Postprint of Soil Salinity Retrieval in Cultivated Land Using Combined Optical and Microwave Remote Sensing

Authors: LIU Ruiliang

Date: 2024-04-01T00:00:00+00:00

Abstract

Farmland protection is crucial to national food security and sustainable socioeconomic development, plays an important role in ecological environment protection, and rapid and accurate acquisition of farmland soil salinity content and spatial distribution information is an essential requirement for farmland protection. Taking Pingluo County, Ningxia as the study area, using Landsat 9 OLI and Sentinel-1 remote sensing images, spectral indices and radar polarization combination indices were extracted, characteristic variables were screened based on the Variable Importance in Projection method and Grey Relational Analysis method, then three machine learning algorithms—Back Propagation Neural Network, Support Vector Machine, and Random Forest—were employed to construct models, and the optimal model was used to invert the spatial distribution of farmland soil salinity content. The results show that: (1) The validation set coefficient of determination (R^2) of the model established using variables screened by the Variable Importance in Projection method is greater than that of the model established using variables screened by the Grey Relational Analysis method. (2) Using the Random Forest algorithm, the synergistic inversion model combining spectral indices and radar polarization combination indices achieved the best performance, with a modeling set R^2 of 0.791 and Root Mean Square Error (RMSE) of 1.016, representing improvements of 0.065 and 0.085 in R^2 and reductions of 0.147 and 0.189 in RMSE compared to single-data-source models, respectively; the validation set R^2 was 0.780 and RMSE was 1.132, representing improvements of 0.091 and 0.237 in R^2 and reductions of 0.175 and 0.377 in RMSE compared to single-data-source models, respectively. (3) In Pingluo County, lightly salinized and moderately salinized soils are widely distributed across farmland, accounting for 23.77% and 33.54% respectively, while heavily salinized soil reaches 15.37%. The study findings indicate that modeling with combined multi-source remote sensing data can effectively improve the accuracy of soil salinity inversion, providing an effective technical reference for

soil salinity inversion in arid region farmland and local agricultural sustainable development.

Full Text

Preamble

Inversion of Soil Salt Content by Combining Optical and Microwave Remote Sensing in Cultivated Land

LIU Ruiliang[1], JIA Keli[1], LI Xiaoyu[1], CHEN Ruihua[1], WANG Yijing[1], ZHANG Junhua[2]

[1]College of Geographical Sciences and Planning, Ningxia University, Yinchuan 750021, Ningxia, China
[2]College of Ecology, Ningxia University, Yinchuan 750021, Ningxia, China

Abstract: The protection of cultivated land is crucial for national food security, sustainable economic and social development, and ecological environmental conservation. Rapid and accurate acquisition of soil salinity content and spatial distribution information in cultivated land represents an essential requirement for cultivated land protection. This study takes Pingluo County in Ningxia as the research area and explores the feasibility of combining optical and microwave remote sensing to improve the accuracy of soil salt content prediction compared with single-source remote sensing data. Using Landsat 9 OLI and Sentinel-1 remote sensing imagery, we extracted spectral indices and radar polarization combination indices. Feature variables were screened based on the Variable Importance in Projection (VIP) method and gray correlation degree method. Machine learning algorithms including back propagation neural network, support vector machine, and random forest were then employed to construct predictive models, with the optimal model used to invert the spatial distribution of soil salt content in cultivated land. The results demonstrate that: (1) Models validated using variables screened by the VIP method generally exhibited higher determination coefficients (R^2) in the validation set than those established using the gray correlation method. (2) Using the random forest algorithm, the synergistic inversion model combining spectral indices and radar polarization combination indices achieved the best performance, with the modeling set achieving R^2 = 0.791 and root mean square error (RMSE) = 1.016 $g \cdot kg^{-1}$. This represents R^2 improvements of 0.065 and 0.085 compared with single-source models, with corresponding RMSE reductions of 0.147 and 0.189 $g \cdot kg^{-1}$. The validation set achieved R^2 = 0.780 and RMSE = 1.132 $g \cdot kg^{-1}$, indicating respective R^2 improvements of 0.091 and 0.237, and RMSE reductions of 0.175 and 0.377 $g \cdot kg^{-1}$ compared with single-source models. (3) Mildly and moderately salinized soils in Pingluo County' s cultivated land were widely distributed, accounting for 23.77% and 33.54% respectively, while severely salinized soil constituted 15.37%. These findings reveal that modeling with combined multi-source remote sensing data can effectively improve the inversion accuracy of soil salt content, providing an effective technical reference for soil salt content inversion in arid region cultivated land and supporting local agricultural sustainable development.

Keywords: optical and microwave remote sensing; machine learning; cultivated land; soil salt content inversion

Introduction

Soil salinization, whether naturally occurring or human-induced, constitutes a significant environmental hazard that occurs extensively in inland arid and semi-arid regions, severely impacting agricultural production and regional sustainable development. It is estimated that by 2050, approximately 50% of global arable land will face salinization, becoming a worldwide ecological problem. Therefore, employing scientific methods to accurately invert and monitor the dynamics of cultivated land soil salinization, and implementing timely remediation measures, holds great significance for regional food production and agricultural sustainable development.

Satellite remote sensing enables rapid and macroscopic acquisition of soil spectral characteristics, and through the construction of remote sensing monitoring models, large-scale soil salinization monitoring and assessment can be realized. Optical remote sensing provides multi-band spectral information that is highly sensitive to soil salinity. Researchers have achieved soil salinity inversion by constructing spectral indices. For instance, Chen et al. improved vegetation indices using Landsat 8 OLI data in the 2100–2300 nm range, substantially enhancing model inversion accuracy. However, single optical data are susceptible to imaging time constraints and weather conditions such as clouds and rain, making salinization information extraction based solely on spectral characteristics limited.

Microwave remote sensing offers all-weather operational capability and is less affected by meteorological conditions and solar illumination levels. Moreover, since the study area has vegetation cover that affects spectral reflectance, microwave remote sensing can penetrate vegetation and possesses the ability to detect subsurface targets. Ma et al. utilized Sentinel-1 cross-polarization mode (VH) and vertical polarization mode (VV) radar imagery to analyze the relationship between backscattering coefficients from VV and VH polarization combinations and soil salinity, demonstrating that the VH backscattering coefficient could effectively separate soils with different salinity levels. Zhang et al. investigated multi-depth soil salinity in the Hetao Irrigation District of Inner Mongolia using Sentinel-1 radar data. Although these studies achieved favorable results, most were based on single remote sensing data, with limited research on combining optical and microwave remote sensing for soil salinity inversion. Xiao et al. combined optical imagery and radar backscattering characteristics to verify the feasibility of optical-radar data combination for soil salinization monitoring. However, their study only used the random forest classification algorithm for salinization level classification without inverting specific salinity values for different regions. In contrast, this study employs the random forest regression algorithm to invert soil salinity at the pixel scale in cultivated land, thereby preserving more salinity information.

Soil salinization causes and salt composition are complex, and results vary across different regions in feature variable selection. Therefore, different methods are needed to screen characteristic variables. Commonly used methods include correlation coefficient method, stepwise regression, gray correlation degree (GCD), and Variable Importance in Projection (VIP). Wang et al. used three methods to screen variables for model construction and found that model accuracy differed under different screening methods. Li et al. used VIP and random forest (RF) for variable screening and established models using partial least squares regression (PLSR), showing that the VIP-PLSR model had higher R^2 than the RF-PLSR model. Different modeling methods yield different inversion effects. Liu et al. used multiple linear regression and back propagation neural network (BPNN) to invert soil salinization in the Xiaokaihe Yellow River irrigation district, finding that BPNN accuracy was superior to traditional multiple linear regression. Chen et al. used support vector machine (SVM) in a soil salinity inversion model based on Landsat 8 OLI data, achieving favorable results.

This study takes Pingluo County, Ningxia as the research area. Based on optical and microwave remote sensing imagery, we extract spectral indices and radar polarization combination indices, use VIP and GCD methods to screen salinity characteristic variables, and then employ three machine learning algorithms to construct cultivated land soil salinity inversion models. We evaluate the inversion accuracy under different variable inputs and modeling methods, and invert the soil salinity distribution in the study area to provide theoretical basis and technical support for potential identification and prevention of cultivated land soil salinization in the Yinchuan Plain.

1.1 Study Area Overview

The study area is located in Pingluo County, Shizuishan City, Ningxia (39°14 09 N–38°26 60 N, 105°57 40 E–106°52 52 E), situated between the alluvial fan at the eastern foot of Helan Mountain and the Yellow River alluvial plain in northern Ningxia Plain, representing an important irrigated agricultural region. The average annual precipitation is 150–203 mm, with an evaporation-to-precipitation ratio reaching 10:1. The main land use types include cultivated land, forest land, grassland, and saline wasteland, with total cultivated land area of 9.5$×10^{4}$ hm^2, primarily divided into paddy fields and dry land. The experimental plots (Fig. 1) were distributed across five major geomorphic units in Pingluo County' s cultivated land: the piedmont alluvial fan area, Xidatan saucer-shaped depression area, alluvial plain area, Lingyan platform area, and river beach area. Sampling was conducted before spring plowing when human disturbance was minimal.

1.2 Data Sources and Processing

1.2.1 Soil Sample Collection and Salinity Measurement Soil samples were collected on March 25-26, 2023. During sampling, the five-point method was used to collect 0-20 cm surface soil, which was bagged, numbered, and brought back to the laboratory. After natural air-drying and grinding, the soil samples were filtered through a 2 mm sieve. A 1:5 soil-water ratio extract was prepared, and the electrical conductivity method was used to calculate sample salinity content. Each sample was measured three times, with the average value taken as the point' s salinity content. Based on salinity levels, sample salinization degrees were classified. The overall coefficient of variation was 79.154%, indicating moderate intensity variation, which suggests the samples were widely distributed and representative. Samples were sorted by salinity content, with every fourth sample selected for validation (50 samples) and the remaining samples used for modeling (150 samples). The distribution trends of both modeling and validation sets were consistent with the total sample distribution (Fig. 2), indicating reasonable sample division suitable for model construction and validation.

1.2.2 Remote Sensing Image Acquisition and Processing Landsat 9 OLI optical remote sensing imagery was obtained from the United States Geological Survey (USGS, https://earthexplorer.usgs.gov), with path/row 129/33, acquisition date of March 27, 2023, and cloud cover of 4.23%. ENVI 5.3 software was used for radiometric correction, cropping, and other preprocessing. Salinity indices and vegetation indices were calculated as shown in Table 2.

Sentinel-1 data were obtained from the European Space Agency data portal (https://scihub.copernicus.eu), with acquisition date of March 26, 2023, data level of Level 1, including vertical polarization (VV) and cross-polarization (VH) modes. The data were processed using SNAP software for thermal noise removal, orbit file correction, radiometric calibration, and other procedures. Previous studies have shown that for single-polarization radar data, the relatively small amount of extracted soil information can affect research results. Therefore, this study combined radar image polarization modes to improve soil salinity inversion accuracy. The indices used are listed in Table 3.

1.2.3 Land Use Data Acquisition and Processing Land use data were obtained from the Chinese Academy of Sciences Resource and Environmental Science Data Center (http://www.resdc.cn). According to classification standards, land use in the study area was categorized into cultivated land, forest land, grassland, water bodies, unused land, and urban/industrial/residential land. Cultivated land was extracted as the research object.

1.3 Research Methods

The research process involved: (1) data acquisition and processing, including calculation of spectral indices, radar polarization combination indices, and soil salinity measurement; (2) screening of soil salinity characteristic variables using VIP and GCD methods; and (3) constructing cultivated land soil salinity inversion models using three machine learning algorithms with screened variables, conducting accuracy verification, and selecting the optimal model for regional inversion. The specific research framework is shown in Fig. 3.

1.3.1 Soil Salinity Characteristic Variable Screening Variable Importance in Projection (VIP) Method: VIP is a variable screening method based on partial least squares regression. For a given independent variable, VIP represents not only its direct effect on the dependent variable but also its indirect effect considering other independent variables. The calculation formula is:

$$VIP_j = \sqrt{\frac{p}{SSY_{total}} \sum_{f=1}^{F} (SSY_f \times W_{jf}^2)}$$

where VIP_j is the importance of variable j; p is the number of independent variables; F is the total number of principal components; f is the principal component; SSY_f is the sum of squares explained by principal component f; SSY_{total} is the total sum of squares for the dependent variable; and W_{jf} is the importance of variable j in principal component f. Larger VIP_j values indicate stronger explanatory power of the independent variable on the dependent variable. When $VIP_j > 1$, the independent variable is considered important.

Gray Correlation Degree (GCD) Analysis: GCD analysis is a multi-factor statistical method that determines the main relationships among system factors, using gray correlation degree to characterize the strength and order of relationships between factors to identify the most influential factors. The calculation formula is:

$$GCD = \frac{1}{n} \sum_{t=1}^{n} \gamma(x_0(t), x_i(t))$$

where GCD is the gray correlation degree; t is the variable; n is the number of variables; $x_0(t)$ is the reference sequence; $x_i(t)$ is the comparison sequence; and $\gamma(x_0(t), x_i(t))$ is the correlation coefficient calculated as:

$$\gamma(x_0(t), x_i(t)) = \frac{\min_i \min_t |x_0(t) - x_i(t)| + \rho \max_i \max_t |x_0(t) - x_i(t)|}{|x_0(t) - x_i(t)| + \rho \max_i \max_t |x_0(t) - x_i(t)|}$$

where $\min_i \min_t |x_0(t) - x_i(t)|$ and $\max_i \max_t |x_0(t) - x_i(t)|$ are the minimum and maximum 极差 values, respectively; and ρ is the resolution coefficient between [0, 1], set to 0.5 in this study.

1.3.2 Machine Learning Models **Back Propagation Neural Network (BPNN):** BPNN trains datasets through backpropagation of errors to minimize error. This model provides global approximation of nonlinear mapping with strong adaptive and self-learning capabilities. In this study, the model training target minimum error was set to 0.001, iteration times to 1000, and learning rate to 0.01.

Support Vector Machine (SVM): SVM is based on structural risk minimization principle, relying on limited samples to search for global optimal solutions with good generalization effects on unknown points. This study used Radial Basis Function (RBF) kernel with penalty factor (C) and kernel function parameter (γ) obtained through sample training as 100 and 0.01, respectively.

Random Forest (RF): RF is an ensemble learning method based on multiple regression trees. By combining multiple decision trees and averaging their results, RF makes the generalization error of decision trees converge to produce better predictions. RF excels at handling nonlinear relationships between variables, with prediction performance affected by parameters such as number of regression trees, maximum depth, and minimum leaf size. In this study, the number of regression trees was set to 100, maximum depth to 10, and minimum leaf size to 1.

1.3.3 Model Evaluation To quantitatively compare the accuracy of different algorithms for soil salinity inversion, this study selected two common metrics: coefficient of determination (R^2) and root mean square error (RMSE). Larger R^2 values and smaller RMSE values indicate better model fitting effects.

---

2 Results and Analysis

2.1 Characteristic Variable Screening for Soil Salinity

The VIP values and gray correlation degrees of indices were calculated (Fig. 4 and Fig. 5) to screen characteristic variables for modeling. Results showed that spectral indices SI-T, SI-2, and radar polarization combination indices VV, VH, and VV×VH had VIP values greater than 1, indicating these variables are important indicators of soil salinity and could serve as characteristic variables for modeling.

Fig. 5 shows that the gray correlation degree between spectral indices and soil salinity ranged from 0.65 to 0.71, with SI-T having the highest value (0.71) and EVI the lowest (0.65). The gray correlation degree between radar polarization combination indices and soil salinity ranged from 0.71 to 0.75, with VV×VH showing the highest value (0.75). To implement feature variable screening, this study set the gray correlation degree threshold at 0.69 for both spectral and radar polarization combination indices. Statistics on salinity-sensitive indices based on gray correlation analysis revealed that more spectral feature variables than radar feature variables were screened. Among them, SI-T, SI-2, SI-3, NDVI, and EVI indices were selected as spectral characteristic variables, while VV, VH, and VV×VH were selected as radar polarization combination indices for modeling.

2.2 Model Construction and Validation

Single Remote Sensing Data Models: In the spectral index model, soil salinity was used as the dependent variable and selected spectral indices as independent variables to construct BPNN, SVM, and RF models. In the radar polarization combination index model, selected VV, VH, and VV×VH indices were used as independent variables to construct corresponding models. Results (Table 4) showed that spectral index models generally achieved higher accuracy than radar polarization combination index models, likely because this study' s salt sampling focused on 0–20 cm surface soil, while radar data are more suitable for detecting subsurface targets and optical data provide rich surface reflectance information.

Among the three models constructed with VIP-selected variables, the VIP-RF model achieved the highest accuracy with modeling set $R^2 = 0.726$ and validation set $R^2 = 0.689$, demonstrating strong model stability. The three models constructed with GC-selected variables showed lower overall validation set R^2, with GC-RF performing best, followed by GC-SVM, and GC-BPNN showing the largest validation set error.

Comparison of Variable Screening Methods: Comparing model accuracies from the two variable screening methods (Table 4) revealed that in spectral index models, validation set accuracy followed the pattern VIP-BPNN > GC-BPNN, VIP-SVM > GC-SVM, and VIP-RF > GC-RF. In radar polarization combination index models, except for the SVM model where VIP-SVM < GC-SVM, the other two machine learning models showed the same pattern where VIP-selected variables produced higher modeling and validation set R^2 than GC-selected variables. This indicates that VIP method is more suitable for soil salinity inversion in the study area.

Multi-source Remote Sensing Data Models: To verify the impact of multi-source remote sensing data on soil salinity inversion accuracy, VIP-selected spectral indices and radar polarization combination indices were combined as independent variables for modeling. Results (Table 5) showed that the synergistic model combining spectral and radar indices significantly improved inversion accuracy compared with single-source models. Specifically, the VIP-RF model showed validation set R^2 improvements of 0.091 and 0.237 over single spectral index and single radar index models, respectively, with corresponding RMSE reductions of 0.175 and 0.377 $g \cdot kg^{-1}$. The modeling set R^2 reached 0.791 and validation set R^2 reached 0.780, with both R^2 values greater than 0.780, indicating stronger model accuracy and learning performance.

Scatter plots of measured versus predicted values from the validation set for different data source variables using the VIP-RF model (Fig. 6) demonstrated that the synergistic model combining spectral indices and radar polarization combination indices achieved the best fitting effect. Therefore, this model was selected to invert soil salinity in the study area' s cultivated land.

2.3 Model Application

The optimal synergistic model (VIP-RF) was applied to invert soil salinity across Pingluo County' s cultivated land, producing a salinity grade distribution map (Fig. 7). Statistical analysis of different salinization grades (Table 6) revealed that salinization in Pingluo County' s cultivated land was severe, with salinized soil accounting for 86.81% of total cultivated land area. Moderately salinized soil was most widespread, covering 33.54% of cultivated land, followed by mildly salinized soil at 23.77%. Severely salinized soil and saline soil were mainly distributed in the central-western and eastern Yellow River irrigation areas. The inversion results were consistent with field sampling conditions, demonstrating that the multi-source remote sensing data-based inversion model is effective for soil salinity inversion in the study area.

---

3 Discussion

Combining multi-source remote sensing data for soil salinity inversion has become a research hotspot in recent years. However, how to efficiently combine different data sources while reducing the loss and distortion of spectral features and radar information during image processing remains to be further studied. Yu et al. used product transformation, Brovey transformation, and wavelet transformation to fuse Landsat 8 OLI and SAR imagery, finding that Brovey transformation could better achieve the fusion of spaceborne SAR and optical images. However, these methods still face issues where increased information content and spatial resolution are accompanied by loss of spectral features. Jiang et al. found that combining radar backscattering coefficients and modified temperature vegetation drought index as SVM input parameters significantly improved soil moisture monitoring accuracy. Therefore, this study combined VIP-selected spectral indices and radar polarization combination indices for modeling, fully leveraging the advantages of optical and microwave remote sensing to better preserve soil spectral characteristics and radar information.

Optimal variable screening methods combined with optimal machine learning algorithms can effectively improve inversion accuracy. This study compared the 优劣 of VIP and GCD methods, finding that VIP produced more reliable results, consistent with Wang et al. This is because VIP optimizes model accuracy by assessing the impact of features on model accuracy and the change in model accuracy before and after adding noise interference. However, Li found that GCD-based variable screening could also effectively improve model fitting accuracy. Some studies have noted that soil composition significantly affects inversion model performance, so whether the variable screening methods selected in this study are applicable to other research areas requires further investigation.

This study selected three traditional machine learning algorithms for soil salinity inversion, finding that BPNN achieved lower accuracy while RF achieved the highest accuracy, meeting the requirements for soil salinity inversion in the study area. This aligns with conclusions from Zhang et al. These optimization algorithms demonstrate that scholars are increasingly focusing on the impact of feature selection and hyperparameter adjustment on machine learning model performance, as well as the complex dependencies and interactions between input features and hyperparameters, thereby improving prediction accuracy from the model level. Future research could consider introducing different optimization algorithms combined with feature variable screening methods to further improve model inversion accuracy and universality.

This study used a synergistic model of spectral indices and radar polarization combination indices to invert soil salinity in the study area' s cultivated land. Results showed that cultivated land salinization was severe, with salinized soil accounting for 86.81% of total cultivated land area, and moderately salinized soil being most widespread at 33.54%. These findings align with previous research. Pingluo County is located in an inland arid region with dry climate and poor soil aeration and permeability, making it prone to salinization. The central-western area comprises piedmont alluvial fan and saucer-shaped depression areas with poor drainage conditions. The eastern Yellow River irrigation area suffers from unreasonable irrigation practices, including excessive water diversion and flood irrigation, where irrigation water leakage causes groundwater level rise and intense evaporation, exacerbating salinization and severely affecting agricultural production. Therefore, remote sensing monitoring combined with salinization improvement mechanisms such as physical regulation, chemical conditioning, irrigation and drainage management, and biological remediation should be implemented according to local conditions to propose feasible management measures and reduce the harm caused by salinization to local agricultural development and the natural environment.

A limitation of this study is the temporal inconsistency between soil sample collection and image acquisition, as soil salinity exhibits short-term variations, making it difficult for image data to truly reflect surface conditions and introducing data uncertainty. Future research could consider using UAV technology to collect soil samples and remote sensing data within the same time period. With UAV data resolution reaching the centimeter level, model inversion accuracy would be higher.

4 Conclusion

This study used VIP and GCD methods to screen characteristic variables and constructed spectral index models, radar polarization combination index models, and synergistic models based on spectral and radar index combinations using three machine learning algorithms. The following conclusions were drawn:

1. The VIP method identified markedly different numbers of spectral indices and radar polarization combination indices compared with the GCD method. Models built with VIP-selected variables achieved relatively higher accuracy, while those built with GCD-selected variables showed poorer performance. In the spectral index model, the VIP-RF model performed best, while in the radar polarization combination index model, all three machine learning models showed better performance with VIP-selected variables, indicating that the VIP method is more suitable for soil salinity inversion in this study area.

2. Among the three machine learning algorithms, BPNN showed the worst performance while RF achieved the best accuracy, making it more suitable for soil salinity inversion in the study area' s cultivated land.

3. The synergistic model combining multi-source remote sensing data achieved validation set $R^2 = 0.780$, representing improvements of 0.091 and 0.237 over single spectral index and single radar index models, respectively. This demonstrates that combining multi-source remote sensing data can effectively improve soil salinity prediction accuracy, contributing to more precise regional soil salinity distribution studies and enhanced soil salinization monitoring.

4. Inversion results revealed that cultivated land salinization in Pingluo County is severe, with non-salinized soil accounting for only 13.19% of total cultivated land area. Mildly and moderately salinized soils are predominant, accounting for 23.77% and 33.54% respectively, while 15.37% of cultivated land has deteriorated to severely salinized soil and saline soil.

References

[1] Abdelaziz M E, Abdelsattar M, Abdeldaym E A, et al. Piriformospora indica alters Na^+ homeostasis, antioxidant enzymes and LeNHX1 expression of greenhouse tomato grown under salt stress[J]. Scientia Horticulturae, 2019, 256: 108532, doi: 10.1016/j.scienta.2019.05.059.

[2] Singh A. Soil salinity: A global threat to sustainable development[J]. Soil Use and Management, 2022, 38(1): 39-67.

[3] Li X Y, Zhang F, Wang Z. Present situation and development trend in building remote sensing monitoring models of soil salinization[J]. Remote Sensing for

Natural Resources, 2022, 34(4): 11-21.

[4] Chen H Y, Zhao G X, Chen J C, et al. Remote sensing inversion of saline soil salinity based on modified vegetation index in estuary area of Yellow River[J]. Transactions of the Chinese Society of Agricultural Engineering, 2015, 31(5): 107-114.

[5] Wang J Z, Ding J L, Ma X K, et al. Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China[J]. Geoderma, 2019, 353: 172-187.

[6] Zhao M Y. Study of clouds removal methods on remote sensing images[D]. Tianjin: Tianjin University of Science and Technology, 2016.

[7] Li B, Wang Y Q. Radar inversion and simulation of salty soil salinization[J]. Journal of Arid Land Resources and Environment, 2015, 29(8): 180-184.

[8] Wei H M, Jia K L, Zhang X, et al. Prediction of soil salinity based on machine learning and multispectral remote sensing in Yinchuan Plain[J]. Arid Land Geography, 2023, 46(1): 103-114.

[9] Li X Y, Jia K L, Wei H M, et al. Prediction of soil salt content based on random forest algorithm[J]. Arid Zone Research, 2023, 40(8): 1258-1267.

[10] Ma C. Quantitative retrieval of soil salt content based on Sentinel-1 dual polarization radar image[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(2): 153-158.

[11] Zhang Z T, Du Y Y, Lao C C, et al. Inversion model of soil salt content in different depths based on radar remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(10): 243-251.

[12] Khan N M, Rastoskuev V V, Sato Y, et al. Assessment of hydrosaline land degradation by using a simple approach of remote sensing indicators[J]. Agricultural Water Management, 2005, 77(1-3): 96-109.

[13] Fan Y G, Zhang W K, Liu J Y. Remote sensing monitoring model of soil salinization in the Yellow River Delta zone based on vegetation index-salt index feature space[J]. Shandong Agricultural Sciences, 2016, 48(5): 137-141.

[14] Xiao S T, Nurmemet I, Muhetaer N, et al. Spatial and temporal analysis of soil salinity in Yutian Oasis by combined optical and radar multi-source remote sensing[J]. Arid Zone Research, 2023, 40(1): 59-68.

[15] Wang F, Ding J L, Wei Y, et al. Sensitivity analysis of soil salinity and vegetation indices to detect soil salinity variation by using Landsat series images: Applications in different oases in Xinjiang, China[J]. Acta Ecologica Sinica, 2017, 37(15): 5007-5022.

[16] Wang H F. Hyperspectral remote sensing based models for soil moisture and salinity prediction[D]. Xianyang: Northwest A & F University, 2019.

[17] Sun Y N, Li X Y, Shi H B, et al. Remote sensing inversion of soil salinity and seasonal difference analysis based on multi-source data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 169-180.

[18] Liu E, Wang J T, Chang B H, et al. Quantitative remote sensing inversion of soil salinization in Xiaokaihe Yellow River irrigation district[J]. China Rural Water and Hydropower, 2019(12): 20-24.

[19] Sahana M, Rehman S, Patel P P, et al. Assessing the degree of soil salinity in the Indian Sundarban Biosphere Reserve using measured soil electrical conductivity and remote sensing data derived salinity indices[J]. Arabian Journal of Geosciences, 2020, 13(24): 1289, doi: 10.1007/s12517-020-06310-w.

[20] Nguyen K A, Liou Y A, Tran H P, et al. Soil salinity assessment by using near infrared channel and vegetation soil salinity index derived from Landsat 8 OLI data: A case study in the Tra Vinh Province, Mekong Delta, Vietnam[J]. Progress in Earth and Planetary Science, 2020, 7(1): 1-16.

[21] Liu Y Q. Quantitative estimation of soil salinity and its composition based on spectra in Yellow River Delta[D]. Tai' an: Shandong Agricultural University, 2017.

[22] Cao X Y, Ding J L, Ge X Y, et al. Estimation of soil conductivity based on spectral simulation of different satellites[J]. Arid Land Geography, 2020, 43(1): 172-181.

[23] Guo X J. Response analysis of multi-polarization radar on freeze-thaw period soil salt in Inner Mongolia Hetao irrigation district[D]. Hohhot: Inner Mongolia Agricultural University, 2014.

[24] Yu H Y, Wang Z J, Li B, et al. Response analysis of multi-polarization radar on soil salt during melting period in Hetao irrigation district of Inner Mongolia[J]. Journal of Yangtze River Scientific Research Institute, 2015, 32(11): 19-24.

[25] Li S Y, Huo Y Z, Wang H X. Response analysis of multi-polarization radar on soil salinity during freezing period in Hetao irrigation district of Inner Mongolia Autonomous Region[J]. Jiangsu Agricultural Sciences, 2016, 44(3): 348-352.

[26] Jiang H, Rusuli Y, Abudoukerimu B, et al. Soil moisture retrieval by synergizing optical and microwave remote sensing data based on support vector machine regression algorithm[J]. Geography and Geo-information Science, 2017, 33(6): 30-36.

[27] Li M L. Study on the grey relation degree estimation model of soil organic matter based on hyperspectral[D]. Tai' an: Shandong Agricultural University, 2017.

[28] Oussama A, Elabadi F, Platikanov S, et al. Detection of olive oil adulteration using FT-IR spectroscopy and PLS with variable importance of projection

(VIP) scores[J]. Journal of the American Oil Chemists' Society, 2012, 89(10): 1807-1812.

[29] Maimaitiyiming M, Ghulam A, Bozzolo A, et al. Early detection of plant physiological responses to different levels of water stress using reflectance spectroscopy[J]. Remote Sensing, 2017, 9(7): 745, doi: 10.3390/rs9070745.

[30] Liu S F, Cai H, Yang Y J, et al. Advance in grey incidence analysis modelling[J]. Systems Engineering Theory and Practice, 2013, 33(8): 2041-2046.

[31] Yang S E, Huang L. Financial crisis warning model based on BP neural network[J]. Systems Engineering Theory & Practice, 2005, 25(1): 12-18.

[32] Zhao Y Q. Forest fire recognition algorithm based on fuzzy neural network[J]. Computer Simulation, 2015, 32(2): 369-373.

[33] Wang H Y, Li J H, Yang F L. Overview of support vector machine analysis and algorithm[J]. Application Research of Computers, 2014, 31(5): 1281-1286.

[34] Biau G, Scornet E. A random forest guided tour[J]. Test, 2016, 25(2): 197-227.

[35] Yang L B, Chen C B, Zheng H W, et al. Retrieval of soil salinity content based on optimized random forest regression model[J]. Journal of Geo-information Science, 2021, 23(9): 1662-1674.

[36] Zhang Z T, Wei G F, Yao Z H, et al. Soil salt inversion model based on UAV multispectral remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(12): 151-160.

[37] Chen B L, Zheng H W, Luo G P, et al. Adaptive estimation of multi-regional soil salinization using extreme gradient boosting with Bayesian TPE optimization[J]. International Journal of Remote Sensing, 2022, 43(3): 778-811.

[38] Wang L Y, Hu P, Zheng H W, et al. Integrative modeling of heterogeneous soil salinity using sparse ground samples and remote sensing images[J]. Geoderma, 2023, 430: 116321, doi: 10.1016/j.geoderma.2022.116321.

[39] Xu H T, Chen C B, Zheng H W, et al. AGA based selection of feature subsets and optimization of parameter in regional soil salinization monitoring[J]. International Journal of Remote Sensing, 2020, 41(12): 4470-4495.

[40] Zhang X. Study on spatial distribution and driving factors of soil salinity in Yinchuan Plain[D]. Yinchuan: Ningxia University, 2022.

[41] Yang J S, Yao R J, Wang X P, et al. Research on salt-affected soils in China: History, status quo and prospect[J]. Acta Pedologica Sinica, 2022, 59(1): 10-27.

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