Incorporating spatial autocorrelation into soil salinity models: Insights from the Minqin Oasis and its desert-oasis transition zone in Northwest China

ZHAO Dan^1,2,3, YANG Xiya^1,3, GAO Yukun^1,2,3, PAN Jing^1,3, YOU Qiangang^1,3, XUE Xian^1,3

2026-07-16

ChinaRxiv

Abstract

Remote sensing-based soil salinity inversion serves as a crucial approach for monitoring and assessment in arid regions. However, most existing models rarely account for the spatial autocorrelation (SAC) of soil salinity, which limits both their predictive accuracy and ability to capture spatial patterns. To address this gap, this study investigated the Minqin Oasis and its adjacent desert-oasis transition zone in Northwest China. Based on collected field soil samples and concurrently acquired Landsat-8 OLI remote sensing images in 2024, we incorporated characteristic bands reflecting SAC into conventional spectral indices. Through multi-band combination optimization and comparison of different models' predictive performance, we constructed an optimal soil salinity inversion model for the Minqin Oasis and its adjacent desert-oasis transition zone. The results demonstrated that incorporating SAC of soil salinity markedly improved model performance, with the Gradient Boosting Regression Trees (GBRT) model incorporating SAC (GBRT_SAC) achieving the best accuracy. Compared with the traditional spectral index-based GBRT model, the coefficient of determination (R^2) increased by 7.320%, the root mean square error (RMSE) decreased by 20.230%, and the mean absolute percentage error (MAPE) decreased by 121.01% using the GBRT_SAC model. The soil salinity distribution derived from the GBRT_SAC model revealed pronounced spatial heterogeneity, with salinized areas covering approximately 1256.75 km^2 (36.170% of the total area). Soil salinity was jointly influenced by natural and anthropogenic factors. At the regional scale, soil type and vegetation type emerged as the dominant drivers shaping soil salinity patterns. In contrast, within the oasis interior, soil salinity was primarily driven by groundwater table regulated by irrigation, leading to surface salt accumulation through capillary rise. In the 1000 m desert-oasis transition zone, the explanatory power (q-value) of all environmental factors for spatial variation of soil salinity significantly increased, indicating a sensitive interface where hydrological and aeolian processes interact. Notably, although soil salinity was relatively lower in sandy areas, sand content emerged as the most influential factor in this region (q-value=0.483), effectively serving as a key indicator of the transitional environment. By introducing SAC-based features into soil salinity inversion models, this study provides a robust methodological framework and valuable data to support understanding and management of soil salinization in arid desert-oasis ecotone systems.

Full Text

J Arid Land (2026) 18(7): 1135-1158
doi: 10.1016/j.jaridl.2026.05.010; CSTR: 32276.14.JAL.20250595

Incorporating spatial autocorrelation into soil salinity models: Insights from the Minqin Oasis and its desert-oasis transition zone in Northwest China

ZHAO Dan^1,2,3, YANG Xiya^1,3, GAO Yukun^1,2,3, PAN Jing^1,3, YOU Qiangang^1,3, XUE Xian^1,3*

^1 State Key Laboratory of Ecological Safety and Sustainable Development in Arid Lands, Lanzhou 730000, China;

^2 University of Chinese Academy of Sciences, Beijing 101408, China;

^3 Drylands Salinization Research Station, Northwest Institute of Eco-environment & Resources, Chinese Academy of Sciences, Lanzhou 730000, China

Abstract: Remote sensing-based soil salinity inversion serves as a crucial approach for monitoring and assessment in arid regions. However, most existing models rarely account for the spatial autocorrelation (SAC) of soil salinity, which limits both their predictive accuracy and ability to capture spatial patterns. To address this gap, this study investigated the Minqin Oasis and its adjacent desert-oasis transition zone in Northwest China. Based on collected field soil samples and concurrently acquired Landsat-8 OLI remote sensing images in 2024, we incorporated characteristic bands reflecting SAC into conventional spectral indices. Through multi-band combination optimization and comparison of different models' predictive performance, we constructed an optimal soil salinity inversion model for the Minqin Oasis and its adjacent desert-oasis transition zone. The results demonstrated that incorporating SAC of soil salinity markedly improved model performance, with the Gradient Boosting Regression Trees (GBRT) model incorporating SAC (GBRT_SAC) achieving the best accuracy. Compared with the traditional spectral index-based GBRT model, the coefficient of determination (R^2) increased by 7.320%, the root mean square error (RMSE) decreased by 20.230%, and the mean absolute percentage error (MAPE) decreased by 121.01% using the GBRT_SAC model. The soil salinity distribution derived from the GBRT_SAC model revealed pronounced spatial heterogeneity, with salinized areas covering approximately 1256.75 km^2 (36.170% of the total area). Soil salinity was jointly influenced by natural and anthropogenic factors. At the regional scale, soil type and vegetation type emerged as the dominant drivers shaping soil salinity patterns. In contrast, within the oasis interior, soil salinity was primarily driven by groundwater table regulated by irrigation, leading to surface salt accumulation through capillary rise. In the 1000 m desert-oasis transition zone, the explanatory power (q-value) of all environmental factors for spatial variation of soil salinity significantly increased, indicating a sensitive interface where hydrological and aeolian processes interact. Notably, although soil salinity was relatively lower in sandy areas, sand content emerged as the most influential factor in this region (q-value=0.483), effectively serving as a key indicator of the transitional environment. By introducing SAC-based features into soil salinity inversion models, this study provides a robust methodological framework and valuable data to support understanding and management of soil salinization in arid desert–oasis ecotone systems.

Keywords: soil salinization; spatial autocorrelation (SAC); machine learning; Gradient Boosting Regression Trees (GBRT); spatial heterogeneity; GeoDetector; desert–oasis transition zone

*Corresponding author: XUE Xian (E-mail: xianxue@lzb.ac.cn)
Received 2025-11-28; revised 2026-04-12; accepted 2026-05-14

http://jal.xjegi.com; https://www.keaipublishing.com/en/journals/journal-of-arid-land/

This version posted 2026-07-16.

Citation: ZHAO Dan, YANG Xiya, GAO Yukun, PAN Jing, YOU Quangang, XUE Xian. 2026. Incorporating spatial autocorrelation into soil salinity models: Insights from the Minqin Oasis and its desert–oasis transition zone in Northwest China. Journal of Arid Land, 18(7): 1135–1158. https://doi.org/10.1016/j.jaridl.2026.05.010; https://cstr.cn/32276.14.JAL.20250595

1 Introduction

Soil salinization is one of the most widespread and severe forms of land degradation (Hassani et al., 2021; Liu et al., 2024; Aihaiti et al., 2025). In recent years, it has intensified and expanded under the dual impacts of climate change and human activities, posing a serious threat to land productivity and food security (Zaman et al., 2018; Guo et al., 2023a). Oases and their adjacent desert–oasis transition zones in arid regions are facing an escalating risk of soil salinization (Hassani et al., 2024; Lu et al., 2025b). On the one hand, climate warming increases regional evapotranspiration (Tariq et al., 2024). Under conditions of shallow groundwater or brackish/saline water irrigation with inadequate drainage, this process enhances the upward migration of soil moisture and promotes salt accumulation in surface layers (Hao et al., 2017). On the other hand, rapid economic development and population growth have accelerated agricultural expansion, intensive cultivation, and irrational irrigation practices (Cui et al., 2025). Therefore, it is essential to identify the spatial distribution patterns and driving factors of soil salinization and apply this knowledge to optimize irrigation systems, plan drainage projects, and develop early warning services. Undertaking these efforts is not only a prerequisite for mitigating soil salinization and preventing land degradation but also represents a critical scientific issue that current research urgently needs to address (Zhang et al., 2022).

Traditional studies on soil salinization primarily rely on field sampling (Chen et al., 2021; Wu et al., 2021), which provides only localized information and fails to meet the demands of large-scale monitoring (Du et al., 2024). In contrast, remote sensing technology enables large-scale, continuous, and cost-effective observations (Xiong et al., 2025), largely overcoming the limitations of *in situ* measurements (Weiss et al., 2020). Saline-affected soils exhibit unique spectral reflectance characteristics, serving as key indicators for remote sensing monitoring and Geographic Information System (GIS)-based analysis (Kumar et al., 2024). In recent years, remote sensing-based soil salinity inversion models have shown great potential in mapping and dynamic monitoring (Ivushkin et al., 2019a; Zeng et al., 2021). Specifically, machine learning models utilizing spectral indices have been widely adopted to capture the nonlinear relationships between remote sensing data and soil salinity. This advancement has markedly enhanced inversion accuracy (Wang et al., 2023a; Ghasempour et al., 2024). Consequently, such models have become a major research focus in this field (Mahmoud et al., 2025; Zeng et al., 2025). Therefore, integrating remote sensing imagery with machine learning models is a crucial approach for achieving efficient and dynamic monitoring of soil salinization at regional and even global scales (Lu et al., 2025a).

Despite significant progress in this field of research, soil salinity inversion models based on remote sensing and machine learning still have obvious limitations (Mahmoud et al., 2025). Firstly, model accuracy depends not only on the selected algorithm but also heavily on the quality of feature composition and input spectral indices (Ivushkin et al., 2019b). Relying on a single or limited number of empirical spectral indices makes it difficult to comprehensively characterize salinity features under complex surface conditions. Consequently, this limitation often leads to information loss and inversion bias (Li et al., 2023). Moreover, model performance is highly susceptible to interference from soil moisture content, gypsum or carbonate content, surface crusting, and vegetation cover (Yin et al., 2022). Secondly, as a typical geographical process, soil salinization exhibits pronounced spatial autocorrelation (SAC). This characteristic aligns with Tobler' s First Law of Geography, which states that salt content between adjacent sampling points often shows a high degree of consistency (Tobler, 1970; Yan et al., 2018). This SAC characteristic reflects the intrinsic structural continuity of soil salinity distribution, serving as a crucial information source for enhancing soil salinity inversion accuracy. However, traditional machine

learning inversion models typically treat samples as independent observations; they often fail to model spatial dependencies explicitly (Chen et al., 2024b). As a result, these models cannot fully explore the global spatial structure of soil salinity patterns. Neglecting this feature may lead to misinterpretation of spatial patterns and an increase in prediction errors. Such neglect limits model accuracy and the representation of spatial patterns (Cao et al., 2025). It also weakens cross-regional generalization and the identification of scale-dependent relationships (Yin et al., 2022; Yang et al., 2025). Furthermore, the scale mismatch between ground sample points and remote sensing pixels means that a single pixel may contain highly heterogeneous surface information. This mismatch prevents pixel-based observations from accurately representing the spatial variability within the area. It may also obscure the authentic spatial relationships between neighboring sample points. Consequently, the detectability and effective use of SAC are weakened, ultimately limiting further improvement of soil salinity inversion models.

As a distinctive coupled ecological-human settlement unit in arid regions, oases host the majority of the population and economic activities, playing a pivotal role in maintaining regional socio-economic development and ecological security (Hao et al., 2025). In such environments, the oasis and its adjacent desert-oasis transition zone form a mosaic of artificial and natural ecosystems shaped by both natural and anthropogenic forces. These complex interactions make the region highly susceptible to soil salinization (Li et al., 2025). Although the oasis interior and its adjacent desert-oasis transition zone experience similar climatic conditions, their contrasting land use patterns lead to distinct responses in water use, salt migration, and ecosystem stability. Previous studies have primarily examined oasis ecotone systems at an integrated or regional scale (Zhang et al., 2025); however, such studies often lack stratified or gradient-based frameworks to explicitly reveal spatial variations. Consequently, the mechanisms governing interactions between natural and human factors across zonal units remain insufficiently understood (Wang et al., 2023b). Moreover, the transition from the oasis interior to the desert-oasis transition zone exhibits pronounced spatial gradients and potential nonlinearities in ecological functions, hydrological processes, and human activity intensity (Jia et al., 2024b). These variations further complicate the understanding of soil salinity patterns and their drivers.

To fill these research gaps, this study selected the Minqin Oasis and its adjacent desert-oasis transition zone in Northwest China as a representative area. The specific objectives of this study are to: (1) develop a high-precision soil salinity inversion model by integrating multi-source spectral features with SAC information; (2) apply the optimal model to map regional soil salinity and construct continuous spatial gradients to quantify the spatial heterogeneity from the oasis interior to the desert-oasis transition zone; and (3) identify and compare the dominant driving factors of soil salinization between the oasis interior and its adjacent desert-oasis transition zone. The results are expected to provide new evidence on the ecological effects and management thresholds associated with oases expansion, offering a scientific basis for controlling soil salinization and for sustainable land management in arid regions.

2 Materials and methods

2.1 Study area

Using the established boundaries of the Minqin Oasis, this study generated buffer zones at distances of 100, 500, and 1000 m. For the primary analytical extent, we defined the Minqin Oasis and its 1000 m buffer zone (Fig. 1). The selection of these specific buffer distances was guided by previous studies on desert-oasis transition zones (Li et al., 2018; Chang et al., 2019; Sun et al., 2023) and aimed to characterize the spatial heterogeneity of the transition zone at multiple scales. Specifically, the 100 m buffer zone captures the immediate edge effects and anthropogenic disturbances at the fringes of the oasis (Liu et al., 2025). The 500 m buffer zone corresponds to the critical threshold where species richness and vegetation cover have been shown to decline sharply in the desert-oasis transition zone of Minqin (Zhao et al., 2019; Han et

Fig. 1 Overview of the study area (Minqin Oasis and its adjacent desert-oasis transition zone) and distribution of sampling points (a), and photos showing different land use/cover types in the study area (b)

al., 2023). Finally, the 1000 m buffer zone represents the broader gradient of biophysical interactions, covering the maximum extent of the oasis cold island effect and the primary wind-breaking influence of the desert-oasis transition zone (Hao et al., 2016; Lu et al., 2023).

Minqin Oasis and its adjacent desert-oasis transition zone (38°41$ '00' '–39°27' 38"$N, 101°49'38"–104°11'55"E), located in Gansu Province of China, covers approximately 3455.39 km^{2}$ and ranges in elevation from 1303 to 1524 m. The region has a typical temperate continental desert climate, characterized by a mean annual temperature of 8.5°C, a mean annual precipitation of 115.11 mm, a mean annual evaporation of 2643.30 mm, and a mean annual sunshine duration of 3079.6 h. The study area is bordered by the Tengger Desert and Badain Jaran Desert on its eastern, northern, and western sides, forming a typical desert-encircled oasis landscape (Chang et al., 2024).

Land use/cover is dominated by cropland, forest, grassland, and desert, reflecting a fragmented oasis ecosystem embedded within an arid landscape. Irrigated agriculture prevails in the oasis, with major crops including maize, wheat, sunflower, fennel, alfalfa, and many kinds of melons. In the study area, natural vegetation is primarily composed of xerophytic shrubs, including *Haloxylon ammodendron*, *Nitraria tangutorum*, *Reaumuria soongarica*, and *Kalidium foliatum*. Vegetation cover is generally low, leading to low ecosystem stability. Soils exhibit significant spatial heterogeneity: gray desert soils dominate the desert areas, whereas irrigated farmland soils, saline-alkaline soils, aeolian sandy soils, and meadow soils are distributed in the oasis and the desert-oasis transition zone (Yang et al., 2025).

2.2 Data sources and pre-processing

2.2.1 Soil sampling collection and pre-processing

Field surveys including soil sampling were conducted from 24 June to 9 July, 2024, synchronized with satellite image acquisition. This specific temporal window was selected to capture optimal conditions for soil salinity detection: (1) high evaporation rates during this period promote surface salt accumulation, thereby enhancing spectral signals in satellite imagery (Gao et al., 2025b); and (2) local vegetation is in its peak growing season, where spectral characteristics can most effectively reflect soil salinity stress (Zhan et al., 2025).

The sampling layout adopted a strategy combining preliminary regular-grid planning with multi-factor collaborative optimization. First, sampling points were initially distributed on a 10 km$×$10 km regular grid to ensure unbiased spatial coverage (Guo et al., 2023b). Subsequently, adjustments were made on the basis of accessibility and spatial uniformity to balance representativeness and field feasibility (Ge et al., 2022). Finally, a total of 274 sampling points were selected, covering both the oasis interior and the desert–oasis transition zone.

At each sampling point, a five-point composite sampling method was employed. Specifically,

surface soils (0–20 cm) were collected from the four corners and the center of a 10 m$×$10 m plot, thoroughly mixed into a composite sample, resulting in a total of 274 soil samples. These samples were stored in sealed polyethylene bags in the field, and then air-dried, ground, and passed through a 1-mm sieve in the laboratory after transport. For soil salinity analysis, each composite sample was precisely weighed to take 20 g of sieved soil, which was mixed with 100 mL of distilled water (1:5 soil:water ratio), and electrical conductivity (EC; dS/m) was measured using a conductivity meter (S400-B; Shanghai INESA Scientific Instruments Company Limited, Shanghai, China). To facilitate spatial analysis of soil salinity at the regional scale, we classified soil salinity levels into five categories according to the scheme of Ma et al. (2023) (Table 1).

Table 1 Classification of soil salinity based on electrical conductivity (EC)

Level	Soil salinity class	EC (dS/m)
Level 1	Non-salinization	<2.000
Level 2	Light salinization	2.000–4.000
Level 3	Moderate salinization	4.000–8.000
Level 4	Severe salinization	8.000–16.000
Level 5	Extreme salinization (saline soil)	>16.000

2.2.2 Remote sensing data acquisition

This study used Landsat-8 OLI surface reflectance (SR) products acquired between 24 June and 9 July, 2024, as a remote sensing data source to maintain consistency with the sampling collection time. The images have a spatial resolution of 30 m. A total of six original multispectral bands were used: Blue band (B2), Green band (B3), Red band (B4), Near-Infrared band (NIR; B5), Shortwave Infrared 1 band (SWIR1; B6), and Shortwave Infrared 2 band (SWIR2; B7). Additionally, to ensure temporal consistency, we analyzed the land use/cover based on a dataset with 30 m resolution from the same period, consisting of nine land use/cover classes: cropland, forest, shrubland, grassland, water body, snow/ice, barren land, impervious surface, and wetland. This dataset was obtained from the China Land Cover/Land Use (CLCD) dataset provided by Wuhan University, China (Yang et al., 2024). All data acquisition and preprocessing were performed on the Google Earth Engine (GEE) cloud platform.

2.2.3 Ancillary data acquisition

The digital elevation model (DEM) was acquired from the Geospatial Data Cloud (http://www.gscloud.cn/). The normalized difference vegetation index (NDVI) and key climatic factors—including precipitation, wind speed, potential evapotranspiration, evapotranspiration, and land surface temperature—were obtained from the National Earth System Science Data Center (http://www.geodata.cn). Additionally, LandScan population distribution data were accessed via the Oak Ridge National Laboratory (https://landscan.ornl.gov/). The vegetation type, soil type, clay content, and sand content data were obtained from the Resource and Environment Science and Data Platform (RESDC; https://www.resdc.cn/). Specifically, the vegetation type and soil type were derived from the spatial distribution datasets for China' s vegetation type and soil type published in 2001 and 1995, respectively, at a scale of 1:1,000,000. The obtained vegetation type and soil type were reclassified in ArcGIS 10.2, and the newly generated vegetation type map was used for further spatial analysis. Soil hydrological parameters were obtained from the 0.0–2.0 m soil hydrological parameter dataset for the arid region of Northwest China developed by Niu et al. (2024). This dataset includes six standard depths (0.0, 0.1, 0.3, 0.6, 1.0, and 2.0 m) and provides parameters such as field capacity, saturated porosity, saturated hydraulic conductivity, active water content, wilting point, saturated moisture content, and residual moisture content. Groundwater table data were obtained from the Wuwei Municipal Water Affairs Bureau (https://www.gswwuwei.gov.cn/). To ensure temporal consistency with the field survey, this study utilized ancillary data in July 2024. All datasets were resampled and unified to a consistent spatial

resolution of 30 m and projection system prior to modeling.

2.3 Data analysis

2.3.1 Feature extraction of soil salinity

In this study, we constructed a total of 23 soil salinity-related spectral indices, based on the original multispectral bands of Landsat-8 OLI, referencing salinity-sensitive indices from the previous literature. Detailed information is provided in Table 2. To eliminate the influence of differences in dimension and numerical range among indices, we normalized all features using min-max normalization and scaled to the [0, 1] interval. For further mitigation of multicollinearity and selection of practical features, we first log-transformed the measured soil EC (lnEC) to improve its normality. Subsequently, Pearson correlation analysis was conducted between each spectral index and the measured EC. Indices showing no significant correlation ($P \geq 0.05$) were excluded, and only statistically significant features were retained for subsequent modeling (Lu et al., 2014).

Table 2 Formulas used for soil salinity inversion

Land surface parameter	Abbreviation	Formula	Reference
Intensity index 1	Int1	$(B3 + B4)/2$	Jia et al. (2024b)
Intensity index 2	Int2	$(B3+B4+B5)/2$	Guo et al. (2023a)
Vegetation soil salinity index	VSSI	$2 \times B3 - 5 \times (B4 + B5)$	Jiang et al. (2022)
Normalized difference salinity index	NDSI	$(B4 - B5)/(B4 + B5)$	Wang et al. (2023a)
Salinity index 1	S1	$B2/B4$	Guo et al. (2023b)
Salinity index 2	S2	$(B2 - B4)/(B2 + B4)$	Han et al. (2021)
Salinity index 3	S3	$(B3 \times B4)/B2$	Ma et al. (2023)
Salinity index 4	S4	$(B2 \times B4)/B3$	Khosravichenar et al. (2023)
Salinity index 5	S5	$(B4 \times B5)/B3$	Khosravichenar et al. (2023)
Salinity index 6	S6	$B6/B7$	Wang et al. (2023b)
Salinity index 7	S7	$(B6 - B7)/(B6 + B7)$	Jiang et al. (2022)
Salinity index 8	S8	$B6 - B7$	Jiang et al. (2022)
Salinity index 9	S9	$(B6 \times B7 - B7 \times B7)/B6$	Jiang et al. (2022)

Land surface parameter	Abbreviation	Formula	Reference
Salinity index I-T	SI-T	$\frac{B4}{B5} \times 100$	Allbed et al. (2014)
Salinity index	SI	$\sqrt{B4 \times B2}$	Ge et al. (2022)
Salinity index I	SI1	$\sqrt{B4 \times B3}$	Yu et al. (2023)
Salinity index II	SI2	$\sqrt{B5^2 + B4^2 + B3^2}$	Xiao et al. (2024)
Salinity index III	SI3	$\sqrt{B4^2 + B3^2}$	Yu et al. (2023)
Canopy response salinity index	COSRI	$\sqrt{\frac{B5 \times B4 - B3 \times B2}{B5 \times B4 + B3 \times B2}}$	Jia et al. (2024a)
Modified soil adjusted vegetation index	MSAVI	$\frac{(2 \times B5 - 1) - \sqrt{(2 \times B5 + 1)^2 - 8 \times (B5 - B4)}}{2}$	Wang et al. (2023a)
Modified normalized difference water index	MNDWI	$\frac{B3 - B6}{B3 + B6}$	Jia et al. (2024a)
Green atmospherically resistant vegetation index	GARI	$\frac{B5 - (B3 + 1.7 \times (B2 - B3))}{B5 + (B3 + 1.7 \times (B2 - B3))}$	Xiong et al. (2025)
Enhanced vegetation index	EVI	$2.5 \times \frac{B5 - B4}{B5 + 6 \times B4 - 7.5 \times B2 + 1}$	Wang et al. (2023a)

Note: B2–B7 represent surface reflectance (dimensionless) of the Blue, Green, Red, Near-infrared, Shortwave Infrared 1, and Shortwave Infrared 2 bands from Landsat-8 OLI, respectively.

2.3.2 Construction of SAC feature

To explicitly incorporate SAC in soil salinity inversion, this study constructed a factor that included SAC to quantify the influence of EC values from neighboring sampling points on soil salinity at each target point (Chen et al., 2024a). This factor was introduced as an additional input into the soil salinity inversion model to explicitly capture spatial dependencies, thereby enhancing the model's ability to identify soil salinity accumulation patterns. Importantly, when calculating SAC for the sampling points, the self-point was excluded to avoid numerical instability (i.e., infinite weights at zero distance) and to strictly capture the influence of neighbors. The generation of the continuous SAC covariate layer for mapping followed a two-step workflow: (1) calculating the SAC values at all sampling points as described above; and (2) spatially interpolating these point-based values to the entire study area.

Accordingly, three specific implementations were designed (Qin et al., 2021)—(1) baseline models: models without adding SAC, using only spectral indices, serving as the performance benchmark; (2) global spatial autocorrelation (SAC_{all}): calculated based on the soil salinity (EC) values using the inverse distance-weighted average of EC values from a sampling point to all other sampling points, representing the global spatial trend; and (3) local spatial autocorrelation (SAC_8): calculated based on the soil salinity (EC) values using the inverse distance-weighted average of EC values from the sampling point to its eight nearest neighbors, capturing local salinity clustering patterns. These three features were separately input into the inversion models, and their predictive accuracy and spatial consistency were compared to evaluate the effectiveness of different SAC features. The formulas are as follows:

$$WS_i = \frac{1}{DS_i^2}, \tag{1}$$

$$SAC_k = \frac{\sum_{i=1}^{n} WS_i EC_i}{\sum_{i=1}^{n} WS_i}, \tag{2}$$

where WS_i is the spatial weighting coefficient for the i^{th} neighboring sampling point; DS_i is the distance between the current sampling point and the i^{th} neighboring sampling point (m); SAC_k is the SAC value calculated based on the neighborhood strategy k; n is the number of neighboring sample points corresponding to the neighborhood strategy k (specifically, $n = 8$ when $k = 8$, and n equals the total number of samples when $k = \text{all}$); and EC_i is the electrical conductivity value at the i^{th} neighboring sampling point (dS/m).

2.3.3 Development and performance evaluation of soil salinity inversion models

In this study, we selected a total of eight models to address the inherent nonlinearity and multivariate characteristics in soil salinity modeling, and to evaluate the applicability of different machine learning and regression models for soil salinity inversion (Mahmoud et al., 2025), including four regression and four machine learning models. The regression models included multiple linear regression (MLR), partial least squares regression (PLS), ridge regression (Ridge), and least absolute shrinkage and selection operator regression (LASSO); and the machine learning models included categorical boosting (CatBoost), gradient boosting regression trees (GBRT), eXtreme Gradient Boosting (XGBoost), and random forest (RF). Regression analysis, as a classical approach for exploring variable relationships, offers theoretical principles and high interpretability, making it suitable for identifying linear associations between soil salinity and spectral indices. In contrast, machine learning models exhibit greater flexibility and adaptability in handling complex nonlinear relationships, enabling them to capture higher-order interactions between remote sensing features and soil salinity.

To ensure sample representativeness and reliable model evaluation, we randomly divided the soil samples into a training set and a validation set at a 2:1 ratio. This division provided sufficient

data for model training while retaining an independent validation set to evaluate the models' predictive performance and generalization. On this basis, three statistical metrics were used to assess the inversion accuracy of each model: coefficient of determination (R^2), root mean square error (RMSE), and mean absolute percentage error (MAPE). Specifically, R^2 measures the correlation between predicted and observed values, ranging from 0.000 to 1.000, with values closer to 1.000 indicating better model fit; RMSE reflects the absolute magnitude of prediction errors, with smaller values indicating higher model accuracy; and MAPE provides a dimensionless measure of error by expressing it as a percentage of the observed values, thereby facilitating comparison across datasets with different scales. Finally, the performance of all models on the validation set was comprehensively compared using R^2, RMSE, and MAPE, and the best-performing soil salinity inversion model was selected for subsequent spatial mapping and driving mechanism analysis.

2.4 GeoDetector

The GeoDetector is an effective statistical method for analyzing spatial heterogeneity. It not only quantifies the individual contribution of each influencing factor to soil salinity distribution (Factor Detector), but also evaluates the interactive effects between different factors on spatial patterns (Interaction Detector) (Wang et al., 2010; Zhao et al., 2020). Since the GeoDetector requires independent variables to be categorical, we performed data preprocessing based on the variable type. Fifteen continuous environmental variables were discretized using the Natural Breaks (Jenks) classification method into six categories, namely DEM, groundwater table, climatic factors (precipitation, temperature, wind speed, potential evapotranspiration, evapotranspiration, and land surface temperature), soil properties (clay content, sand content, residual moisture content, active water content, saturated hydraulic conductivity, and saturated porosity), LandScan, and NDVI. This method minimizes intra-class variance while maximizing inter-class variance. Meanwhile, three categorical variables—soil type, vegetation type, and land use/cover—were incorporated into the models after reclassification. The discretization process was implemented using the 'GD' and 'classInt' packages in the RStudio 4.5.3 statistical environment.

We collected 23 potential environmental drivers of soil salinity. After screening for multicollinearity and redundancy, 18 independent variables were selected for GeoDetector analysis, including: soil type, vegetation type, groundwater table, saturated porosity, clay content, DEM, sand content, potential evapotranspiration, saturated moisture content, precipitation, wind speed, land surface temperature, saturated hydraulic conductivity, evapotranspiration, LandScan, available water content, land use/cover, and NDVI. This approach was used to quantify the individual explanatory power (q-value) and interaction-enhancement effects of these factors on soil salinity patterns across the study area. To identify the dominant drivers and their interaction mechanisms underlying soil salinity variation in the study area, we applied both the Factor Detector and Interaction Detector to systematically analyze the explanatory power of multiple environmental variables, thereby enhancing understanding of the complex interrelationships among these factors (Zhu et al., 2024).

$$q = 1 - \frac{1}{N\delta^2} \sum_{h=1}^{L} N_h \delta_h^2, \tag{3}$$

where q is the explanatory power of the factor influencing soil salinity patterns; N is the total number of study units; δ^2 is the variance of soil salinity across the study area; L is the number of strata; h is the stratum index (h=1, 2, ⋯, L); N_h is the number of units in stratum h; and δ_h^2 is the variance of soil salinity within stratum h. The value of q ranges from 0.000 to 1.000, where a larger value indicates a stronger explanatory power of the factor influencing soil salinity patterns.

2.5 Workflow for soil salinity modeling and driver identification

This study proposed a spatially explicit framework for soil salinity inversion and multi-scale analysis by integrating SAC, aiming to reveal the scale-dependent evolution characteristics of soil

salinity spatial heterogeneity and its driving mechanisms during the transition from the oasis interior to the desert-oasis transition zone (Fig. 2). The workflow consists of three stages: (1) selection of key predictors and model evaluation; (2) spatial inversion using the selected optimal model; and (3) interpretation of spatial patterns and driving mechanisms across the oasis interior and the desert-oasis transition zone. This approach enables high-accuracy mapping and reveals scale-dependent controls on soil salinity distribution in arid ecosystems (Guo et al., 2023b; Mahmoud et al., 2025).

Visible text in Fig. 2. Model building: Data—soil sample measurements of EC; Landsat-8 (29 feature variables); influencing factors including topography, climate, soil, vegetation, human, and land use/cover. Methods—algorithms and features; Strategy I—SAC_s; Strategy II—SAC_{all}; Strategy III—baseline model (without SAC); trained different models, including GBRT, RF, CatBoost, XGBoost, MLR, Ridge, Lasso, and PLS, grouped as machine learning and regression; comparison of R^2, RMSE, and MAPE under different strategies; selecting the optimal model; $GBRT\_SAC_{\text{all}}$ model. Prediction and analysis: optimal model—soil salinity spatial prediction based on the optimal $GBRT\_SAC_{\text{all}}$ model; model performance (R^2 = 0.870, $RMSE$ = 0.808 dS/m, $MAPE$ = 1.862%). Geospatial analysis—spatial distribution of soil salinity; soil salinity mapping; spatial patterns of soil salinity; soil salinity class; land use/cover; oasis interior and its periphery. Driver analysis—influencing factors; input to GeoDetector model; regional heterogeneity; local heterogeneity.

Fig. 2 Workflow of this study. EC, electrical conductivity; SAC, spatial autocorrelation; SAC_s, local spatial autocorrelation; SAC_{all}, global spatial autocorrelation; MLR, multiple linear regression; PLS, partial least squares regression; Ridge, ridge regression; Lasso, least absolute shrinkage and selection operator regression; CatBoost, categorical boosting; GBRT, gradient boosting regression trees; XGBoost, eXtreme Gradient Boosting; RF, random forest; R^2, coefficient of determination; RMSE, root mean square error; MAPE, mean absolute percentage error.

3 Results

3.1 Soil samples and variable characteristics

3.1.1 Statistics of measured data on soil salinity Table 3 presents descriptive information on soil EC and provides the number of soil samples across different salinity ranges. For non-salinization, 169 sampling points were recorded, with EC values ranging from 0.049 to 1.993 dS/m. The standard deviation was 0.526 dS/m, and the coefficient of variation (CV) reached 107.276%, indicating the highest degree of variability. Although the mean EC was low, the differences between data points were relatively significant. The CVs for light salinization, moderate salinization, severe salinization, and extreme salinization

were 20.336%, 17.014%, 20.590%, and 17.782%, respectively. The CVs of these sampling points (105 points) within these salinity classes were less than 21.000%, indicating that the differences between the data points were small. Overall, soil salinity in the study area spanned all classes from non-salinization to extreme salinization, with an overall CV of 152.288%, reflecting pronounced spatial variability and complexity across soil salinity classes.

Table 3 Descriptive statistics of soil EC

Soil salinity class	Number of samples	Mean (dS/m)	SD (dS/m)	CV (%)	Min. (dS/m)	Max. (dS/m)	Median (dS/m)
Non-salinization	169	0.490	0.526	107.276	0.049	1.993	0.238

Soil salinity class	Number of samples	Mean (dS/m)	SD (dS/m)	CV (%)	Min. (dS/m)	Max. (dS/m)	Median (dS/m)
Light salinization	33	2.947	0.599	20.336	2.040	3.980	2.830
Moderate salinization	36	5.953	1.013	17.014	4.150	7.920	5.860
Severe salinization	26	11.270	2.320	20.590	8.010	15.590	11.345
Extreme salinization (saline soil)	10	21.285	3.785	17.782	16.030	25.600	21.775
All samples	274	3.286	5.004	152.288	0.049	25.600	1.086

Note: SD, standard deviation; CV, coefficient of variation; Min., minimum; Max., maximum.

3.1.2 Constructing and selecting variables of soil salinity inversion

Based on Pearson correlation analysis and feature importance evaluation, we selected two SAC and nine spectral indices that showed significant correlations with EC ($P < 0.05$). These variables included: SAC_8, SAC_{all}, salinity index 9 (S9), salinity index 1 (S1), salinity index 2 (S2), salinity index 8 (S8), salinity index 7 (S7), salinity index 6 (S6), modified normalized difference water index (MNDWI), salinity index I-T (SI-T), and Shortwave Infrared 2 band (B7) (Fig. 3). Among these, SAC_8 and SAC_{all} ranked the first ($r = 0.549$) and second ($r = 0.542$), respectively, and substantially outperformed the other variables. This indicated that SAC structures have strong explanatory potential for soil salinity distribution. Notably, although both SAC_8 and SAC_{all} were retained in the candidate variable pool, they represented distinct SAC characteristics.

To clarify the independent contributions of spatial features and ensure a fair model comparisons, we adopted a controlled variable strategy. Specifically, the set of spectral indices (S9, S1, S2, S7, S6, MNDWI, SI-T, and B7) served as the common baseline inputs for all three modeling paradigms. The three types of models were then differentiated by their specific spatial configurations: (1) the baseline models used only these spectral indices; (2) the models incorporated the SAC_{all} alongside these spectral indices; and (3) the models incorporated the SAC_8 alongside these spectral indices. This approach effectively isolated and evaluated the impact of different spatial neighborhood strategies. Furthermore, multicollinearity diagnostics revealed that all selected variables had variance inflation factor (VIF) values below 10, confirming their stability and suitability for subsequent modeling.

3.2 Establishing soil salinity inversion models based on Landsat-8 data

3.2.1 Regression analysis methods for modeling soil salinity inversion

Figure 4 presents the soil salinity inversion performance of four regression models (MLR, Ridge, Lasso, and PLS) on the validation set under three feature strategies (baseline models without adding SAC, models with adding SAC_8, and models with adding SAC_{all}). The results indicated that, among all four models, Lasso achieved the highest R^2 value when SAC was incorporated, demonstrating its strong synergy with SAC-enhanced predictors. Specifically, the Lasso_SAC_{all} model exhibited the highest R^2 of 0.612, slightly outperforming the Lasso_SAC_8 model ($R^2 = 0.551$). Meanwhile, the Lasso model without adding SAC achieved the lowest R^2 of only 0.080. Compared to the baseline model without adding SAC, incorporating SAC_{all} improved R^2 by 0.532, and incorporating SAC_8 improved it by 0.471, demonstrating the critical role of spatial information in improving model accuracy. Overall, incorporating SAC substantially improved the regression models' ability to accurately capture the relationship between remote sensing imagery and soil salinity, underscoring the critical importance of explicitly considering spatial structure in soil salinity mapping.

Fig. 3 Importance ranking of soil salinity feature variables based on Pearson correlation analysis (ranked by absolute correlation coefficients). Variables with a negative correlation are marked with a "(−)". S1–S9, salinity index 1–salinity index 9, respectively; MNDWI, modified normalized difference water index; SI-T, salinity index I-T; B7, Shortwave Infrared 2 band; EVI, enhanced vegetation index; MSAVI, modified soil adjusted vegetation index; B5, Near-infrared band; B2, Blue band; SI2, salinity index II; Int2, intensity index 2; B3, Green band; SI, salinity index; SI1, salinity index I; Int1, intensity index 1; SI3, salinity index III; VSSI, vegetation soil salinity index; B4, Red band; GARI, green atmospherically resistant vegetation index; B6, Shortwave Infrared 1 band; COSRI, canopy response salinity index; NDSI, normalized difference salinity index. Variables were grouped by significance: correlated (significantly correlated with soil salinity at $P < 0.05$ level) and not correlated (not significantly correlated with soil salinity at $P \geq 0.05$ level).

3.2.2 Machine learning methods for modeling soil salinity inversion

Figure 5 displays the soil salinity inversion results of four machine learning models (RF, CatBoost, GBRT, and XGBoost) on the validation set under three feature strategies (baseline models without adding SAC, models with adding SAC_8, and models with adding SAC_{all}). Models incorporating SAC consistently achieved higher R^2 values and lower prediction errors, demonstrating that spatial structural information enhances machine learning models' ability to capture the relationship between remote sensing spectral data and soil salinity.

Among all four models, the GBRT model had the highest R^2 on the test set, with 0.838 and 0.870 for incorporating SAC_8 and SAC_{all}, respectively, outperforming the other models (Fig. 5a). The GBRT model incorporating SAC_{all} achieved the highest R^2 (0.870) and the lowest RMSE (0.808 dS/m), highlighting the role of SAC in enhancing predictive precision. The results of the optimal GBRT model revealed that, compared with the traditional spectral index-based GBRT model, the GBRT model incorporating SAC_{all} demonstrated significantly improved performance. The R^2 increased by 7.320%, the RMSE decreased by 20.230%, and the MAPE decreased by 121.01%. The next-best-performing model, RF, achieved R^2 values of 0.811 and 0.858 with the addition of SAC_8 and SAC_{all}, respectively. In contrast, the XGBoost model performed relatively poorly. Moreover, models incorporating SAC_{all} had the most significant impact on GBRT and CatBoost, whereas SAC_{all} had a minor effect on XGBoost, indicating that different models respond differently to SAC.

Fig. 4 Scatter plots of predicted and measured soil EC using four regression models with different feature strategies. (a1–a4) MLR, Ridge, Lasso, and PLS models with SAC_{all} added, respectively; (b1–b4) MLR, Ridge, Lasso, and PLS models with SAC_8 added, respectively; (c) MLR, Ridge, Lasso, and PLS models without adding SAC, respectively. R^2, RMSE, and MAPE denote the performance evaluation metrics.

Visible panel annotations:

- **(a1) MLR_SAC$_{all}$**: RMSE: 3.742 dS/m; MAPE: 2.166; R^2: 0.523.
- **(a2) Ridge_SAC$_{all}$**: RMSE: 4.390 dS/m; MAPE: 2.040; R^2: 0.485.
- **(a3) Lasso_SAC$_{all}$**: RMSE: 3.263 dS/m; MAPE: 2.857; R^2: 0.612.
- **(a4) PLS_SAC$_{all}$**: RMSE: 3.845 dS/m; MAPE: 2.014; R^2: 0.496.
- **(b1) MLR_SAC$_8$**: RMSE: 3.937 dS/m; MAPE: 2.738; R^2: 0.507.
- **(b2) Ridge_SAC$_8$**: RMSE: 4.439 dS/m; MAPE: 2.065; R^2: 0.436.
- **(b3) Lasso_SAC$_8$**: RMSE: 3.962 dS/m; MAPE: 2.248; R^2: 0.551.
- **(b4) PLS_SAC$_8$**: RMSE: 4.027 dS/m; MAPE: 2.244; R^2: 0.447.
- **(c1) MLR**: RMSE: 3.571 dS/m; MAPE: 2.507; R^2: 0.328.
- **(c2) Ridge**: RMSE: 4.034 dS/m; MAPE: 7.097; R^2: 0.226.
- **(c3) Lasso**: RMSE: 4.637 dS/m; MAPE: 10.589; R^2: 0.080.
- **(c4) PLS**: RMSE: 3.471 dS/m; MAPE: 8.273; R^2: 0.419.

Axes and legend visible in the figure: y-axis, Predicted EC (dS/m); x-axis, Measured EC (dS/m); legend, 1:1 line and fitting line.

Despite the CatBoost model' s relatively poor performance, R^2 improved by 0.100% and 12.850% with the additions of SAC_8 and SAC_{all}, respectively, further demonstrating the effectiveness of adding SAC in improving model accuracy. Furthermore, the validity of this conclusion was corroborated through a comparative assessment of the RMSE and MAPE performance metrics across the evaluated models (Fig. 5b and c). Consequently, the GBRT_SAC_{all} model, which exhibited optimal performance, was selected as the definitive inversion model for the spatial mapping of regional soil salinity.

3.3 Spatial variation characteristics of soil salinity

3.3.1 Soil salinity mapping based on the optimal GBRT_SAC_{all} modeling

The spatial distribution of land use/cover, soil pH, EC, and soil salinity classes in the study area is shown in Figure 6. The oasis interior was dominated by cropland, grassland, and forest. In contrast, the surrounding areas were primarily barren land (Fig. 6a). Soil pH values were predominantly weakly alkaline (Fig. 6b), and soil salinity was primarily sulfate salts, followed by chloride salts. High concentrations of sodium ions may cause soil dispersion, thereby reducing aggregate stability and permeability. The spatial distribution of soil salinization exhibited considerable heterogeneity, with EC values ranging from 0.097 to 23.854 dS/m (Fig. 6c). Although areas with non-salinization accounted for 63.830% of the total study area, approximately 1256.75 km^2 (36.170%) was affected by varying degrees of salinization. The majority of the affected regions predominantly exhibited light to moderate salinization, whereas severe salinization was concentrated in the southern and southeastern parts of the study area,

Fig. 5 Comparison of the predictive performance of different machine learning models with different feature strategies (models without adding SAC, with adding SAC_8, and with adding SAC_{all}). (a), R^2; (b), RMSE; (c), MAPE.

Visible figure labels: (a) R^2; (b) RMSE; (c) MAPE. Axis labels: Model; R^2; RMSE (dS/m); MAPE. Legend: GBRT; RF; XGBoost; CatBoost. Model labels: GBRT_SAC_{all}; RF_SAC_{all}; GBRT_SAC_8; RF_SAC_8; GBRT; RF; XGBoost_SAC_{all}; XGBoost_SAC_8; XGBoost; CatBoost_SAC_8; CatBoost_SAC_{all}; CatBoost.

Visible figure labels for **Fig. 6**: (a) Land use/cover; (b) pH; (c) EC; (d) Soil salinity class; (e) Clay content; (f) Sand content. Legend items: Minqin County; Non-study area. Land use/cover classes: Barren land; Cropland; Forest; Grassland; Impervious surface; Water body; Wetland. pH: High: 9.047; Low: 7.041. EC (dS/m): High: 23.854; Low: 0.097. Soil salinity class: Non-salinization; Light salinization; Moderate salinization; Severe salinization; Extreme salinization. Clay content (%): High: 29; Low: 4. Sand content (%): High: 93; Low: 42.

Fig. 6 Spatial distribution of land use/cover (a), soil pH (b), EC (c), soil salinity class (d), clay content (e), and sand content (f) in the study area.

encompassing 7.220% of the total area (Fig. 6d). In contrast, the areas surrounding the oasis were dominated by barren land (desert and Gobi landscapes), which are characterized by coarse-textured sandy soils with high gravel content, low cation exchange capacity, and limited adsorption capacity. Consequently, the overall soil salinity in these transitional areas remained comparatively low. As illustrated in Figure 6e, clay content attained its highest levels in the central and southwestern parts of the oasis, particularly within cropland areas, and declined progressively toward the oasis margins. Furthermore, sand content remained low in the oasis interior but increased markedly in the desert–oasis transition zone (Fig. 6f). The spatial distribution of clay content was largely complementary to that of sand content, with high-clay and high-sand areas exhibiting minimal overlap.

3.3.2 Spatial patterns of soil salinity and its relationship with land use/cover change in different buffer zones

Soil salinization exhibited a distinct spatial gradient from the interior of the oasis toward the desert–oasis transition zone (Fig. 7a). The interior of the oasis was dominated by non-salinized soil (60.890%) and lightly salinized soil (36.640%). Areas with moderately and severely salinized soils combined accounted for less than 2.470% of the study area, while extremely salinized soil was absent, indicating low salinization levels and a stable soil environment. Nevertheless, the areal extent of severe salinization increased progressively with distance from the oasis boundary. Within the 100 m buffer zone from the oasis boundary, soil salinization level intensified, the area with light salinization declined sharply, moderate salinization remained relatively stable, and the percentage of severe salinization rose to 10.500%, reflecting the onset of significant salt accumulation. At the 500 m buffer distance, the area percentage of severe salinization peaked at 11.720%, becoming the dominant salinity class and suggesting a transition toward a moderately to severely salinized ecotone. At the 1000 m buffer distance, soil salinization declined slightly but remained substantially higher than that in the oasis interior area, while moderate salinization continued to decrease, suggesting potential surface changes, such as stabilized dunes or salt crust formation.

Fig. 7 Area distribution of soil salinity classes with distance from the oasis boundary (a), and soil salinity variation across different land use/cover types with distance from the oasis boundary (b). In the right panel, the colored filled surfaces represent the continuous trend surfaces of EC for each land use/cover type, illustrating the spatial variation of soil salinity.

Significant spatial differentiation in soil salinity was evident between the oasis interior and the desert-oasis transition zone, attributable to variations in land use/cover types (Fig. 7b). The results indicated that soil salinity was highest in barren land, followed by grassland and cropland, and lowest in impervious surface. As a typical indicator of ecological degradation, barren land showed a sharp increase in soil salinity, rising from 2.624 dS/m within the interior of the oasis to

5.599 dS/m in the 1000 m buffer zone. In contrast, impervious surface exhibited pronounced salinity fluctuations, rising sharply from 1.148 dS/m within the interior of the oasis to 2.618 dS/m in the 1000 m buffer zone.

3.4 Drivers of soil salinization

3.4.1 Drivers affecting soil salinity across the entire study area Factor Detector analysis indicated that soil type exhibited the highest q-value (0.186), identifying it as the dominant driver of soil salinity spatial distribution (Fig. 8a). Vegetation type also showed strong explanatory power, with q-values of 0.150. In addition, groundwater table (q-value = 0.146), saturated porosity (q-value = 0.131), and clay content (q-value = 0.112) showed moderate contributions to the spatial variability of soil salinity. In contrast, factors such as actual evapotranspiration (q-value = 0.026), LandScan (q-value = 0.023), active water content (q-value = 0.021), land use/cover (q-value = 0.019), and NDVI (q-value = 0.004) exhibited relatively low q-values, indicating limited influence on soil salinity variation.

Fig. 8 Analysis of the driving mechanism of soil salinity using the GeoDetector model. (a), explanatory power (q-value) of individual influencing factors derived from the Factor Detector analysis; (b), interactive effects between pairs of influencing factors derived from the Interaction Detector analysis. The color gradient represents the magnitude of the interaction strength. NDVI, normalized difference vegetation index.

Interaction Detector analysis suggested that the combined effect of drivers significantly enhanced the explanatory power for soil salinity spatial variation, exhibiting non-linear enhancement and indicating the importance of integrated factor effects (Fig. 8b). Among these, the combinations of vegetation type and saturated porosity (q-value=0.393), vegetation type and DEM (q=0.385), and clay content and saturated porosity (q-value=0.379) exhibited relatively strong interactive effects. These findings highlighted pronounced non-linear coupling among hydrological conditions, soil physical properties, and vegetation dynamics, indicating that soil salinity patterns are strongly governed by the integrated influence of multiple factors. Notably, although LandScan and land use/cover showed weak individual explanatory power (q-value<0.050), their interactions with soil type yielded substantially higher q-values (0.210 and 0.245, respectively). This suggested that anthropogenic influences can amplify natural controls through synergistic interactions. Overall, the spatial distribution of soil salinity is regulated by complex, non-linear interactions among multiple environmental and human drivers.

3.4.2 Differences in driving factors of soil salinity along the oasis boundary gradient

The explanatory power of driving factors for soil salinity spatial variation exhibited pronounced spatial gradients across buffer zones at increasing distances from the oasis boundary (Fig. 9). Within the oasis interior area, groundwater table emerged as the dominant driving factor (q-value=0.277), underscoring the pivotal role of irrigation practices and groundwater regulation in shaping soil salinity dynamics in this intensively managed environment. Moving outward into

Fig. 9 Explanatory power (q-value) of influencing factors affecting soil salinity changes at different distances from the oasis boundary based on Factor Detection analysis. (a), oasis interior (the original Minqin Oasis boundary, i.e., only the oasis interior area); (b), 100 m buffer zone (the oasis interior plus a 100 m buffer zone extending outward from the oasis perimeter); (c) 500 m buffer zone (the oasis interior plus a 500 m buffer zone extending outward from the oasis perimeter); (d), 1000 m buffer zone (the oasis interior plus a 1000 m buffer zone extending outward from the oasis perimeter.

the desert-oasis transition zone, the dominant driving drivers shifted progressively from hydrological processes to soil physical properties. At the 1000 m buffer zone from the oasis boundary, sand content emerged as the most influential factor with a q-value of 0.483, followed closely by clay content (q-value = 0.479) and saturated hydraulic conductivity (q-value = 0.409). This pattern indicated that soil texture and hydraulic properties exert primary control over salt accumulation in this ecotonal region, revealing a scale-dependent amplification of salinization processes in coarse-textured soils. Although soil type showed the highest explanatory power (q-value = 0.186) in the overall single-factor analysis across the study area, its constituent attributes—particularly textural components—demonstrated stronger explanatory capacity at finer spatial scales within the buffer zones. This finding highlighted a shift in dominance from categorical soil classification to quantitative soil physical properties at local scales. Notably, most influencing factors attained their peak q-values within the desert-oasis transition zone (1000 m buffer), indicating intensified water–salt interactions and heightened ecosystem sensitivity to environmental perturbations in this transition zone. In contrast, climatic variables, including evapotranspiration and precipitation, consistently exhibited weak explanatory power (q-value < 0.050), suggesting that regional climate exerts a limited direct influence on salinity patterns relative to local soil-hydrological conditions. Overall, the driving mechanisms of soil salinity exhibit distinct spatial differentiation. Within the oasis interior, anthropogenic factors—particularly irrigation and groundwater management—play a dominant role in regulating salt distribution. Beyond the oasis boundary, the influence of soil texture, soil type, and hydrological properties becomes increasingly pronounced, ultimately governing salinity dynamics within the desert-oasis transition zone.

4 Discussion

4.1 Advantages of improving the accuracy of soil salinity estimation

Previous studies on soil salinity inversion have primarily relied on empirical spectral indices, demonstrating good effectiveness and reproducibility (Allbed et al., 2014; Khosravichenar et al., 2023; Jia et al., 2024b). While these spectral indices are effective in characterizing soil salinity at specific sampling points, they often treat each location as an independent entity, thereby failing to capture the broader spatial patterns. Furthermore, most models have overlooked the spatial continuity and autocorrelation characteristics of soil salinity distribution, focusing solely on spectral information while lacking explicit geospatial representation (Xiao et al., 2024). Although spectral indices remain fundamental predictors, systematic efforts to incorporate and quantify SAC have been limited.

In this context, to the best of our knowledge, this study represented one of the first attempts to incorporate SAC variables as direct input features together with spectral indices in machine learning models for soil salinity mapping. Importantly, SAC here was not used to build an independent spatial statistical model, but served as an auxiliary feature to represent neighborhood dependence and contextual consistency. The findings showed that incorporating SAC_{all} information markedly improves model performance, with the $GBRT\_SAC_{\text{all}}$ model achieving the highest R^2 (0.870) and the lowest RMSE (0.808 dS/m), highlighting the role of SAC in enhancing the predictive precision of soil salinity inversion models.

A comparative assessment of different modeling strategies (Fig. 10) provided further insight into the source of this predictive performance. The spectral-only baseline model (GBRT), which explicitly excluded all SAC inputs, achieved a relatively high R^2 of 0.725 (Fig. 10a). This result confirmed that the primary predictive strength derives from the intrinsic physical relationship between spectral signals and soil salinity, rather than from spatial memorization. The incorporation of SAC further improved the R^2 of the full-variable model with an increase from 0.725 to 0.768 (Fig. 10a and b), while that of the feature-optimized model improved from 0.797 to 0.831 (Fig. 10c and d). These findings indicated that effective feature selection can reduce

redundancy and interference among predictors, and that combining optimized variables with SAC can jointly enhance model accuracy and stability (Li, 2018; Wadoux et al., 2021; Yang et al., 2024). Regarding the validation strategy, although spatial cross-validation is valuable for testing transferability, validation choices should be aligned with the inference target (within-domain map accuracy versus spatial transferability), and no single CV scheme is universally optimal under spatial dependence (Metternicht et al., 2003; Hengl et al., 2018). Accordingly, the performance improvements observed in Figure 10 can be interpreted primarily as enhancements in within-domain regional mapping accuracy under the adopted validation framework, rather than definitive evidence of long-range spatial extrapolation capability.

Figure 10 panels

(a) Full variables without SAC_{all}

- $y = 0.842x + 0.709$
- RMSE: 1.175 dS/m
- R^2: 0.725
- Y-axis: Predicted EC (dS/m)
- Legend: Sample; 95% CI; Fitting line; 1:1 line

(b) Full variables with SAC_{all}

- $y = 1.001x + 0.414$
- RMSE: 1.079 dS/m
- R^2: 0.768

(c) Filtered variables without SAC_{all}

- $y = 0.793x + 0.580$
- RMSE: 1.011 dS/m
- R^2: 0.797
- Y-axis: Predicted EC (dS/m)
- X-axis: Observed EC (dS/m)

(d) Filtered variables with SAC_{all}

- $y = 0.824x + 0.446$
- RMSE: 0.922 dS/m
- R^2: 0.831
- X-axis: Observed EC (dS/m)

Fig. 10 Prediction accuracy comparison of the GBRT model under different modeling scenarios. (a), full variables without SAC_{all}; (b), full variables with SAC_{all}; (c), filtered variables without SAC_{all}; (d), filtered variables with SAC_{all}.

In parallel with spatial feature integration, the findings indicated the importance of feature selection in reducing predictor redundancy. Although Pearson correlation measures only linear dependence and its P-value is sensitive to sample size, we used the significance test merely as a conservative coarse-screening step. Specifically, we removed variables for which our data provided insufficient evidence of linear association ($P \geq 0.05$), thereby reducing dimensionality and redundancy given the limited sample size. The observed improvement in model performance (higher R^2 and lower RMSE) indicated that, in the current dataset, these excluded variables did not contribute stable predictive gains. Thus, Pearson correlation-based filtering proved to be a pragmatic strategy for optimizing the feature space without compromising predictive performance.

Nevertheless, reliance solely on spectral information presents inherent limitations in arid environments, a challenge widely recognized in remote sensing applications. The 30 m spatial resolution of Landsat imagery may introduce mixed-pixel effects in fragmented, mosaic-like oasis landscapes, while the spatial heterogeneity of salinization patches further undermines the stability of spectral–salinity relationships (Qin et al., 2025). Furthermore, variations in surface conditions such as physical soil crusting or dry sandy surfaces may cause spectral confusion, thereby increasing the error in salinity estimation based solely on spectral data. Although these factors were not quantitatively separated in this study, the significant performance improvement after integrating SAC into the model suggests that SAC, as a spatial-context feature, can provide neighborhood-dependence constraints for inversion. By incorporating neighborhood-level spatial dependence, SAC enables the model to utilize both pixel-level spectral signals and surrounding spatial structure, thereby enhancing prediction stability and improving the spatial coherence of salinity maps.

Overall, integrating SAC selection represents a promising framework for high-accuracy and interpretable soil salinity inversion. However, the current models primarily incorporate spectral and spatial features. While the influence of topographic and climatic factors has been examined (Xiong et al., 2025), these variables have not yet been directly integrated into the predictive framework. Future research should therefore integrate multi-source environmental data and explore advanced non-linear feature selection techniques (e.g., Mutual Information or SHapley Additive exPlanations-based ranking). In addition, although the present validation confirms strong within-domain performance, the application of spatial or block cross-validation methods would further strengthen the evaluation of model transferability under stricter spatial independence constraints. The integrated modeling framework, established based on the distribution characteristics of soil salinization in the Minqin Oasis, holds broad application potential in other arid and semi-arid oasis systems facing similar issues, provided that region-specific recalibration is conducted to mitigate spatial leakage in cross-regional implementations.

4.2 Shifts in dominant factors controlling soil salinity from oasis interior to its desert-oasis transition zone

At the regional scale, soil type and vegetation type—representing the integrated outcomes of long-term water–salt interactions and anthropogenic activities—establish the macro-level framework governing the spatial distribution of soil salinity. In the Minqin Oasis, irrigated agriculture is predominantly concentrated on alluvial plains characterized by fine-textured soils with high water-holding capacity. Sustained artificial irrigation has resulted in dense vegetation cover, while irrigation-induced fluctuations in the groundwater table have become the principal control on salt transport and accumulation. Frequent irrigation promotes downward leaching of salts and facilitates their removal via drainage systems, leading to predominantly non-saline or slightly saline conditions (Gao et al., 2025a). This reflects the effectiveness of irrigation-driven hydrological regulation in mitigating salinity accumulation within the oasis interior.

In contrast, the desert-oasis transition zone is characterized by sandy soils and sparse vegetation cover. These marginal areas function as zones of salt accumulation, typically located at considerable distances from rivers and settlements. They are often associated with terminal or poorly flushed environments that receive agricultural drainage, where land use/cover is dominated by bare land or abandoned cropland (Jin et al., 2017). Importantly, the role of soil texture should be distinguished from the underlying hydrological drivers. Although deep groundwater and weak capillarity in sandy soils would theoretically limit upward salt transport (Fang and Fan, 2020), salinization in this zone is primarily driven by lateral input of saline drainage water from irrigated areas and intense evaporative demand under hyper-arid conditions, which drives salt concentration near the surface following infiltration.

Within this framework, sand content should not be interpreted as a direct source of salinity; rather, it serves mainly as an environmental indicator that co-varies with the oasis-margin hydrological context conducive to salt retention and accumulation. Once saline inputs reach these

areas, the hydraulic properties of sandy soils—such as high infiltration rates and limited capillarity—facilitate rapid water redistribution and loss, thereby enhancing near-surface salt retention. In this sense, sandy marginal zones function as an "evaporation filter" (Cao et al., 2025), whereby water is preferentially removed through evaporation while salts are progressively concentrated within the shallow soil profile. This highlights the oasis margin as a critical hydrological-ecological interface where natural processes and human interventions interact.

As the influence of the groundwater table diminishes and sand content increases toward the oasis margin, soil salinity exhibits pronounced spatial variability, reaching its maximum within 1000 m of the oasis boundary. This ecotonal zone represents the area where environmental controls on soil salinity variation are strongest, marking it as a key transition region where hydrological, salt, and ecological processes jointly govern soil salt redistribution. Here, soil salinity primarily originates from agricultural drainage; after discharge through alkaline channels (Ma et al., 2023), intense evaporation leads to substantial salt accumulation (Ngabire et al., 2022), forming highly heterogeneous patches of severe salinization. These patterns indicate that anthropogenic impacts extend beyond cultivated land, as drainage networks transport salts to the desert-oasis transition zone, thereby shaping the high-salinity mosaics observed in the surrounding landscape (Velilla et al., 2025).

Ecological water replenishment projects have further altered these dynamics by artificially recharging surface and subsurface water in recent years. Such interventions can replenish and raise groundwater levels, thereby altering the water-salt exchange between soil and groundwater. A moderate increase in groundwater level enhances soil moisture and promotes salt leaching, reducing surface salinity and facilitating vegetation recovery (Ding et al., 2025). However, when the groundwater table approaches the capillary zone, upward migration of saline water through evaporation becomes intensified, leading to secondary salinization and salt accumulation in the topsoil (Huang et al., 2021; Lu et al., 2025a). Consequently, ecological water replenishment exerts a dual effect. While it may alleviate salinization pressure in well-drained areas, it exacerbates salt accumulation in poorly drained or low-lying regions. Overall, groundwater depth serves as a crucial intermediary linking anthropogenic water management, hydrological processes, and the spatial heterogeneity of soil salinity.

4.3 Implications for soil salinization management and control

Under conditions where macroscale factors such as uniform topography and static indicators of human activities exert limited influence due to their inherent constraints (Xiao et al., 2024; Cao et al., 2025), the regulatory role of soil texture becomes more prominent. These findings emphasize that, under macroscopically homogeneous conditions typical of arid regions, microscale heterogeneity in soil texture plays a decisive role in regulating water-salt migration processes. Therefore, traditional homogeneous land management strategies may be insufficient to control the expansion of soil salinization.

A spatially differentiated management strategy is therefore required. Within the Minqin Oasis, optimizing irrigation regimes is essential to minimize excess drainage and reduce salt mobilization. In contrast, high-risk marginal zones should be managed through the establishment of ecological drainage corridors or saline-alkali restoration areas to intercept and regulate salt transport pathways. Additionally, the functional roles of different land use/cover types must be explicitly considered. Cropland and impervious surface, typically subject to intensive human disturbance, exhibit soil salinity patterns strongly influenced by irrigation practices and land management. Conversely, natural ecosystems such as grasslands provide important buffering and regulatory functions that contribute to maintaining regional water–salt balance. Barren land, which exhibits pronounced salt accumulation, should be prioritized for ecological restoration and salinity mitigation interventions (Javed et al., 2025). In summary, sustainable soil salinity management requires an integrated approach that combines source reduction, process regulation, and end-of-pipe remediation. Such a framework is essential for maintaining long-term water–salt balance and enhancing the ecological resilience of the desert–oasis ecotone systems.

5 Conclusions

This study systematically evaluated the performance of multiple regression and machine learning models for soil salinity inversion in the Minqin Oasis and its desert–oasis ecotone, a representative inland arid region of Northwest China. By integrating spectral indices with SAC, the proposed framework significantly improved model accuracy, generalization capacity, and interpretability. Among the tested models, the GBRT_SAC_{all} configuration achieved the best overall performance, demonstrating strong capability for both salinity inversion and spatial mapping.

Soil salinity inversion results indicated that light salinization predominates within the oasis interior, whereas moderate to severe salinization is concentrated in the desert-oasis transition zone. Soil type and vegetation type emerge as the dominant drivers shaping soil salinity patterns. However, the relative importance of these drivers exhibits a transitional shift from the oasis interior to its desert-oasis transition zone. Within the oasis interior, salinity accumulation is governed by hydrological processes, specifically irrigation-regulated groundwater tables. In contrast, the transition zone exhibits a pattern mediated by soil texture, in which sand content becomes the primary factor governing salt redistribution at the dynamic interface between hydrological and aeolian processes.

Overall, this study demonstrates that incorporating SAC effectively captures the inherent spatial dependence of soil salinity patterns. The proposed modeling framework not only improves the characterization of soil salinization patterns but also provides a scientific basis for differentiated land management and ecological restoration strategies in arid regions, offering methodological guidance for future large-scale salinization monitoring and risk assessment.

Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Central Government Guiding Funds for Local Scientific and Technological Development (23ZYQHO298), the Science and Technology Program of Gansu Province (21JR7RA070), and the Open Fund of the National Cryosphere Desert Data Center of China (2024NCDC003).

Author contributions

Conceptualization: ZHAO Dan, XUE Xian; Data curation: ZHAO Dan, YANG Xiya, GAO Yukun; Formal analysis: ZHAO Dan, YANG Xiya, PAN Jing, YOU Quangang; Investigation: ZHAO Dan, YANG Xiya, GAO Yukun; Methodology: ZHAO Dan, GAO Yukun; Resources: ZHAO Dan, YANG Xiya, GAO Yukun, PAN Jing, YOU Quangang, XUE Xian; Visualization: ZHAO Dan; Writing - original draft preparation: ZHAO Dan, XUE Xian; Writing - review and editing: XUE Xian; Funding acquisition: XUE Xian; Supervision: PAN Jing, YOU Quangang, XUE Xian. All authors approved the manuscript.

References

Aihaiti A, Nurmemet I, Yu X R, et al. 2025. An enhanced soil salinity estimation method for arid regions using multisource remote sensing data and advanced feature selection. CATENA, 256: 109116, doi: 10.1016/j.catena.2025.109116.

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, 230–231: 1–8.

Cao Q Q, Liu T, Xia J B, et al. 2025. Shifting sands to sustainable soils: spatial dynamics of soil water and salinity in a desert oasis ecotone. Agricultural Water Management, 316: 109562, doi: 10.1016/j.agwat.2025.109562.

Chang L, Yi S H, Qin Y, et al. 2024. Exploring main soil drivers of vegetation succession in abandoned croplands of Minqin Oasis, China. PeerJ, 12: e17627, doi: 10.7717/peerj.17627.

Chang X L, Ji S X, Qiao R R, et al. 2019. NDVI-based identification of oasis-desert transitional zone wideness a case study in the central Hexi corridor. Acta Ecologica Sinica, 40(15): 5327–5336. (in Chinese)

Chen H F, Wu J W, Xu C. 2024a. Monitoring soil salinity classes through remote sensing-based ensemble learning concept: considering scale effects. Remote Sensing, 16(4): 642, doi: 10.3390/rs16040642.

Chen H Y, Ma Y, Zhu A X, et al. 2021. Soil salinity inversion based on differentiated fusion of satellite image and ground spectra. International Journal of Applied Earth Observation and Geoinformation, 101: 102360, doi: 10.1016/j.jag.2021.102360.

Chen Y N, Fang G H, Li Z, et al. 2024b. The crisis in oases: research on ecological security and sustainable development in arid regions. Annual Review of Environment and Resources, 49(1): 1–20.

Cui G, Liu Y, Li X J, et al. 2025. Impacts of groundwater storage variability on soil salinization in a semi-arid agricultural plain. Geoderma, 454: 117162, doi: 10.1016/j.geoderma.2024.117162.

Ding Y D, Lü H S, Xu L G, et al. 2025. Estimating the groundwater table threshold for mitigating soil salinization in the Songnen Plain of China. Journal of Hydrology: Regional Studies, 59: 102326, doi: 10.1016/j.ejrh.2025.102326.

Du R Q, Xiang Y Z, Chen J Y, et al. 2024. Potential of solar-induced chlorophyll fluorescence (SIF) to access long-term dynamics of soil salinity using OCO-2 satellite data and machine learning method. Geoderma, 444: 116855, doi: 10.1016/j.geoderma.2024.116855.

Fang H Y, Fan Z M. 2020. Assessment of soil erosion at multiple spatial scales following land use changes in 1980–2017 in the black soil region, (NE) China.

International Journal of Environmental Research and Public Health, 17(20): 7378, doi: 10.3390/ijerph17207378.

Gao L R, Xu E Q. 2025a. Mapping cropland soil salinity using multi-cycle classification to mitigate retrieval errors. Computers and Electronics in Agriculture, 231: 110055, doi: 10.1016/j.compag.2025.110055.

Gao Z R, Li X J, Zuo L J, et al. 2025b. Unveiling soil salinity patterns in soda saline-alkali regions using Sentinel-2 and SDGSAT-1 thermal infrared data. Remote Sensing of Environment, 322: 114708, doi: 10.1016/j.rse.2025.114708.

Ge X Y, Ding J L, Teng D X, et al. 2022. Updated soil salinity with fine spatial resolution and high accuracy: The synergy of Sentinel-2 MSI, environmental covariates and hybrid machine learning approaches. CATENA, 212: 106054, doi: 10.1016/j.catena.2022.106054.

Ghasempour R, Aalami M T, Saghebian S M, et al. 2024. Analysis of spatiotemporal variations of drought and soil salinity via integrated multiscale and remote sensing-based techniques (case study: Urmia Lake Basin). Ecological Informatics, 81: 102560, doi: 10.1016/j.ecoinf.2024.102560.

Guo B, Liu Y F, Fan J F, et al. 2023a. The salinization process and its response to the combined processes of climate change-human activity in the yellow river delta between 1984 and 2022. CATENA, 231: 107301, doi: 10.1016/j.catena.2023.107301.

Guo B, Yang X C, Yang M L, et al. 2023b. Mapping soil salinity using a combination of vegetation index time series and single-temporal remote sensing images in the Yellow River Delta, China. CATENA, 231: 107313, doi: 10.1016/j.catena.2023.107313.

Han H J, Yan X, Xie H W, et al. 2023. Incorporating a new landscape intensity indicator into landscape metrics to better understand controls of water quality and optimal width of riparian buffer zone. Journal of Hydrology, 625: 130088, doi: 10.1016/j.jhydrol.2023.130088.

Han L J, Ding J L, Zhang J Y, et al. 2021. Precipitation events determine the spatiotemporal distribution of playa surface salinity in arid regions: Evidence from satellite data fused via the enhanced spatial and temporal adaptive reflectance fusion model. CATENA, 206: 105546, doi: 10.1016/j.catena.2021.105546.

Hao X M, Li W H. 2016. Oasis cold island effect and its influence on air temperature: A case study of Tarim Basin, Northwest China. Journal of Arid Land, 8(2): 172–183.

Hao X M, Sun F, Zhang J J, et al. 2025. Vapor pressure deficit governs oasis cooling efficiency and drought intensified water-heat tradeoffs in arid regions. Journal of Hydrology, 661: 133804, doi: 10.1016/j.jhydrol.2025.133804.

Hao Y Y, Xie Y W, Ma J H, et al. 2017. The critical role of local policy effects in arid watershed groundwater resources sustainability: a case study in the Minqin Oasis, China. Science of The Total Environment, 601–602: 1084–1096.

Hassani A, Azapagic A, Shokri N. 2021. Global predictions of primary soil salinization under changing climate in the 21st century. Nature Communications, 12(1): 6663, doi: 10.1038/s41467-021-26907-3.

Hassani A, Smith P, Shokri N. 2024. Negative correlation between soil salinity and soil organic carbon variability. Proceedings of the National Academy of Sciences of the United States of America, 121(18): e2317332121, doi: 10.1073/pnas.2317332121.

Hengl T, Nussbaum M, Wright M N, et al. 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 6: e5518, doi: 10.7717/peerj.5518.

Huang F, Ochoa C G, Chen X. 2021. Assessing environmental water requirement for groundwater-dependent vegetation in arid inland basins by combining the copula joint distribution function and the dual objective optimization: An application to the Turpan Basin, China. Science of The Total Environment, 799: 149323, doi: 10.1016/j.scitotenv.2021.149323.

Ivushkin K, Bartholomeus H, Bregt A K, et al. 2019a. UAV based soil salinity assessment of cropland. Geoderma, 338:

502–512.

Ivushkin K, Bartholomeus H, Bregt A K, et al. 2019b. Global mapping of soil salinity change. Remote Sensing of Environment, 231: 111260, doi: 10.1016/j.rse.2019.111260.

Javed T, Zhang J Z, Li W H, et al. 2025. Unlocking soil salinity prediction with remote sensing indices and environmental insights. Land Degradation & Development, 36(15): 5168–5183.

Jia P P, He W, Hu Y, et al. 2024a. Inversion of coastal cultivated soil salt content based on multi-source spectra and environmental variables. Soil & Tillage Research, 241: 106124, doi: 10.1016/j.still.2024.106124.

Jia P P, Zhang J H, Liang Y N, et al. 2024b. The inversion of arid-coastal cultivated soil salinity using explainable machine learning and Sentinel-2. Ecological Indicators, 166: 112364, doi: 10.1016/j.ecolind.2024.112364.

Jiang X F, Duan H C, Liao J, et al. 2022. Estimation of soil salinization by machine learning algorithms in different arid regions of Northwest China. Remote Sensing, 14(2): 347, doi: 10.3390/rs14020347.

Jin Y Z, Zeng J J, Hu X Q, et al. 2017. Research on the response of the water sources to the climatic change in Shiyang River Basin. In: Proceedings of the 3rd International Conference on Water Resource and Environment. Qingdao: IOP Publishing, doi: 10.1088/1755-1315/82/1/012093.

Khosravichenar A, Aalijahan M, Moaazeni S, et al. 2023. Assessing a multi-method approach for dryland soil salinization with respect to climate change and global warming-the example of the Bajestan region (NE Iran). Ecological Indicators, 154: 110639, doi: 10.1016/j.ecolind.2023.110639.

Kumar P, Tiwari P, Biswas A, et al. 2024. Spatio-temporal assessment of soil salinization utilizing remote sensing derivatives, and prediction modeling: Implications for sustainable development. Geoscience Frontiers, 15(6): 101881, doi: 10.1016/j.gsf.2024.1018811674-9871.

Li D H, Wang Z, Li H, et al. 2025. Accelerated disappearance of desert oases: The impact of climate warming and farmland expansion. Agricultural Water Management, 319: 109798, doi: 10.1016/j.agwat.2025.109798.

Li K, Chi G Q, Wang L, et al. 2018. Identifying the critical riparian buffer zone with the strongest linkage between landscape characteristics and surface water quality. Ecological Indicators, 93: 741–752.

Li Y S, Chang C Y, Wang Z R, et al. 2023. Upscaling remote sensing inversion and dynamic monitoring of soil salinization in the Yellow River Delta, China. Ecological Indicators, 148: 110087, doi: 10.1016/j.ecolind.2023.110087.

Li Z. 2018. GBDT-SVM credit risk assessment model and empirical analysis of peer-to-peer borrowers under consideration of audit information. Open Journal of Business and Management, 6: 362–372.

Liu B S, Huang L H, Jiang X T, et al. 2024. Quantitative evaluation and mechanism analysis of soil chemical factors affecting rice yield in saline-sodic paddy fields. Science of The Total Environment, 929: 172584, doi: 10.1016/j.scitotenv.2024.172584.

Liu X X, Gou Q Q, Wang G H. 2025. Analysis of the width dynamics of the oasis-desert ecotone at the northern margin of the Tarim Basin. Arid Zone Research, 42(3): 534–544. (in Chinese)

Lu B, Castillo I, Chiang L, et al. 2014. Industrial PLS model variable selection using moving window variable importance in projection. Chemometrics and Intelligent Laboratory Systems, 135: 90–109.

Lu H Y, Mokarram M. 2025a. Forecasting climate change effects on saline lakes through advanced remote sensing and deep learning. Science of The Total Environment, 980: 179582, doi: 10.1016/j.scitotenv.2025.179582.

Lu J Q, Zhang X F, Liang S M, et al. 2023. Spatiotemporal dynamics of vegetation index in an oasis-desert transition zone and relationship with environmental factors. Sustainability, 15(4): 3503, doi: 10.3390/su15043503.

Lu T, Luo P P, Wang J C, et al. 2025b. Soil salinity accumulation and groundwater degradation due to overexploitation over recent 40-year period in Yaoba Oasis, China. Soil and Tillage Research, 248: 106398, doi: 10.1016/j.still.2024.106398.

Ma S L, He B Z, Ge X Y, et al. 2023. Spatial prediction of soil salinity based on the Google Earth Engine platform with multitemporal synthetic remote sensing images. Ecological Informatics, 75: 102111, doi: 10.1016/j.ecoinf.2023.102111.

Mahmoud M F, Arabi M, Pallickara S. 2025. Harnessing ensemble machine learning models for improved salinity prediction in large river basin scales. Journal of Hydrology, 652: 132691, doi: 10.1016/j.jhydrol.2025.132691.

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

Ngabire M, Wang T, Xue X, et al. 2022. Soil salinization mapping across different sandy land-cover types in the Shiyang River Basin: A remote sensing and multiple linear regression approach. Remote Sensing Applications: Society and Environment, 28: 100847, doi: 10.1016/j.rsase.2022.100847.

Niu L T, Jia X X, Li X D, et al. 2024. Developing novel ensemble models for predicting soil hydraulic properties in China's arid region. Journal of Hydrology, 636: 131354, doi: 10.1016/j.jhydrol.2024.131354.

Qin C Z, An Y M, Liang P, et al. 2021. Soil property mapping by combining spatial distance information into the soil land inference model (SOLIM). Pedosphere, 31(4): 638–644.

Qin Y, Nurmemet I, Aihaiti A, et al. 2025. Synergistic retrieval of soil moisture in arid regions using GF-3 SAR and Sentinel-2 optical data. Land Degradation & Development, 37(9): 4274–4294.

Sun L, Li H, Wang J, et al. 2023. Impacts of climate change and human activities on NDVI in the Qinghai-Tibet Plateau.

Remote Sensing, 15(3): 587, doi: 10.3390/rs15030587.

Tariq A, Sardans J, Zeng F J, et al. 2024. Impact of aridity rise and arid lands expansion on carbon-storing capacity, biodiversity loss, and ecosystem services. Global Change Biology, 30(4): e17292, doi: 10.1111/gcb.17292.

Tobler W R. 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography, 46: 234–240.

Velilla E, Snethlage J, Poelman M, et al. 2025. Too salty to farm: rethinking coastal land use in response to soil salinization. Restoration Ecology, 33(4): e70006, doi: 10.1111/rec.70006.

Wadoux A M J C, Heuvelink G B M, de Bruin S, et al. 2021. Spatial cross-validation is not the right way to evaluate map accuracy. Ecological Modelling, 457: 109692, doi: 10.1016/j.ecolmodel.2021.109692.

Wang D Y, Yang H C, Qian H, et al. 2023a. Minimizing vegetation influence on soil salinity mapping with novel bare soil pixels from multi-temporal images. Geoderma, 439: 116697, doi: 10.1016/j.geoderma.2023.116697.

Wang J F, Li X H, Christakos G, et al. 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region, China. International Journal of Geographical Information Science, 24(1): 107–127.

Wang W Y, Yang P, Xia J, et al. 2023b. Impact of land use on water quality in buffer zones at different scales in the Poyang Lake, middle reaches of the Yangtze River Basin. Science of The Total Environment, 896: 165161, doi: 10.1016/j.scitotenv.2023.165161.

Weiss M, Jacob F, Duveiller G. 2020. Remote sensing for agricultural applications: A meta-review. Remote Sensing of Environment, 236: 111402, doi: 10.1016/j.rse.2019.111402.

Wu C W, Zheng Y, Yang H, et al. 2021. Effects of different particle sizes on the spectral prediction of soil organic matter. CATENA, 196: 104933, doi: 10.1016/j.catena.2020.104933.

Xiao S T, Nurmemet I, Zhao J. 2024. Soil salinity estimation based on machine learning using the GF-3 radar and Landsat-8 data in the Keriya Oasis, southern Xinjiang, China. Plant and Soil, 498(1–2): 451–469.

Xiong J, Ge X Y, Ding J L, et al. 2025. Optimal time-window for assessing soil salinity via Sentinel-2 multitemporal synthetic data in the arid agricultural regions of China. Ecological Indicators, 176: 113642, doi: 10.1016/j.ecolind.2025.113642.

Yan D, Lei Y L, Shi Y K, et al. 2018. Evolution of the spatiotemporal pattern of $PM_{2.5}$ concentrations in China–A case study from the Beijing-Tianjin-Hebei region. Atmospheric Environment, 183: 225–233.

Yang J, Huang X. 2024. The 30 m annual land cover datasets and its dynamics in China from 1985 to 2023. Earth System Science Data, 13: 3907–3925.

Yang X Y, Xue X, Zhao D, et al. 2025. Spatial distribution and ecological significance of halophyte communities in the desert-oasis transition zone of the Minqin Lake area. Acta Ecologica Sinica, 45(15): 7500–7512. (in Chinese)

Yin X W, Feng Q, Li Y, et al. 2022. An interplay of soil salinization and groundwater degradation threatening coexistence of oasis-desert ecosystems. Science of The Total Environment, 806(Part 2): 150599, doi: 10.1016/j.scitotenv.2021.150599.

Yu H, Wang Z M, Mao D H, et al. 2023. Spatiotemporal variations of soil salinization in China' s West Songnen Plain. Land Degradation & Development, 34(8): 2366–2378.

Zaman M, Shahid S A, Heng L. 2018. Guideline for Salinity Assessment, Mitigation and Adaptation Using Nuclear and Related Techniques. Cham: Springer.

Zeng T, Li Y Z, Ma L, et al. 2025. Unveiling latent interaction mechanisms influencing the spatial pattern of soil salinity in arid oases: Insights from integrated modeling. CATENA, 250: 108769, doi: 10.1016/j.catena.2025.108769.

Zeng Y L, Hao D L, Badgley G, et al. 2021. Estimating near-infrared reflectance of vegetation from hyperspectral data. Remote Sensing of Environment, 267: 112723, doi: 10.1016/j.rse.2021.112723.

Zhan D X, Liu Y T, Yang W H, et al. 2025. Spatial variability of soil salinity in coastal saline-alkali farmlands: A novel approach integrating a stacked model with the reconstructed *in-situ* hyperspectral feature. Computers and Electronics in Agriculture, 235: 110376, doi: 10.1016/j.compag.2025.110376.

Zhang L, Jia X, Zhao Y H, et al. 2025. Long time-series change characteristics and suitable scale of oasis in arid and semi-arid regions of China. Applied Geography, 180: 103663, doi: 10.1016/j.apgeog.2025.103663.

Zhang Y T, Hou K, Qian H, et al. 2022. Characterization of soil salinization and its driving factors in a typical irrigation area of Northwest China. Science of The Total Environment, 837: 155808, doi: 10.1016/j.scitotenv.2022.155808.

Zhao P, Qu J J, Xu X Y, et al. 2019. Desert vegetation distribution and species-environment relationships in an oasis-desert ecotone of northwestern China. Journal of Arid Land, 11(3): 461–476.

Zhao R, Zhan L P, Yao M X, et al. 2020. A geographically weighted regression model augmented by Geodetector analysis and principal component analysis for the spatial distribution of $PM_{2.5}$. Sustainable Cities and Society, 56: 102106, doi: 10.1016/j.scs.2020.102106.

Zhu M F, Yu X X, Chen K, et al. 2024. Spatiotemporal characteristics and driving factors of chemical oxygen demand emissions in China' s wastewater: An analysis based on spatial autocorrelation and Geodetector. Ecological Indicators, 166: 112308, doi: 10.1016/j.ecolind.2024.112308.

Note: Figure translations are in progress. See original paper for figures.