Spatiotemporal evolution characteristics and driving factors of soil salinization in the Zhundong mining area (postprint)

Authors: He Yangyang, Du Huadong, Bi Yinli, Liu Yunlong

Date: 2026-01-30T16:53:17+00:00

Abstract

Open-pit coal mining in arid regions, while disturbing regional hydrological processes, soil physicochemical conditions, and vegetation patterns, also exacerbates the risk of secondary salinization; however, its spatiotemporal evolution characteristics and driving mechanisms remain insufficiently understood. Taking the Zhunge' er Dong (Zhunge' er East) mining area in Xinjiang, an extremely arid region, as the study area, this research uses Landsat remote sensing imagery from 2000–2024 and 220 in situ soil salinity measurements, and constructs multiple salinity inversion models by combining BP neural networks, support vector machines, and random forests. Model accuracy is further improved by selecting sensitive spectral variables through the variable importance in projection (VIP) method, and the driving effects of climate and human activities on salinization intensity are analyzed. The results show that: (1) The VIP-RF model performs best under the complex surface background conditions of an extremely arid mining area. (2) Spatially, non-saline or slightly salinized soils dominate the central desert and northern alluvial plain, moderate salinization prevails in the transition zone, and the gravel desert-gobi areas in the northwest and southeast constitute concentrated zones of severe salinization. (3) Temporally, the period 2000–2010 is dominated by slight salinization; after mining development began in 2010, the areas of moderate and severe salinization increased by approximately 79% and 84%, respectively, while the area of non-saline soils decreased by 62.2%. (4) In terms of driving mechanisms, salinization during 2000–2010 was mainly controlled by natural factors such as climate and geomorphology, whereas during 2010–2024 human activities gradually became dominant. In summary, integrating spectral variable selection with the random forest method can effectively improve the accuracy of salinization inversion in arid mining areas, and can robustly reveal the evolution patterns of salinization under the interactive influence of natural and anthropogenic factors, providing important

reference value for ecological environment monitoring and sustainable land resource utilization in mining areas of arid regions.

Full Text

Introduction

Soil salinization, as a typical land degradation process, represents one of the most serious environmental hazards in semi-arid regions, directly threatening regional ecological security, agricultural production, and sustainable land use. Against the backdrop of intensified climate change and superimposed human activities, the combined disturbance of land cover, hydrological processes, and microclimate accelerates surface salt accumulation. Particularly in extremely arid regions with strong evaporation and scarce precipitation, salt migration and accumulation processes become more intense, leading to continuous expansion in both the extent and degree of salinization, which has become a critical factor constraining ecosystem stability. Therefore, conducting spatio-temporal dynamic monitoring of salinization and revealing its natural and anthropogenic driving mechanisms hold important theoretical and practical significance for ecological environmental protection and rational resource development in arid mining areas.

The causes of salinization in arid regions exhibit dual attributes of natural and anthropogenic factors. Natural causes include the fact that annual evaporation far exceeds precipitation in arid areas, resulting in dominant vertical soil water movement and salt enrichment at the surface through capillary action. Meanwhile, low vegetation coverage in arid regions, combined with freeze-thaw and wet-dry cycles, leads to severe soil weathering. Under wind erosion, weathering of salt-bearing strata and aeolian salt dust input further increase soil base ion content. Anthropogenic factors primarily involve the overdevelopment of irrigated agriculture, where large-scale flood irrigation raises groundwater levels and fertilizer misuse leads to accumulation of ions such as Na^+ and Cl^-. Mining activities cause particularly significant surface disturbance, with open-pit mining destroying natural aquitards and allowing deep, high-mineralization groundwater to ascend to the vadose zone along fractures. With intense evaporation, salts continuously accumulate at the soil surface. Additionally, subsidence pits and steep slopes from open-pit mining alter surface runoff pathways, forming local water accumulation zones and inducing ring-shaped salt enrichment areas. Frequent open-pit mining activities, superimposed on extreme arid climatic conditions, make the salinization problem more prominent and pose a serious threat to regional ecological environments and sustainable mining development.

Current research on soil salinization after mining in arid and semi-arid regions mostly focuses on point studies of specific functional areas such as mine dumps and gangue stacking zones, or small-scale studies of particular mining areas. However, research at the regional scale on the spatio-temporal variation patterns of salinization during the development of concentrated and contiguous mining areas remains urgently needed. Such research is of great significance for mining area ecological security assessment and sustainable development planning.

To achieve regional-scale salinization monitoring, remote sensing technology has been widely applied due to its advantages of large coverage, long time series, and multi-scale capabilities, yielding abundant results in revealing salinization spatial distribution, evolution trends, and driving factors in northwest China, Central Asia, and Mediterranean arid regions. Currently, there are three main approaches for remote sensing inversion of salinization: first, spectral index-based methods; second, statistical modeling methods; and third, machine learning methods. Spectral index methods are computationally simple and highly versatile but rely on single remote sensing indices and are significantly affected by spectral confusion among ground objects, leading to insufficient monitoring accuracy, especially in arid areas where spectral features of saline soil, bare land, sand, and low-biomass vegetation overlap substantially. Statistical modeling methods (such as multiple linear regression) can reveal correlations between salinity and spectral features but often struggle to capture nonlinear characteristics. Machine learning methods (such as neural networks, support vector machines, and random forest) demonstrate outstanding performance in handling nonlinear, multi-dimensional data, showing strong adaptability to the combined effects of salinization, climate, vegetation, and human activities. However, most deep learning studies still have deficiencies in model comparison, validation, long-term monitoring, and driving mechanism exploration. Moreover, in strongly human-disturbed environments like mining areas, mining activities destroy original landscape structures, alter topography, and trigger ecological environmental problems such as soil erosion and degradation. Open-pit mining creates large areas of bare land, making saline soil, bare land, sand, and mining-disturbed surfaces highly similar in spectral response, particularly difficult to distinguish in visible to near-infrared bands, resulting in typical spectral confusion phenomena. Therefore, how to achieve satellite-ground collaborative verification and improve model robustness in the process of salinization inversion in mining areas remains an urgent problem to be solved.

In summary, this study focuses on the soil salinization problem during the development of typical energy bases in extremely arid regions, taking the Zhundong mining area in Xinjiang as the research object. First, it utilizes long time series and multi-source data fusion to reconstruct the spatio-temporal evolution sequence of soil salinization in the mining area over more than 20 years of development, compensating for previous deficiencies in large-scale, long-term dynamic monitoring of soil salinization in mining areas. Simultaneously, through multi-model comparison and measured data validation, the optimal model is selected for regional-scale quantitative inversion of salinization, thereby improving the robustness and prediction accuracy of the mining area soil salinization inversion model. Finally, based on revealing the spatio-temporal pattern of mining area salinization, multidimensional factors including climate (precipitation, evaporation), hydrology (groundwater depth), vegetation cover, and human activities (mining intensity) are integrated to quantitatively analyze the natural and anthropogenic driving mechanisms of salinization. Theoretically, this deepens understanding of salinization causes in mining areas, while practically providing scientific basis for scientific development, ecological restoration, and environmental monitoring system construction in coal mining areas.

1.1 Study Area Overview

The Zhundong mining area is located within Changji Hui Autonomous Prefecture in Xinjiang (44.22°~45.01°N, 88.91°~90.82°E), situated in the desert and Gobi region on the southern edge of the Junggar Basin and the southern foothills of the Kalamaili Mountains. The mining area has a fragile ecological environment with a temperate continental climate, characterized by scarce precipitation and dry conditions. The region exhibits significant seasonal temperature variations, with an annual average temperature of 6.9°C, extreme high temperatures concentrated in July, and extreme minimum temperatures reaching -45.6°C in January. Surface vegetation is sparse, with extensive deserts and few biological species. The Zhundong mining area features flat terrain with shallow and thick coal seams. Since its development in 2005, open-pit mining has been the primary method, and its efficient and large-scale development has made it an important energy base in Xinjiang. However, open-pit mining pits and dumps have altered the original topography and landforms, causing land damage through excavation and external dump construction, as well as damage to surface ecosystems including original hydrological cycles, soil structure, and surface vegetation, further intensifying land desertification and soil erosion.

1.2 Remote Sensing Image Sources and Processing

Based on the study area boundary, Landsat remote sensing images with low cloud cover and coinciding with sampling times from 2000 to 2024 were selected. Image data were obtained from the United States Geological Survey (USGS) website (http://earthexplorer.usgs.gov/). Landsat 5 TM data were used for 2000-2010, and Landsat 8 OLI data for 2010-2024. Landsat 5 TM has seven bands, while Landsat 8 OLI has nine bands, as detailed in Table 1.

To reduce errors during image data acquisition, original remote sensing images require preprocessing before analysis. ENVI 5.3 software was used for radiometric calibration, atmospheric correction, mosaicking, and clipping.

1.4 Data Sources and Processing of Influencing Factors

Natural and anthropogenic factors are the two main factors affecting soil salinization in the study area. By referencing existing research and combining it with the current status of soil salinization in the Zhundong mining area, domestic product (GDP), population (Population sources), groundwater resources (Groundwater Resources), soil pH (pH), organic matter content (Organic Matter Content), normalized difference vegetation index (NDVI), precipitation (Precipitation), annual average temperature (Annual Mean Temperature), surface wind speed (Surface Wind Speed), and digital elevation model (DEM) were selected. Considering that evaporation has a significant impact on salt accumulation but lacks long-term continuous monitoring station data, meteorological factors such as temperature, precipitation, and wind speed were used to indirectly reflect evaporation effects. Among these, GDP and population data came from the National Bureau of Statistics, NDVI was calculated from Landsat image bands, and meteorological data came from the National Tibetan Plateau Science Data Center. For missing and abnormal values in various datasets, linear interpolation was used for filling.

1.5.1 Spectral Indices

The selected spectral indices include Salinity Index (SI), Salinity Index 1 (SI1), Salinization Index 1, Salinization Index 2, Salinization Index 5, Extended Difference Vegetation Index (Extended DVI), Extended Ratio Vegetation Index (Extended RVI), Atmospherically Resistant Vegetation Index (ARVI), Canopy Reflectance Salinity Index (CRSI), and Extended Enhanced Vegetation Index (Extended EVI).

1.5.2 Variable Screening Methods

Variable Importance in Projection (VIP) is a variable screening method used in multivariate regression analysis, often combined with Partial Least Squares Regression (PLSR). By calculating each independent variable' s contribution in the projection direction, it assesses their explanatory power for the dependent variable. Higher importance values indicate more critical variables for model prediction. Grey Relational Analysis (GRA) measures the correlation between sequences by calculating their geometric similarity (such as consistency in change trends), making it suitable for small-sample, incomplete information system analysis.

1.5.3 Inversion Models and Accuracy Verification

Back Propagation Neural Network (BPNN) is a multilayer feedforward neural network based on the backpropagation algorithm, continuously optimizing weights to minimize the loss function and achieve approximation of inversion values. Support Vector Machine (SVM) is a supervised learning classification model based on optimal hyperplanes, with main parameters including penalty coefficient C, kernel function type, and its parameters. Random Forest Regression (RFR) is an ensemble learning method composed of multiple decision trees, suitable for classification and regression tasks, with strong robustness and generalization ability.

To comprehensively evaluate model accuracy, coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) were selected. R^2 reflects the model' s ability to explain data variability, while RMSE and MAE reflect errors between predicted and true values. RMSE is more sensitive to large errors, while MAE reflects the typical level of errors. The calculation formulas are:

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

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

where y_i is the measured value, $\hat{y}_i$ is the predicted value, and n is the sample size.

1.5.4 Driving Forces of Soil Salinization in Mining Areas

To further quantify the relative impact of various driving factors on the evolution of soil salinization in the Zhundong mining area, representative years were selected to construct a socioeconomic-natural environment composite driving force index system. The index system covers 15 factors including economy, population, hydrology, soil, climate, and terrain. To eliminate dimensional effects and ensure comparability among different variables, all data were standardized before statistical analysis. Based on this, Multiple Linear Regression (MLR) was used to quantitatively analyze the relationship between each driving factor and changes in salinization area in different years, calculating each factor' s relative contribution rate to reflect its dominance and sensitivity in the evolution process. Finally, heatmaps were used to visually present the driving forces of soil salinization in different years.

Since the units of various influencing factors differ, original data require dimensionless processing to standardize variable values. Multiple linear regression was used to quantitatively calculate the relative contribution rate of each influencing factor to changes in soil salinization area. The calculation formula is:

$$Z_Y = \sum_{i=1}^{n} a_{iX} i$$

$$P_i = \frac{a_i}{\sum_{i=1}^{n} a_i} \times 100\%$$

where Z_Y is the standardized value of the dependent variable after dimensionless processing; $X_1, X_2, \ldots, X_n$ are standardized values of independent variables after dimensionless processing; i is the number of driving factors; a_i is the regression coefficient corresponding to the i-th driving factor after dimensionless processing; and P_i is the relative contribution rate of the i-th driving factor.

2 Results and Analysis

2.1 Spectral Information Screening

Using Variable Importance in Projection and Grey Relational Analysis methods, the spectral indices most highly correlated with soil salinization in the Zhundong mining area were screened to construct salinization inversion models. The screening results are shown in Figure 2.

[Figure 2: see original paper]

The VIP method identified 8 key indices with VIP > 1, while the GRA method identified 9 key indices with GRA ≥ 0.75.

2.2 Model Accuracy Comparison

According to measured soil salinity content data, 220 soil samples were collected in the study area from May to June. Each sampling point' s latitude and longitude were recorded (Figure 1). To ensure sampling accuracy, sampling points were laid out along the east-west provincial highway across the Zhundong mining area, combined with the distribution of open-pit coal mines, using the Zhundong mining area' s "one large and eight small" mining areas as the center and key sampling zones. The concentric circle method was employed, with the mining area as the center, setting sampling points at 0-100 m, 200-300 m, 300-400 m, 400-500 m, 1-2 km, 2-3 km, 3-4 km, 4-5 km, and 5-10 km distance intervals, with 5 sampling points in each distance range. Additional sampling points were uniformly distributed along the road, and further adjustments were made based on already laid-out sampling points and mining area boundaries, with uniform distribution across different geomorphic regions to cover all landform types in the study area. A total of 220 sampling points were established. Collected soil samples were naturally air-dried in the laboratory, impurities removed, ground and sieved, and thoroughly mixed for analysis. A 1:5 water suspension was prepared at room temperature. Soil sample electrical conductivity was measured using a DDS-308F conductivity meter. Referencing existing research results, soil salinity content (SSC) data were converted from electrical conductivity values. Currently, the widely used soil salinization degree classification standard comes from Mr. Wang Zunqin' s book "China' s Saline Soil," and this evaluation system was adopted to classify salinization degree in the study area.

Based on the salinization classification standard, combined with field survey data and visual interpretation of remote sensing images, soil salinization degree in the study area was divided into 4 levels (Table 2).

According to measured soil salinity content data, 220 samples were randomly divided, with one sample selected for every 5 samples for model validation and the remaining data used for model training, resulting in 176 samples as the training set and 44 samples as the validation set. Using spectral information screened by VIP and GRA methods, three machine learning algorithms were applied to build models: BPNN with input layer nodes of 8, training steps of

1000, and learning rate of 0.01; SVM with Gaussian radial basis kernel function, penalty factor C of 10, and gamma of 0.1; and RFR with 100 decision trees, random seed of 42, and loss function of MSE. Inversion model accuracy results are shown in Table 3.

The models constructed using variables screened by the VIP method performed significantly better than those using GRA. The VIP-RF model showed the best performance, with training and validation set R^2 values of 0.85 and 0.78, respectively, and RMSE of 1.163 $g \cdot kg^{-1}$ and 1.307 $g \cdot kg^{-1}$, with all error indicators being the lowest and showing the best stability. The VIP-SVM model also performed relatively well (validation set $R^2 = 0.71$), while the VIP-BPNN model showed obvious overfitting with the largest validation set error ($R^2 = 0.327$, RMSE = 2.158 $g \cdot kg^{-1}$, MAE = 1.748 $g \cdot kg^{-1}$). The GRA-RF model performed the worst. The comparison shows that screening variables through VIP can more effectively improve model accuracy. Therefore, VIP-RF was selected as the model for salinization inversion in the Zhundong mining area.

2.3 Spatio-temporal Characteristics of Soil Salinization in the Zhundong Mining Area

The VIP-RF model was used to reconstruct the spatio-temporal differentiation pattern of soil salinization in the Zhundong mining area from 2000 to 2024. Spatial inversion results show (Figure 3) that the degree of salinization in the study area over the past 24 years exhibits significant spatial differentiation characteristics. Non-salinized and slightly salinized soils are the main types, accounting for 55%-76% in recent years, mainly distributed in the central aeolian sandy area and northern alluvial fan groups. Moderately salinized soils account for about 18%-30%, concentrated in the desert-Gobi transition zone. Severely salinized soils and saline soils are relatively less, accounting for 6%-15% and 2%-6% respectively, scattered in the northwestern edge and southeastern gravel desert areas of the study area.

[Figure 3: see original paper]

The land use transfer matrix method was used to quantitatively analyze dynamic transfer rules from 2000 to 2024 (Figure 4). Based on dynamic transfer matrix analysis of land salinization, results show that different salinized soil types exhibit significant heterogeneity in area transfer intensity. Before 2010, non-salinized land showed an obvious expansion trend, with a net loss of 224.7 km^2. The transfer of saline soil was the most limited, with a transfer-out of 130.9 km^2 and transfer-in of 122.0 km^2, maintaining a dynamic balance. Moderately salinized soil showed net loss, with transfer-out (1110.7 km^2) significantly larger than transfer-in (977 km^2). Severely salinized soil decreased slightly during this stage, with transfer-out of 611.3 km^2 and transfer-in of 542.3 km^2.

[Figure 4: see original paper]

After 2010, the transformation trend of non-salinized land reversed, with transfer-out increasing to 707.9 km^2 and transfer-in decreasing to 550.7 km^2, showing a net loss. The migration scale of slightly salinized soil expanded significantly, with transfer-out and transfer-in increasing to 1611.5 km^2 and 1597.5 km^2 respectively, maintaining high-frequency two-way flow. Moderately salinized soil shifted from net loss to significant net growth, with transfer-in reaching 1798 km^2, far exceeding the transfer-out area of 1374.6 km^2. The migration volume of severely salinized soil also increased, with transfer-out of 837 km^2 and transfer-in of 789.9 km^2, showing a net loss. Saline soil continued its previous decreasing trend, although transfer-out increased to 324.6 km^2, transfer-in was only 182.3 km^2, remaining relatively passive.

Specifically, the transfer flux between salinized land types in descending order is: slightly salinized soil → moderately salinized soil → non-salinized soil → severely salinized soil → saline soil, a sequence that remained consistent across the two monitoring periods.

2.4 Driving Factors of Salinization in Mining Areas

Multiple linear regression analysis shows (Figure 5) that the driving mechanism of soil salinization in the Zhundong mining area exhibits significant stage differences on the temporal scale. In the early stage (2000-2010), the salinization process was mainly dominated by natural environmental factors, with NDVI and precipitation contribution rates of 22.3% and 19.5% respectively, indicating that vegetation cover and regional precipitation conditions played key regulatory roles in salt accumulation and migration. During this stage, climate fluctuations and ecosystem self-regulation capacity jointly maintained the natural balance of salinization.

[Figure 5: see original paper]

However, with the continuous advancement of mining area resource development activities and rapid changes in land use patterns, the influence of socioeconomic factors gradually increased. By 2010-2024, the contribution rates of GDP and groundwater resources rose to 24.4% and 16.9% respectively, becoming the main controlling factors driving salinization evolution. This change indicates that human activities such as industrial development, infrastructure construction, and over-exploitation of groundwater have significantly interfered with surface water-salt balance. Meanwhile, the contribution rates of natural factors such as NDVI and precipitation showed a continuous decline, reflecting that regional ecosystems are increasingly sensitive to human disturbance and the dominant role of natural environment is gradually weakening.

Overall, the driving mechanism of soil salinization in the Zhundong mining area has shifted from "nature-dominated" to "human-dominated." Economic growth and groundwater utilization have become key factors affecting the spatial expansion and intensity changes of salinization, while the weakening of ecological restoration and vegetation regulation capacity has further intensified salt accumulation risks. This evolution characteristic reveals the complex feedback relationship between economic activities and environmental processes under high-intensity mining development, providing important guidance for future regional ecological restoration and water resource regulation.

3 Discussion

This study conducted remote sensing inversion of soil salinization in the Zhundong mining area based on multiple spectral index combinations and machine learning methods. Results show that the VIP-RF model performed best in terms of training accuracy and generalization ability, adapting to complex and diverse surface conditions in arid regions with strong stability and reliability. This result not only demonstrates the applicability of the RF model but also highlights the critical role of variable screening in improving model performance. Traditional salinization inversion studies often rely on full-band input, which easily introduces redundant features and amplifies noise interference. By screening sensitive spectral variables through VIP, spectral confusion between saline soil, bare land, sand, and low-biomass vegetation can be effectively reduced, thereby improving model accuracy in highly heterogeneous mining area environments. Although this study’ s modeling framework represents an extension of existing mature methods, its high stability and accuracy in the complex mining area background demonstrate methodological improvement significance.

Model comparisons further illustrate the advantages of the VIP-RF framework. The BPNN model is significantly affected by input feature noise when dealing with high-reflectance bare land and mixed gravel backgrounds, easily falling into local optima and producing unstable results. Although SVM performs well under small sample conditions, it is sensitive to kernel functions and data scales in mining area environments with diverse surface types and significant spectral differences, making it difficult to comprehensively consider different ground object types. In contrast, the RF model, relying on its ensemble learning mechanism, can adaptively handle high-dimensional nonlinear features and achieve variable screening through feature importance ranking, effectively capturing complex nonlinear relationships between salinity and spectral signals. This characteristic is particularly suitable for salinization monitoring in highly spatially variable arid mining areas, making the method selection and optimization in this study reasonable and innovative under the mining area application background.

From the perspective of spatio-temporal evolution characteristics, the overall salinization level in the Zhundong mining area is relatively high, showing obvious zonal and geomorphic differences. This pattern is closely related to arid continental climate conditions. With annual precipitation of only 50-200 mm and annual evaporation as high as 1500-3000 mm, extreme water deficits cause rapid surface water evaporation, driving continuous salt migration to the surface through capillary water movement and forming “salt crust” or “salt frost” phenomena. While this aligns with findings from other arid region salinization studies, the Zhundong mining area shows stronger differences in salinization across geomorphic units due to its complex landform types. For example, seasonal river channels become concentrated salt accumulation areas due to shallow groundwater and frequent flood recharge; gravel desert areas have limited salt leaching due to strong evaporation enhanced by surface black gravel layers, leading to more severe salinization. In contrast, sand and some gravel areas have relatively lower salinization due to strong permeability facilitating downward salt migration. This indicates that, against an extremely arid background, the regional salinization pattern is controlled by both general climate conditions and significant micro-geomorphic environments.

From a temporal evolution perspective, salinization changes in the Zhundong mining area were relatively slow from 2000-2010, mainly driven by climate and topographic factors. After 2010, with the launch of large-scale coal open-pit mining, salinization degree intensified significantly, manifested by rapid increases in moderately salinized area and expansion trends in severely salinized zones. Natural factors still played a role during this period, but human activities became the dominant driving force. On one hand, open-pit mining cut the continuity of phreatic aquifers, and mine drainage changed original groundwater levels and salt distribution, forming new salt accumulation zones. On the other hand, large-scale stripping operations destroyed surface vegetation, substantially increasing bare land proportion, enhancing evaporation, and eliminating vegetation' s regulatory role in salt cycling, promoting salt migration to the surface. Additionally, the formation of dumps and subsidence areas changed original runoff pathways, causing local water-salt balance disorder and creating new secondary salinization hotspots. Urban expansion and over-exploitation of groundwater further intensified water deficits, driving surface salt accumulation. It is evident that the evolution of salinization in the Zhundong mining area results from the coupling of natural factors and human disturbance, but under high-intensity human intervention, salinization expansion shows an accelerating trend.

In comparison with existing research, this study' s conclusions first verify the feasibility and advantages of combining spectral indices with machine learning methods for salinization monitoring in arid mining areas, particularly the adaptability of the VIP-RF framework in complex backgrounds. Second, it reveals the driving mechanism of salinization under open-pit mining conditions, providing new case support for understanding the interaction between salinization and human activities in arid regions. Compared with other regional studies, this research emphasizes the mechanism of special human disturbance in mining areas, which represents an important distinction from salinization studies in general agricultural or natural arid regions. Future salinization inversion in mining areas can be further developed in several aspects: first, in terms of remote sensing data, although this study mainly relied on Landsat series images with a long time span, the limited spatial resolution makes it difficult to precisely identify small-scale salinization patches. Future integration of Sentinel-2 or high-resolution commercial satellite data could improve spatial characterization capability. Second, in terms of modeling frameworks, although VIP-RF achieved accuracy improvement, uncertainties remain in time-series inversion. Future exploration could combine deep learning methods (such as convolutional neural networks) with multi-source data fusion (such as optical, radar, and hydrological observations) to further improve inversion accuracy and generalization ability. Finally, in driving mechanism analysis, this study mainly used statistical correlation and comparative analysis. Future work could combine numerical simulation and process models to deeply explore the dynamic mechanisms of groundwater level changes, evaporation processes, and salt migration, achieving a breakthrough from "correlation" to "causality."

4 Conclusions

Based on Landsat remote sensing images and measured soil salinity data from 2000-2024, this study conducted a comprehensive analysis of the evolution characteristics and driving mechanisms of soil salinization in the Zhundong mining area of Xinjiang. The main conclusions are:

(1) Under the background of highly complex surface types in extremely arid mining areas, the random forest model based on variable importance projection screening demonstrates high stability and reliability in salinization inversion, indicating that constraining key spectral variables can effectively characterize the nonlinear response relationship between mining area soil salinity and remote sensing signals.

(2) The soil salinization pattern in the Zhundong mining area shows significant spatial differentiation, with its distribution and evolution jointly controlled by geomorphic units and water-salt migration conditions. Differences in salt accumulation processes under different geomorphic backgrounds determine the aggregation characteristics and distribution patterns of salinization at the regional scale.

(3) The evolution of soil salinization in the Zhundong mining area shows obvious stage characteristics, with its dominant mechanism gradually shifting from natural environmental factors to human activity dominance. Open-pit coal mining has intensified water-salt imbalance processes by altering surface structure, hydrological processes, and vegetation cover, becoming the key driving factor for salinization aggravation in the past decade.

(4) The research results indicate that the technical path based on "spectral variable screening + machine learning inversion + driving mechanism analysis" can be applied to long-term dynamic monitoring and evolution diagnosis of soil salinization in arid mining areas, providing scientific basis for mining area development intensity control, salinization risk identification, and ecological restoration zoning.

References

[1] Sahbeni G, Ngabire M, Musyimi P K, et al. Challenges and opportunities in remote sensing for soil salinization mapping and monitoring: A review[J]. Remote Sensing, 2023, 15(10): 2540.

[2] Wang Yubao, Li Tao, Jin Yingchun. Analysis on control approaches of secondary soil salinization in Yanqi Basin, Xinjiang[J]. Hydrogeology & Engineering Geology, 2004, 22(6): 99-101.

[3] Zhang Weizheng. Secondary salinization of grassland soil: Study on the genesis of secondary saline alkaline patches in Songnen Plain[J]. Acta Pedologica Sinica, 1993, 30(2): 182-190.

[4] Li Xia, Qiao Mu, Zhou Shengbin. Spatio-temporal variation and causes of soil salinization in Manas River Irrigation District of Xinjiang during 1985-2014[J]. Bulletin of Soil and Water Conservation, 2016, 36(3): 152-158, 370.

[5] Wang Li. Research on microbial remediation technology for ecological environment pollution in open-pit coal mining areas[J]. Environmental Science and Management, 2025, 50(2): 150-154.

[6] Han Xiuna, Dong Ying, Geng Yuqing, et al. Effects of coal gangue cover on soil nutrient and salinity characteristics in mining area[J]. Ecology and Environmental Sciences, 2021, 30(11): 2251-2256.

[7] Yang Yuping, Xu Li. Study on salinity characteristics of weathered materials in waste dump of Pinggou coal mine, Wuhai City[J]. Journal of Inner Mongolia Agricultural University (Natural Science Edition), 2022, 43(1): 35-39.

[8] Han Yong, Yu Xiangyu, Jiang Kaisheng, et al. Remote sensing monitoring of soil salinization in Hongshaquan open-pit coal mine[J]. Journal of China Coal Society, 2023, 48(Suppl.): 704-712.

[9] Wang Nan, Chen Songchao, Huang Jingyi, et al. Global soil salinity estimation at 10 m using multi-source remote sensing[J]. Journal of Remote Sensing, 2024, 4(1): 0130.

[10] Huang Shuai, Ding Jianli, Li Xiang, et al. Hyperspectral characteristics analysis and modeling of soil salinization[J]. Chinese Journal of Soil Science, 2016, 47(5): 1042-1048.

[11] Cheng Yongxiang, Zhang Fenghua, Wang Jiangli, et al. Research on control of secondary salinization of oasis soil in Xinjiang[J]. China Resources Comprehensive Utilization, 2023, 41(4): 135-139.

[12] Chen Zhongxin, Ren Jianqiang, Tang Huajun, et al. Research progress and prospects in agricultural remote sensing applications[J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(8): 168-173.

[13] Huang Shuai, Tan Hongjing, Fu Shangke, et al. Hyperspectral inversion modeling of soil salinity content in oasis[J]. Hubei Agricultural Sciences, 2024, 63(9): 196-203.

[14] Fu Guoshuang. Application of variable importance in projection analysis in screening independent variables[J]. Modern Preventive Medicine, 2012, 39(22): 5813-5815.

[15] Bian Zhengfu, Yu Haochen, Hou Jing, et al. Influencing factors and evaluation of land degradation of 12 coal mine areas in western China[J]. Journal of China Coal Society, 2020, 45(1): 338-350.

[16] Wu Zhenhua, Che Mingliang, Zhang Shutao, et al. Remote sensing monitoring and driving force analysis of salinized soil in grassland mining area[J]. Sustainability, 2022, 14(2): 741.

[17] Zhao Xiaofan, Zeng Qiang. Study on soil moisture distribution characteristics in Zhundong mining area based on Landsat images[J]. Mining Safety & Environmental Protection, 2023, 50(5): 137-143.

[18] Zhou Yamin, Zhang Rongqun, Ma Hongyuan, et al. Hyperspectral inversion of mineral ion content in Salt Lake based on BP neural network[J]. Remote Sensing for Land & Resources, 2016, 28(2): 34-40.

[19] Shan Wei, Jin Xiaobin, Meng Xiansu, et al. Dynamical monitoring of ecological environment quality of land consolidation based on multi-source remote sensing data[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(1): 234-242.

[20] Yu Haochen, Chen Fu, Yin Dengyu, et al. Impacts of mining and climate change on the land ecosystem in Gobi mining areas: A case study of the Zhundong coal base[C]// Abstracts of the 3rd International Symposium on Land Reclamation and Ecological Restoration. China Coal Society, China University of Mining and Technology, China Coal Society, 2021: 1.

[21] Yuan Jie, Cao Guangchao, Yang Dengxing, et al. Temporal and spatial variation characteristics and influencing factors of vegetation NDVI in Heihe source region of Qilian Mountains[J]. Ecological Science, 2021, 40(5): 172-182.

[22] Yu Haochen, Bian Zhengfu, Mu Shouguo, et al. Effects of climate change on land cover change and vegetation dynamics in Xinjiang, China[J]. International Journal of Environmental Research and Public Health, 2020, 17(13): 4865.

[23] Gorji T, Yildirim A, Hamzehpour N, et al. Soil salinity analysis of Urmia Lake Basin using Landsat 8 OLI and Sentinel-2A based spectral indices and electrical conductivity measurements[J]. Ecological Indicators, 2020, 112: 106173.

[24] Liu Zepeng. Spatiotemporal Characteristics and Influencing Factors of Soil Salinization in Cangzhou City based on Landsat Data[D]. Shijiazhuang: Hebei Normal University, 2022.

[25] Li Xingyou, Zhang Fei, Wang Zheng. Current status and development trends of methods for constructing remote sensing monitoring models of soil salinization[J]. Remote Sensing for Natural Resources, 2022, 34(4): 11-21.

[26] Feng Yiming, Wu Bo, Zhou Na, et al. Gobi classification system based on remote sensing image recognition[J]. Journal of Desert Research, 2013, 33(3): 635-641.

[27] Khan N M, Sato Y. Monitoring hydro-salinity status and its impact in irrigated semi-arid areas using IRS-1B LISS-II data[J]. Asian Journal of Geoinformatics, 2001, 1(3): 63-73.

[28] Qiang Xinhuan, Gao Wenwen, Wang Bo, et al. Risk assessment and evolution pattern of soil salinization based on remote sensing[J]. Arid Zone Research, 2025, 42(3): 431-444.

[29] Wang Fei, Ding Jianli, Wu Manchun. Remote sensing monitoring model of soil salinization based on the NDVI-SI feature space[J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26(8): 168-173.

[30] Wu Xia, Wang Changjun, Fan Liqin, et al. An application analysis of salination evaluation index based on multispectral remote sensing to soil salinity prediction in Yinbei irrigation area, Ningxia[J]. Remote Sensing for Land & Resources, 2021, 33(2): 124-133.

[31] Wang Jingping, Wu Xiaodan, Ma Dujun, et al. Remote sensing inversion based on machine learning: Analysis of uncertainty factors[J]. Journal of Remote Sensing, 2023, 27(3): 790-801.

[32] Huang Quanzhong, Xu Xu, Lyu Lingjiao, et al. Distribution of soil salinity retrieved from remote sensing and its impact on crop growth in Hetao Irrigation District[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(1): 102-109.

[33] Zhang Ziguang, Bai Jiyuan. Analysis on influencing factors of slope stability in external dump of coal mine[J]. Opencast Mining Technology, 2018, 33(5): 51-53, 58.

[34] Zhou Shixun, Yin Juan, Wang Juntao, et al. Hyperspectral estimation and spatial distribution of soil salt content in irrigation district[J]. Journal of Irrigation and Drainage, 2025, 44(2): 72-82.

[35] Shi Yanzi, Li Wenjuan, Li Xiaobin, et al. Research progress and prospects of soil salinization monitoring based on multi-source remote sensing data[J]. China Agricultural Informatics, 2024, 36(5): 28-41.

[36] Liu Fengli, Yan Dong, Li Wenjie. Current situation and improvement countermeasures of secondary soil salinization in protected areas of southern Hebei[J]. Hebei Agricultural Machinery, 2025, 5(5): 94-96.

[37] Guo Yanping, Wang Xuemei, Zhao Feng, et al. RF hyperspectral inversion of oasis topsoil salinity based on optimal mathematics and wavelet transform[J]. Transactions of the Chinese Society of Agricultural Engineering, 2025, 41(3): 83-93.

[38] Shen Deyou, Li Sheng, Gao Yuan, et al. Distribution characteristics and spatiotemporal evolution analysis of soil salinization in the Kashgar River Basin, Xinjiang[J]. Science Technology and Engineering, 2022, 22(24): 10461-10469.

[39] Bai Xuejiao, Wang Pengxin, Xie Yi, et al. Spatial distribution characteristics of drought in Guanzhong Plain based on structural similarity[J]. Transactions of the Chinese Society of Agricultural Machinery, 2025, 46(11): 345-351.

[40] Song Ziling, Fan Junfu, Wang Laigui, et al. Impact analysis on mining status and ecological environment in open-pit coal mine[J]. Opencast Mining Technology, 2016, 31(9): 1-4.

[41] Cui Yan, Li Siyang. Impact of open-pit coal mine development on Gobi in Xinjiang and ecological restoration countermeasures[J]. Coal Engineering, 2024, 56(11): 24-28.

[42] Lei Qingqing, Zhang Xiaoyi, Liu Yongbing, et al. Evaluation of soil quality in saline-alkali sandy wasteland after application of soil conditioner by grey correlation analysis[J]. Journal of Agricultural Sciences, 2024, 45(2): 71-76.

[43] Wei Huimin, Jia Keli, Zhang Xu. Spatio-temporal characteristics and driving factors of soil salinization in Yinchuan Plain from 2008 to 2020[J]. Journal of Tianjin Normal University (Natural Science Edition), 2023, 43(3): 46-52.

[44] Liu Ying, Yue Hui. The Temperature-vegetation-dryness index (TVDI) based on bi-parabolic NDVI-Ts space and gradient-based structural similarity (GSSIM) for long-term drought assessment across Shaanxi Province, China (2000-2016)[J]. Remote Sensing, 2018, 10(6): 959.

[45] Chen Maosheng, Wang Lei, Niu Zilu, et al. Characteristics of soil salinization in Hetao Irrigation District, Inner Mongolia[J]. Journal of Agricultural Sciences, 2024, 45(4): 40-48.

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