad (5)$$

$$UA = \frac{TP}{TP+FP} \quad (6)$$

$$PA = \frac{TP}{TP+FN} \quad (7)$$

In these formulas, TP, TN, FP, and FN represent the counts of true positives, true negatives, false positives, and false negatives, respectively; PA is the accuracy of the model on the actual data, PE is the expected accuracy under random classification.

2.6.4 Cotton yield prediction

High correlation between cotton yield and NDVI, which has been extensively validated in practice (Liu et al., 2015; Meng et al., 2019). Therefore, in this study, raster image estimation is used to expand county-level yield maps to pixel-level spatial distributions. By combining data from statistical yearbooks and government bulletins on cotton economic yield, the average cotton yield at the pixel level for three county-level divisions in the study area is obtained (Cai and Sharma, 2010) (the average yield of Xinhe County and Shayar County is used to replace the missing cotton yield per mu in Kuqa City in 2022). Then, the cotton yield of individual pixels is expanded using the NDVI index of cotton. The calculation formula is as follows:

$$Y_P = Y_{avg} \times \frac{NDVI_P}{NDVI_{avg}} \qquad (8)$$

Where Y_P represents the cotton yield of a single pixel (kg·hm^{-2}); Y_{avg} is the average cotton yield at the county level, obtained from statistical yearbooks and government bulletins; $NDVI_P$ is the NDVI value during the flowering period of cotton; $NDVI_{avg}$ is the average NDVI value during the flowering period at the county level.

2.7 MGWR model

2.7.1 Model principle

The MGWR model is a geographically weighted regression model used to explore spatial heterogeneity in geospatial data

(Fotheringham et al., 2017). In contrast to traditional global regression models, the MGWR model is capable of capturing local spatial correlations in spatial data, thereby enhancing the model's explanatory power and prediction accuracy. This model allows regression coefficients to vary with spatial location, thus better understanding the nonlinear relationships and spatial heterogeneity in spatial data.

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{p} \beta_j(u_i, v_i)x_{ij} + \varepsilon_i \qquad (9)$$

Where $\beta_0(u_i, v_i)$ and $\beta_j(u_i, v_i)$ represent the locally varying regression coefficients, and ε_i denotes the error term. Typically, $\beta_0(u_i, v_i)$ and $\beta_j(u_i, v_i)$ can be weighted averaged using an appropriate spatial weighting function to reflect the spatial heterogeneity of spatial data.

2.7.2 Accuracy validation of MGWR model

We use the adjusted coefficient of determination $R^2_{adjusted}$ to validate the rationality of the MGWR model. The formula for adjusted $R^2_{adjusted}$ is as follows:

$$R^2_{adjusted} = 1 - \frac{(1-R^2)(n-1)}{n-p-1} \qquad (10)$$

Where: $R^2_{adjusted}$ is the adjusted coefficient of determination; n is the sample size; p is the number of independent variables used in the model.
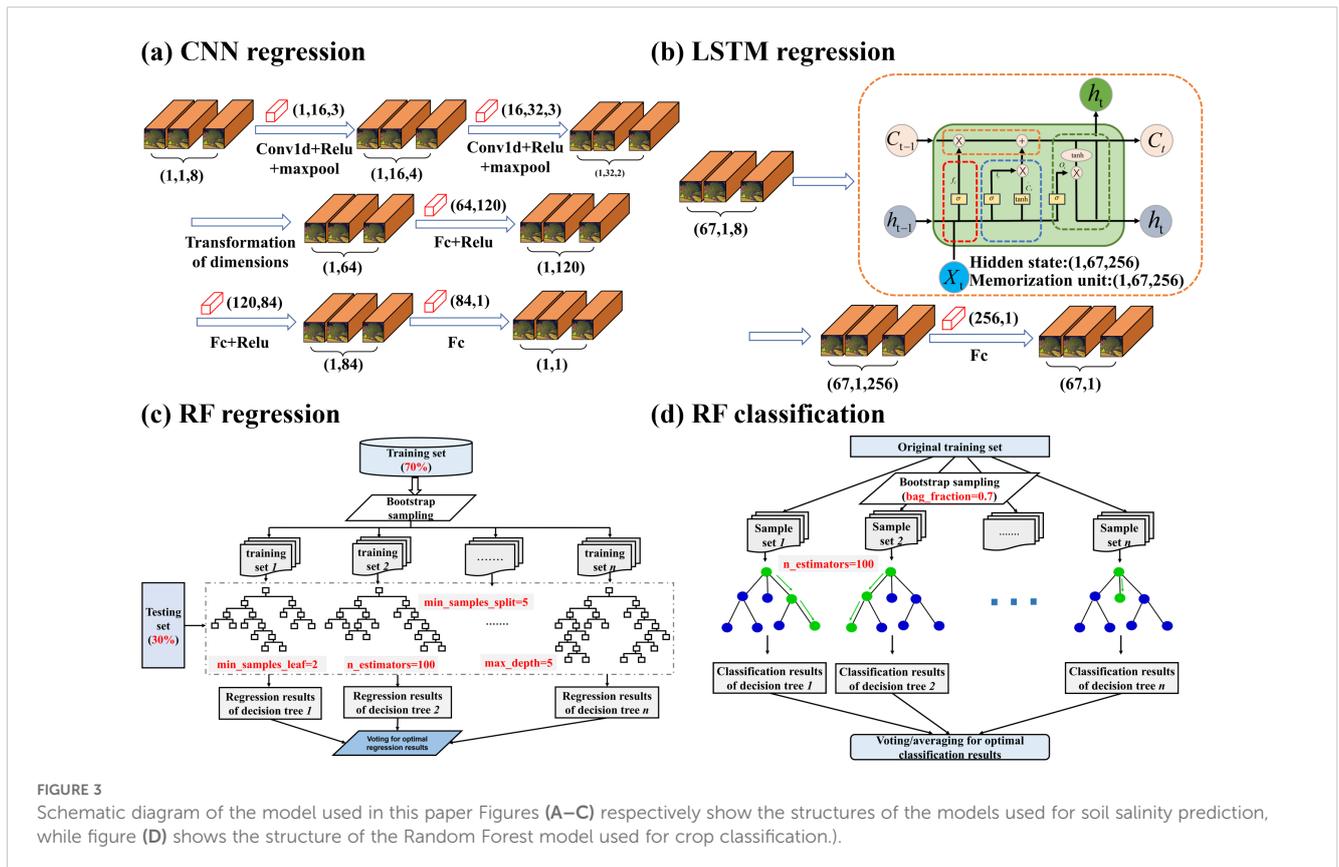
FIGURE 3
Schematic diagram of the model used in this paper Figures **(A–C)** respectively show the structures of the models used for soil salinity prediction, while figure **(D)** shows the structure of the Random Forest model used for crop classification.).

3 Results

3.1 Digital mapping of soil salinization inversion

3.1.1 Descriptive statistics of soil salinity

The EC in the study area exhibits a considerable range, spanning from 0.17 to 117.9 dS/m, with a mean of 18.3 dS/m (Table 5). It is generally considered that a coefficient of variation (CV) less than 10% indicates weak variability, values between 10% and 100% represent moderate variability, and values greater than 100% indicate strong variability. The coefficient of variation is greater than 100%, indicating a high spatial variability of soil salinity (Hu et al., 2023). The pH value ranges from 6.463 to 8.644, with relatively low variability.

3.1.2 Importance screening of soil salinity and environmental variables

We obtained the importance ranking of environmental variables by iterating 100 times using the random forest model (Figure 4), and selected the top eight important environmental variables (NLI, GVMI, S2, S1, S3, EVI, B4, GYEX) for modeling. This indicates that apart from salinity index, nonlinear vegetation index can also explain the EC content within the oasis, and soil index is also an important variable for predicting soil EC.

Furthermore, the correlation coefficients between each variable and the soil EC value were obtained through Pearson correlation coefficient analysis. The Pearson correlation coefficients calculated using all feature variables were further validated, showing that the Pearson correlation coefficients for the eight feature variables used to construct the model were all above 0.5, indicating a strong correlation with soil salinity and improving the accuracy of the modeling.

3.1.3 Comparison of soil salinity mapping and accuracy among different models

Based on the selection of eight feature variables, CNN, LSTM, and RF models were established for soil salinization inversion. From the perspective of modeling accuracy of the three different models, all three models have achieved a good level of performance. RF is prone to limitations due to the step size when processing data, which may result in suboptimal performance in long-term prediction tasks and does not exhibit higher stability compared to deep learning models. Nevertheless, machine learning models represented by RF have a significantly lower training time cost compared to deep learning models, allowing for rapid predictions. LSTM aims to model temporal dependencies in sequential data, emphasizing global features. In contrast, CNN can capture subtle variations in data features, which allows it to better identify local features, resulting in superior performance in surface soil salinity prediction, with R^2

TABLE 5 Descriptive statistics of soil salinity.

Soil properties	Minimum	Maximum	Mean	Standard deviation	Variable Coefficient%
EC/(ds·m^{-1})	0.17	117.9	18.3	24.71	135%
pH value	6.463	8.644	7.582	0.364	4.8%

FIGURE 4
Importance and correlation of feature variables Figure **(A)** represents the feature importance selection used for model construction, while Figure **(B)** shows the Pearson correlation analysis of all feature variables. This analysis is used to further validate the selected features by checking if their correlation with soil salinity exceeds 0.5.).
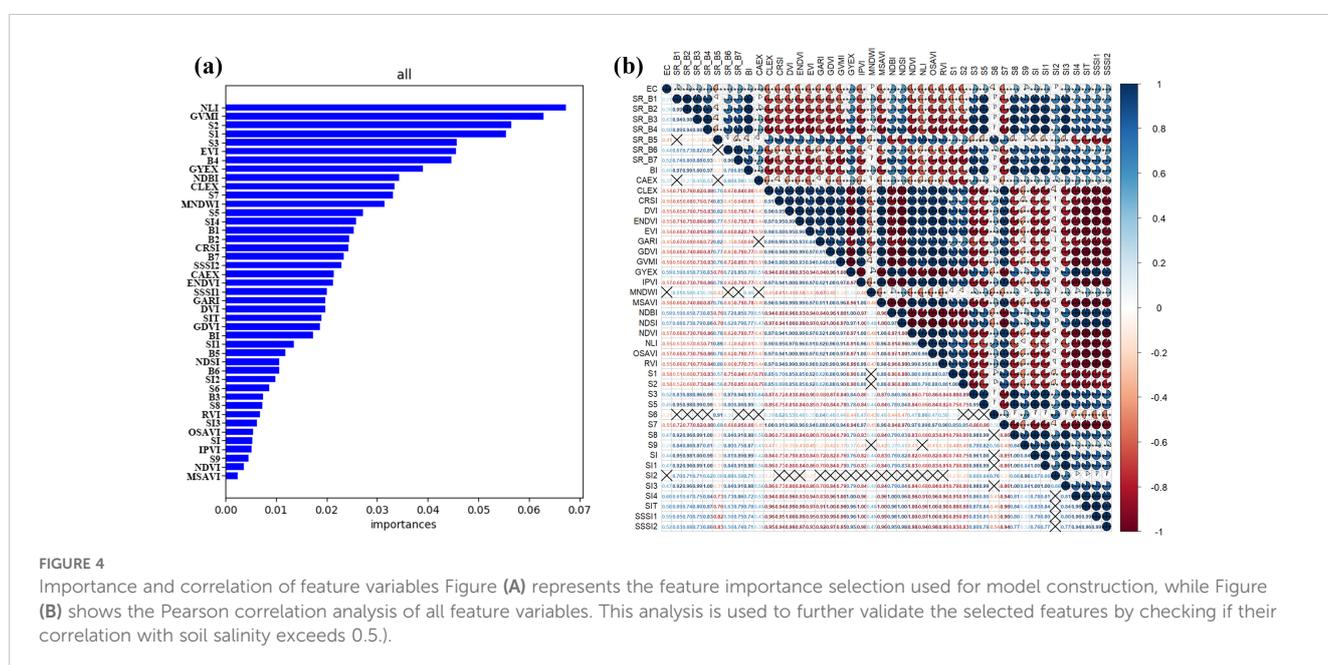
values of 0.84 and 0.73 for the training and testing sets, respectively (Figure 5). Additionally, the MAE and RMSE values for the training set are 6.56 dS/m and 10.16 dS/m, respectively, and for the testing set, they are 5.17 dS/m and 11.36 dS/m, respectively.

Soil salinity is classified into five categories based on the criteria for soil salinity levels: non-salinized (EC< 2), slight salinization (2< EC< 4), moderate salinization (4< EC< 8), severe salinization (8< EC< 16), and saline soil (EC > 16) (Scudiero et al., 2015). The three models previously described were applied to Landsat-8 images in order to obtain digital maps of soil EC for the Wei-Ku Oasis in 2013 and 2022. As shown in Figure 6, the spatial distribution characteristics derived from the three models are similar. Soils within the oasis have low salinity, while soils outside the oasis generally have high salinity. The RF model can only show a general trend in predicting soil salinity in the study area, which also demonstrates its tendency for instability in long-term prediction tasks. The LSTM model shows minimal local variation in predicting salinity within oases and deserts, which aligns with its focus on global features and neglects the depiction of local characteristics. In contrast, the CNN model can capture more spatial details and leverage its advantage in identifying local features of surface soil salinity, consistent with the results of Wang et al. (2023b), and effectively reveals the spatial heterogeneity of soil salinity in the study area.

According to CNN prediction results, non-salinized and slightly salinized areas are mainly concentrated within the oasis. The degree of salinization increases in the desert areas extending outward from the oasis edge's desert interleaved belt. The deserts in the eastern part of the oasis are primarily moderately and severely salinized, while the central and western desert areas are mainly characterized by severe salinization and saline soil, consistent with the results of He et al. (2019). Analyzing over the time scale from 2013 to 2022, the saline-alkali area in the study region has decreased by 484.3 km² over the past decade. The areas of severe and moderate salinization have shown slight changes, while the area of mild salinization has increased by 164.59 km². The area of moderate salinization has not changed significantly because moderate salinization has decreased in desert areas, while some non-salinized and lightly salinized soils within the oasis have transitioned to moderate salinization. Overall,

the soil salinization problem in the Wei-Ku Oasis has been alleviated and improved due to a series of government management measures.

3.2 Expansion of cotton fields and changes in soil EC values

3.2.1 Cotton field identification

The overall accuracy of the cotton field recognition model constructed using RF reached 94.4%, and the Kappa coefficient reached 88.88%. In order to ascertain the veracity of the model in 2013, 500 validation points were randomly generated within the classified results utilizing the ArcGIS 10.8 software (Figure 7). The validation points were imported into Google historical imagery for visual interpretation, and combined with historical on-site sampling images from the laboratory, achieving an accuracy of 90.6%, demonstrating good applicability.

3.2.2 Changes in soil EC values in cotton fields

We calculated the changes in cotton field area using a land use transfer matrix (Figure 8). In 2022, 1334.68km² of cotton fields remained unchanged, 500.78km² of cotton fields converted to non-cotton fields, and another 1679.85km² of non-cotton fields were converted to cotton fields through changes in crops and cultivation of barren land. Among the areas converted to cotton fields, 1145.18km² of land had varying degrees of salinization in 2013, with severe salinization and saline-alkali areas reaching 177.91km² and 381.46km², respectively. However, by 2020, the non-salinized area had reached 1256.53km², with only 19.49km² and 1.12km² remaining for severe salinization and saline-alkali areas, respectively. This significant change demonstrates the positive role of cotton field expansion in addressing soil salinization issues.

3.3 Cotton field yield prediction and spatial heterogeneity response of soil salinity

First, the average cotton yield of three counties and NDVI index were used to plot the predicted cotton yield maps of the study area
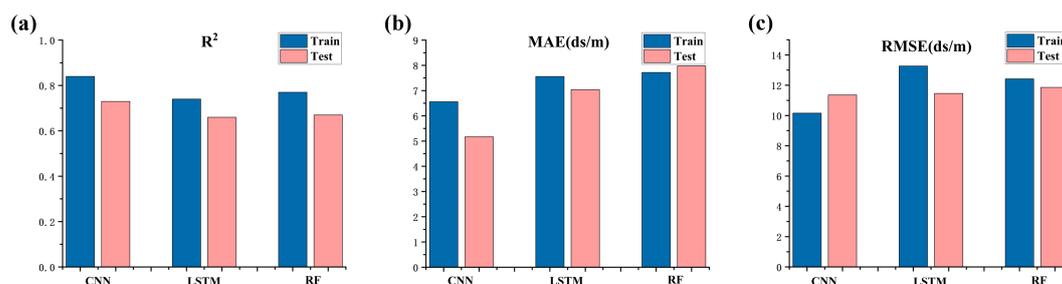
FIGURE 5
Comparison of model accuracy (**A** represents the model determination coefficient, **B** represents the mean absolute error, **C** represents the root mean square error, blue represents the accuracy verification of the soil salinity inversion model on the training set, and pink represents the accuracy verification on the test set).
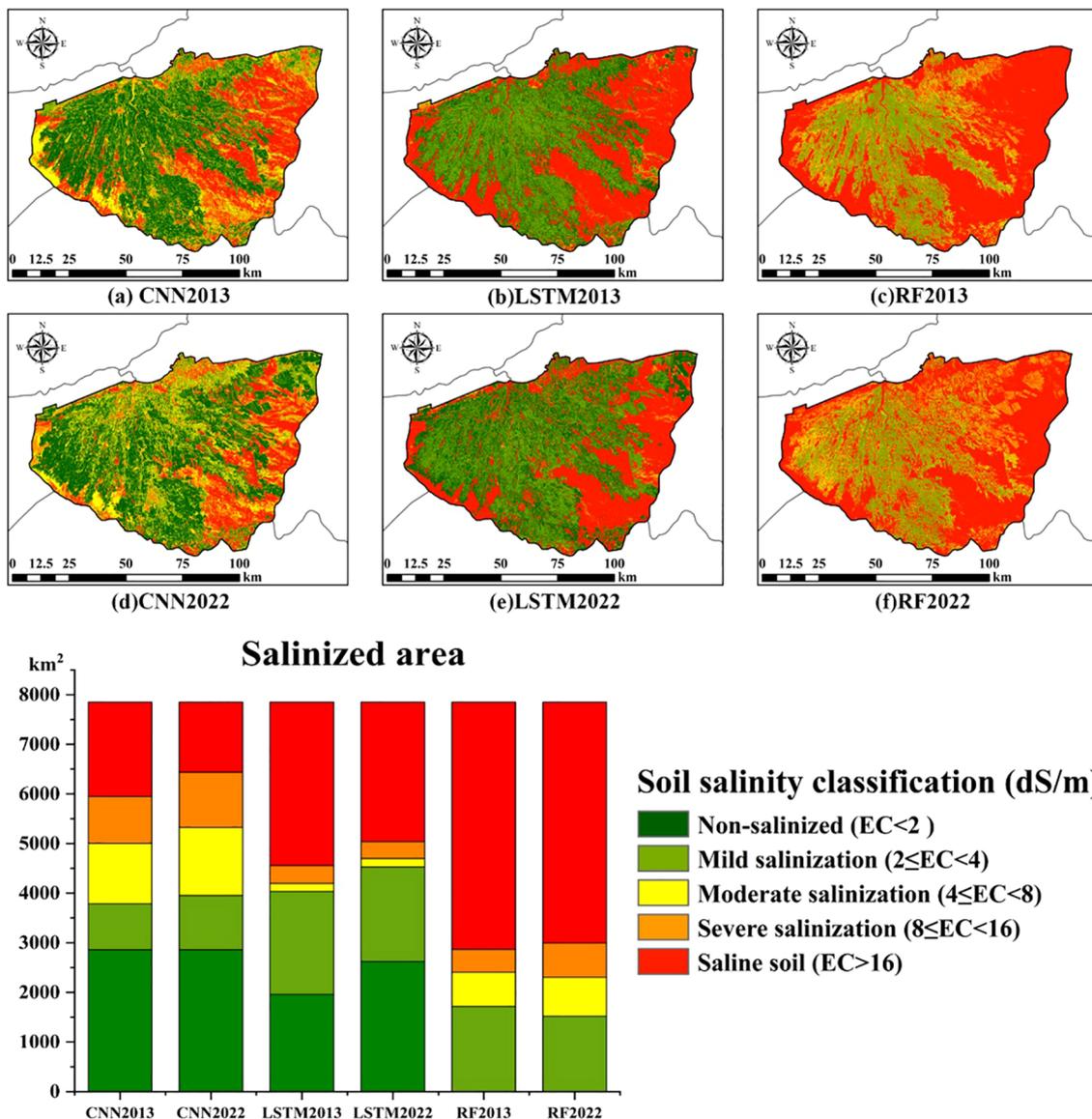
FIGURE 6

Soil salinity maps for 2013 and 2022 generated by different models, along with the proportions of different degrees of saline-alkali soil Figures
(A–C) represent the soil salinity distribution maps for the year 2013, generated using the CNN, LSTM, and RF models, respectively. Figures
(D–F) represent the soil salinity distribution maps for the year 2022, generated using the CNN, LSTM, and RF models, respectively.).

for 2013 and 2022 to demonstrate the growth status of cotton (Figure 9). In 2013, areas with low average cotton yields were mainly concentrated in the western and peripheral regions of the cotton fields. In 2022, areas with low average cotton yields were mainly concentrated in the western, southern, and newly reclaimed northeastern parts of the cotton fields.

The predicted cotton yield map and soil salinity level map were resampled to 500m, and then the MGWR coefficients of soil EC value and cotton yield were calculated (Figure 10). According to the accuracy verification results, the $R^2_{adjusted}$ of the MGWR model in 2013 was 0.75, and in 2022 it was 0.86, reaching the expected level. Areas sensitive to soil salinity are mainly distributed in the western part of the Wei-Ku oasis, in Xinhe County and Shaya County, as well as in the northeastern part of Kuqa City, which has been newly reclaimed.

4 Discussion

4.1 Impact of environmental variables on soil salinity

Currently, a large number of research findings indicate that various environmental factors are highly effective in constructing soil salinity models. Zhu et al (2021) studied the relationship between vegetation indices and salinity, finding that soil salinity stresses crop growth, and vegetation greenness can indirectly reflect soil salinity levels. The highest contribution of vegetation index NLI in modelling soil salinity in arid zone was also found in the results of this study. Analysis of vegetation indices at a finer field scale (Polivova and Brook, 2021) showed that NDVI, EVI, GARI, and

FIGURE 7
Cotton field identification and accuracy validation random points for 2013 and 2022 Figure **(A)** shows the distribution of cotton fields in 2013, figure **(B)** shows the distribution of cotton fields in 2022, and figure **(C)** shows the random validation points used for the cotton field distribution in 2013.).
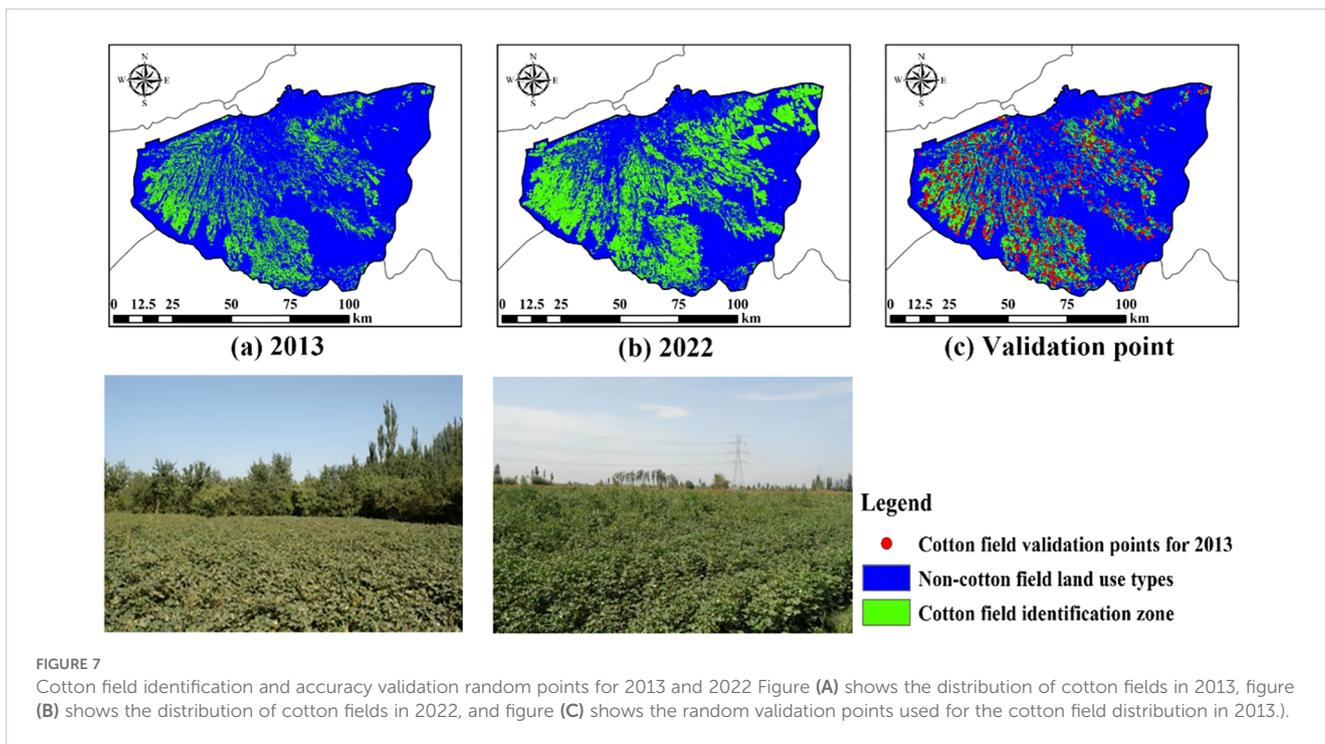
GDVI indices are highly sensitive to changes in vegetation. In this study, the EVI vegetation index also showed good sensitivity in the study area. The EVI index can eliminate the influence of soil and atmospheric aerosols, especially suitable for densely vegetated areas in oases, to obtain more accurate spectral information of vegetation

canopy and assist in determining the soil salinity information under vegetation cover. A study conducted by Nguyen et al. (2021) in the Mekong Delta of Vietnam revealed that 13 indices, including EVI, CRSI, B5, B3, and B7, are the most effective in measuring soil salinity. The results of this study indicate that vegetation factors

FIGURE 8
Cotton field expansion and improvement of saline-alkali soil area Figure **(A)** represents the distribution changes of cotton fields from 2013 to 2022, figure **(B)** shows the soil salinity levels in newly added cotton fields in 2013, and figure **(C)** shows the soil salinity levels in newly added cotton fields in 2022.

FIGURE 9
Cotton field yield (per pixel) in 2013 and 2022. **(A)** represents cotton field yield in 2013, and **(B)** represents cotton field yield in 2022.

play an important role in improving model accuracy and monitoring soil salinity.

Salt index and soil index have been proven to be key variables in soil prediction in arid areas in previous studies. Wang et al. (2022a) pointed out in their another research that soil salinity directly changes the spectral reflectance in various bands. GVMI, GYEX, and DEM are among the factors with the highest contribution rates. The reason why DEM was not considered as a feature variable in this study is that the altitude fluctuation in the Wei-Ku oasis is not significant, and the differences are small, resulting in a low contribution rate to the prediction of soil salinity. The favorable performance of the salt index in this study corroborates the findings of Peng et al. (2019), namely, the salt index is the optimal and effective method for soil salinity monitoring in southern Xinjiang. Racetin et al. (2020) pointed out that salt indices S1 and S2 have the highest positive correlation, which can characterize the degree of salt influence in irrigated agricultural areas. In this study, S1 and S2, two salinity indices, also have relatively large contribution rates. S1 and S2 use the ratio of red band and blue band information, highlighting that in vegetation-covered areas, the red band has low absorption reflectance and a large ratio, while the opposite is true in desert areas.

The spectral characteristics of saline encrustations are correlated with soil roughness. The reflectance values increase with increasing soil salinity (Günal et al., 2021). The high reflectance characteristic of saline soils is readily discernible by satellite remote sensing, rendering

it an invaluable tool for the remote monitoring of large-scale soil salinization. Gypsum desert soils are widely distributed in Xinjiang, with almost zero surface vegetation cover. Under intense wind erosion conditions, the gypsum layer often approaches or exposes the surface, making it prone to salt accumulation and the formation of salt crusts. Shortwave infrared produces strong reflections in these areas, resulting in relatively high gypsum index values. This suggests that the GVMI and GYEX indices are more effective at detecting soil salinity in arid areas with minimal vegetation cover or bare soil surfaces compared to other indices.

4.2 The impact of different models on soil electrical conductivity values

There is currently no universally acknowledged parameter solution for evaluating the effectiveness of machine learning (including deep learning) results, thus requiring a case-by-case analysis (Domingos, 2012). The performance of most algorithms on specific datasets highly depends on the learning parameters used to train them, yet parameter settings that yield optimal performance on one dataset may not generalize well to another (Bengio, 2012). From a single-model analysis perspective, LSTM is suitable for handling sequential data, capable of capturing long-term dependencies within the data, thereby aiding in better future

FIGURE 10
Spatial heterogeneity relationship between cotton field yield and soil salinity in 2013 and 2022. **(A)** represents cotton field yield and soil salinization spatial heterogeneity in 2013, and **(B)** represents cotton field yield and soil salinization spatial heterogeneity in 2022.
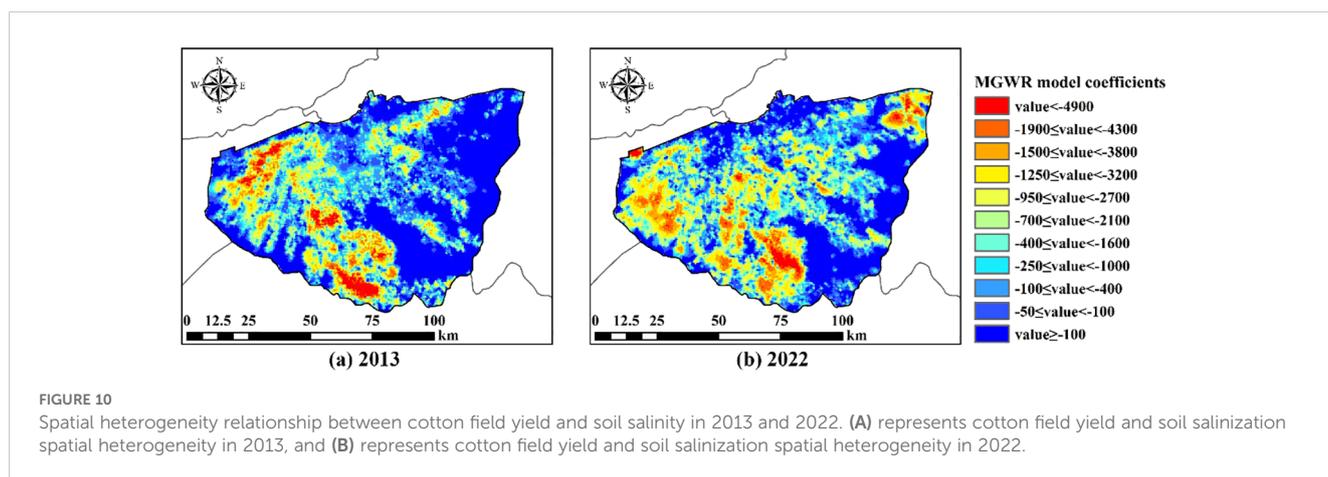
value prediction. Hence, we made the initial attempt to predict soil salinity using LSTM in this study. Random forests, on the other hand, can handle a large number of features and exhibit a certain degree of robustness to missing values and outliers. They are effective in illustrating trends between low and high salinity areas, yet their ability to predict local soil EC values is not as prominent as CNN. Therefore, as a relatively mature deep learning model, CNN presents potential for exploring local information extraction across different study areas in future research, facilitating a better integration of soil salinity research with deep learning techniques.

4.3 Spatial heterogeneity response of soil salinization in cotton fields

The research findings indicate that the expansion of cotton fields can ameliorate soil salinization issues in the Wei-Ku oasis, particularly evident in reclaimed wasteland. To investigate the spatial heterogeneity of soil salinity and cotton field growth, the MGWR model was selected as the optimal method to capture these features.

The MGWR coefficients elucidate the spatial heterogeneity response of soil salinization to cotton growth, offering more detailed insights into the spatial relationship between the two factors. Based on the regression coefficients of soil EC values and cotton yield, we found a close correlation between soil EC values and cotton growth. Xinjiang has become an important cotton production base in China (Yang et al., 2020). For over two decades, cotton cultivation has advocated and employed plastic film-covered drip irrigation (Yang et al., 2022). It has been demonstrated that the duration of drip irrigation is a significant factor in the accumulation of soil salinity in cotton fields (Guan et al., 2019). In winter without irrigation, a large amount of salt concentrated in the subsoil will be redistributed upward, which leads to the accumulation of salt in the topsoil again in the spring (Qin et al., 2021), which is also the reason for the increase in salt in part of the inner oasis in result 3.1.3 (Zong et al., 2022). This in turn leads to reduced crop yields and reduced soil tillage capacity (Ajay, 2015). Consequently, cotton fields require substantial irrigation water. The prolonged percolation effects of long-term drip or surface irrigation cause soil salinity to migrate to the periphery of the oasis with the river flow. Thus, reducing irrigation water in the region may significantly increase soil salinization levels, leading to soil quality degradation, consistent with the findings of Yang et al (2022). Localized analysis reveals a widespread high negative correlation in Shaya County and the western part of Xinhe County. This area conforms to the aforementioned irrigation characteristics, with cotton yields relatively lower than in other regions. This suggests that changes in soil salinization levels are more likely to affect cotton growth in this area. The main reasons may be (1) the quality of irrigation water is partly from shallow groundwater, whose dissolved salt ions exacerbate the accumulation of salts. (2) The groundwater table is shallow and salt ions accumulate upwards through evaporation. (3) Drip irrigation not only fails to ensure sufficient irrigation water, but also keeps the soil in a relatively wet state, which is more conducive to the collection of salts in the upper layers. Therefore, a large amount of irrigation water is also needed to wash salt. Additionally, in newly cultivated cotton fields such as the northwest and northeast regions of the study area, the high

sensitivity to soil EC values may be attributed to sparse surrounding vegetation cover, arid climate, and imperfect irrigation and salinity control measures, resulting in a strong response of cotton fields to soil salinity (Wang et al., 2023a). For cotton fields that are highly responsive to soil salinity, the following measures can be taken to reduce the impact of soil salinity on cotton growth. (1) Application of different types of organic, inorganic, and mixed amendments (Wang et al., 2022b). Organic amendments mainly include compost, humus, and plant residues (Scotti et al., 2015). They can increase soil organic matter (OM) content, thereby enhancing soil permeability (Chen et al., 2020). Additionally, they are rich in nitrogen and phosphorus, which can enhance the diversity of soil bacterial communities and strengthen interactions among microorganisms (Mao et al., 2022). Inorganic amendments mainly include gypsum, desulfurized gypsum, and aluminum sulfate, which are minerals or synthetically prepared compounds (Zhou et al., 2019). They alter the chemical properties of saline-alkaline soils through neutralization or ionic balance adjustments, thus improving the soil environment (Nan et al., 2016). Mixing organic and inorganic amendments can combine their advantages to achieve more effective results (Wang et al., 2024a). (2) Intercropping of cotton with halophytes. The main halophytes include Suaeda salsa and alfalfa (Medicago) (Díaz et al., 2018; Song and Wang, 2015). These plants can improve the physical and chemical properties of the soil by reducing bulk density, increasing soil porosity, and enhancing hydraulic conductivity (Ashraf et al., 2010),thereby promoting the leaching of soil salinity in saline-alkaline soils and reducing soil salt accumulation (Liang and Shi, 2021). (3) Adopt the Deep Vertical Rotary Tillage technique for cotton cultivation (Bai et al., 2024). Deep tillage creates a more favorable soil environment for root growth and development (Wang et al., 2020). It increases soil porosity, enhances the downward migration of salts, and limits the upward movement of water in soil capillaries, thereby reducing soil salinity and achieving the goal of increasing crop yields (Wei et al., 2020).(4) The impact of winter irrigation and water resource management (Li et al., 2020; Yang et al., 2022; Zhao et al., 2021). Winter irrigation effectively conserves soil moisture and inhibits salinity by pushing soil salts deeper into the soil layers, preventing salt return due to evaporation in the following year, thus providing favorable soil moisture and salinity conditions for crop growth (Xiao et al., 2018). Water resource management measures using subsurface drainage can reduce soil salinity and improve cotton emergence rates (Feng et al., 2019). Meanwhile, combining subsurface drainage with surface drainage can maximize the effectiveness of subsurface drainage. Therefore, improving the construction of drainage canals and enhancing water and fertilizer management (Wang et al., 2024b) can not only reduce the impact of soil salinization on cotton fields but also ameliorate soil salinization issues, improve soil quality and land productivity, and achieve scientific sustainable development.

4.4 Research limitations and future research directions

This study collected a total of 97 field samples; however, there is still a lack of soil salinization information validation in areas that are difficult to reach due to limited manpower (such as uninhabited

areas or the interior of deserts). Additionally, due to the limited explanatory power of the model, these uninhabited areas still present some uncertainty. Furthermore, the model was constructed using primarily summer remote sensing imagery data, rendering it more suitable for mapping the spatial distribution of soil salinity for the same season across different years. Since the soil salt data we used were sampled in summer, further research will be conducted in different periods of the same year, and more measured samples will be collected to establish soil salinization inversion models respectively, and explore the spatial distribution characteristics of soil salt in different seasons of the year. Additionally, due to the limitations of the study period, higher-resolution satellite imagery, such as the 10-meter resolution Sentinel-1 and Sentinel-2 satellites and the 3-meter resolution Planet satellite, could not be utilized for soil salinity prediction and cotton field classification. Future research could consider the integration of multiple remote sensing images to enhance the accuracy of soil salinity prediction models and cotton field classification. To clarify the factors influencing cotton field yield, future research will consider the spatial heterogeneity of various influencing factors, including soil moisture content, precipitation, temperature, and groundwater level. Data-driven and inductive machine learning can capture information and extract patterns from the ever-growing geospatial data streams, demonstrating strong data adaptability; however, it lacks theoretical support and has weak interpretability. Future research could explore how to combine remote sensing physical process modeling with flexible data-driven modeling, aiming to develop a dual-driven quantitative remote sensing model that couples physical mechanisms with machine learning models, which may help solve the challenges of quantitative analysis of remote sensing data.

5 Conclusions

Based on the measured sample data and remote sensing image data, various deep learning and machine learning inversion models were constructed to map soil salinity in the Wei-Ku oasis, and the changes in soil EC values and spatial responses were investigated in relation to the expansion of cotton fields. The main results obtained are as follows:

(1) The CNN model showed higher accuracy as well as applicability, with an R^2 of 0.84 for the training set and 0.73 for the test set. it can better mine the spatially localized information of soil salinity compared with other models.

(2) The expansion of cotton fields also significantly improved the soil salinity problem in the study area. The area of new cotton fields with heavy salinity and saline soils decreased from $177.91km^2$ and $381.46km^2$ to $19.49km^2$ and $1.12km^2$ from 2013 to 2022.The saline soil in the arid zone can be fully utilized by planting salt-tolerant crops.

(3) In 2013, the low yielding cotton fields were mainly located in the western part of the Wei-Ku oasis. In 2022, the low

yielding areas of the cotton fields appeared in the southern and northeastern part of the newly reclaimed cotton fields in addition to the western part of the study area.

(4) The areas of cotton fields that are more sensitive to salinity alteration were explored by the MGWR model. The highly negatively correlated areas are located in the western and northeastern newly reclaimed areas of the study area, which provides a more scientific theoretical basis for the zonal management of salinization and promotes the sustainable development of oasis ecology and agriculture.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

JZ: Conceptualization, Methodology, Software, Visualization, Writing – original draft. JD: Funding acquisition, Methodology, Writing – review & editing. JW: Supervision, Writing – review & editing. ZZ: Writing – review & editing. JT: Formal analysis, Writing – review & editing. XG: Data curation, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbas, A., Khan, S., Hussain, N., Hanjra, M. A., and Akbar, S. (2013). Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Phys. Chem. Earth Parts A/B/C* 55, 43–52. doi: 10.1016/j.pce.2010.12.004

Abulaiti, A., Nurmemet, I., Muhetaer, N., Xiao, S., and Zhao, J. (2022). Monitoring of soil salinization in the keriya oasis based on deep learning with PALSAR-2 and landsat-8 datasets. *Sustainability* 14, 2666. doi: 10.3390/su14052666

Ajay, S. (2015). Soil salinization and waterlogging: a threat to environment and agricultural sustainability. *Ecol. Indicators.* 57, 128–130. doi: 10.1016/j.ecolind.2015.04.027

Al-Ali, Z. M., Bannari, A., El-Battay, A., and Hameid, N. (2021). "Potionential of spectral indices for halophyte vegetation cover detection in arid and salt-affected landscape," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.* 4632–4635 (IEEE). doi: 10.1109/IGARSS47720.2021.9554101

Allbed, A., Kumar, L., and Aldakheel, Y. Y. (2014). Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region. *Geoderma* 230, 1–8. doi: 10.1016/j.geoderma.2014.03.025

Amarasinghe, K., Marino, D. L., and Manic, M. (2017). "Deep neural networks for energy load forecasting," in *2017 IEEE 26th international symposium on industrial electronics (ISIE).* 1483–1488 (IEEE). doi: 10.1109/ISIE.2017.8001465

Ashraf, M. Y., Ashraf, M., Mahmood, K., Akhter, J., Hussain, F., and Arshad, M. (2010). Phytoremediation of saline soils for sustainable agricultural productivity. *Plant adaptation phytoremediation*, 335–355. doi: 10.1007/978-90-481-9370-7

Bai, Z., Li, Z., Li, L., Li, P., Gong, P., Wang, T., et al. (2024). Deep vertical rotary tillage reduced soil salinity and improved seed cotton yield and water productivity under limited irrigation in saline-alkaline fields. *Ind. Crops Products* 218, 118943. doi: 10.1016/j.indcrop.2024.118943

Bannari, A., Guedon, A., El-Harti, A., Cherkaoui, F., and El-Ghmari, A. (2008). 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.* 39, 2795–2811. doi: 10.1080/00103620802432717

Bengio, Y. (2012). *Practical recommendations for gradient-based training of deep architectures, Neural networks: Tricks of the trade: Second edition* (Berlin, Heidelberg: Springer), 437–478.

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Cai, X., and Sharma, B. R. (2010). Integrating remote sensing, census and weather data for an assessment of rice yield, water consumption and water productivity in the Indo-Gangetic river basin. *Agric. Water Manage.* 97, 309–316. doi: 10.1016/j.agwat.2009.09.021

Chakraborty, S., Datta, H. N., and Chakraborty, S. (2023). Grey relational analysis-based optimization of machining processes: A comprehensive review. *Process Integration Optimization Sustainability* 7, 609–639. doi: 10.1007/s41660-023-00311-4

Chen, A., He, H., Wang, J., Li, M., Guan, Q., and Hao, J. (2019). A study on the arable land demand for food security in China. *Sustainability* 11, 4769. doi: 10.3390/su11174769

Chen, X., Ding, J., Ge, X., Wang, F., and Wang, J. (2021). Scale ef-fect on soil salinization simulation in arid oasis based on ma-chine learning methods. *Trans. Chin. Soc. Agric. Machinery* 52, 312–320. doi: 10.6041/j.issn.1000-1298.2021.09.035

Chen, X., Yaa, O.-K., and Wu, J. (2020). Effects of different organic materials application on soil physicochemical properties in a primary saline-alkali soil. *Eurasian Soil Sci.* 53, 798–808. doi: 10.1134/S1064229320060034

Chengzhi, F. A. N., Ziwen, W. A. N. G., Xingchao, Y. A. N. G., Yongkai, L. U. O., Xuexin, X. U., Bin, G. U. O., et al. (2022). Machine learning inversion model of soil salinity in the yellow river delta based on field hyperspectral and UAV multispectral data. *Smart Agriculture* 4, 61–73. doi: 10.12133/j.smartag.SA202212001

Díaz, F., Grattan, S. R., Reyes, J. A., de la Roza-Delgado, B., Benes, S. E., Jiménez, C., et al. (2018). Using saline soil and marginal quality water to produce alfalfa in arid climates. *Agric. Water Manage.* 199, 11–21. doi: 10.1016/j.agwat.2017.12.003

Domingos, P. (2012). A few useful things to know about machine learning. *Commun. ACM* 55, 78–87. doi: 10.1145/2347736.2347755

Du, R., Chen, J., Zhang, Z., Xu, Y., Zhang, X., Yin, H., et al. (2021). Inversing soil salinity under vegetation cover using Sentinel-2 multispectral satellite remote sensing. *Trans. Chin. Soc Agric. Eng.* 37, 107–115. doi: 10.11975/j.issn.1002-6819.2021.17.012

Feng, G., Zhang, Z., and Zhang, Z. (2019). Evaluating the sustainable use of saline water irrigation on soil water-salt content and grain yield under subsurface drainage condition. *Sustainability* 11, 6431. doi: 10.3390/su11226431

Fotheringham, A. S., Yang, W., and Kang, W. (2017). Multiscale geographically weighted regression (MGWR). *Ann. Am. Assoc. Geographers* 107, 1247–1265. doi: 10.1080/24694452.2017.1352480

Ghazali, M. F., Wikantika, K., and Harto, A. B. (2022). Potential use of spectral analysis to delineate coastal boundary of a landmass based on estimation soil salinity and salt water intrusion: a preliminary result. *J. Coast. Conserv.* 26, 1–16. doi: 10.1007/s11852-022-00871-6

Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. doi: 10.1016/S0034-4257(96)00072-7

Gogtay, N. J., and Thatte, U. M. (2017). Principles of correlation analysis. *J. Assoc. Physicians India* 65, 78–81.

Gong, L., Han, L., Ren, M., and Gui, D. (2012). Spatial variability of soil water-salt in a typical oasis on the upper reaches of the Tarim River. *J. Soil Water Conserv.* 26, 251–255. doi: 10.13870/j.cnki.stbcxb.2012.04.047

Guan, Z., Jia, Z., Zhao, Z., and You, Q. (2019). Dynamics and distribution of soil salinity under long-term mulched drip irrigation in an arid area of northwestern China. *Water* 11, 1225. doi: 10.3390/w11061225

Günal, E., Wang, X., Kılıç, O. M., Budak, M., Al Obaid, S., Ansari, M. J., et al. (2021). Potential of Landsat 8 OLI for mapping and monitoring of soil salinity in an arid region: a case study in Dushak, Turkmenistan. *PloS One* 16, e0259695. doi: 10.1371/journal.pone.0259695

Guo, G., Wu, Z., Xiao, R., Chen, Y., Liu, X., and Zhang, X. (2015). Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landscape Urban Plann.* 135, 1–10. doi: 10.1016/j.landurbplan.2014.11.007

Guo, B., Lu, M., Fan, Y., Wu, H., Yang, Y., and Wang, C. (2023). A novel remote sensing monitoring index of salinization based on three-dimensional feature space model and its application in the Yellow River Delta of China. *Geomatics Natural Hazards Risk* 14, 95–116. doi: 10.1080/19475705.2022.2156820

Hao, J., Lin, Y., Ren, G., Yang, G., Han, X., Wang, X., et al. (2021). Comprehensive benefit evaluation of conservation tillage based on BP neural network in the Loess Plateau. *Soil Tillage Res.* 205, 104784. doi: 10.1016/j.still.2020.104784

Haq, Y.U., Shahbaz, M., Asif, H. M. S., Al-Laith, A., and Alsabban, W. H. (2023). Spatial mapping of soil salinity using machine learning and remote sensing in kot addu, Pakistan. *Sustainability* 15, 12943. doi: 10.3390/su151712943

Hardie, M., and Doyle, R. (2012). Measuring Soil Salinity. In: Shabala, S., and Cuin, T. (eds) *Plant Salt Tolerance. Methods in Molecular Biol*, vol 913. Humana Press, Totowa, NJ. doi: 10.1007/978-1-61779-986-0_28

He, B., Ding, J., Liu, B., and Wang, J. (2019). Spatiotemporal variation of soil salinization in Weigan-Kuqa River Delta Oasis. *Scientia Silvae Sinicae* 55, 185–196. doi: CNKI:SUN:LYKE.0.2019-09-020

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Hong, G., Bai, T., Wang, X., Li, M., Liu, C., Cong, L., et al. (2023). Extraction and analysis of soil salinization information in an alar reclamation area based on spectral index modeling. *Appl. Sci.* 13, 3440. doi: 10.3390/app13063440

Hu, M., Miao, Q., Shi, H., Feng, W., Hou, C., Yu, C., et al. (2023). Spatial variations and distribution patterns of soil salinity at the canal scale in the hetao irrigation district. *Water* 15, 3342. doi: 10.3390/w15193342

Huang, J., Wang, R., Zhao, Z., and Zhang, H. (2006). Effects of Climate change on soil salinization in Xinjiang Oasis, China. (eds.) Principles and practices of desertification control——*Proceedings of the International Specialty Conference on Science and Technology for Desertification Control*. Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences; Graduate School, Chinese Academy of Sciences. 1, 312–319.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213. doi: 10.1016/S0034-4257(02)00096-2

Jia, P., Zhang, J., He, W., Hu, Y., Zeng, R., Zamanian, K., et al. (2022). Combination of hyperspectral and machine learning to invert soil electrical conductivity. *Remote Sens.* 14, 2602. doi: 10.3390/rs14112602

Khan, N. M., Rastoskuev, V. V., Sato, Y., and Shiozawa, S. (2005). Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators. *Agric. Water Manage.* 77, 96–109. doi: 10.1016/j.agwat.2004.09.038

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 2278–2324. doi: 10.1109/5.726791

Li, Y., Chang, C., Wang, Z., and Zhao, G. (2023). Upscaling remote sensing inversion and dynamic monitoring of soil salinization in the Yellow River Delta, China. *Ecol. Indic.* 148, 110087. doi: 10.1016/j.ecolind.2023.110087

Li, L., Liu, H., He, X., Lin, E., and Yang, G. (2020). Winter irrigation effects on soil moisture, temperature and salinity, and on cotton growth in salinized fields in northern Xinjiang, China. *Sustainability* 12, 7573. doi: 10.3390/su12187573

Liang, J., and Shi, W. (2021). Cotton/halophytes intercropping decreases salt accumulation and improves soil physicochemical properties and crop productivity in saline-alkali soils under mulched drip irrigation: A three-year field experiment. *Field Crops Res.* 262, 108027. doi: 10.1016/j.fcr.2020.108027

Liu, H., Meng, L., Zhang, X., Susan, U., Ning, D., and Sun, S. (2015). Estimation model of cotton yield with time series Landsat images. *Trans. Chin. Soc. Agric. Eng.* 31, 215–220. doi: 10.11975/j.issn.1002-6819.2015.17.028

Ma, G., Ding, J., Han, L., Zhang, Z., and Ran, S. (2021). Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms. *Regional Sustainability* 2, 177–188. doi: 10.1016/j.regsus.2021.06.001

Ma, S., He, B., Ge, X., and Luo, X. (2023). Spatial prediction of soil salinity based on the Google Earth Engine platform with multitemporal synthetic remote sensing images. *Ecol. Inf.* 75, 102111. doi: 10.1016/j.ecoinf.2023.102111

Ma, Y., and Tashpolat, N. (2023). Current status and development trend of soil salinity monitoring research in China. *Sustainability* 15, 5874. doi: 10.3390/su15075874

Mao, X., Yang, Y., Guan, P., Geng, L., Ma, L., Di, H., et al. (2022). Remediation of organic amendments on soil salinization: Focusing on the relationship between soil salts and microbial communities. *Ecotoxicology Environ. Saf.* 239, 113616. doi: 10.1016/j.ecoenv.2022.113616

Meng, L., Liu, H., Zhang, X., Ren, C., Ustin, S., Qiu, Z., et al. (2019). Assessment of the effectiveness of spatiotemporal fusion of multi-source satellite images for cotton yield estimation. *Comput. Electron. Agric.* 162, 44–52. doi: 10.1016/j.compag.2019.04.001

Moisa, M. B., Gabissa, B. T., Hinkosa, L. B., Dejene, I. N., and Gemeda, D. O. (2022). Analysis of land surface temperature using geospatial technologies in Gida Kiremu, Limu, and Amuru District, Western Ethiopia. *Artif. Intell. Agric.* 6, 90–99. doi: 10.1016/j.aiia.2022.06.002

Moradi, F., Darvishsefat, A. A., Pourrahmati, M. R., Deljouei, A., and Borz, S. A. (2022). Estimating aboveground biomass in dense hyrcanian forests by the use of sentinel-2 data. *Forests* 13, 104. doi: 10.3390/f13010104

Muhetaer, N., Nurmemet, I., Abulaiti, A., Xiao, S., and Zhao, J. (2022). An efficient approach for inverting the soil salinity in Keriya Oasis, northwestern China, based on the optical-radar feature-space model. *Sensors* 22, 7226. doi: 10.3390/s22197226

Nan, J., Chen, X., Wang, X., Lashari, M. S., Wang, Y., Guo, Z., et al. (2016). Effects of applying flue gas desulfurization gypsum and humic acid on soil physicochemical properties and rapeseed yield of a saline-sodic cropland in the eastern coastal area of China. *J. soils sediments* 16, 38–50. doi: 10.1007/s11368-015-1186-3

Nguyen, T. G., Tran, N. A., Vu, P. L., Nguyen, Q. H., Nguyen, H. D., and Bui, Q. T. (2021). 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. doi: 10.1016/j.geodrs.2021.e00424

Nicolas, H., and Walter, C. (2006). Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data. *Geoderma* 134, 217–230. doi: 10.1016/j.geoderma.2005.10.009

Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., et al. (2019). Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma* 337, 1309–1319. doi: 10.1016/j.geoderma.2018.08.006

Polivova, M., and Brook, A. (2021). Detailed investigation of spectral vegetation indices for fine field-scale phenotyping. *Vegetation index dynamics*, 103. doi: 10.5772/INTECHOPEN.96882

Prout, J. M., Shepherd, K. D., McGrath, S. P., Kirk, G. J., and Haefele, S. M. (2021). What is a good level of soil organic matter? An index based on organic carbon to clay ratio. *Eur. J. Soil Sci.* 72, 2493–2503. doi: 10.1111/ejss.13012

Qin, Y., Bai, Y., Chen, G., Liang, Y., Li, X., Wen, B., et al. (2021). The effects of soil freeze–thaw processes on water and salt migrations in the western Songnen Plain, China. *Sci. Rep.* 11, 3888. doi: 10.1038/s41598-021-83294-x

Racetin, I., Krtalic, A., Srzic, V., and Zovko, M. (2020). Characterization of short-term salinity fluctuations in the Neretva River Delta situated in the southern Adriatic Croatia using Landsat-5 TM. *Ecol. Indic.* 110, 105924. doi: 10.1016/j.ecolind.2019.105924

Rafik, A., Ibouh, H., El Alaoui El Fels, A., Eddahby, L., Mezzane, D., Bousfoul, M., et al. (2022). Soil salinity detection and mapping in an environment under water stress between 1984 and 2018 (Case of the largest oasis in africa-Morocco). *Remote Sens.* 14, 1606. doi: 10.3390/rs14071606

Rongjiang, Y., and Jingsong, Y. (2009). Robustness analysis on spatial variability of apparent soil electrical conductivity in Yellow River Delta. *J. Liaoning Tech. University (Natural Science).* 28, 326–329. doi: 10.3969/j.issn.1008-0562.2009.02.046

Scotti, R., Bonanomi, G., Scelza, R., Zoina, A., and Rao, M. (2015). Organic amendments as sustainable tool to recovery fertility in intensive agricultural systems. *J. Soil Sci. Plant Nutr.* 15, 333–352. doi: 10.4067/S0718-95162015005000031

Scudiero, E., Skaggs, T. H., and Corwin, D. L. (2015). Regional-scale soil salinity assessment using Landsat ETM+ canopy reflectance. *Remote Sens. Environ.* 169, 335–343. doi: 10.1016/j.rse.2015.08.026

Shabiti, M., Maimaiti, Y., and Jiang-Ling, H. U. (2008). Study on ecological water consumption in the ogan-kuqa river delta oasis in xinjiang. *Arid Zone Res.* 3, 325–330. doi: 10.13866/j.azr.2008.03.014

Sheikh, A. T., Mugera, A., Pandit, R., Burton, M., and Davies, S. (2022). What determines the time to gypsum adoption to remediate irrigated salt-affected agricultural lands? Evidence from Punjab, Pakistan. *Soil Tillage Res.* 217, 105266. doi: 10.1016/j.still.2021.105266

Song, B., and Park, K. (2020). Detection of aquatic plants using multispectral UAV imagery and vegetation index. *Remote Sens.* 12, 387. doi: 10.3390/rs12030387

Song, J., and Wang, B. (2015). Using euhalophytes to understand salt tolerance and to develop saline agriculture: Suaeda salsa as a promising model. *Ann. Bot.* 115, 541–553. doi: 10.1093/aob/mcu194

Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., and Malone, B. (2014). Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma* 213, 15–28. doi: 10.1016/j.geoderma.2013.07.020

Tian, Q., and Min, X. (1998). Advances in study on vegetation indices. *Advance In Earth Sci.* 4, 10–16. doi: 10.3321/j.issn:1001-8166.1998.04.002

Wang, Y.-X., Chen, S. P., Zhang, D. X., Li, Y. A. N. G., Tao, C. U. I., Jing, H. R., et al. (2020). Effects of subsoiling depth, period interval and combined tillage practice on soil properties and yield in the Huang-Huai-Hai Plain, China. *J. Integr. Agric.* 19, 1596–1608. doi: 10.1016/S2095-3119(19)62681-X

Wang, N., Peng, J., Xue, J., Zhang, X., Huang, J., Biswas, A., et al. (2022a). A framework for determining the total salt content of soil profiles using time-series Sentinel-2 images and a random forest-temporal convolution network. *Geoderma* 409, 115656. doi: 10.1016/j.geoderma.2021.115656

Wang, Y., Xing, W., Liang, X., Xu, Y., Wang, Y., Huang, Q., et al. (2022b). Effects of exogenous additives on wheat Cd accumulation, soil Cd availability and physicochemical properties in Cd-contaminated agricultural soils: a meta-analysis. *Sci. Total Environ.* 808, 152090. doi: 10.1016/j.scitotenv.2021.152090

Wang, Y., Shi, L., Hu, Y., Hu, X., Song, W., and Wang, L. (2023b). A comprehensive study of deep learning for soil moisture prediction. *Hydrology Earth System Sci. Discussions* 2023, 1–38. doi: 10.5194/hess-2023-177

Wang, X., Ding, J., Wang, J., Han, L., Tan, J., and Ge, X. (2024a). Ameliorating saline-sodic soils: A global meta-analysis of field studies on the influence of exogenous amendments on crop yield. *Land Degradation Dev.* 35, 3330–3343. doi: 10.1002/ldr.v35.10

Wang, J., Hu, B., Liu, W., Luo, D., and Peng, J. (2023a). Characterizing soil profile salinization in cotton fields using landsat 8 time-series data in southern xinjiang, China. *Sensors* 23, 7003. doi: 10.3390/s23157003

Wang, Y., Qu, Z., Yang, W., Chen, X., and Qiao, T. (2024b). Inversion of soil salinity in the irrigated region along the southern bank of the yellow river using UAV multispectral remote sensing. *Agronomy* 14, 523. doi: 10.3390/agronomy14030523

Wang, F., Yang, S., Yang, W., Yang, X., and Jianli, D. (2019). Comparison of machine learning algorithms for soil salinity predictions in three dryland oases located in Xinjiang Uyghur Autonomous Region (XJUAR) of China. *Eur. J. Remote Sens.* 52, 256–276. doi: 10.1080/22797254.2019.1596756

Wang, W., Zhang, X., Zhao, Y., and Wang, S. (2017). Cotton extraction method of integrated multi-features based on multitemporal Landsat 8 images. *J. Remote Sens* 21, 115–124. doi: 10.11834/jrs.20175317

Wei, B., Shen, Z. Y., Zhou, J., Zhou, L. Z., Hu, P., and Zhang, X. (2020). Study on effect and mechanism of improving saline-alkali soil by Fenlong tillage. *Soils* 52, 699–703. doi: 10.13758/j.cnki.tr.2020.04.007

Wu, W., Zucca, C., Muhaimeed, A. S., Al-Shafie, W. M., Fadhil Al-Quraishi, A. M., Nangia, V., et al. (2018). Soil salinity prediction and mapping by machine learning regression in C entral M esopotamia, I raq. *Land degradation Dev.* 29, 4005–4014. doi: 10.1002/ldr.v29.11

Xiao, C., Wu, Y., and Zhu, X. (2023). Evaluation of the monitoring capability of 20 vegetation indices and 5 mainstream satellite band settings for drought in spring wheat using a simulation method. *Remote Sens.* 15, 4838. doi: 10.3390/rs15194838

Xiao, Y. Q., Yang, P. N., Wu, B., Liu, J., and Guo, T. (2018). Analysis of arid oasis areas soil nitrogen accumulation and winter irrigation effect. *Water Sav. Irrig* 2, 71–76.

Xueping, H. (2009). Si-albedo space-based extraction of salinization information in arid area. *Acta Pedologica Sinica.* 46, 381–390. doi: 10.1186/1471-2318-9-49

Yang, Y., Zhou, X., Gao, Y., Li, D., Wang, X., and Li, Z. (2022). Factors influencing usage of subsurface drainage to improve soil desalination and cotton yield in the Tarim Basin oasis in China. *Irrigation Sci.* 40, 857–871. doi: 10.1007/s00271-022-00825-8

Yang, G., Li, F., Tian, L., He, X., and Ren, F. (2020). Soil physicochemical properties and cotton (Gossypium hirsutum L.) yield under brackish water mulched drip irrigation. *Soil Tillage Res.* 199, 104592. doi: 10.1016/j.still.2020.104592

Yu, M.-Y., Xu, Y., Li, J. Q., Lu, X. C., Xing, H. Q., and Ma, M. L. (2021). Geographic detector-based spatiotemporal variation and influence factors analysis of PM2. 5 in shandong, China. *Polish J. Environ. Stud.* 30, 463–475. doi: 10.15244/pjoes/120522

Zhang, H., Fu, X., Zhang, Y., Qi, Z., Zhang, H., and Xu, Z. (2023). Mapping multi-depth soil salinity using remote sensing-enabled machine learning in the yellow river delta, China. *Remote Sens.* 15, 5640. doi: 10.3390/rs15245640

Zhang, Z., Fan, Y., Zhang, A., and Jiao, Z. (2022). Baseline-based soil salinity index (BSSI): A novel remote sensing monitoring method of soil salinization. *IEEE J. Selected*

Topics Appl. Earth Observations Remote Sens. 16, 202–214. doi: 10.1109/JSTARS.2022.3223935

Zhao, L., Heng, T., Yang, L., Xu, X., and Feng, Y. (2021). Study on the farmland improvement effect of drainage measures under film mulch with drip irrigation in saline–alkali land in arid areas. *Sustainability* 13, 4159. doi: 10.3390/su13084159

Zhou, M., Liu, X., Meng, Q., Zeng, X., Zhang, J., Li, D., et al. (2019). Additional application of aluminum sulfate with different fertilizers ameliorates saline-sodic soil of Songnen Plain in Northeast China. *J. Soils Sediments* 19, 3521–3533. doi: 10.1007/s11368-019-02311-9

Zong, R., Han, Y., Tan, M., Zou, R., and Wang, Z. (2022). Migration characteristics of soil salinity in saline-sodic cotton field with different reclamation time in non-irrigation season. *Agric. Water Manage.* 263, 107440. doi: 10.1016/j.agwat.2021.107440