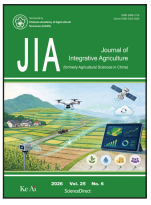
Journal of Integrative Agriculture 2026, 25(6): 2607–2622

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

# A classification modeling strategy based on dominant factors of salinization to enhance remote sensing inversion accuracy

**Mengchao Zheng[1, 2], Jianjun Zhang[2#], Weini Wang[4], Zhigang Qiao[6], Junmei Liu[4], Min Gong[1], Xiaobin Li[1, 3#], Hongyuan Zhang[1, 3], Yuyi Li[1, 3], Ningning Li[5], Lin Yang[2], Wenjuan Li[1]**

[1] *State Key Laboratory of Efficient Utilization of Arid and Semi-arid Arable Land in Northern China, Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China*
[2] *School of Land Science and Technology, China University of Geosciences (Beijing), Beijing 100083, China*
[3] *National Center of Technology Innovation for Comprehensive Utilization of Saline-alkali Land, Dongying 257000, China*
[4] *Ordos Agriculture and Animal Husbandry Ecology and Resource Protection Center, Ordos 017001, China*
[5] *College of Geography and Remote Sensing Sciences, Xinjiang University, Urumqi 830046, China*
[6] *Inner Mongolia Autonomous Region Agriculture and Animal Husbandry Technology Extension Center, Hohhot 010010, China*

**Highlights**

- Internal partition modeling, driven by key factors, enhances model accuracy.
- Among all scenarios, the random forest (RF) model exhibits superior overall performance.
- Drainage conditions and soil texture significantly influence the correlation between modeling factors and soil salt content (SSC).
- A high-precision prediction methodology based on primary salinization drivers is proposed.

**Abstract**

Soil salinization represents a primary manifestation of land degradation and presents a significant threat to sustainable agricultural development. Remote sensing-based methodologies currently constitute the preferred approach for salinization monitoring. Environmental factors' spatial heterogeneity substantially constrains the modeling process in accurately capturing the soil salt content (SSC)-modeling factor relationship, thereby affecting monitoring accuracy. This study proposes a classification modeling framework based on dominant salinization factors, establishing distinct remote sensing inversion models through categorization of soil texture and surface drainage conditions. Results indicate that classification modeling substantially improves the capture of SSC-modeling factor relationships. The efficacy of identical modeling indicators and methods varies significantly across different classification scenarios. Among the three modeling approaches, random forest demonstrates superior overall robustness. Of the three variable selection methods, light gradient boosting machine (LightGBM) shows the strongest compatibility with the modeling approaches. The classification strategy significantly enhances model accuracy: compared to non-classified modeling ($R^2_V$=0.62), the testing set $R^2$ increases by up to 24% ($R^2_V$=0.77). Models under poor surface drainage category demonstrate optimal performance, with coupled models achieving $R^2_C$=0.82 (training set) and $R^2_V$=0.77 (testing set). This research provides valuable insights for remote sensing monitoring of soil salinization in precision agriculture contexts.

**Keywords:** soil salinization, remote sensing, spatial heterogeneity, machine learning, food security

## 1. Introduction

Soil salinization refers to the accumulation of salts on the soil surface, causing alterations in soil physical and chemical properties (Rengasamy 2006). This phenomenon represents a global land degradation issue that significantly threatens the ecological environment and sustainable agricultural

Received 1 April 2025; Received in revised form 13 May 2025; Accepted 2 July 2025; Available online 21 August 2025
Mengchao Zheng, E-mail: zhengmengchao1999@163.com; #Correspondence Xiaobin Li, E-mail: lixiaobin01@caas.cn; Jianjun Zhang, E-mail: j.zhang@cugb.edu.cn

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2025.08.016

development, particularly in arid and semi-arid regions (Daliakopoulos *et al.* 2016; Scudiero *et al.* 2016; Wang *et al.* 2021; Wang N *et al.* 2024). Approximately 20% of global irrigated farmland experiences soil salinization effects (Wang *et al.* 2025). Thus, acquiring accurate and efficient information on soil salinization is crucial for supporting comprehensive saline-alkali land management to ensure food security and promote sustainable agricultural development (Farifteh *et al.* 2007; Wang *et al.* 2025).

Conventional soil salinization monitoring methods depend extensively on manual sampling, limiting large-scale temporal measurements and requiring substantial human and material resources (Chen *et al.* 2022; Cui *et al.* 2023). In contrast, machine learning and remote sensing technologies have rapidly integrated into soil salinization monitoring due to their advantages in monitoring timeliness, cost-effectiveness, and capacity to capture and process nonlinear relationships (Scudiero *et al.* 2015; Wang *et al.* 2019, 2025; Li 2021). Multispectral remote sensing data, characterized by small data volumes and high spatiotemporal resolution (e.g., Sentinel-2 and Landsat satellite series data), have mostly been utilized for soil salinization monitoring (Scudiero *et al.* 2015; Wang *et al.* 2019; Kumar *et al.* 2024). Several studies have enhanced monitoring effectiveness by integrating various remote sensing datasets to fully utilize the advantages of multiple data sources (Chen *et al.* 2021; Li *et al.* 2023; Jia *et al.* 2024). Additionally, researchers have developed and implemented diverse spectral-derived indices, including vegetation indices (VIs) and salinity indices (SIs), to optimize spectral information utilization (Chen *et al.* 2022; Cui *et al.* 2023; Wang *et al.* 2023).

However, spectral indices demonstrate lower stability compared to terrain and texture features in practical applications due to environmental heterogeneity (Vermeulen and Van 2017; Yang *et al.* 2023; He *et al.* 2024), yet texture features have rarely been used. Furthermore, the relationships between certain indicators and soil salt content (SSC) exhibit non-linear (Chen *et al.* 2022). Consequently, various variable selection methods capable of managing nonlinear relationships, including variable importance projection (VIP), light gradient boosting machine (LightGBM), and principal component analysis (PCA), have been implemented (Wang *et al.* 2019; Jiang *et al.* 2022). Similarly, machine learning algorithms adept at handling nonlinear relationships, such as random forest (RF), extreme gradient boosting (XGBoost), support vector machine (SVM), or extremely randomized trees (ET), have been increasingly utilized (Wang N *et al.* 2024; Wang *et al.* 2025). However, the optimal modeling method likely varies by study area (Jiang *et al.* 2022), and limited research has comprehensively evaluated the performance of coupling various machine learning algorithms and variable selection methods under differing environmental conditions.

Environmental heterogeneity results in poor model transferability across study areas. This limitation arises because optical remote sensing-based salinization monitoring primarily relies on surface spectral reflectance obtained by satellites (Chen *et al.* 2022), which is affected by the composite surface environment (Metternicht and Zinck 2003; Farifteh *et al.* 2007). Given that both soil salinization and environmental driving factors display significant spatial heterogeneity (Lu 1999; Wang *et al.* 2023), spectral reflectance varies across different environments (Jiang *et al.* 2022). This variation indirectly causes differences in the relationships between spectral indices and SSC across regions, leading to discrepancies in modeling results and reduced accuracy in salinization monitoring (Cui *et al.* 2023). Chen *et al.* (2022) demonstrated that no single fixed set of modeling features suits all environments, necessitating the selection of modeling features specific to particular environmental conditions for salinization monitoring.

Addressing environmental heterogeneity is crucial for ensuring accurate remote sensing-based salinization monitoring (Salcedo *et al.* 2022; Wang *et al.* 2023). Zhang *et al.* (2021) addressed this challenge by considering regional variations in vegetation coverage, finding that separate classification of vegetation coverage and modeling significantly improved the correlation between SSC and modeling variables, thereby enhancing monitoring accuracy. Cui *et al.* (2023) improved farmland salinization monitoring accuracy by classifying crop types and developing separate models, accounting for spatial heterogeneity in soil salinization and its driving factors in agricultural areas. Although environmental drivers of soil salinization vary by region, research emphasizes that soil texture and surface drainage conditions are fundamental determinants of water-salt transport processes and significantly influence salinization development and spatial distribution in most agricultural areas (Wang J J *et al.* 2016; Ramos *et al.* 2023; Mandal 2024). While previous studies have addressed environmental factors' impact on salinization monitoring, few have specifically examined the influence of key regional salinization factors, such as soil texture and drainage, particularly in agricultural contexts. With the increasing demand for thecomprehensive utilization of saline land and precision agriculture (Cisternas *et al.* 2020; Sanaeifar *et al.* 2023), high-accuracy monitoring ofsalinization is urgently needed to support sustainable agricultural development.

This study proposes a classification modeling strategy based on the dominant factors of salinization, validated in a typical salinized area within the Yellow River irrigation region of Inner Mongolia, China. The study objectives are to: (1) evaluate how classifying soil texture and surface drainage conditions affects SSC remote sensing monitoring accuracy and whether the classification modeling strategy improves monitoring performance; (2) examine relationships between SSC and modeling features, identify optimal features for model construction in each category, and assess texture features' applicability in SSC prediction; and (3) assess effects of coupling different machine learning methods and variable selection techniques across various categories, and establish the optimal monitoring model for each category. This research aims to provide valuable reference for high-quality soil salinization monitoring in agricultural regions.

## 2. Materials and methods

### 2.1. Study area

Wangaizhao Town (40°08′–40°33′N, 110°5′–110°24′E),

a subordinate town of Ordos City, Inner Mongolia, China, features terrain that is elevated in the south and descends northward bordering the Yellow River in the north (Fig. 1-A). The predominant soil texture is loamy clay. The region experiences irregular and sparse annual rainfall, primarily concentrated in summer, with average annual precipitation ranging from 240–360 mm and annual evaporation between 2,100–2,500 mm. The combination of consistently low precipitation, high annual evaporation, shallow water table with high mineralization, and extensive Yellow River lateral erosion, along with widespread irrigation and drainage practices, has resulted in significant soil salinization in this area (Fig. 1-B and C) (Pereira *et al.* 2007; Ren *et al.* 2019).

## 2.2. Materials

**Soil sample data collection** Soil surface samples (0–20 cm) were systematically collected throughout the study area from October 15–18, 2022. The sampling unit was established at 10 m×10 m to correspond with satellite remote sensing spatial resolution. Sample collection utilized the "five-point sampling method," wherein each unit's final sample comprised a composite of five individual subsamples. The methodology is illustrated in Fig. 1-D. The central coordinates of each sampling unit were recorded using a handheld GPS device, and the surrounding environmental conditions were documented *via* mobile devices. The collection yielded 168 soil composite samples, with their spatial distribution shown in Fig. 1-A.

The collected samples underwent air-drying and were sieved through a 2 mm mesh before laboratory analysis. Soil electrical conductivity (EC) and soil texture measurements were conducted. EC (1:5) measurements were obtained from

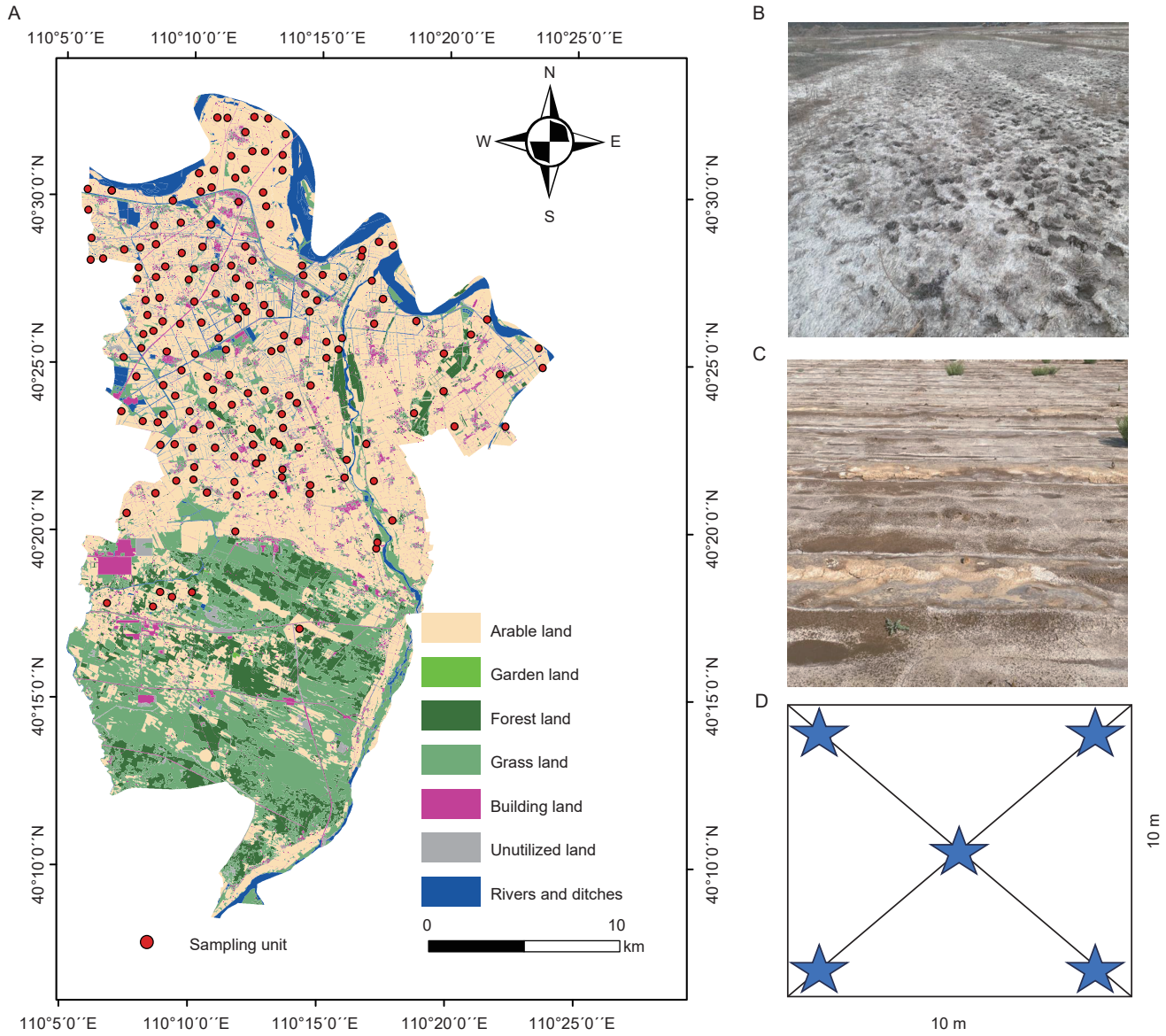



**Fig. 1 Overview of the study area of Wangaizhao Town, Ordos City, Inner Mongolia, China.** A, location characteristics of the study area, land use status of the study area in 2022 (provided by the government of Ordos, Inner Mongolia, China), and distribution of sampling units. B and C, local salinization status. D, details of the five-point sampling method.

saturated mud using a soil to pure water mass ratio of 1:5. To align with local salinity classification criteria, the EC values were converted to SSC (g $kg^{-1}$) using the empirical formula SSC=3.609EC–0.523 (Wang Y *et al.* 2024). Soil texture determination employed the hydrometer method. Given the region's predominant loamy soil type and insufficient clay and sand content in the collected samples for modeling purposes, 89 loam samples were selected from the total 168 samples for further analysis.

**Acquisition and preprocessing of remote sensing images** Sentinel-2 data encompasses 13 bands spanning the visible band, near-infrared to short-wave infrared region, with spatial resolutions ranging from 10–60 m per band (Table 1). Its multi-band and high spatial resolution capabilities enable widespread application in agricultural monitoring, forestry management, natural disaster monitoring, and other fields (Wang and Atkinson 2018; Astola *et al.* 2019; Wang *et al.* 2019).

This study utilized L2A class Sentinel-2 remotely sensed imagery data from October 15, 2022, with less than 10% cloud cover in the study area. According to ESA definitions, L2A class data undergoes orthorectification, geometric refinement, and atmospheric correction (details available at https://senti.com/.copernicus.eu). The remote sensing data underwent resampling to 10 m using bilinear interpolation for non-10 m band images in SNAP software to ensure consistency between image band information and sampling units. B1, B9, and B10 bands, primarily used for atmospheric feature detection (Wang Q *et al.* 2016; Wang *et al.* 2019; Song *et al.* 2024), were excluded from subsequent analysis.

**Table 1 Spectral band characteristics of Sentinel-2 imagery**

| Band | Band center (nm) | Band width (nm) | Spatial resolution (m) |
|---|---|---|---|
| B1 | 443 | 433–453 | 60 |
| B2 | 490 | 458–523 | 10 |
| B3 | 560 | 543–578 | 10 |
| B4 | 665 | 650–680 | 10 |
| B5 | 705 | 698–713 | 20 |
| B6 | 740 | 733–748 | 20 |
| B7 | 783 | 773–793 | 20 |
| B8 | 842 | 785–900 | 10 |
| B8A | 865 | 849–880 | 20 |
| B9 | 945 | 935–955 | 60 |
| B10 | 1,380 | 1,360–1,390 | 60 |
| B11 | 1,610 | 1,565–1,655 | 20 |
| B12 | 2,190 | 2,100–2,280 | 20 |

## 2.3. Methods

This study proposes a classification modeling strategy based on salinity-governing factors and empirically investigates a typical saline area of Inner Mongolia, China. The research methodology, illustrated in Fig. 2, encompasses data acquisition and processing, characterization factor selection, and construction and evaluation of various coupled salinization inversion models.

**Sample category classification** Soil texture classification

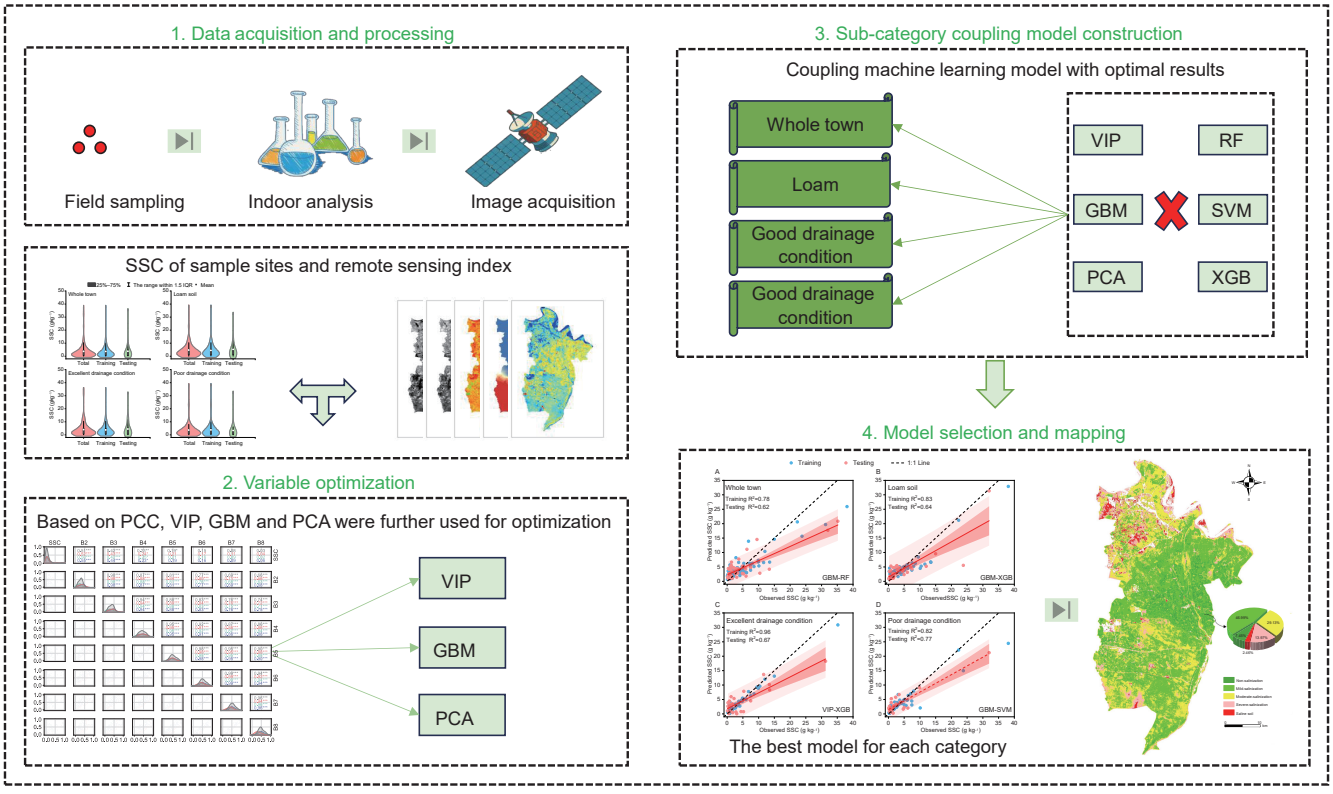



**Fig. 2 Research flowchart.** SSC, soil salt content; PCC, pearson correlation coefficient; VIP, variable importance in projection; GBM, gradient boosting machine; PCA, principal component analysis; RF, random forest model; SVM, support vector machine model; XGB, extreme gradient boosting machine model.

was based on soil texture measurements as described in Section 2.2. Surface drainage condition classification (excellent or poor) relied on environmental information collected during sampling. Sampling points on ridge tillage land or flat arable land with comprehensive drainage facilities were classified as having excellent drainage conditions. Conversely, points on flat cropland lacking drainage facilities were classified as having poor drainage conditions. Seven samples remained unclassified due to the presence of both ridge and flat cultivation in their locations. Notably, some samples categorized as loamy soil also qualified for either good or poor surface drainage condition categories.

**Modeling factors** This study utilized processed remote sensing images to obtain modeling metrics commonly used in previous studies, including image band reflectance, vegetation indices, salinity indices, and terrain factors. The mean values of texture features of the image principal component bands under different extraction windows based on the gray-level covariance matrix (GLCM) were also included. Research has demonstrated that image texture features and terrain factors exhibit greater stability than optical indicators and can substantially enhance model accuracy (Vermeulen and Van Niekerk 2017; Yang *et al.* 2023; He *et al.* 2024). Table 2 presents the specific formulas and references for these indicators.

**Modeling methods** First, we used Pearson correlation analysis (PCC) to select the modeling factors highly correlated with the SSC. Subsequently, variable optimization was performed using variable importance in projection (VIP), distributed gradient boosting machine (LightGBM), and principal component analysis (PCA). VIP calculations were executed using SIMCA 14.1, with variables having VIP values exceeding one considered beneficial for modeling (Wang *et al.* 2019). In LightGBM, higher scores indicate greater variable importance (Jiang *et al.* 2022). To minimize human influence, the LightGBM model parameters remained at default settings during feature optimization. For PCA dimensionality reduction, principal components with a cumulative contribution of 90% were selected (Greenacre *et al.* 2022).

The model construction employed machine learning algorithms that have demonstrated superior performance in existing studies: RF, SVR, and XGBoost. RF algorithm incorporates bootstrap techniques and random feature selection, exhibiting robust performance and generalization capabilities (Wang *et al.* 2019). Support vector machine regression, an application of the SVM, effectively manages highly nonlinear problems in low-dimensional spaces without requiring direct data mapping to high-dimensional spaces (Taghizadeh-Mehrjardi *et al.* 2021). XGBoost represents an efficient regression model based on gradient boosting decision tree (GBDT), incorporating advanced regularization terms and second-order gradient information to achieve enhanced training speed and performance on large datasets compared to traditional GBDT (Niazkar *et al.* 2024).

**Modeling strategy and evaluation** The dataset for each scenario was randomly divided into training and testing sets at a 3:1 ratio during data preparation. Model optimization employed grid search combined with 5-fold cross-validation to determine optimal parameters, ensuring model stability. Model evaluation utilized the root mean square error (root mean square error (RMSE), mean absolute error (mean

**Table 2 Calculation of the metrics utilized in this study and their corresponding references**

| Indicator types | Indicator[1)] | Calculation formula | Reference |
|---|---|---|---|
| Band reflectance | B2–B8A, B11, B12 | – | Wang *et al.* (2019) |
| Vegetation index | NDVI | NDVI=(B8–B4)/(B8+B4) | Jiang *et al.* (2022); Bai *et al.* (2023); Cui *et al.* (2023) |
| | NDWI | NDWI=(B3–B8)/(B3+B8) | Jiang *et al.* (2022) |
| | EVI | EVI=2.5[(B8–B4)/(B8+6B4–7.5B2+1)] | Jiang *et al.* (2022); Cui *et al.* (2023) |
| | SAVI | SAVI=1.5[(B8–B4)/(B8+B4+0.5)] | Jiang *et al.* (2022); Bai *et al.* (2023) |
| | GRVI | GRVI=B8/B3 | Akhtar and Tsuyuzaki (2024) |
| | GNDVI | GNDVI=(B8–B3)/(B8+B3) | Bai *et al.* (2023) |
| Salinity index | SI-T | SI-T=100(B4/B8) | Allbed *et al.* (2014) |
| | SI | $SI=(B2B4)^{0.5}$ | Wang *et al.* (2019); Wang N *et al.* (2024) |
| | SI1 | $SI1=(B3B4)^{0.5}$ | Wang *et al.* (2019); Cui *et al.* (2023) |
| | SI2 | $SI2=(B3B3+B4B4+B8B8)^{0.5}$ | Wang *et al.* (2019); Cui *et al.* (2023) |
| | SI3 | $SI3=(B3B3+B4B4)^{0.5}$ | Wang *et al.* (2019); Cui *et al.* (2023) |
| | S1 | S1=B2/B4 | Wang *et al.* (2019); Wang N *et al.* (2024) |
| | S2 | S2=(B3–B4)/(B3+B4) | Wang *et al.* (2019); Wang N *et al.* (2024) |
| | CRSI | $CRSI=[(B8B4–B3B2)/(B8B4+B3B2)]^{0.5}$ | Cui *et al.* (2023); Wang N *et al.* (2024) |
| | NDSI | NDSI=(B4–B8)/(B4+B8) | Wang *et al.* (2019); Cui *et al.* (2023) |
| Terrain factor | DEM, Slope, Aspect, TWI | – | Vermeulen and Van Niekerk (2017); Jiang *et al.* (2022) |
| Texture | Mean, Variance, Homogeneity, Contrast, Dissimilarity, Entropy, Second Moment, Correlation | – | Yang *et al.* (2023); He *et al.* (2024) |

[1)] B2–B8A, B11, B12, the reflectance of each band; NDVI, normalized difference vegetation index; NDWI, normalized difference water index; EVI, enhanced vegetation index; SAVI, soil adjusted vegetation index; GRVI, green-red vegetation index; GNDVI, green normalized difference vegetation index; SI-T, salinity index-thermal; SI, salinity index; SI1, salinity index 1; SI2, salinity index 2; SI3, salinity index 3; S1, salinity 1; S2, salinity 2; CRSI, canopy response salinity index; NDSI, normalized difference salinity index; DEM, digital elevation model; TWI, topographic wetness index.

absolute error (MAE), and coefficient of determination ($R^2$). Eqs. (1–3) express these calculations (Wang *et al.* 2019).

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i)^2} \quad (1)$$

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-x_i| \quad (2)$$

$$R^2=\frac{\sum_{i=1}^{n}(y_i-\overline{x}_i)^2}{\sum_{i=1}^{n}(x_i-\overline{x}_i)^2} \quad (3)$$

## 3. Results

### 3.1. Statistical analysis of salinity characteristics of soil samples for each category

We classified the collected soil salinity data into five SSC gradients according to the grading criteria of (Wang Y *et al.* 2024): non-saline soils (SSC<1 g kg$^{-1}$), mildly saline (1 g kg$^{-1}$<SSC<3 g kg$^{-1}$), moderately saline (3 g kg$^{-1}$<SSC<6 g kg$^{-1}$), heavily saline (6 g kg$^{-1}$<SSC<15 g kg$^{-1}$), and extremely severely saline (SSC>15 g kg$^{-1}$). Four category scenarios were delineated: whole town, loamy soil, good drainage, and poor drainage. Among the four scenarios, the drainage conditions were classified as good according to whether the surrounding drainage facilities could fully satisfy or fulfill the drainage requirements, and poor according to whether the surrounding drainage facilities could basically satisfy or not fulfill the drainage requirements. The dataset was divided according to the dataset division method described in Section 2.3. The results of the division and its statistical analysis are shown in Fig. 3 and Table 3, respectively.

As shown in Fig. 3, non-saline and mildly saline samples predominated across all four categories, while heavy and extremely severe samples were least frequent. The violin distribution demonstrated that both training and testing sets maintained consistent data distribution with the total set across all four scenarios, indicating a reasonable dataset division. Table 3 reveals that the standard deviation exceeded 5.4 under all four scenarios, with coefficients of variation exceeding 132%, indicating substantial sample variability and strong spatial heterogeneity of soil salinization. The coefficients of variation differed notably among scenarios, with poor drainage conditions showing the highest values and loamy soil displaying the lowest overall values. These findings suggest that surface drainage conditions and loamy soil significantly influenced SSC distribution.

### 3.2. Correlation analysis between factors

The correlation strength between modeling factors and soil salinity significantly influenced model performance. The correlations between various modeling factors and SSC were calculated under the four scenarios, with results presented in Fig. 4 and Appendix A. The analysis revealed varying correlations between vegetation index, salinity index, topography factor, and SSC across the four scenarios, with notable differences in band reflectance. Among the 10 bands reflectance, B2, B3, and B4 demonstrated significant correlations with SSC across all scenarios; B12 showed

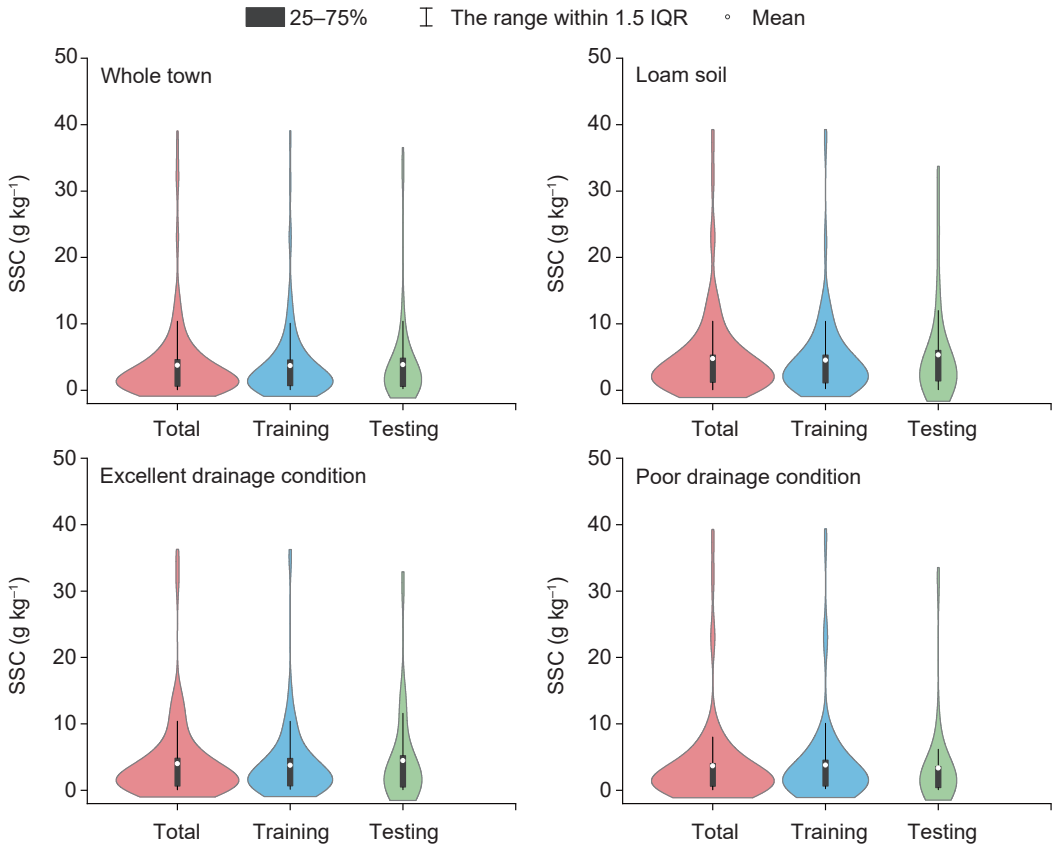



**Fig. 3 Violin plot distribution of each data set under different categories.** SSC, soil salt content.

**Table 3 Descriptive statistics of soil salt content (g kg$^{-1}$) in different data sets under different categories**

| Category | Dataset | Sample size | Mean | Standard deviation | Minimum | Maximum | Variation (%) |
|---|---|---|---|---|---|---|---|
| Whole town | Total | 168 | 3.76 | 6.06 | 0.09 | 38.10 | 161 |
| | Training set | 117 | 3.71 | 5.78 | 0.09 | 38.10 | 156 |
| | Testing set | 51 | 3.86 | 6.72 | 0.23 | 35.20 | 174 |
| Loam soil | Total | 89 | 4.77 | 6.44 | 0.10 | 38.10 | 135 |
| | Training set | 62 | 4.54 | 5.96 | 0.24 | 38.10 | 132 |
| | Testing set | 27 | 5.31 | 7.50 | 0.10 | 32.00 | 141 |
| Excellent drainage condition | Total | 81 | 3.99 | 5.88 | 0.10 | 35.20 | 147 |
| | Training set | 56 | 3.78 | 5.42 | 0.16 | 35.20 | 143 |
| | Testing set | 25 | 4.48 | 6.90 | 0.10 | 31.30 | 154 |
| Poor drainage condition | Total | 80 | 3.68 | 6.48 | 0.09 | 38.10 | 176 |
| | Training set | 56 | 3.81 | 6.51 | 0.23 | 38.10 | 171 |
| | Testing set | 24 | 3.36 | 6.54 | 0.09 | 32.00 | 195 |

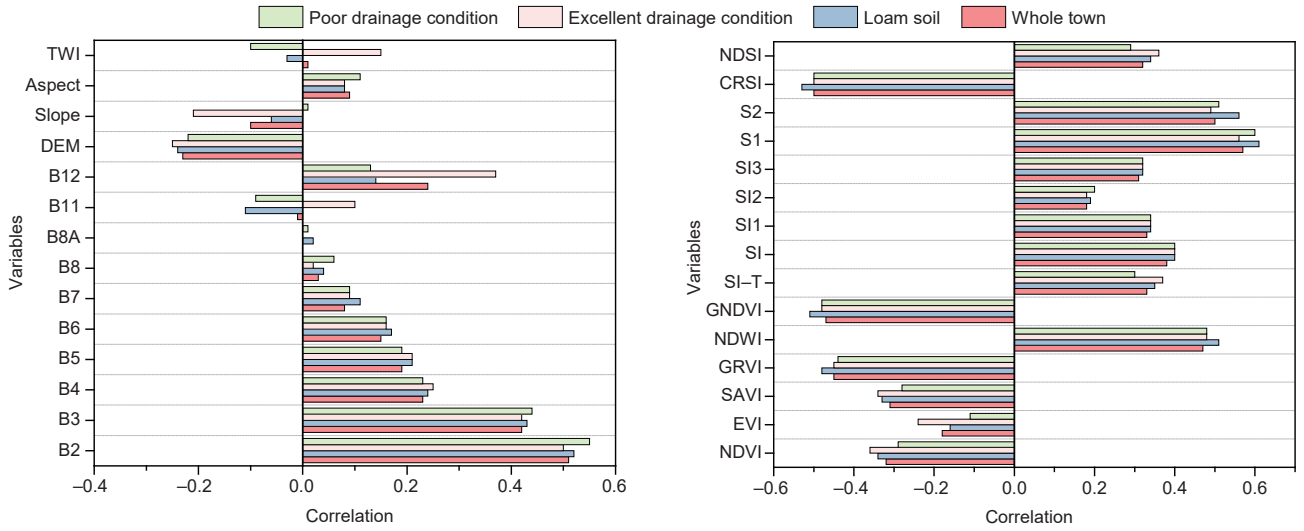



**Fig. 4 Correlation between soil salt content (SSC) and variables under different categories.** TWI, topographic wetness index; DEM, digital elevation model; B12, B11, B8A, B8, B7, B6, B5, B4, B3, B2, the reflectance of each band; NDSI, normalized difference salinity index; CRSI, canopy response salinity index; S2, salinity 2; S1, salinity 1; SI3, salinity index 3; SI2, salinity index 2; SI1, salinity index 1; SI, salinity index; SI-T, salinity index-thermal; GNDVI, green normalized difference vegetation index; NDWI, normalized difference water index; GRVI, green-red vegetation index; SAVI, soil adjusted vegetation index; EVI, enhanced vegetation index; NDVI, normalized difference vegetation index.

significant correlation in the whole town and good drainage scenarios; B5 correlated significantly in the whole town scenario; while remaining bands showed no significant correlation across scenarios. All vegetation indices except EVI significantly correlated with SSC across scenarios, though correlation magnitudes varied. Similarly, salinity indices followed the vegetation indices' pattern, with S1 showing the strongest correlation with SSC, reaching 0.61 under poor drainage, while SI2 consistently showed the weakest correlation. Among topographic factors, only DEM exhibited a low significant correlation with SSC (Fig. 4-A; Appendix A).

Fig. 5 illustrates the broken-line trend of correlations between mean texture features and soil salinity under different windows across scenarios. Overall, the correlation of the same texture features with soil salinity varied significantly across scenarios; for instance, Correlation (7×7, 9×9, 11×11) showed significant correlation with soil salinity in whole town and good drainage scenarios but not in loamy soil and poor drainage scenarios. In poorly drained scenarios, the correlation between mean textural features and soil salinity showed gradual increase after the 11×11 window, contrasting with good drainage scenarios. This difference likely stems from varying rates of surface water accumulation and discharge affecting temporal surface landscape distribution. Well-drained areas exhibited faster water absorption and discharge with minimal short-term surface landscape variation, differing from poorly drained areas. The eight texture features under each window showed consistent correlation patterns with soil salinity: mean, variance, contrast, and dissimilarity demonstrated positive correlations, while the remaining features showed negative correlations.

### 3.3. Optimization of modeling factors

While stronger correlations between modeling factors and target variables generally improve model performance, excessive variables can create information redundancy due

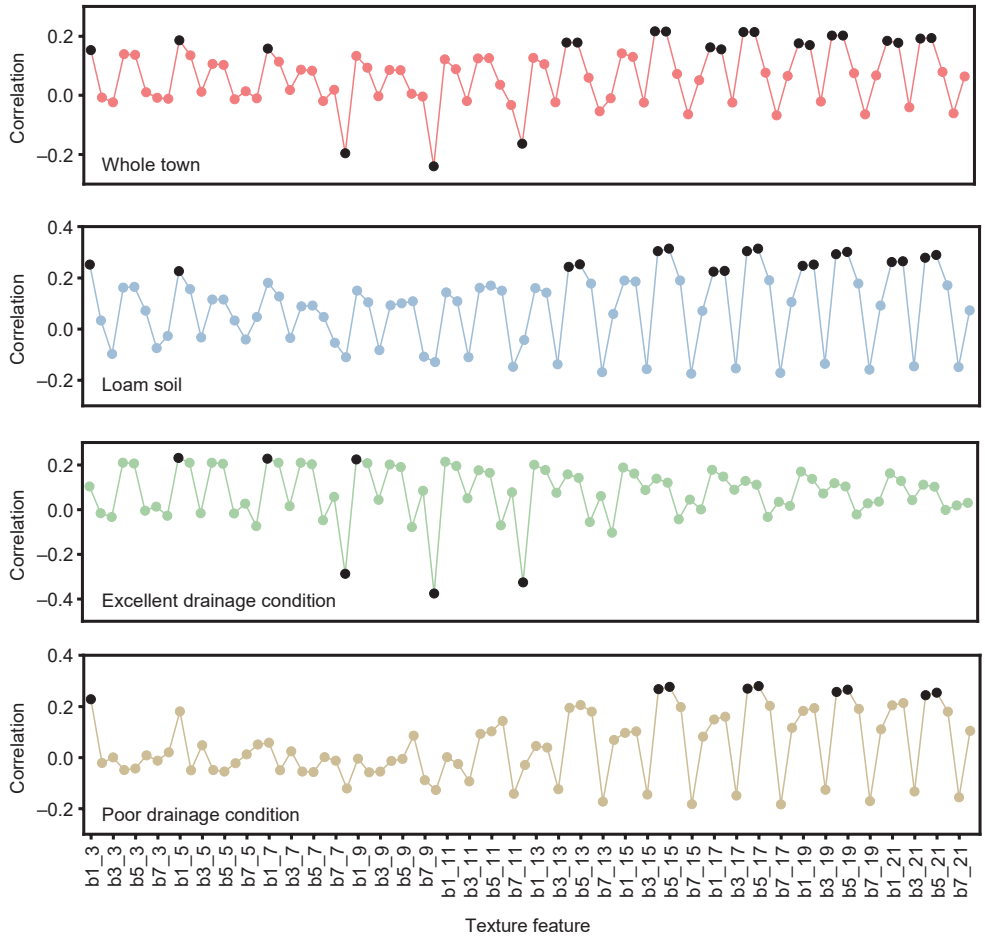



**Fig. 5 Line chart depicting the correlation between mean texture feature values and soil salt content across varying window sizes and categories.** b1–b8 represent eight distinct texture features: Mean, Variance, Homogeneity, Contrast, Dissimilarity, Entropy, Second Moment and Correlation. The numerical suffix (e.g., "3" in "b1_3") indicates the window size used during texture extraction; this notation remains consistent throughout. Black indicates significance at least at the 0.05 level.

to inter-variable collinearity, potentially hampering model construction. Furthermore, the relationship between soil salt content and various indicators may not be linear (Farifteh *et al.* 2007). Therefore, following the correlation analysis, VIP, PCA, and LightGBM were employed to further optimize the modeling factors.

As shown in Figs. 6–8, from the perspective of the optimization of the modeling factors, according to the screening method introduced in Section 2.3, the number of factors optimized by the three variable optimization methods decreased significantly. However, there were significant differences in the strength of the optimization, where PCA was the best, VIP was the second best, and GBM was the worst. Second, there were large differences in the scores of each factor under the four scenarios, whether in the VIP or GBM importance score. For example, in the GBM score, the DEM score reached 46 in the whole town scenario, and in the loamy soil and good/poor drainage conditions, the highest score did not exceed 20. These results suggest that the suitability of all the factors selected for modeling differ in different scenarios and needs to be specifically analyzed and screened during the modeling process.

### 3.4. Model construction and optimization

Based on the factor optimization results presented in Section 3.3, three machine learning models were input, and the coupled model was constructed according to the modeling strategy described in Section 2.3 under different categories. The optimal model for each category was then selected using evaluation metrics and scatter plots (Fig. 9). The model construction outcomes are detailed in Table 4 and Appendices B–E.

Table 4 reveals significant variations in coupled model accuracy across the four scenarios. According to model metrics, GBM-RF emerged as the superior inversion model for the whole town, achieving $R^2_C$=0.788 and $R^2_V$=0.629 for training and testing sets, respectively, demonstrating robust stability. Following intra-regional zonal modeling, the best model accuracy metrics for each scenario were improved relative to the whole town's best model, with the highest improvement of 0.18 for the training set $R^2_C$ and 0.15 for the testing set $R^2_V$. Notable distinctions emerged between models with good and poor drainage conditions. Under good drainage conditions, the training set $R^2_C$ achieved 0.961, though testing set results were less satisfactory, with $R^2_V$ generally below 0.676 and substantial accuracy variations

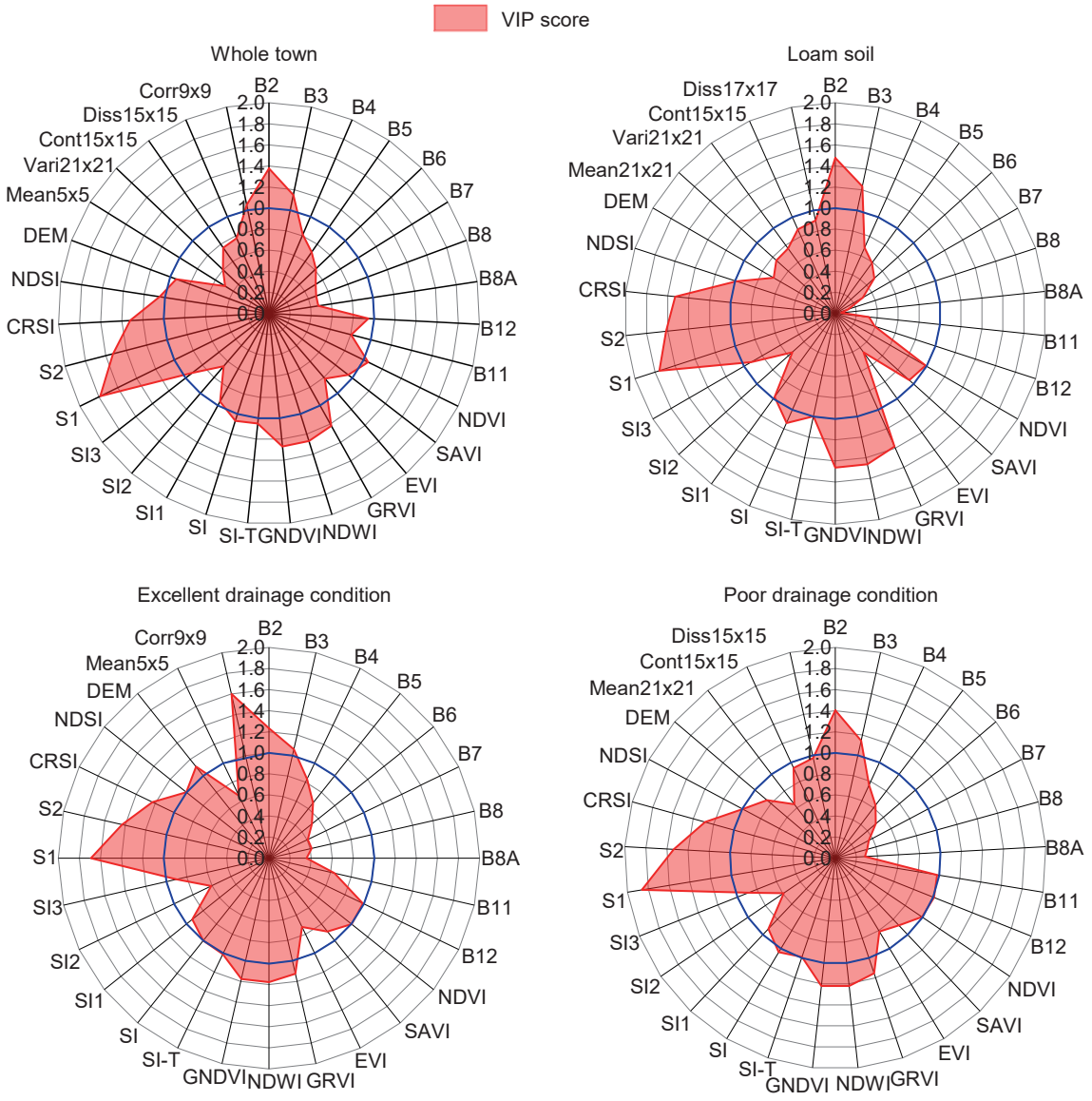



**Fig. 6 Variable importance in projection (VIP) scores for modeling factors within each category.** B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12 represent the reflectance of each band. NDVI denotes normalized difference vegetation index; SAVI, soil adjusted vegetation index; EVI, enhanced vegetation index; GRVI, green-red vegetation index; NDWI, normalized difference water index; GNDVI, green normalized difference vegetation index; SI-T, salinity index-thermal; SI, salinity index; SI1, salinity index 1; SI2, salinity index 2; SI3, salinity index 3; S1, salinity 1; S2, salinity 2; CRSI, canopy response salinity index; NDSI, normalized difference salinity index; DEM, digital elevation model. Mean5×5 indicates mean texture extracted in 5×5 window size; Vari21×21, variance texture extracted in 21×21 window size; Cont15×15, contrast texture extracted in 15×15 window size; Diss15×15, dissimilarity texture extracted in 15×15 window size; Corr9×9, correlation texture extracted in 9×9 window size; Mean21×21, mean texture extracted in 21×21 window size; Diss17×17, dissimilarity texture extracted in 17×17 window size.

between models. Under poor drainage conditions, the accuracy of each coupled model varied less, and the overall stability was more stable. Moreover, among the machine learning models and factor optimization methods, GBM exhibited the most effective coupling with each machine learning method, followed by VIP and PCA.

Fig. 9 displays scatter plots comparing predicted *vs*. measured values for the optimal models across four scenarios. The plots reveal that predictions from all four best models closely aligned with the 1:1 line at SSC<15g kg$^{-1}$, while generally underestimating at SSC>15 g kg$^{-1}$.

To further evaluate the performance of each optimal model after classification modeling, a unified test was conducted on the testing set in the whole town category. The results, presented in Fig. 10 and Table 5, demonstrate that the optimal models for loam ($R^2_T$=0.71) and excellent drainage conditions ($R^2_T$=0.82) outperformed the whole town model ($R^2_T$=0.63) on the unified testing set. This indicates that the best model for classification modeling does not only have a reliable predictive power on samples within the same category, but also a certain degree of generalization ability with respect to across categories. Although the optimal model for poor drainage conditions ($R^2_T$=0.56) showed lower performance in the unified testing set, it demonstrated superior performance

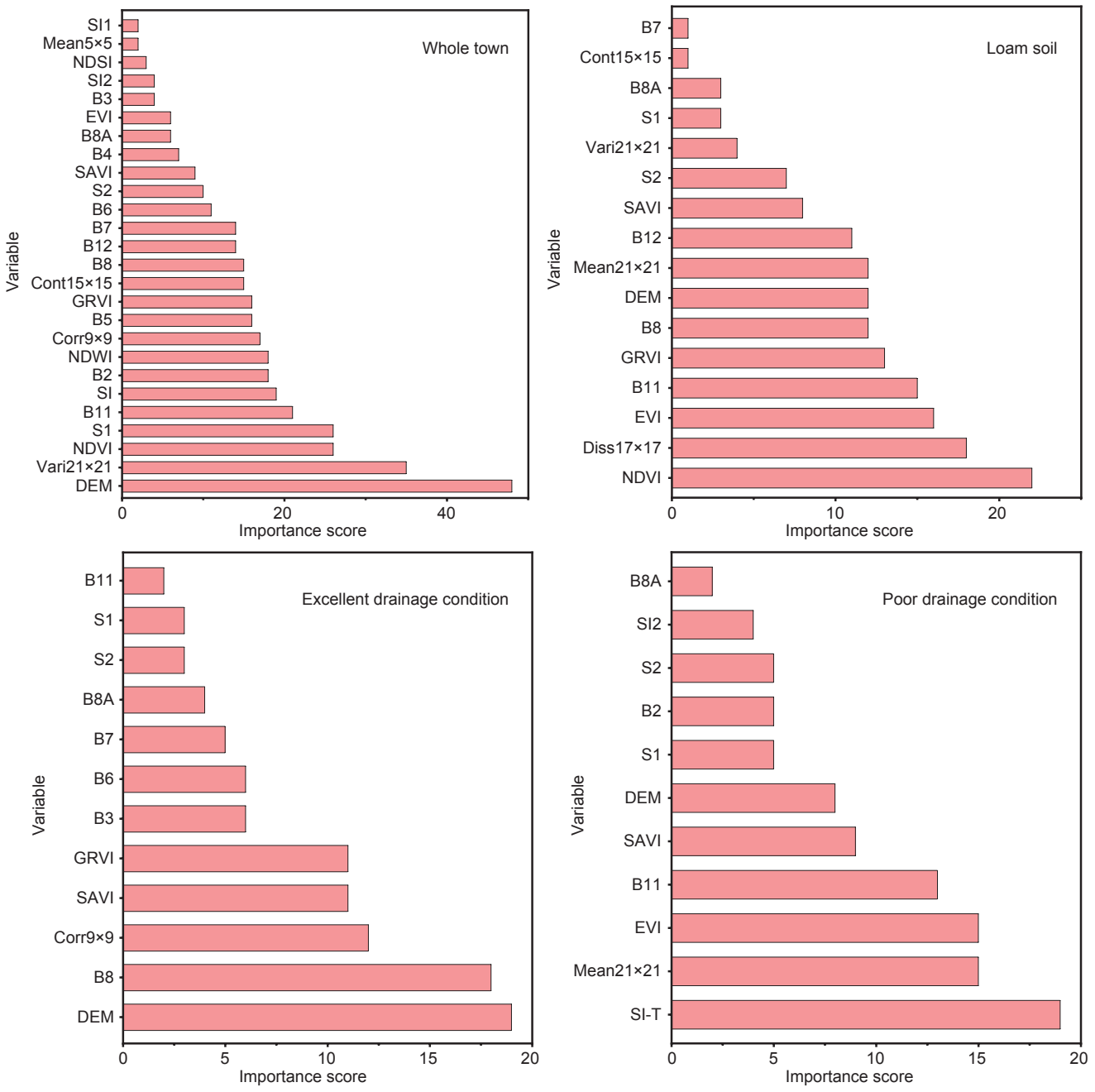



**Fig. 7 Importance scores of the modeling factors under each category.** SI1, salinity index 1; Mean5×5, mean texture is extracted in 5×5 window size; NDSI, normalized difference salinity index; SI2, salinity index 2; B3, the reflectance of band 3; EVI, enhanced vegetation index; B8A, the reflectance of band 8A; B4, the reflectance of band 4; SAVI, soil adjusted vegetation index; S2, salinity 2; B6, the reflectance of band 6; B7, the reflectance of band 7; B12, the reflectance of band 12; B8, the reflectance of band 8; Cont15×15, contrast texture is extracted in 15×15 window size; GRVI, green-red vegetation index; B5, the reflectance of band 5; Corr9×9, correlation texture is extracted in 9×9 window size; NDWI, normalized difference water index; B2, the reflectance of band 2; SI, salinity index; B11, the reflectance of band 11; S1, salinity 1; NDVI, normalized difference vegetation index; Vari21×21, variance texture is extracted in 21×21 window size; DEM, digital elevation model; Mean21×21, mean texture is extracted in 21×21 window size; Diss17×17, dissimilarity texture is extracted in 17×17 window size; SI-T, salinity index-thermal.

within its category (Fig. 9-D). Overall, comparative testing confirmed that classification modeling strategies enhanced the performance of remote sensing inversion monitoring models for salinization.

### 3.5. Soil salinity mapping

Utilizing the optimal GBM-RF inversion model for the whole town category, a distribution map of soil salinization classes was generated for the study area on October 15, 2022 (Fig. 11). The map reveals predominantly mild salinization throughout the study area, with the most severe salinization occurring in the northern region near the river coast. The central and southern regions exhibit lower salinization levels due to reduced river erosion vulnerability

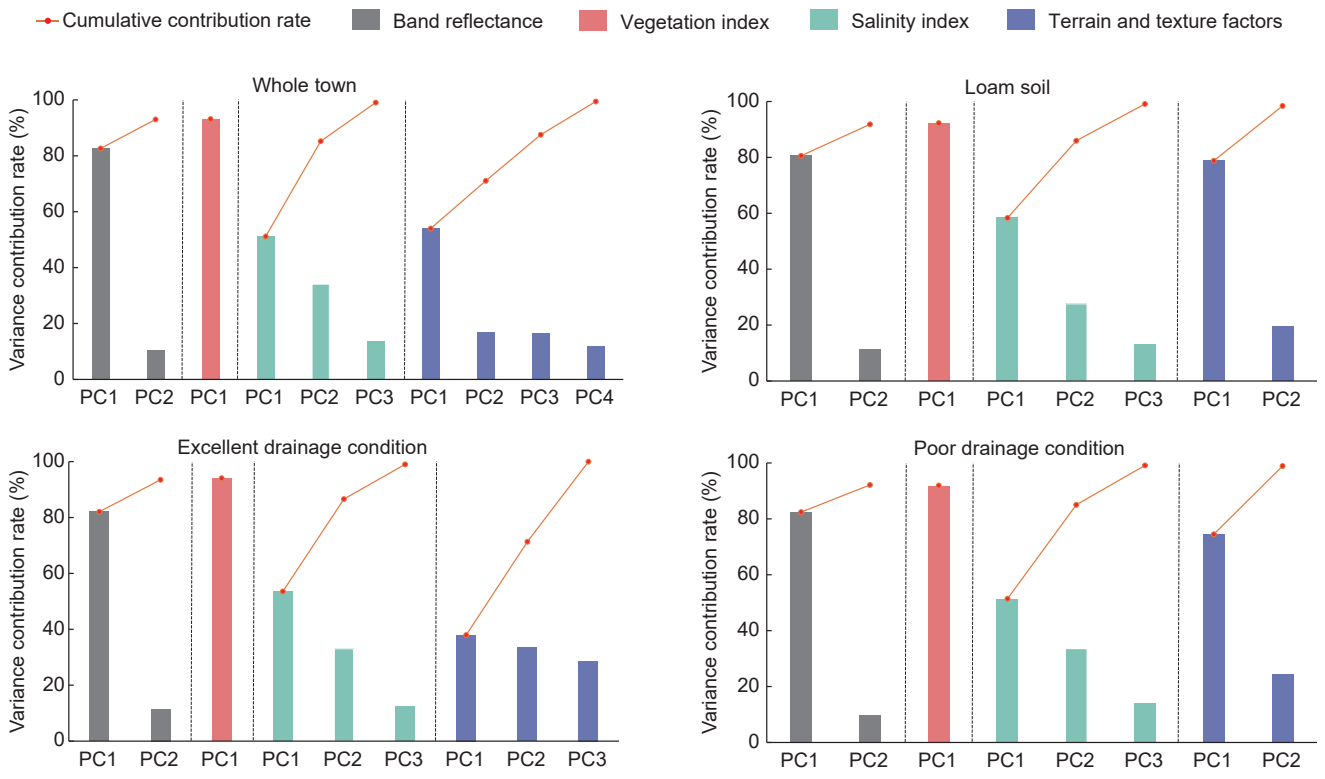



**Fig. 8 Principal component analysis (PCA) dimensionality reduction results of various modeling factors in each category.**

and denser forest coverage.

## 4. Discussion

### 4.1. The necessity of classification modeling

Factors affecting soil salinity formation within a region, including vegetation cover, soil type, and cropping patterns, also influence soil salinity monitoring accuracy to varying degrees (Zhang *et al.* 2021; Salcedo *et al.* 2022; Cui *et al.* 2023). Soil texture and surface drainage conditions represent two primary factors affecting soil salt transport and accumulation (Daliakopoulos *et al.* 2016).

This study evaluated the impact of varying soil textures and drainage conditions on salinity modeling in Wangaizhao Town, Erdos, Inner Mongolia, China. The correlation analysis between modeling factors and SSC revealed significant variations in the relationships among 10 band reflectance factors, six vegetation indices, nine salinity indices, and the mean values of texture features across different modeling scenarios. The vegetation indices demonstrated stronger correlations with SSC under favorable surface drainage conditions compared to poor drainage, while salinity indices exhibited the opposite relationship. This phenomenon may be attributed to poor drainage conditions leading to localized moisture accumulation, similar to salinization patterns in coastal areas. During water evaporation, salts migrate upward to the surface, causing soil salinization (Stavi *et al.* 2021). Additionally, vegetation struggles to survive due to ionic toxicity, osmotic stress, and oxidative stress (Rengasamy 2006; Daliakopoulos *et al.* 2016), resulting in reduced vegetation cover. Conversely, good drainage conditions proved more conducive to vegetation growth, yielding higher vegetation coverage compared to areas with poor drainage.

In the whole town scenario, the correlation between certain factors and SSC showed variations, with some factors like GNDVI and SI2 showing decreased correlation, while others such as B5 demonstrated increased correlation. The reason for this is that the whole town synthesis included sample sites with different soil textures and different drainage conditions, which possibly caused the importance of some factors to decrease with an increase in sample size and the importance of some factors to increase with an increase in sample size, as evidenced in the factor optimization of each scenario. When the town remained undivided into zones, the influence of various sample types either obscured important information or amplified irrelevant data. Zone division mitigated this effect, facilitating optimal modeling factor selection for different scenarios and enhancing modeling quality.

### 4.2. Model errors of soil salinity distribution

Machine learning algorithms improve as the management of the nonlinear relationship between independent and dependent variables improves (Bouasria *et al.* 2022), making the former preferable to traditional linear means of prediction. The prediction results indicate that the RF model exhibited optimal generalization and stability, performing effectively across scenarios, aligning with Wang N *et al.* (2024). However, based on the scatter plots of the best models for each scenario, the prediction accuracies of each model were very close to the 1:1 line for SSC<15 g $kg^{-1}$. However, prediction accuracy decreased beyond this threshold. This

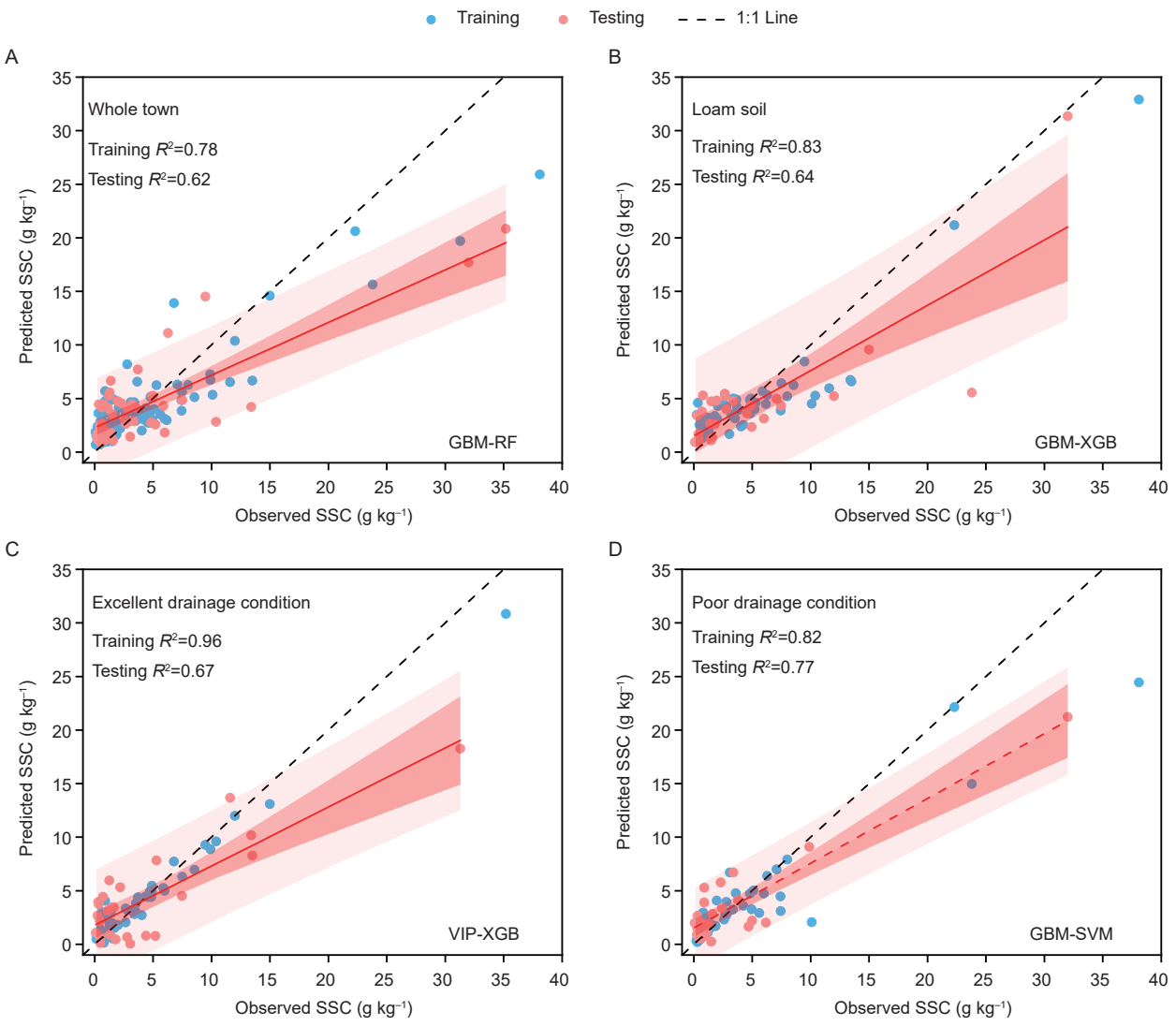



**Fig. 9 Scatter plots of observed and predicted values of the best inversion model under each category.** SSC, soil salt content; GBM-RF, coupled model of gradient boosting machine and random forest; GBM-XGB, coupled model of gradient boosting machine and extreme gradient boosting machine; VIP-XGB, coupled model of variable importance in projection and extreme gradient boosting machine; GBM-SVM, coupled model of gradient boosting machine and support vector machine.

limitation likely stems from insufficient samples with SSC>15 g kg$^{-1}$, preventing the models from adequately capturing predictive information, consistent with findings by Cui *et al.* (2023). Improving prediction accuracy beyond this threshold requires future research to expand the sample size for SSC>15 g kg$^{-1}$.

The study area exhibits a soil salinization pattern characterized by higher levels in the north and lower levels in the south. This distribution results from combined natural and anthropogenic factors. Wangaizhao Town's topography features higher elevation in the north and lower elevation in the south. The northern region borders a river and primarily consists of cropland, while the central and southern regions predominantly comprise sandy land, forests, and grassland. Extended agricultural irrigation practices and river water lateral infiltration, combined with terrain differences, have resulted in elevated groundwater levels and mineralization in the northern region compared to the south (Daliakopoulos *et al.* 2016; Deng and Bailey 2020). Consequently, in arid and semi-arid regions, where evaporation is strong,and precipitation is low, areas with high groundwater levels are more prone to soil salinization.

### 4.3. Limitations and suggestions

Environmental condition heterogeneity at small and medium scales has generated substantial variations in regional salinization remote sensing inversion accuracy (Zhang *et al.* 2021; Salcedo *et al.* 2022; Cui *et al.* 2023). This research examined the influence of soil texture and surface drainage conditions within a region on soil salinization modeling and proposed strategies for improving SSC prediction accuracy. However, environmental characteristics vary across regions, and this study's focus on soil texture and surface drainage conditions as factors influencing regional salinization modeling provides reference value for subsequent studies but may not be applicable in areas where other factors (e.g.,

**Table 4 Results of coupling model construction under different modeling categories[1)]**

| Modeling category | Optimization of variables[2)] | Model[3)] | $R^2_C$ | $RMSE_C$ | $MAE_C$ | $R^2_V$ | $RMSE_V$ | $MAE_V$ | $R^2_V/R^2_C$ |
|---|---|---|---|---|---|---|---|---|---|
| Whole town | VIP | RF | 0.790 | 2.639 | 1.573 | 0.624 | 4.080 | 2.703 | 0.790 |
| | | SVM | 0.499 | 4.072 | 2.169 | 0.582 | 4.302 | 2.576 | 1.165 |
| | | XGB | 0.660 | 3.353 | 2.056 | 0.486 | 4.768 | 3.095 | 0.736 |
| | GBM | RF | 0.788 | 2.650 | 1.728 | 0.629 | 4.054 | 2.704 | 0.798 |
| | | SVM | 0.655 | 3.380 | 1.901 | 0.516 | 4.625 | 2.723 | 0.789 |
| | | XGB | 0.626 | 3.520 | 2.239 | 0.481 | 4.792 | 2.882 | 0.768 |
| | PCA | RF | 0.587 | 3.697 | 2.162 | 0.408 | 5.117 | 2.842 | 0.695 |
| | | SVM | 0.654 | 3.387 | 1.673 | 0.357 | 5.335 | 3.004 | 0.546 |
| | | XGB | 0.644 | 3.432 | 1.999 | 0.332 | 5.436 | 2.872 | 0.515 |
| Loam soil | VIP | RF | 0.682 | 3.335 | 2.341 | 0.536 | 5.017 | 3.471 | 0.785 |
| | | SVM | 0.735 | 3.045 | 1.608 | 0.474 | 5.340 | 3.271 | 0.645 |
| | | XGB | 0.688 | 3.306 | 2.044 | 0.411 | 5.650 | 3.879 | 0.597 |
| | GBM | RF | 0.641 | 3.548 | 2.386 | 0.484 | 5.286 | 3.384 | 0.756 |
| | | SVM | 0.638 | 3.562 | 2.039 | 0.421 | 5.601 | 3.087 | 0.661 |
| | | XGB | 0.833 | 2.417 | 1.841 | 0.645 | 4.387 | 2.728 | 0.774 |
| | PCA | RF | 0.787 | 2.733 | 1.809 | 0.447 | 5.476 | 3.584 | 0.568 |
| | | SVM | 0.481 | 4.261 | 2.417 | 0.463 | 5.392 | 2.919 | 0.963 |
| | | XGB | 0.636 | 3.572 | 2.153 | 0.379 | 5.800 | 3.862 | 0.597 |
| Excellent drainage condition | VIP | RF | 0.871 | 1.932 | 1.208 | 0.167 | 6.166 | 3.508 | 0.192 |
| | | SVM | 0.575 | 3.503 | 1.943 | 0.183 | 6.110 | 3.352 | 0.318 |
| | | XGB | 0.961 | 1.068 | 0.725 | 0.676 | 3.848 | 2.998 | 0.704 |
| | GBM | RF | 0.506 | 3.775 | 2.050 | 0.161 | 6.191 | 3.582 | 0.318 |
| | | SVM | 0.648 | 3.186 | 1.966 | 0.209 | 6.009 | 3.555 | 0.323 |
| | | XGB | 0.476 | 3.889 | 2.428 | 0.367 | 5.378 | 3.308 | 0.771 |
| | PCA | RF | 0.544 | 3.625 | 1.898 | 0.271 | 5.769 | 3.437 | 0.498 |
| | | SVM | 0.942 | 1.291 | 0.521 | 0.325 | 5.550 | 3.847 | 0.345 |
| | | XGB | 0.897 | 1.724 | 1.265 | 0.567 | 4.446 | 2.883 | 0.632 |
| Poor drainage condition | VIP | RF | 0.770 | 3.094 | 1.730 | 0.563 | 4.236 | 2.791 | 0.731 |
| | | SVM | 0.670 | 3.705 | 1.637 | 0.652 | 3.777 | 2.375 | 0.973 |
| | | XGB | 0.742 | 3.275 | 1.993 | 0.637 | 3.859 | 2.706 | 0.858 |
| | GBM | RF | 0.769 | 3.100 | 1.619 | 0.700 | 3.511 | 2.308 | 0.909 |
| | | SVM | 0.828 | 2.675 | 1.106 | 0.772 | 3.056 | 2.130 | 0.933 |
| | | XGB | 0.842 | 2.566 | 1.694 | 0.710 | 3.448 | 2.474 | 0.844 |
| | PCA | RF | 0.727 | 3.369 | 1.816 | 0.578 | 4.160 | 2.680 | 0.795 |
| | | SVM | 0.767 | 3.115 | 1.387 | 0.730 | 3.329 | 2.357 | 0.952 |
| | | XGB | 0.731 | 3.348 | 1.672 | 0.626 | 3.916 | 2.477 | 0.857 |

[1)] $R^2_C$, $RMSE_C$ and $R^2_V$ represent the performance indicators of the model during the training set; $R^2_V$, $RMSE_V$ and $MAE_V$ indicate the performance indicators of the model during the testing set.
[2)] VIP, variable importance in projection; GBM, gradient boosting machine; PCA, principal component analysis.
[3)] RF, random forest model; SVM, support vector machine model; XGB, extreme gradient boosting machine model.

climate, and topography) predominate. Future research should therefore expand the study scale to incorporate additional factors affecting salinity modeling.

Additionally, owing to future monitoring needs at large scales, heterogeneity within small and medium scales may be weakened if large-scale internal partitioning is performed, which should also be considered in further research. The surface-UAV-satellite three-level correction cooperative inversion technique likely represents the key solution to addressing both this challenge and fulfilling this requirement. Manually deployed sample points effectively account for surface environment heterogeneity; however, this approach offers limited monitoring range and requires substantial time (Cui *et al.* 2023). UAVs provide broader monitoring coverage than manual sampling with centimeter-level spatial resolution; however, implementing monitoring at larger scales remains challenging (Zhang and Zhao 2019; Jia *et al.* 2024). Satellite monitoring enables extensive coverage; however, most satellites offer relatively coarse monitoring accuracy and struggle to account for surface heterogeneity (Zhang and Zhao 2019). From a comprehensive view, the three methods can compensate for each other's deficiencies, and the UAV is an intermediate means; combining satellite technology upward can consider the heterogeneity of small and medium scales in large-scale internal partitioning, and combining manual sampling downward can further improve the accuracy of UAV monitoring. Therefore, further research should aim to develop a set of surface-UAV-satellite three-

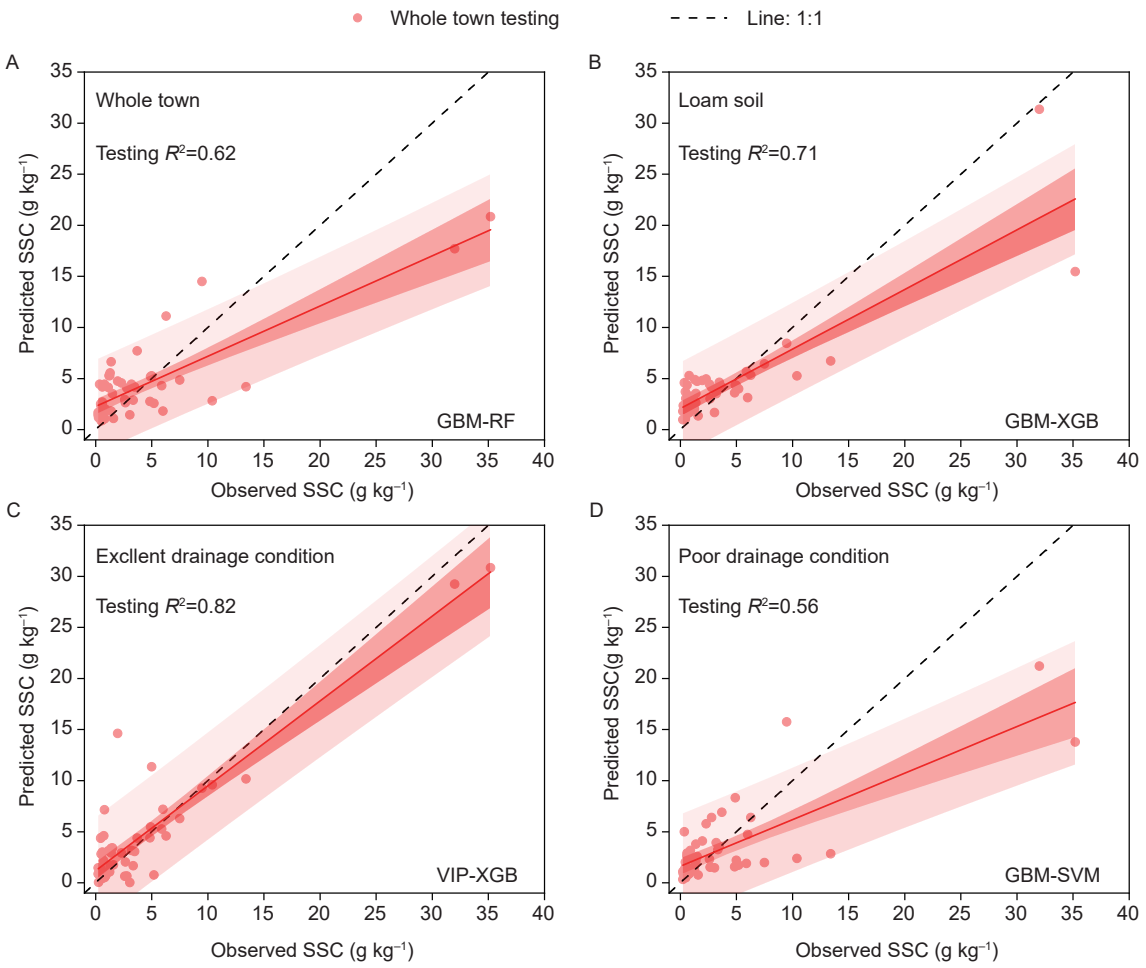



**Fig. 10 Scatter plots of observed and predicted values for the four best inversion models on the whole town testing set.** SSC, soil salt content; GBM-RF, coupled model of gradient boosting machine and random forest; GBM-XGB, coupled model of gradient boosting machine and extreme gradient boosting machine; VIP-XGB, coupled model of variable importance in projection and extreme gradient boosting machine; GBM-SVM, coupled model of gradient boosting machine and support vector machine.

**Table 5 Results of coupling model construction under different scenarios[1)]**

| Modeling scenario | Coupling model | $R^2_C$ | $RMSE_C$ | $MAE_C$ | $R^2_T$ | $RMSE_T$ | $MAE_T$ |
|---|---|---|---|---|---|---|---|
| Whole town | GBM-RF | 0.788 | 2.650 | 1.728 | 0.629 | 4.054 | 2.704 |
| Loam | GBM-XGB | 0.833 | 2.417 | 1.841 | 0.714 | 3.558 | 2.163 |
| Excellent drainage condition | VIP-XGB | 0.961 | 1.068 | 0.725 | 0.819 | 2.830 | 1.776 |
| Poor drainage condition | GBM-XVM | 0.828 | 2.675 | 1.106 | 0.563 | 4.396 | 2.488 |

[1)] $R^2_C$, $RMSE_C$ and $R^2_C$ represent the model performance indicators during the training set across different categories; $R^2_T$, $RMSE_T$ and $MAE_T$ indicate the model performance indicators during the comprehensive town testing set.

level corrected inversion systems for soil salinity monitoring.

## 5. Conclusion

In this study, we constructed a coupled SSC prediction model for different categories (whole town, loamy soil, and good and poor surface drainage) at small and medium scales using Sentinel-2 data in combination with field sampling data. The results demonstrated that the spatial heterogeneity of salinity dominant factors constrains the generalizability of both modeling indicators and modeling approaches. Independent modeling of different environmental factor categories significantly enhances the ability to capture the relationship between SSC and modeling indicators, thereby improving the effectiveness of SSC remote sensing monitoring. Texture features, due to their relative stability, exhibit strong applicability in SSC monitoring. Future research should examine the influence of other environmental factors on SSC monitoring and emphasize developing modeling indicators and methods with high generalizability to enhance monitoring accuracy.

## Acknowledgements

This work was supported by the National Key Research and Development Program Young Scientists Project, China (2023YFD1901900), the Key Research and Development Program of China (NK2022180406-01), the Key Science and

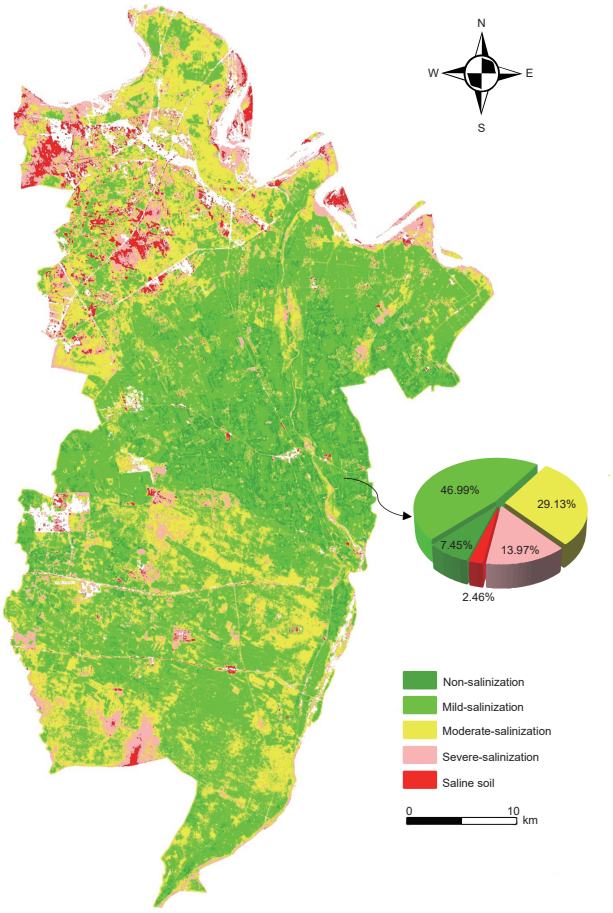



**Fig. 11 Distribution of soil salinity conditions across the study area of Wangaizhao Town, Ordos City, Inner Mongolia, China.**

Technology Tasks of the Chinese Academy of Agricultural Sciences (CAAS-ZDRW202407-02), the Ordos Saline-alkali Land Comprehensive Utilization of Agriculture and Animal Husbandry Science and Technology Innovation Test Demonstration Project, China (2023-KJCX-01), the Ordos City Bureau of Agriculture and Animal Husbandry Major Scientific and Technological Innovation Project, China (2023ZDKJCX), and the Taishan Scholars Program, China (TSQN202312322).

## Declaration of competing interest

The authors declare that they have no conflict of interest.

## Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

**Appendices** associated with this paper are available at https://doi.org/10.1016/j.jia.2025.08.016

## References

Akhtar N, Tsuyuzaki S. 2024. Effects of disturbances on the spatiotemporal patterns and dynamics of coastal wetland vegetation. *Ecological Indicators*, **166**, 112430.

Allbed A, Kumar L, 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.

Astola H, Häme T, Sirro L, Molinier M, Kilpi J. 2019. Comparison of Sentinel-2 and landsat 8 imagery for forest variable prediction in boreal region. *Remote Sensing of Environment*, **223**, 257–273.

Bai J, Wang N, Hu B, Feng C, Wang Y, Peng J, Shi Z. 2023. Integrating multisource information to delineate oasis farmland salinity management zones in southern Xinjiang, China. *Agricultural Water Management*, **289**, 108559.

Bouasria A, Ibno N K, Rahimi A, Ettachfini E M, Rerhou B. 2022. Evaluation of landsat 8 image pansharpening in estimating soil organic matter using multiple linear regression and artificial neural networks. *Geo-spatial Information Science*. **25**, 353–364.

Chen B, Zheng H, Luo G, Chen C, Bao A, Liu T, Chen X. 2022. Adaptive estimation of multi-regional soil salinization using extreme gradient boosting with Bayesian TPE optimization. *International Journal of Remote Sensing*, **43** 778–811.

Chen H, Ma Y, Zhu A, Wang Z, Zhao G, Wei Y. 2021. Soil salinity inversion based on differentiated fusion of satellite image and ground spectra. *International Journal of Applied Earth Observation and Geoinformation*, **101**, 102360.

Cisternas I, Velásquez I, Caro A, Rodríguez A. 2020. Systematic literature review of implementations of precision agriculture. *Computers and Electronics in Agriculture*, **176**, 105626.

Cui X, Han W, Zhang H, Dong Y, Ma W, Zhai X, Zhang L, Li G. 2023. Estimating and mapping the dynamics of soil salinity under different crop types using Sentinel-2 satellite imagery. *Geoderma*, **440**, 116738.

Daliakopoulos I N, Tsanis I K, Koutroulis A, Kourgialas N N, Varouchakis A E, Karatzas G P, Ritsema C J. 2016. The threat of soil salinity: A European scale review. *Science of the Total Environment*, **573**, 727–739.

Deng C, Bailey R T. 2020. Assessing causes and identifying solutions for high groundwater levels in a highly managed irrigated region. *Agricultural Water Management*, **240**, 106329.

Farifteh J, Van der Meer F, Atzberger C, Carranza E J M. 2007. Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, **110**, 59–78.

Greenacre M, Groenen P J, Hastie T, d'Enza A I, Markos A, Tuzhilina E. 2022. Principal component analysis. *Nature Reviews Methods Primers*, **2**, 100.

He Y, Yin H, Chen Y, Xiang R, Zhang Z, Chen H. 2024. Soil salinity estimation based on Sentinel-1/2 texture features and machine learning. *IEEE Sensors Journal*, **24**, 15302–15310.

Jia J, Chen C, Liu Q, Ding B, Ren Z, Jia Y, Bai X, Du R, Chen Q, Wang S, Luo L, Zhang Z, Geng H. 2024. Soil salinity monitoring model based on the synergistic construction of ground-UAV-satellite data. *Soil Use and Management*, **40**, e12980.

Jiang X, Duan H, Liao J, Guo P, Huang C, Xue X. 2022. Estimation of soil salinization by machine learning algorithms in different arid regions of northwest China. *Remote Sensing*, **14**, 347.

Kumar P, Tiwari P, Biswas A, Srivastava P K. 2024 Spatio-temporal assessment of soil salinization utilizing remote sensing derivatives, and prediction modeling: Implications for sustainable development. *Geoscience Frontiers*, **15**, 101881.

Li Y. 2021. Research progress of remote sensing monitoring of soil salinization. *IOP Conference Series* (Earth and Environmental Science), **692**, 042007.

Li Y, Chang C, Wang Z, Zhao G. 2023. Upscaling remote sensing inversion and dynamic monitoring of soil salinization in the Yellow River Delta, China. *Ecological Indicators*, **148**, 110087.

Lu R K. 1999. Microzone soil science — A possible new branch of soil science. *Acta Pedologica Sinica*, **36**, 287–288. (in Chinese)

Mandal A K. 2024. Spatial assessment and chemical characterization of degraded (salt-affected) soils at post-reclamation stage of the Indo-Gangetic Plain in Haryana State. *Environmental Monitoring and Assessment*, **196**, 213.

Metternicht G I, Zinck J A. 2003. Remote sensing of soil salinity: Potentials and constraints. *Remote Sensing of Environment*, **85**, 1–20.

Niazkar M, Menapace A, Brentan B, Piraei R, Jimenez D, Dhawan P, Righetti M. 2024. Applications of XGBoost in water resources

engineering: A systematic literature review. *Environmental Modelling & Software*, **174**, 105971.
Pereira L S, Gonçalves J M, Dong B, Mao Z, Fang S X. 2007. Assessing basin irrigation and scheduling strategies for saving irrigation water and controlling salinity in the upper Yellow River Basin, China. *Agricultural Water Management*, **93**, 109–122.
Ramos T B, Darouich H, Oliveira A R, Farzamian M, Monteiro T, Castanheira N, Paz A, Alexandre C, Goncalves M C, Pereira L S. 2023. Water use, soil water balance and soil salinization risks of Mediterranean tree orchards in southern Portugal under current climate variability: Issues for salinity control and irrigation management. *Agricultural Water Management*, **283**, 108319.
Ren D, Wei B, Xu X, Engel B, Li G, Huang Q, Xiong Y, Huang G. 2019. Analyzing spatiotemporal characteristics of soil salinity in arid irrigated agro-ecosystems using integrated approaches. *Geoderma*, **356**, 113935.
Rengasamy P. 2006. World salinization with emphasis on Australia. *Journal of Experimental Botany*, **57**, 1017–1023.
Salcedo F P, Cutillas P P, Cabañero J J A, Vivaldi A G. 2022. Use of remote sensing to evaluate the effects of environmental factors on soil salinity in a semi-arid area. *Science of the Total Environment*, **815**, 152524.
Sanaeifar A, Guindo M L, Bakhshipour A, Fazayeli H, Li X, Yang C. 2023. Advancing precision agriculture: The potential of deep learning for cereal plant head detection. *Computers and Electronics in Agriculture*, **209**, 107875.
Scudiero E, Skaggs T H, Corwin D L. 2015. Regional-scale soil salinity assessment using Landsat ETM+ canopy reflectance. *Remote Sensing of Environment*, **169**, 335–343.
Scudiero E, Skaggs T H, Corwin D L. 2016. Comparative regional-scale soil salinity assessment with near-ground apparent electrical conductivity and remote sensing canopy reflectance. *Ecological Indicators*, **70**, 276–284.
Song J, Yu D, Wang S, Zhao Y, Wang X, Ma L, Li J. 2024. Mapping soil organic matter in cultivated land based on multi-year composite images on monthly time scales. *Journal of Integrative Agriculture*, **23**, 1393–1408.
Stavi I, Thevs N, Priori S. 2021. Soil salinity and sodicity in drylands: A review of causes, effects, monitoring, and restoration measures. *Frontiers in Environmental Science*, **9**, 712831.
Taghizadeh-Mehrjardi R, Schmidt K, Toomanian N, Heung B, Behrens T, Mosavi A, Band S S, Amirian-Chakan A, Fathabadi A, Scholten T. 2021. Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models. *Geoderma*, **383**, 114793.
Vermeulen D, Van Niekerk A. 2017. Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates. *Geoderma*, **299**, 1–12.
Wang J, Ding J, Yu D, Ma X, Zhang Z, Ge X, Teng D, Li X, Liang J, Lizaga I, Chen X, Yuan L, Guo Y. 2019. Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China. *Geoderma*, **353**, 172–187.
Wang J J, Huang Y F, Long H Y. 2016. Water and salt movement in different soil textures under various negative irrigating pressures. *Journal of Integrative Agriculture*, **15**, 1874–1882.
Wang L, Hu P, Zheng H, Bai J, Liu Y, Hellwich O, Liu T, Chen X, Bao A. 2025. An automated framework for interaction analysis of driving factors on soil salinization in Central Asia and Western China. *Remote Sensing*, **17**, 987.
Wang L, Hu P, Zheng H, Liu Y, Cao X, Hellwich O, Liu T, Luo G, Bao A, Chen X. 2023. Integrative modeling of heterogeneous soil salinity using sparse ground samples and remote sensing images. *Geoderma*, **430**, 116321.
Wang N, Chen S, Huang J, Frappart F, Taghizadeh R, Zhang X, Wigneron J, Xue J, Xiao Y, Peng J, Shi Z. 2024. Global soil salinity estimation at 10 m using multi-source remote sensing. *Journal of Remote Sensing*, **4**, 0130.
Wang Q, Atkinson P M. 2018. Spatio-temporal fusion for daily Sentinel-2 images. *Remote Sensing of Environment*, **204**, 31–42.
Wang Q, Shi W, Li Z, Atkinson P M. 2016. Fusion of Sentinel-2 images. *Remote Sensing of Environment*, **187**, 241–252.
Wang Y, Qu Z, Yang W, Chen X, Qiao T. 2024. Inversion of soil salinity in the irrigated region along the southern bank of the yellow river using UAV multispectral remote sensing. *Agronomy*, **14**, 523.
Wang Z, Zhang F, Zhang X, Chan N W, Ariken M, Zhou X, Wang Y. 2021. Regional suitability prediction of soil salinization based on remote-sensing derivatives and optimal spectral index. *Science of the Total Environment*, **775**, 145807.
Yang H, Wang Z, Cao J, Wu Q, Zhang B. 2023. Estimating soil salinity using Gaofen-2 imagery: A novel application of combined spectral and textural features. *Environmental Research*, **217**, 114870.
Zhang J, Zhang Z, Chen J, Chen H, Jin J, Han J, Wang X, Song Z, Wei G. 2021. Estimating soil salinity with different fractional vegetation cover using remote sensing. *Land Degradation & Development*, **32**, 597–612.
Zhang S, Zhao G. 2019. A harmonious satellite-unmanned aerial vehicle-ground measurement inversion method for monitoring salinity in coastal saline soil. *Remote Sensing*, **11**, 1700.

Executive Editor-in-Chief Keke Yi
Managing Editor Lujuan Sun