

Article

Multi-Model Comprehensive Inversion of Surface Soil Moisture from Landsat Images Based on Machine Learning Algorithms

Weitao Lv ¹, Xiasong Hu ^{1,*}, Xilai Li ², Jimei Zhao ², Changyi Liu ¹, Shuaifei Li ¹, Guorong Li ¹ and Haili Zhu ¹

¹ School of Geological Engineering, Qinghai University, Xining 810016, China; lwt1055541638@163.com (W.L.); liuchangyi1991@sina.com (C.L.); afei2275@163.com (S.L.)

² College of Agriculture and Animal Husbandry, Qinghai University, Xining 810016, China; zhaojimei@qhu.edu.cn (J.Z.)

* Correspondence: huxiasong01@163.com

Abstract: Soil moisture plays an important role in maintaining ecosystem stability and sustainable development, especially for the upper reaches of the Yellow River region. Therefore, accurately and conveniently monitoring soil moisture has become the focus of scholars. This study combines three machine learning algorithms: random forest (RF), support vector machine (SVM), and back propagation neural network (BPNN)—with the traditional monitoring of soil moisture using remote sensing indices to construct a more accurate soil moisture inversion model. To enhance the accuracy of the soil moisture inversion model, 27 environmental variables were screened and grouped, including vegetation index, salinity index, and surface temperature, to determine the optimal combination of variables. The results show that screening the optimal independent variables in the Xijitan landslide distribution area lowered the root mean square error (RMSE) of the RF model by 16.95%. Of the constructed models, the combined model shows the best applicability, with the highest R^2 of 0.916 and the lowest RMSE of 0.877% with the test dataset; the further research shows that the BPNN model achieved higher overall accuracy than the other two individual models, with the test set R^2 being 0.809 and the RMSE 0.875%. The results of this study can provide a theoretical reference for the effective use of Landsat satellite data to monitor the spatial and temporal distribution of and change in soil water content on the two sides of the upper Yellow River basin under vegetation cover.

Keywords: soil moisture inversion; random forest; support vector machine; neural network; combined modeling method; optimal combination of variables



Citation: Lv, W.; Hu, X.; Li, X.; Zhao, J.; Liu, C.; Li, S.; Li, G.; Zhu, H.

Multi-Model Comprehensive Inversion of Surface Soil Moisture from Landsat Images Based on Machine Learning Algorithms.

Sustainability **2024**, *16*, 3509. <https://doi.org/10.3390/su16093509>

Received: 4 January 2024

Revised: 28 March 2024

Accepted: 28 March 2024

Published: 23 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Soil moisture is one of the environmental variables controlling the heat and water exchange cycle between the Earth's surface and the atmosphere, directly affecting soil evaporation and vegetation transpiration [1,2]. Changes in soil moisture affect, to a certain extent, the development of vegetation, which is an indispensable element in maintaining the stability and sustainable development of the ecosystem. Therefore, accurate and convenient monitoring of soil moisture has become a focus of attention for scholars around the world [3]. Furthermore, alterations in soil moisture levels can lead to variations in soil pore water pressure, which may trigger the reactivation of coastal landslides. It is crucial to monitor soil moisture levels with greater precision in regions where landslides are more prevalent [4]. In addition, the development of soil moisture monitoring technology is important for the monitoring and forecasting of global hydrometeorological hazards [5], the development of precision agricultural irrigation technology [6], and the monitoring and forecasting of global geological hazards [7,8].

The Xijitan and Xiazangtan landslides are typical giant landslides located in the upper reaches of the Yellow River. Maintaining the stability of these two landslides is of great significance for maintaining the ecological stability of the upper reaches of the Yellow River.

In addition, there are large hydropower projects such as Longyangxia Hydropower Station, Laxiwa Hydropower Station, Liji Xia Hydropower Station, Gongboxia Hydropower Station, and Jishixia Hydropower Station in the watershed of this section [8], and timely monitoring of the stability of these two landslides is important to ensure the stability of the hydropower stations [8]. Therefore, better and more convenient monitoring of the soil moisture is of great practical significance for effectively preventing and controlling disasters such as landslides. It can promote the sustainable development of the ecological environment in the basin.

Soil moisture data are acquired either in situ or via remotely sensed soil moisture products [9]. In situ soil moisture observations are accurate but costly and time-consuming, and the acquired data generally lack spatial representativeness and coverage [10,11]. These deficiencies do not apply to remotely sensed soil moisture. The main remotely sensed soil moisture products currently in use are the ESA CCI soil moisture, NASA USDA Global Soil Moisture Data, and the Copernicus Global Land Service Surface Soil Moisture (CSSM) [11,12]. These soil moisture data products have the advantages of extensive coverage and easy accessibility, which can overcome the shortcomings of traditional field methods. They can be used to effectively monitor the spatial and temporal distribution and dynamic changes in soil moisture on a large scale [13]. However, this type of data is generally less accurate than in situ collected soil moisture data and tend to have a coarse spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$, or $1 \text{ km} \times 1 \text{ km}$, which makes it challenging to meet the actual needs of landslide disaster management, soil erosion control, and other related applications [13]. At present, there is insufficient research on the acquisition and inversion of high spatial resolution and high accuracy remote sensing soil moisture information. Therefore, it is of great practical significance to further invert soil moisture using remote sensing data with a high spatial resolution and accuracy. The inverted results are essential for monitoring drought and preventing land degradation, soil erosion, landslides, and other disasters.

Remote sensing soil moisture inversion is classified into three types according to the data sources: inversion based on thermal infrared data, inversion based on microwave data, and inversion based on optical data [14,15]. Of these three types, thermal infrared sensors can monitor the thermodynamic characteristics of soils with different water contents [16] and estimate soil moisture through thermal inertia functions or via calculating crop water stress indices [17,18]. Optical remote sensing detects mainly surface reflectance and short-wave radiation at visible and near-infrared wavelengths [19]. It can intuitively obtain much spectral information on the surface [20]. In some studies, drought indices such as temperature vegetation drought index (TVDI), visible and shortwave infrared drought index (VSDI), and vegetation water supply index (VSWI) have been used to assess soil moisture, but these indices could not accurately reflect the actual soil moisture, so how to invert soil moisture more accurately has become the main goal of further research by scientists [9]. Optical remote sensing data have been analyzed using machine learning methods such as neural networks, support vector machine (SVM) models, and random forest (RF) models. They provide reliable means for soil moisture inversion from optical satellite data in recent years [20]. Yao et al. (2022) constructed three kinds of soil moisture inversion models under different depths, using BPNN, SVM, and MLR, based on GF-1 satellite images. Their results showed that the ability of the three models follows the descending order of BPNN > SVM > MLR in inverting vegetation conditions [21–23]. Wang et al. (2023) constructed a soil moisture inversion model based on structural equation modeling (SEM) and neural network method (ANN) using microwave physics, vegetation, temperature, lower mat albedo, topography, and drought index as the input variables [9].

It can be seen that most of the existing studies on soil moisture inversion used Sentinel-1 microwave data. In contrast, fewer studies have used spectral data for constructing soil moisture inversion models. At the same time, most existing studies used a single machine learning model to invert soil moisture, and few have combined different models to invert soil moisture. This study aims to fill these knowledge gaps by inverting soil moisture in

the upper reaches of the Yellow River by combining multiple machine learning models. The specific objectives are (1) to assess the optimal independent environmental variables in the two areas of Xijitan and Xiazangtan that are critical to the inversion of soil moisture based on an RF model; (2) to compare the performance of three machine learning methods (BPNN model, SVM model, and RF model) in inverting surface soil moisture; and (3) to determine whether the combined model can outperform the joint inversion of three single models. The best model was used to invert the distribution of soil moisture in the study area. The results of the study provide a theoretical basis and practical guidance for the maintenance of the stability and sustainability of the ecosystem of the upper Yellow River basin, as well as the prevention and control of soil erosion, landslides, and other disasters.

2. Study Area

The area selected for this study is the Xijitan and Xiazangtan super large scale landslides located in the upper Yellow River from Longyangxia to Jishixia section, as shown in Figure 1 [22]. The Xijitan super large scale landslide is located on the north bank of the Yellow River, in the Guide Basin, with the geographic coordinates of $101^{\circ}24'2.1'' \sim 101^{\circ}29'1.8''$ E, $36^{\circ}02'51.0'' \sim 36^{\circ}07'48.2''$ N. The highest elevation in the area is 3150 m and the lowest is 2176 m, with a standard deviation of 188.93 m. The terrain is hilly [21]. The area is arid with a cool and cold climate and strong solar radiation, with an average annual temperature of 9.03°C , an average annual precipitation of 254.8 mm, a relatively sparse distribution of vegetation, and a relatively high soil salinity [21]. The vegetation distribution in the distribution area of Xijitan landslide is relatively sparse, with herbaceous plants distributed in patches and shrubs distributed sporadically, and the dominant plant species are *Achnatherum inebrians*, *Caragana roborovskyi*, *Salsola passerina*, and so on [23]. These plant species have an average plant height of 2.51–10.33 cm, an average root diameter of 0.42–0.46 mm, and a dominant taproot type [23]. *Achnatherum inebrians* is a perennial herb with pliable fibrous roots, erect culms, few tufts, 60–100 cm tall, and 2.5–3.5 mm in diameter [23]. *Caragana roborovskyi* is a perennial with flexible fibrous roots, upright culms, and few tufts. *Caragana roborovskyi* is a shrub with small obovate or oblong leaves, 0.3–1.0 m long, apex rounded or acute, spiny, and base cuneate [23].

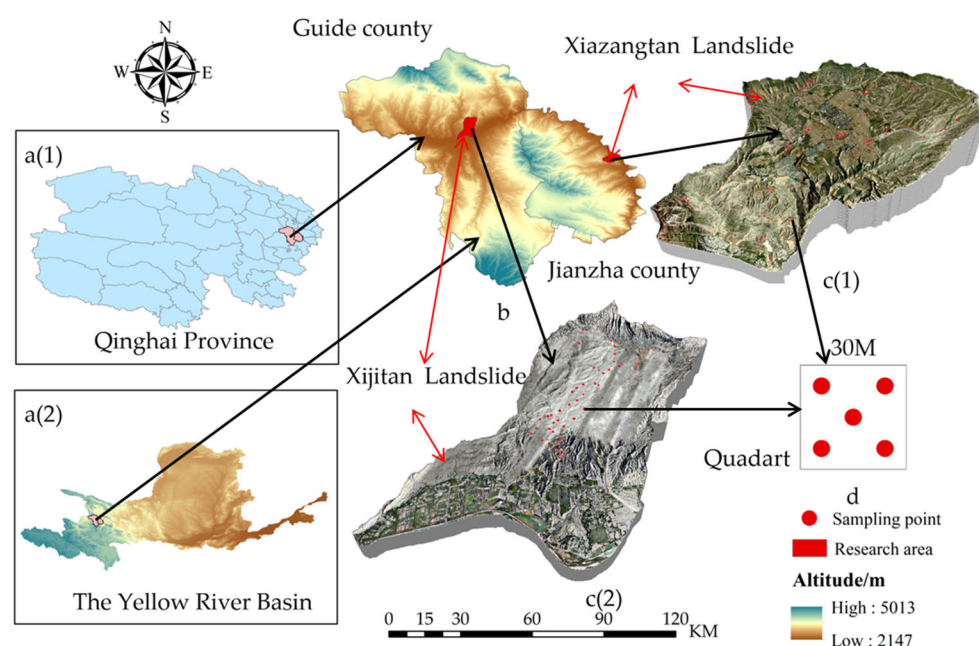


Figure 1. Geographic location of the study area: a(1) is Qinghai Province, a(2) is the Yellow River Basin, b is Guide and Jianzha counties, c(1) is the Xiazangtan landslide distribution area, c(2) is the Xijitan landslide distribution area, and d is the sampling site.

The Xiazangtan super large scale landslide is located in Maketang Town, Jianzha County, Qinghai Province, on the right bank of the Yellow River, with the geographical coordinates of $101^{\circ}57'9.0''\sim 102^{\circ}1'7.5''$ E, $35^{\circ}57'38.16''\sim 35^{\circ}59'54.36''$ N. The trailing-edge elevation is 2820 m, the leading-edge height is 1001 m, and the difference in size between the front and back is 819 m. The average annual temperature is 4.11°C , the average annual precipitation is 347.20 mm, and the average annual evapotranspiration is 1667.80 mm [22]. The overall vegetation cover is high, the vegetation distribution is relatively dense, and the soil salinity is relatively low [22]. The vegetation in the distribution area of Xiazangtan landslide is relatively densely distributed, dominated by herbaceous plants with a few shrubs. Its dominant species are mainly *Stipa purpurea*, *Stipa bungeana*, *Carex parvula*, and so on [23]. The average plant height of these plant species was 2.81–9.12 cm, the average root diameter was 1.38–2.09 mm, and the root type of the plant was mainly fibrous [23]. The fibrous roots of *Stipa purpurea* are fine and tough; the culms are thin, 20–45 cm tall; and the grass is hard, grazing-tolerant, and highly productive [23]. The fibrous roots of *Carex parvula* are coarse and tough with a sandy outer covering, and the culms are erect, hard, with a white pith, forming large dense tufts, 50–250 cm tall, and 3–5 mm in diameter [23].

3. Materials and Methods

3.1. Data

3.1.1. In Situ Soil Moisture Data

The soil moisture sampling was undertaken from 29 April to 31 April 2023 and from 14 May to 16 May 2023, coinciding with the satellite transit time. In total, 400 soil samples were collected, and they were widely distributed in the Xiazangtan and Xijitan landslides (Figure 1c). In this study, 40 sampling areas were set up in each of the two landslides, with a minimum interval of 500 m between adjacent sampling areas; Each sampling area was $30\text{ m} \times 30\text{ m}$, and five sampling points were set up in each plot, and the sampling points in each sampling plot were arranged as shown in Figure 1d. First, the GPS coordinates of each sampling point were recorded, and then soil samples were collected at 0–10 cm below the surface of each sampling point; finally, the soil samples were sealed in aluminum boxes quickly and brought back to the laboratory in time, where soil moisture was determined using the drying method.

3.1.2. Remote Sensing Data

The remote sensing data used in this study are mainly Landsat 9 satellite data, an ASTER digital elevation model (DEM) of 30 m grid size. The Landsat 9 satellite data were obtained from the USGS website, recorded mainly over 29 April 2023–15 May 2023. Chronologically, the image data coincided mostly with the timing of in situ soil moisture data collection. At the same time, the influence of topographic conditions on soil moisture in the study area was addressed via radiometric correction of the selected data using the ASTER DEM, obtained from the Geospatial Data Cloud website. All the collected data were processed using ArcGIS 10.8, ENVI 5.3, and other software such as image mosaicing, cropping, radiometric calibration, and atmospheric correction to obtain the base data of the study area.

3.2. Environmental Variables and Surface Temperature Calculations

3.2.1. Derivation of Environmental Variables

It has been shown that spectral indices such as vegetation index, salinity index, surface temperature, and surface reflectance can provide practical information on soil moisture monitoring [21]. Among these indices, normalized difference vegetation index (NDVI) [24], ratio vegetation index (RVI) [25], difference vegetation index (DVI) [26], soil-adjusted vegetation index (SAVI) [27], enhanced vegetation index (EVI) [28], and greenness vegetation index (GVI) [28] are commonly used for inversion of soil moisture from optical satellite data. In addition, soil moisture and salinity have similar effects on soil reflectance spectra when soil moisture has some dynamic changes in space and time [24]. When plants in a

region are subject to water and salt stress, the difference in their state is reflected in the spectral features [24]. Therefore, in this study, five commonly used salinity indices, such as the specific salinity index (SI-T) [29], salinity index 1 (S1) [30], salinity index 2 (S2) [30], salinity index 3 (S3) [30], and normalized salinity index (NDSI) [31], were selected and simultaneously applied to invert soil moisture in the study area.

In addition, soil moisture distribution is influenced by several factors, such as soil properties, climate, land cover, surface biophysical characteristics, and surface topographic parameters. In this study, to effectively improve the accuracy of the soil moisture inversion model, in addition to the six selected vegetation indices and five salinity indices, five spectral indices such as brightness index (Bright) [32], greenness index (Green) [32], humidity index (Wet) [32], normalized difference water index (NDWI) [21], and normalized difference built-up index (NDBI) [21] were also calculated, along with the topographic parameters such as elevation, slope, and slope direction. The specific formulae for calculating each environmental variable are shown in Table 1.

Table 1. Formula for calculating environmental variables in the study area.

Variable Group	Variable Name	Formulas and Notes
BR	B2, B3, B4, B5, B6, B7	Reflectance in $\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$, and α_7 bands.
VI	NDVI [24]	$NDVI = (\alpha_5 - \alpha_4) / (\alpha_5 + \alpha_4)$
	RVI [25]	$RVI = \alpha_5 / \alpha_4$
	DVI [26]	$DVI = \alpha_5 - \alpha_4$
	SAVI [27]	$SAVI = \frac{(\alpha_5 - \alpha_4) * (1 + 0.5)}{(\alpha_5 + \alpha_4 + 0.5)}$
	EVI [28]	$EVI = \frac{2.5 * (\alpha_5 - \alpha_4)}{\alpha_5 + 6 * \alpha_4 - 7.5 * \alpha_2 + 1}$
	GVI [28]	$GVI = (\alpha_5 - \alpha_3) / (\alpha_5 + \alpha_3)$
SI	SI-T [29]	$SI - T = \alpha_4 / \alpha_5$
	S1 [30]	$S1 = \alpha_2 / \alpha_4$
	S2 [30]	$S2 = (\alpha_2 - \alpha_4) / (\alpha_2 + \alpha_4)$
	S3 [30]	$S3 = (\alpha_3 * \alpha_4) / \alpha_2$
	NDSI [31]	$NDSI = (\alpha_4 - \alpha_5) / (\alpha_4 + \alpha_5)$
GP	Bright [32]	$Bright = 0.3029\alpha_2 + 0.2786\alpha_3 + 0.4733\alpha_4 + 0.5599\alpha_5 + 0.5080\alpha_6 + 0.1872\alpha_7$
	Green [32]	$Green = -0.2941\alpha_2 - 0.2430\alpha_3 - 0.5424\alpha_4 + 0.7276\alpha_5 + 0.7130\alpha_6 - 0.1608\alpha_7$
	Wet [32]	$Wet = 0.1511\alpha_2 + 0.1973\alpha_3 + 0.3283\alpha_4 + 0.3407\alpha_5 - 0.7117\alpha_6 - 0.4559\alpha_7$
	Albedo [21]	$Albedo = 0.356\alpha_2 + 0.130\alpha_4 + 0.373\alpha_5 + 0.085\alpha_6 + 0.072\alpha_7 - 0.0018$
	NDWI [21]	$NDWI = (\alpha_3 - \alpha_5) / (\alpha_3 + \alpha_5)$
	NDBI [21]	$NDBI = (\alpha_6 - \alpha_5) / (\alpha_6 + \alpha_5)$
TF	Elevation, AS, S	Elevation, Slope direction, Gradient

Tables: $\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$, and α_7 are the reflectances of blue, green, red, near-infrared bands, short-wave infrared 1, and short-wave infrared 2 bands after atmospheric correction.

3.2.2. Surface Temperature

Surface temperature indirectly affects soil moisture by influencing surface evapotranspiration [33]. In this study, the thermal infrared band of Landsat 9 satellite data was used to calculate the surface temperature using the thermal radiation transmission model. First, the parameters $L\uparrow$ and $L\downarrow$ were obtained through the NASA website, and second, the surface temperature of the study area was calculated based on ENVI 5.3, and the calculation formula is shown below [33]:

$$FVI = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \quad (1)$$

$$\varepsilon = \begin{cases} 0.973 & NDVI < 0.2 \\ 0.004 \times FVC + 0.986 & 0.2 \leq NDVI \leq 0.5 \\ 0.986 & NDVI > 0.5 \end{cases} \quad (2)$$

$$L_{\lambda} = (\varepsilon B(T_S) + (1 - \varepsilon)L_{\downarrow})\varepsilon + L_{\uparrow} \quad (3)$$

$$T_S = 1321.08 / \ln(774.89 / B(T_S) + 1) \quad (4)$$

where ε is the ground-specific emissivity; T_S is the ground surface temperature, K; $B(T_S)$ is the thermal radiation brightness; $L_{\uparrow}$ is the radiant brightness of the atmosphere upwards, W/m^2 ; $L_{\downarrow}$ is the brilliant brightness of the atmosphere downwards reflected by the ground surface, W/m^2 . The NDVI formula is shown in Table 1.

3.3. Variable Preference

In soil moisture inversion, not all environmental variables provide information favorable for model construction, and models should avoid the influence of unfavorable variables [24]. The results of related studies have shown that removing potentially irrelevant environmental variables is beneficial for improving model accuracy [24]. In this study, the preferred environmental variables are decided based on the RF model, and the process of determining variable preference is described as follows [24]: ① Input each group of variables into the RF model, calculate the importance of all variables, and rank them in the descending order of importance; ② Delete the least important variable in each group, and then input the remaining variables into the RF model to retrain and reorder them, followed by deletion of the last-placed variable; ③ Repeatedly delete the least important variable in each group until there is only one variable remaining, which indicates the end of the cycle; ④ Each model run is to calculate the determination coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE), mean bias error (MBE); ⑤ Use RMSE, R^2 , MAE, and MBE to finally judge the optimal set of variables; ⑥ Repeat the aforementioned steps 1~3 until all subgroups are covered and the optimal set of variables is obtained for the two study areas of Xiazangtan and Xijitan. The specific workflow is shown in Figure 2.

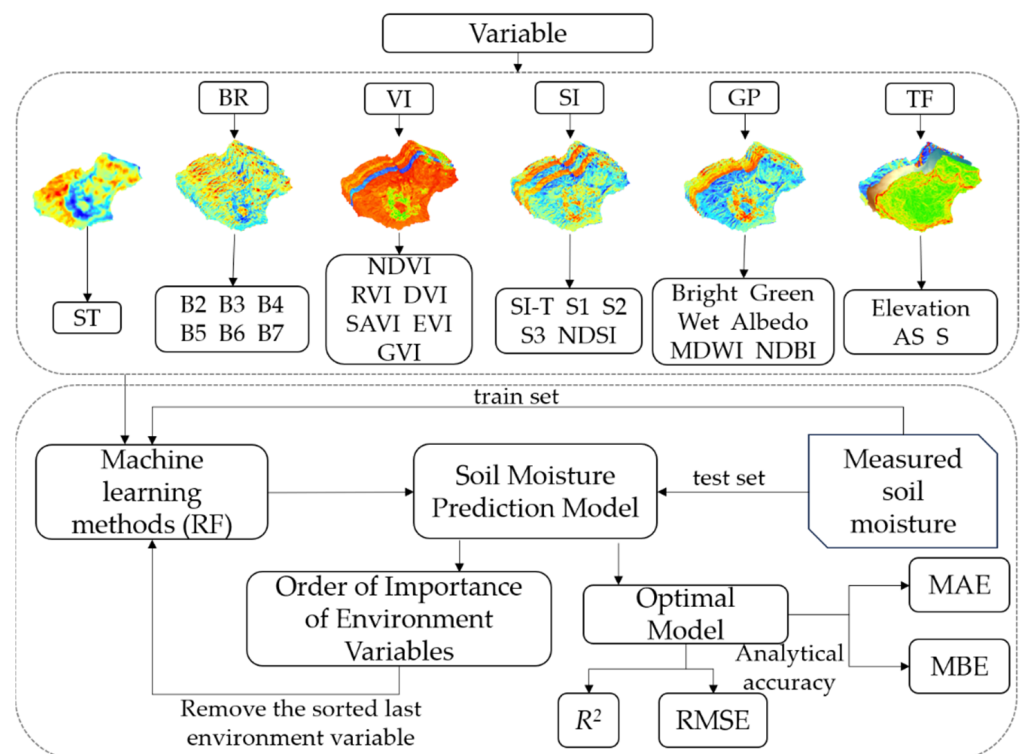


Figure 2. Workflow of variable optimization and model construction.

3.4. Modeling Methodology

3.4.1. Random Forest (RF) Model

The RF model is a multi-decision tree learning method based on categorical regression trees [21]. The classification and regression tree (CART) is used as a weak learner by the bootstrap resampling technique. First, n ($N > n$) samples from the original training sample set N are repeatedly sampled randomly to generate a new training sample set, and m new training sample sets are generated by repeated sampling [21]. A new collection of training samples is used to train decision trees to form a decision forest RF (Yao et al., 2022). Finally, the model is validated against the remaining $N-n$ samples as a test sample set [21]. Through iterative calculations and comparisons, in this study, m was set to 1500, and the number of training sample sets n was chosen to be 80% of the total samples.

3.4.2. Support Vector Machine (SVM) Model

The SVM model is based on classification boundaries, which can be extended for regression analysis [34]. Based on the principle of structural risk minimization, SVM uses the kernel function mapping method to map the points in low-dimensional to high dimensional space through a nonlinear transformation so that the original nonlinear data become linearly divisible. A linear model is established to solve the nonlinear problem further, and the result overcomes the issues of “many discrete values” and “over-learning” to a large extent [34]. The SVM in this study is computed using the radial base function (RBF) kernel function, which is formulated as [34]

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (5)$$

where x is the unknown vector, x' is the support vector, and σ is the function width.

3.4.3. Back Propagation Neural Network (BPNN) Model

The BPNN model is a widely used neural network model [21]. The model's topology consists of three main layers: an input layer, a hidden layer, and an output layer. Among them, the input layer is the whole set of environmental variables (six band reflectance, five salinity indices, six vegetation indices, six spectral indices, one surface temperature layer, and three surface topographic parameters) [21]. The number of hidden layers was set to 5, and the activation function used is the hyperbolic tangent. The output layer is soil moisture.

3.4.4. Combined Model

In this study, the traditional ordinary least squares (OLS) method was used to construct the combined model, and the predicted values of soil moisture obtained through the RF model, the SVM model, and the BPNN model were treated as the input variables to invert the soil moisture. The combinatorial model was calculated as follows [21]:

$$SMC = \sum_{i=1}^{\gamma} (\alpha_i \times SMC_i) + \beta \quad (6)$$

where SMC is the soil moisture content calculated by the combined model in %; α_i is the regression coefficient of the i th soil moisture model; SMC_i is the predicted value of the i th soil moisture model; β is the intercept; and γ takes the value of three.

3.5. Model Construction and Validation

In this study, 200 in surface (0–10 cm) soil moisture samples (29 April–31 May 2023) from the Xijitan landslide area and 200 in surface (0–10 cm) soil moisture samples (14–16 May 2023) from the Xiazangtan landslide area were randomly divided into two datasets, the training set and the test set, at a ratio of 80% (160 sample points) versus 20% (40 sample points). The former was used for training the machine learning models, and

the latter was used to test the accuracy of the constructed soil moisture inversion models. In addition, the accuracy of the inversion models was judged by the following statistical parameters: the coefficient of determination (R^2), the root mean square error (RMSE), the mean absolute error (MAE), and the mean bias error (MBE). The more negligible the RMSE, the smaller the MAE, the smaller the absolute value of MBE, and the larger the R^2 , the more accurate the soil moisture inversion model [21].

4. Results

4.1. Correlation between Environmental Variables and Surface Soil Moisture

In this study, 6 bands of reflectance (B2, B3, B4, B5, B6, B7) and 21 spectral indices derived from the Landsat 9 satellite data were selected to form the independent variable set of soil moisture. The correlation coefficients between the group of independent variables and soil moisture in the two landslide distribution areas are shown in Figure 3. The significance of the independent variable set and soil moisture was tested by using the F-test method. From Figure 3, it can be seen that the significance level of 0.05 was reached when the F value was more significant than the critical value (f) of 1.262, and the significance level of 0.01 was called when it was more important than the critical value (f) of 1.391. In addition, the two landslide distribution areas have significant differences in surface plants, soil physical characteristics, and surface topographic parameters. Hence, the correlation coefficients R between the environmental variables and surface soil moisture in these areas also differ from each other considerably. This difference is mainly manifested in the following way: the correlation between B2 ($R = -0.35$), B3 ($R = -0.35$), B4 ($R = 0.56$), B5 ($R = 0.57$), and SMC all passed the significance test ($p < 0.01$) in Xijitan, while five bands of reflectance, such as B2, B3, B5, B6, and B7, were significantly negatively correlated with soil moisture in Xiazangtan ($R = -0.302, -0.277, -0.291, -0.270, -0.280$, respectively).

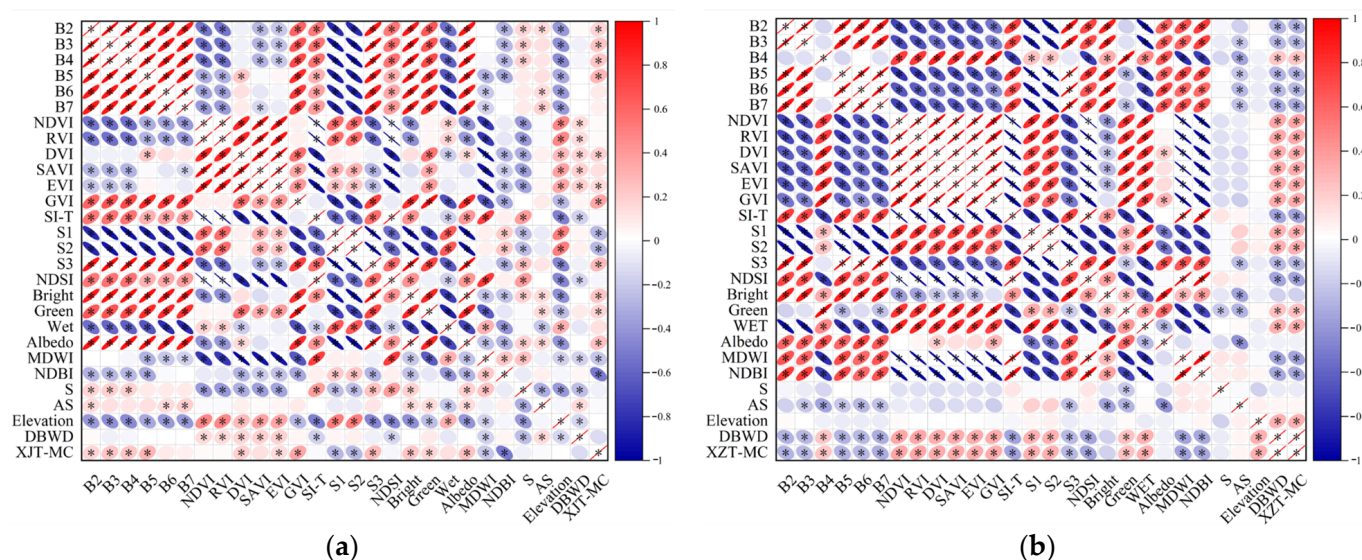


Figure 3. Results of correlation coefficient analysis between environmental variables and surface soil moisture in the Xijitan (a) and Xiazangtan (b) landslide distribution areas. Note: “*” in the graph indicates significance $p > 0.05$.

Further analyses revealed that there was an insignificant positive correlation between soil moisture and three vegetation indices (NDVI, RVI, and SAVI) in the Xiazangtan area, with the correlation coefficients being 0.021, 0.023, and 0.121, respectively, whereas there was a significant positive correlation between soil moisture and six vegetation indices (NDVI, RVI, DVI, SAVI, and EVI) in the Xiazangtan area. GVI and the 6 vegetation indices were all significantly and positively correlated with soil moisture, with the correlation coefficients being 0.346, 0.347, 0.325, 0.332, 0.327, and 0.321, respectively. Further analyses

yielded a significant positive correlation between soil moisture and the humidity index in Xiazangtan, with a correlation coefficient of 0.174. The correlation coefficient between soil moisture and the greenness index is 0.306. Through the above analyses, it can be seen that the correlation between the environmental variables such as vegetation indices, salinity indices, and soil moisture in the two different landslide distribution areas has a big difference between them; therefore, to improve the accuracy of the model further, it is necessary to carry out variable selection of the environmental variables separately for the two landslide distribution areas.

4.2. Importance of Environmental Variables and Optimal Independent Variable Sets

4.2.1. Importance of Environmental Variables

In this study, the importance of environmental variables in the two areas of Xijitan and Xiazangtan is shown in Figure 4. The most essential variables in Xijitan were B5 (importance = 0.181), GVI (importance = 0.340), S1 (importance = 0.345), NDBI (importance = 0.643), and elevation (importance = 0.088). Correspondingly, the most important variables in the Xiazangtan landslide distribution area were B2, NDVI, NDS1, Wet, and elevation, whose importance is 0.296, 0.196, 0.159, 0.252, and 0.545, respectively.

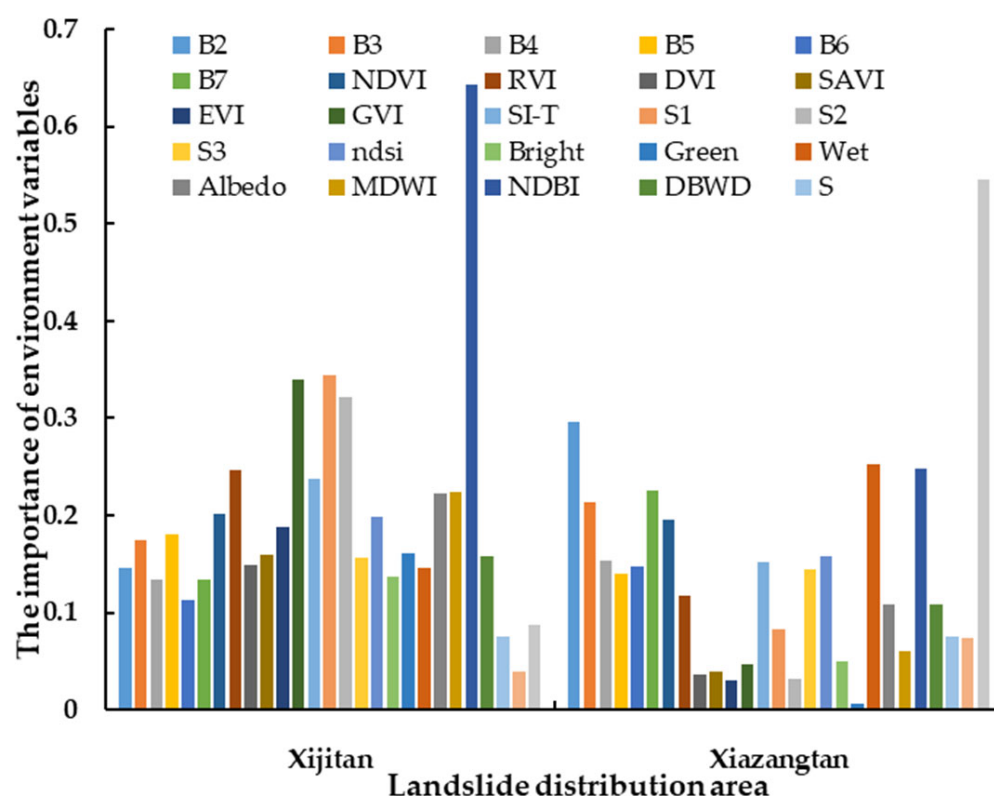


Figure 4. Importance analysis of environmental variables in the two different landslide distribution areas of Xijitan and Xiazangtan.

Therefore, this study was based on the RF algorithm to explore the optimal combination of independent variables for SMC inversion in the two areas (Table 2). Among them, the criteria for the division of variable combinations were mainly based on the importance of individual variables and the variety of each variable group with each other, and a total of 20 combinations were attempted.

Table 2. Results of the best combination of independent variables for soil moisture inversion models in two different landslide distribution areas.

Combination Criteria	Combined Serial Number	Variable Combinations	Preferred Environment Variables	
			Xijitan	Xiazangtan
All variable	1	BR + VI	B2, B3, B4, B5, B6, B7, NDVI, RVI, DVI, SAVI, EVI, GVI, DBWD	
	2	BR + VI + SI	B2, B3, B4, B5, B6, B7, NDVI, RVI, DVI, SAVI, EVI, GVI, SI-T, S1, S2, S3, NDSI, DBWD	
	3	BR + VI + SI + GP	B2, B3, B4, B5, B6, B7, NDVI, RVI, DVI, SAVI, EVI, GVI, SI-T, S1, S2, S3, NDSI, Bright, Green, Wet, Albedo, MDWI, NDBI, DBWD	
	4	BR + VI + SI + GP + TF	B2, B3, B4, B5, B6, B7, NDVI, RVI, DVI, SAVI, EVI, GVI, SI-T, S1, S2, S3, NDSI, Bright, Green, Wet, Albedo, MDWI, NDBI, S, AS, ELEVATION, DBWD	
Top 1 in importance	5	BR + VI	B5, GVI, DBWD	B2, NDVI, DBWD
	6	BR + VI + SI	B5, GVI, S1DBWD	B2, NDVI, NDSI, DBWD
	7	BR + VI + SI + GP	B5, GVI, S1, NDBI, DBWD	B2, NDVI, NDSI, WET, DBWD
	8	BR + VI + SI + GP + TF	B5, GVI, S1, NDBI, E, DBWD	B2, NDVI, NDSI, WET, E, DBWD
Top 2 in importance	9	BR + VI	B5, B3, GVI, RVI, DBWD	B2, B7, NDVI, RVI, DBWD
	10	BR + VI + SI	B5, B3, GVI, RVI, S1, S2, DBWD	B2, B7, NDVI, RVI, NDSI, SI-T, DBWD
	11	BR + VI + SI + GP	B5, B3, GVIRVI, S1, S2, NDBI, MDWI, DBWD	B2, B7, NDVI, RVI, NDSI, SI-T, WET, NDBI, DBWD
	12	BR + VI + SI + GP + TF	B5, B3, GVI, RVI, S1, S2, NDBI, MDWI, E, S, DBWD	B2, B7, NDVI, RVI, NDSI, SI-T, WET, NDBI, E, AS, DBWD
Top 3 in importance	13	BR + VI	B5, B3, B4, GVI, RVI, NDVI, DBWD	B2, B7, B3, NDVI, RVI, GVI, DBWD
	14	BR + VI + SI	B5, B3, B4, GVI, RVI, NDVI, S1, S2, SI-T, DBWD	b2, b7, b3, NDVI, RVI, GVI, NDSI, SI-T, S3, DBWD
	15	BR + VI + SI + GP	B5, B3, B4, GVI, RVI, NDVI, S1, S2, SI-T, NDBI, MDWI, Albedo, DBWD	B2, B7, B3, NDVI, RVI, GVI, NDSI, SI-T, S3, Wet, NDBI, Albedo, DBWD
	16	BR + VI + SI + GP + TF	B5, B3, B4, GVI, RVI, NDVI, S1, S2, SI-T, NDBI, MDWI, Albedo, E, S, AS, DBWD	B2, B7, B3, NDVI, RVI, GVI, NDSI, SI-T, S3, Wet, NDBI, Albedo, E, AS, S, DBWD
Top 4 in importance	17	BR + VI	B5, B3, B4, B2, GVI, RVI, NDVI, EVI, DBWD	B2, B7, B3, B4, NDVI, RVI, GVI, DVI, DBWD
	18	BR + VI + SI	B5, B3, B4, B2, GVI, RVI, NDVI, EVI, S1, S2, SI-T, NDSI, DBWD	B2, B7, B3, B4, NDVI, RVI, GVI, DVI, NDSI, SI-T, S3, S1, DBWD
	19	BR + VI + SI + GP	B5, B3, B4, B2, GVI, RVI, NDVI, EVI, S1, S2, SI-T, NDSI, NDBI, MDWI, Albedo, Green, DBWD	B2, B7, B3, B4, NDVI, RVI, GVI, DVI, NDSI, SI-T, S3, S1, WET, NDBI, Albedo, MDWI, DBWD
	20	BR+VI+SI+GP+TF	B5, B3, B4, B2, GVI, RVI, NDVI, EVI, S1, S2, SI-T, NDSI, NDBI, MDWI, Albedo, Green, E, S, AS, DBWD	B2, B7, B3, B4, NDVI, RVI, GVI, DVI, NDSI, SI-T, S3, S1, WET, NDBI, Albedo, MDWI, E, AS, S, DBWD

4.2.2. Screening for Optimal Combinations of Independent Variables

In this study, six types of variables (BR, VI, SI, GP, ST, and TF) were used as the independent variables input to the RF model to predict soil moisture, the target variable. The RF prediction models were established differently in the two areas of Xiazangtan and Xijitan. The results of the prediction accuracy of the models with other variables are shown in Figure 5. As can be seen from Figure 5, the full-variable inversion scheme includes scenarios 1, 2, 3, and 4, with scheme 3 having the relatively highest overall accuracy. Its accuracy indicators are R^2 of 0.780, RMSE of 1.536%, MAE of 0.746%, and MBE of -0.001% with the training dataset, and with the validation set, they became R^2 of 0.531, RMSE of 1.812%, MAE of 0.882%, and MBE of -0.045% . Further analyses indicate that among the

20 optimal independent variable combinations, scheme 15 has the highest overall accuracy, in that the RF model has an R^2 of 0.825, RMSE of 1.386%, MAE of 0.708%, and MBE of -0.013% with the training dataset. They represent an improvement of 25.762% in R^2 and a decrease of 25.564% in RMSE. With the test set, the accuracy indicators change to an R^2 of 0.769, an RMSE of 1.744%, an MAE of 0.886%, and an MBE of -0.204% , or an improvement of 36.833% in R^2 and a decrease of 16.953% in RMSE compared to the pre-screening scenario (i.e., scenario 4). This is mainly attributed to the interactions between the variables in the joint inversion of the RF prediction model for soil moisture in the two areas [24].

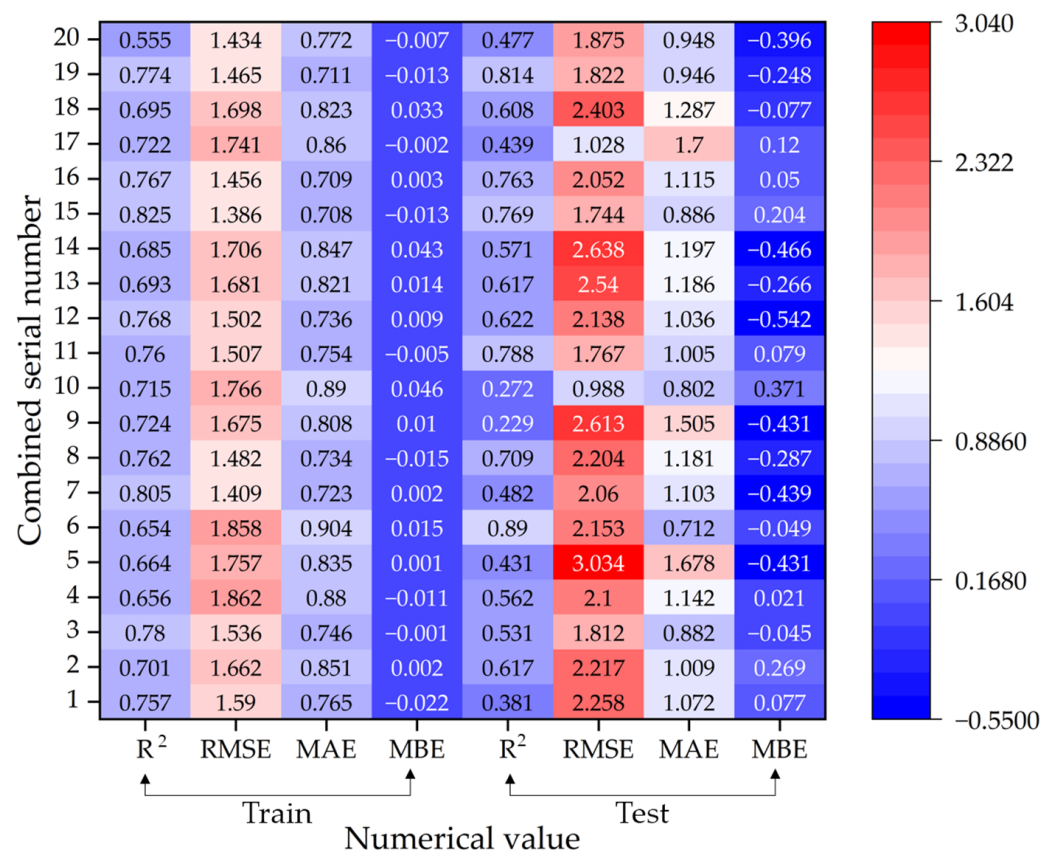


Figure 5. Optimal independent variable combinations and their accuracy validation results in the distribution area of the Xijitan landslide area. Note: In the table, R^2 is the coefficient of determination; RMSE is the root mean square error; MAE is the mean absolute error; MBE is the mean bias error.

In addition, as shown in Figure 6, among the 20 sets of optimal independent variable combinations, scenario 20 has the highest relative overall accuracy. The RF model has an R^2 of 0.690, an RMSE of 1.965%, an MAE of 1.526%, and an MBE of -0.015% with the training dataset. They change to an R^2 of 0.537, RMSE of 2.236%, MAE of 1.649%, and MBE of -0.521% with the test dataset. Through further analysis, it is found that scenario 20 has its overall accuracy improved by 6.646% in terms of R^2 and reduced RMSE by 23.720% for the training set, and improved by 15.456% in terms of R^2 and by 5.614% in terms of RMSE reduction for the testing set in comparison with the pre-screening scenario (i.e., scenario 4). Therefore, the preferred important spectral reflectance variables are B2, B7, B3, and B4; the preferred important vegetation indices include NDVI, RVI, GVI, and DVI; and the important salinity indices are NDSI, SI-T, S3, and S1.

In summary, the screening of the preferred variables for the best combination of independent variables for the soil moisture inversion model showed that there were significant differences between the two areas of Xijitan and Xiazangtan. Because Xijitan has a relatively low vegetation cover, and the addition of the VI joint inversion (scenarios 2, 6, and 10) demonstrated a relatively small improvement in the overall accuracy of the

RF model. In contrast, the Xiazangtan area has a relatively high vegetation cover, and the inclusion of the vegetation indices joint inversion (scenarios 2, 6, and 10) demonstrated a fairly significant improvement in the overall accuracy of the RF model.

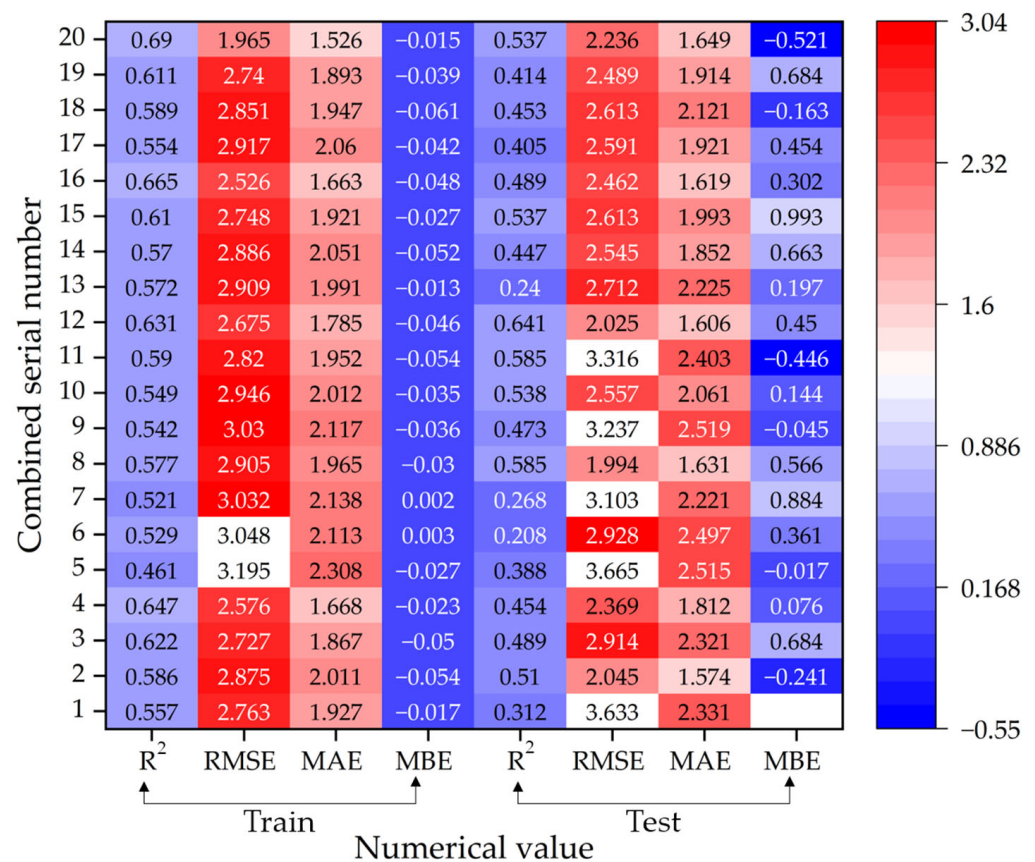


Figure 6. Optimal independent variable combinations and their accuracy validation results in the distribution area of the Xiazangtan landslide area. Note: R^2 , RMSE, MAE, and MBE in the table are the same as in Table 2.

4.3. Predicted Soil Moisture in Different Areas

Figure 7 comprehensively compares the accuracy of the RF model validation set for soil moisture in Xijitan and Xiazangtan. The predicted and measured soil moisture values in Xijitan are generally low and have low fluctuations. In contrast, the predicted and measured values of soil moisture in the Xiazangtan area are relatively high and have significant instability. Through further analyses, the RF model is found to be more accurate in the Xijitan area than in the Xiazangtan area. As can be seen from Figure 8a, the spatial distribution of soil moisture in Xijitan showed a decreasing and then increasing trend from the south to the north. The areas of high soil moisture in the Xijitan are mainly located in the relatively flat terrain.

In addition, it can be seen from Figure 8b that soil moisture is relatively high at the back edge and the center of the landslide in Xiazangtan, with a value of 10–15%. In contrast, soil moisture at the front edge of the landslide is relatively low, with a value ranging from 7 to 9%. It should be noted that the field sampling time of this study is May 2023, which is the growing season of wheat, during which the watering of farmland is more frequent, so the highest soil moisture in the Xiazangtan landslide area is mainly distributed in farmland. In addition, the soil moisture results predicted from the RF model for the two landslide distribution areas were consistent with the actual conditions observed in the field. They showed similar characteristics to the field-measured soil moisture results.

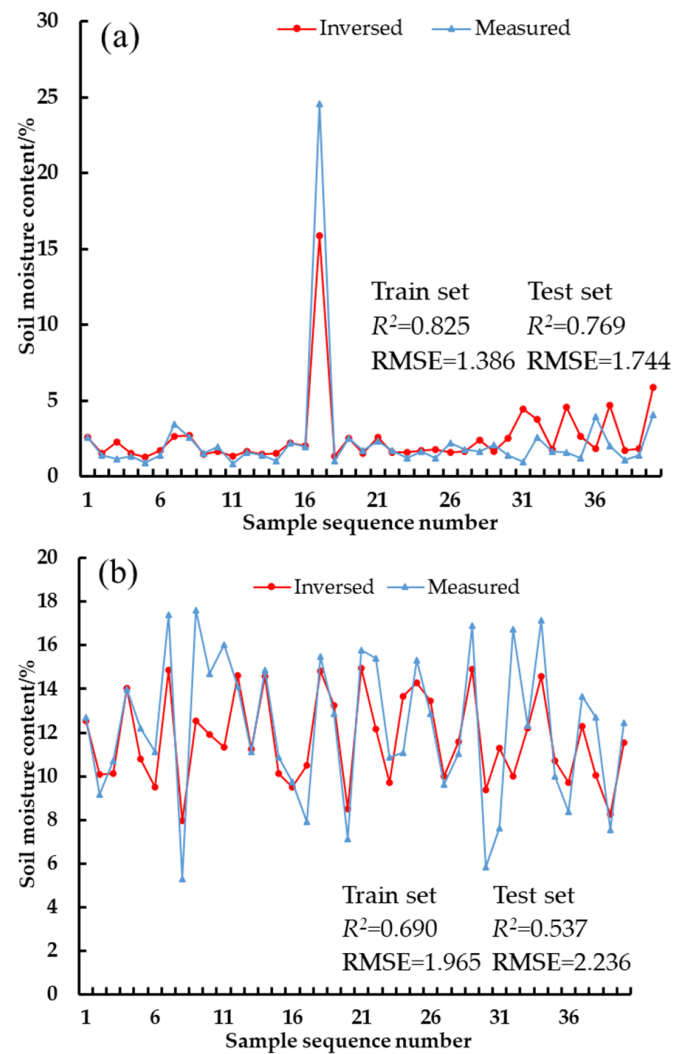


Figure 7. Comprehensive comparison results of the accuracy of the validation set of soil moisture inversion models in two different landslide distribution areas: (a) Xijitan landslide distribution area, (b) Xiazangtan landslide distribution area.

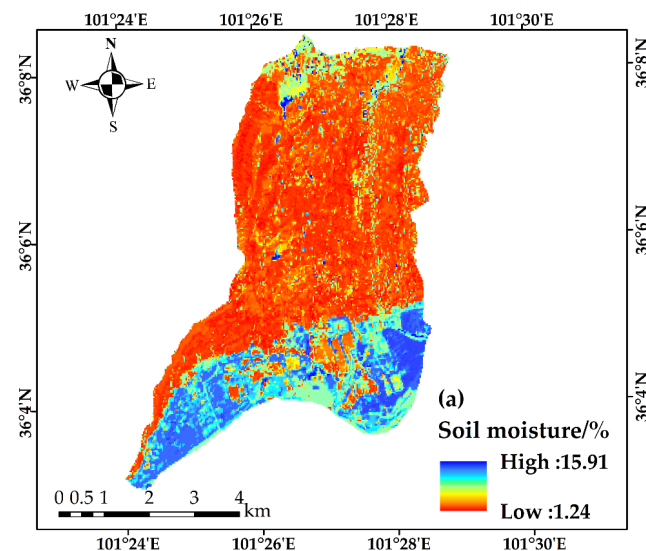


Figure 8. Cont.

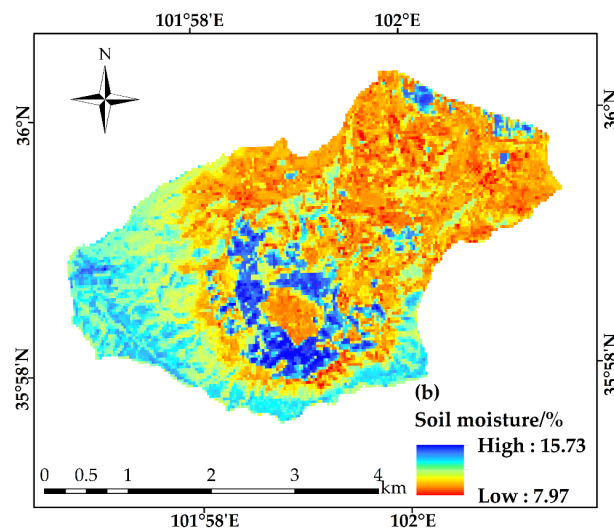


Figure 8. Characteristics of the spatial distribution of surface soil moisture in two different landslide distribution areas in China: (a) Xijitan landslide distribution area, (b) Xiazangtan landslide distribution area.

4.4. Model Construction and Validation of Different Methods

The results of a comprehensive comparison of the predicted and measured soil moisture values obtained from the validation set of the four different soil moisture inversion models are shown in Figure 9. As can be seen in Figure 9, the accuracy of the RF model is relatively low, with a training accuracy of R^2 of 0.656, RMSE of 1.862%, MAE of 0.880%, and MBE of -0.011% and test accuracy of R^2 of 0.562, RMSE of 2.100%, MAE of 1.142%, and MBE of 0.021%. Further analysis shows that the SVM model is more accurate in inverting soil moisture than the RF model, with the latter having an R^2 of 0.956, RMSE of 0.655%, MAE of 0.353%, and MBE of -0.071% for the training set and R^2 of 0.833, RMSE of 1.375%, MAE of 0.980%, and MBE of -0.075% for the test dataset. Furthermore, comparative analysis indicates that the BPNN model has relatively higher accuracy in inverting soil moisture than both the RF and SVM models, with the training set accuracy being R^2 of 0.922, RMSE of 0.952%, MAE of 0.651%, and MBE of -0.082% and the test set accuracy being R^2 of 0.809, RMSE of 0.875%, MAE of 0.695%, and MBE of -0.228% . As can be seen from Figure 9 and Table 3, the soil moisture inversion results obtained from the combined model showed good applicability and the highest estimation accuracy, i.e., the model had an R^2 of 0.886, an RMSE of 1.085%, an MAE of 0.690%, and an MBE of -0.001% with the training dataset, and an R^2 of 0.916%, an RMSE of 0.877%, an MAE of 0.620%, and an MBE of 0.154% with the test dataset.

Table 3. Comprehensive comparison of the accuracy of the training set and test set of different modeling methods in the distribution area of the Xijitan landslide.

Modelling Methodology	Train				Test			
	R^2	RMSE	MAE	MBE	R^2	RMSE	MAE	MBE
RF	0.656	1.862	0.880	-0.011	0.562	2.100	1.142	0.021
SVM	0.956	0.655	0.353	-0.071	0.833	1.375	0.980	-0.075
BPNN	0.922	0.952	0.651	-0.082	0.809	0.875	0.695	-0.228
Combined model	0.886	1.085	0.690	-0.001	0.916	0.877	0.620	0.154

Note: R^2 , RMSE, MAE, and MBE in the table are the same as in Table 2.

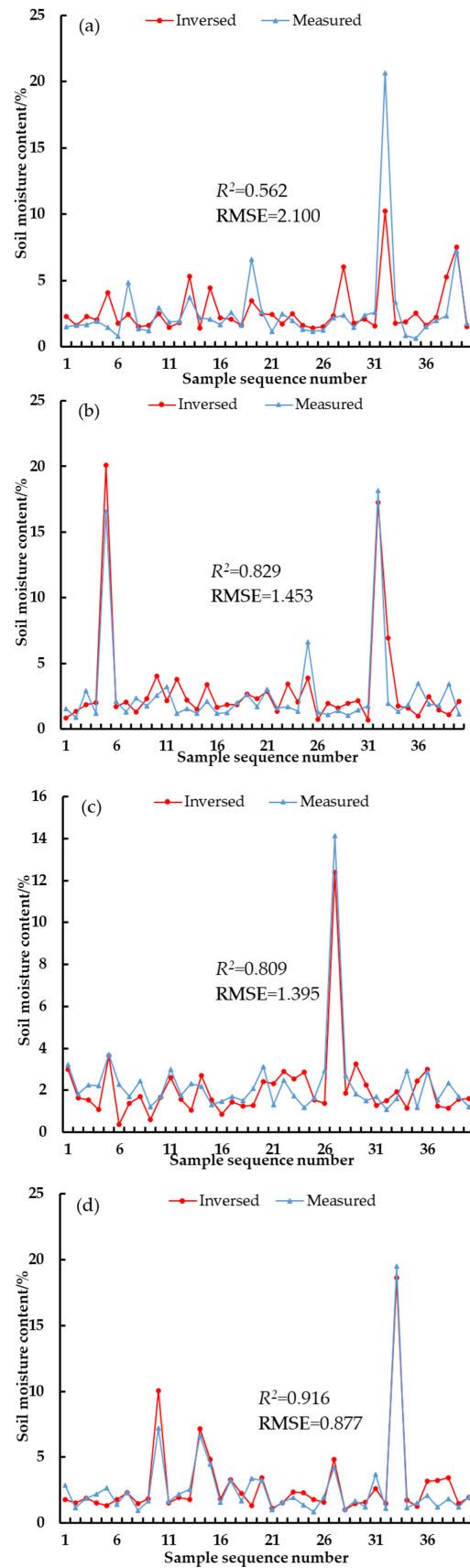


Figure 9. Comprehensive comparison results of the accuracy of different model validation sets in the distribution area of the Xijitan landslide: (a) RF model, (b) SVM model, (c) BPNN model, (d) combined model.

In summary, among the four machine learning models compared in this study, the RF model has a relatively large RMSE between the predicted and measured soil moisture values in areas of high water content while the BPNN model, the SVM model, and the combined model all have a relatively small RMSE between the predicted and measured values. Moreover, the combined model has a higher accuracy than the individual models, and the accuracy of the BPNN model was also relatively high among the three single models. The main reason is that the combined model has the combined advantages of all three single models.

4.5. Spatial Distribution of Soil Moisture

It should be noted that the soil moisture results obtained from the inversion using the four models are consistent with the field measurements, as shown in Figure 10. The soil moisture is relatively high (about 10–15%) in areas such as the back edge of the landslide and Xijitan Village in the distribution area of Xijitan landslide. Soil moisture is relatively low (approximately 0.5–3%) in the leading edge and the middle of the landslide. The soil moisture is below 10% in most (more than 90%) of the total area. In addition, the results obtained by using different models in the distribution area of the Xijitan landslide also showed some differences. The soil moisture obtained from the RF model was 1.35 to 11.43%, while the soil moisture obtained from the BPNN model, SVM model, and the combined model was 0.76 to 23.90%, 0.83 to 30.33%, and 0.23 to 27.01%, respectively. In addition, in Figure 10b, there are only a few areas of high moisture (approximately 10.02–23.90%) in the northern part of Xijitan. Most areas in the south have a low soil moisture of approximately 0.76–3.73%. However, in Figure 10a, the soil moisture is higher (approximately 5.87–8.14%) in much of the south. Furthermore, in Figure 10c, there are large areas of both higher (approximately 10.19–18.37%) and lower (approximately 1.43–7.15%) soil moisture in the southern region. In Figure 10d, most areas in the south have a low moisture level just above the minimum value of 0.23%.

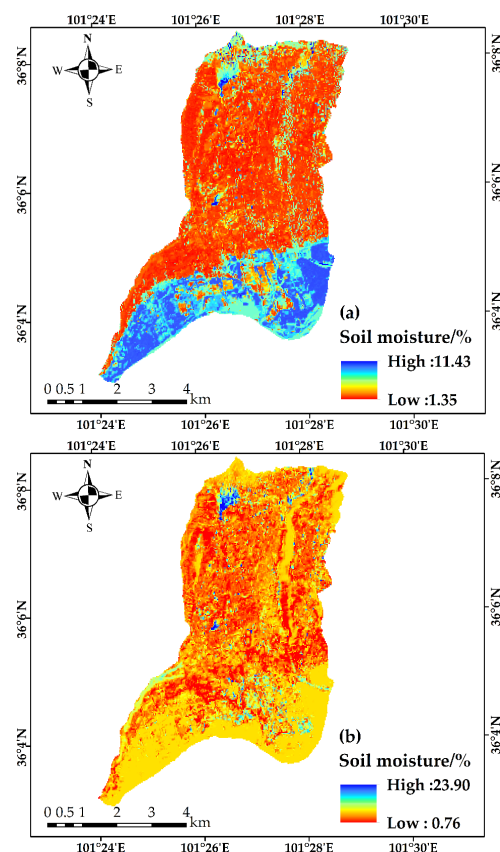


Figure 10. Cont.

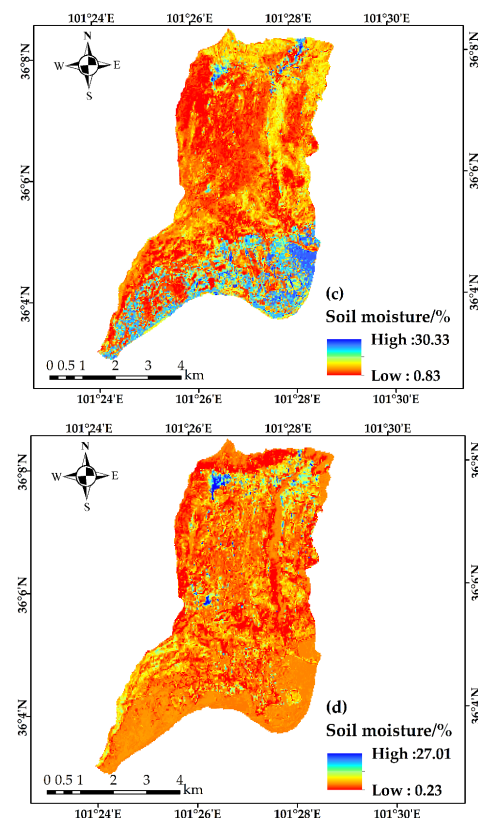


Figure 10. Characteristics of the spatial distribution of soil moisture in the distribution area of Xijitan landslide: (a) RF model, (b) SVM model, (c) BPNN model, (d) combined model.

In summary, the differences in the spatial distribution of soil moisture based on the four model inversions were mainly concentrated in the southern part of the study area, while the spatial distribution of soil moisture in the other regions was similar. The soil moisture in the southern part of the Split River was relatively high in the random forest model. The soil moisture in the southern part of the Xijitan Beach inverted by the SVM model, the BP neural network model, and the combined model was relatively low.

5. Discussion

It has been shown that under vegetative cover, changes in soil moisture affect plant growth, leading to significant changes in plant growth conditions and vigor [35]. At the same time, a correlation is found between the root biomass of plants and soil moisture [9]. Therefore, different types of vegetation (e.g., grasses, shrubs, trees) have different impacts on soil moisture due to their different root lengths and root biomass [9]. In this study, however, only vegetation cover was considered in the selection of input variables, and the correlation between different types of vegetation and soil moisture was not fully considered. In addition, the study area is located in an arid and semi-arid zone where rainfall is much lower than evaporation [35]. The decrease in soil moisture leads to precipitation and accumulation of medium soluble salts in the soil, which makes it possible to indirectly assess soil moisture by interpreting the changing conditions of soil salinity through satellite spectral images [9,36]. However, the vegetation indices and salt indices have completely opposite behavior because the data used are spectral [19]. The high vegetation coverage affects the interpretation of the salt indices to some extent [37]. Therefore, in this study, to investigate the differences in the importance of the vegetation index and soil salinity index for soil moisture inversion modeling in areas with different vegetation cover and vegetation types, Xiazangtan and Xijitan were selected as the study areas. The vegetation in the Xiazangtan region comprises predominantly trees, grasses, and shrubs, with a dense cover and scarce bare areas. The vegetation in the Xijitan region is predominantly grasses

and shrubs with sparse coverage, numerous bare patches, and high soil salinity. The vegetation coverage in the Xiazangtan landslide distribution area is relatively high, and the results also show that the vegetation indices are more important than the salinity indices in the constructed model for this area, with importance values of 0.196 and 0.152. The vegetation coverage in the Xijitan landslide distribution area is relatively low, and the results also show that the importance of the vegetation indices is less than that of the salt indices in the constructed model for this area, with the importance values of 0.202 and 0.237. A single environmental variable cannot be applied to both sites due to the significant differences in plant distribution, soil physical characteristics, and topographic parameters between Xijitan and Xiazangtan. Therefore, the effects of multiple environmental variables on soil moisture should be considered simultaneously in the selection of environmental variables. At the same time, screening the best independent variables according to the environmental characteristics of the study area can effectively improve the inversion accuracy of the soil moisture model.

Other soil properties (e.g., soil porosity, bulk density, and soil organic matter) were not considered in the selection of environmental variables in this study, which may lead to uncertainty in the inversion results. In addition, although 27 different environmental variables were considered simultaneously in this study, the interactions that exist between the variables were not considered. As some of the variables do not always guarantee better modeling results, when new variables are added, the new variables may have a negative impact on the other variables. This is consistent with the results when comparing the effects of different groups of variables on the accuracy of soil moisture predictions. The inversion accuracy was not as good as the spectral index group when the full group of variables was entered, which could be due to overfitting caused by the introduction of too many independent variables. In spite of the best set of environmental variables and the best prediction model used, the inversion accuracy of soil moisture reached the maximum R^2 of only 91.6%. This is because various optical remote sensing indices are predominantly representative of above-surface plant, soil physical, and topographic parameters. However, the soil moisture samples acquired for this study were mostly collected from 0–10 cm below the surface. As a result, the remote-sensing-inverted and the in situ sampled soil moisture results are not exactly identical to each other.

Due to the differences in satellite and sensor characteristics and the applicability of element-specific inversion algorithms (e.g., optical satellite data are only applicable to clear sky conditions), it is difficult to model the accuracy requirements using only data from a single satellite platform as an environmental element. The current study has some shortcomings in this area and only considers data from a single satellite platform in the data selection. Second, the discrepancy between the inverted and observed soil moisture can be attributed to the imprecise match between the satellite transit time and the field soil moisture data sampling time. The transit time of the satellites is short, and the sampling of the soil moisture data in the field often takes 2–3 days to complete. Climate changes (e.g., rainfall, air temperature) during this period may affect the water content of soil samples to some extent. For example, in the Xijitan landslide area, there was no rainfall and no significant change in temperature during the sampling period, so the accuracy of the model is higher. However, in the Xiazangtan landslide area, there was some rainfall during the sampling period, lowering the accuracy of the model in this area.

Finally, it should be noted that there are still some deficiencies in this study. First, only optical data were considered in the selection of input variables. In contrast, microwave data and the effect of climatic factors on soil moisture have not yet been considered. Microwaves can penetrate deeper into soil, vegetation, and clouds than optical data, which makes the remotely sensed moisture more closely match the in situ collected soil moisture within the depth of 0–10 cm, not just the surface soil moisture as inverted from optical remote sensing data. The surface soil moisture is more volatile with radiation and timing of sensing, lowering the inversion accuracy. Second, the optimal combination of independent variables designed in the optimal independent variable screening is still relatively homogeneous.

Therefore, the stability of the soil moisture inversion model for the two areas of Xiazangtan and Xijitan and the improvement of the inversion accuracy need further in-depth study.

6. Conclusions

This study maps the spatial distribution of soil moisture in the areas of Xijitan and Xiazangtan based mainly on Landsat-9-derived variables and compares four machine learning methods in predicting soil moisture in the Xijitan landslide area, including the RF model, SVM model, BPNN model, and combined model. It is concluded that:

- (1) The importance of variables such as SR, VI, and TF is greater than SI and GP in inverting soil moisture in the Xiazangtan landslide area. The importance of variables such as SI and GP is greater than that of SR, VI, and TF in the soil moisture inversion model in the Xijitan landslide area. In addition, surface temperatures are more important in the Xijitan landslide area than in the Xiazangtan landslide area.
- (2) The accuracy of the soil moisture prediction model in the Xijitan landslide distribution area was 36.833% higher in R^2 and 16.953% lower in RMSE in the model test set after screening the optimal independent variable combinations compared with that of the unscreened variables. The accuracy of the soil moisture prediction model for the Xiazangtan landslide distribution area increased by 15.456%, and the RMSE decreased by 5.614% for the model test set after screening with the optimal combination of independent variables compared to the unscreened variables.
- (3) The combined model showed the best relative applicability among the four soil moisture inversion models. The test set R^2 of the integrated model was 0.916, and the RMSE was 0.877%. In addition, the overall accuracy of the BPNN model is also higher than the other two individual models. It had the highest R^2 of 0.809% and an RMSE of 0.875% with the test dataset.

Author Contributions: W.L. is the co-first author; X.H. is the corresponding author. Data curation, formal analysis, and writing—original draft, W.L.; writing—review and editing, and supervision, X.H. and X.L.; investigation, validation, and supervision, J.Z., C.L., S.L., G.L. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The project was financially supported by the National Natural Science Foundation of China (grant no. 42041006), the Natural Science Foundation of Qinghai Province (2020-ZJ-906), and Discipline Innovation and Introducing Talents Program of Higher Education Institutions, the 111 Project of China (D18013).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Seneviratne, S.I.; Corti, T.; Davin, E.L.; Hirschi, M.; Jaeger, E.B.; Lehner, I.; Orlowsky, B.; Teuling, A.J. Investigating Soil Moisture–Climate Interactions in a Changing Climate: A Review. *Earth-Sci. Rev.* **2010**, *99*, 125–161. [\[CrossRef\]](#)
2. Vereecken, H.; Huisman, J.A.; Pachepsky, Y.; Montzka, C.; van der Kruk, J.; Bogen, H.; Weihermüller, L.; Herbst, M.; Martinez, G.; Vanderborght, J. On the Spatio-Temporal Dynamics of Soil Moisture at the Field Scale. *J. Hydrol.* **2014**, *516*, 76–96. [\[CrossRef\]](#)
3. Lei, L.; Zheng, J.; Li, S.; Yang, L.; Wang, W.; Zhang, F.; Zhang, B. Soil Hydrological Properties' Response to Long-Term Grazing on a Desert Steppe in Inner Mongolia. *Sustainability* **2023**, *15*, 16256. [\[CrossRef\]](#)
4. Zhang, Z.Y. *Rapid Measurement Method of Hydraulic Conductivity of Unsaturated Soil and Mechanism of Landslide Induced by Rainfall Infiltration*; Beijing Jiaotong University: Beijing, China, 2023.
5. Lagasio, M.; Pulvirenti, L.; Parodi, A.; Boni, G.; Rommen, B. Effect of the Ingestion in the WRF Model of Different Sentinel-Derived and GNSS-Derived Products: Analysis of the Forecasts of a High Impact Weather Event. *Eur. J. Remote Sens.* **2019**, *52*, 16–33. [\[CrossRef\]](#)
6. Mirus, B.B. HydroMet: A New Code for Automated Objective Optimization of Hydrometeorological Thresholds for Landslide Initiation. *Water* **2021**, *13*, 1752. [\[CrossRef\]](#)

7. Bronstert, A.; Creutzfeldt, B.; Graeff, T.; Hajnsek, I.; Heistermann, M.; Itzerott, S.; Jagdhuber, T.; Kneis, D.; Lück, E.; Reusser, D.; et al. Potentials and Constraints of Different Types of Soil Moisture Observations for Flood Simulations in Headwater Catchments. *Nat. Hazard.* **2012**, *60*, 879–914. [\[CrossRef\]](#)
8. Yin, Z.Q.; Cheng, G.M.; Hu, G.S.; Wei, G.; Wang, Y.Q. Preliminary study on characteristic and mechanism of super large landslide in upper Yellow River since late-pleistocene. *J. Eng. Geol.* **2010**, *8*, 41–51.
9. Wang, S.; Li, R.; Wu, Y.; Wang, W. Estimation of Surface Soil Moisture by Combining a Structural Equation Model and an Artificial Neural Network (SEM-ANN). *Sci. Total Environ.* **2023**, *876*, 162558. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Babaeian, E.; Sadeghi, M.; Jones, B.S. Ground, Proximal, and Satellite Remote Sensing of Soil Moisture. *Rev. Geophys.* **2019**, *57*, 530–616. [\[CrossRef\]](#)
11. Peng, J.; Albergel, C.; Balenzano, A.; Brocca, L.; Cartus, O.; Cosh, M.H.; Crow, W.T.; Dabrowska-Zielinska, K.; Dadson, S.; Davidson, M.W.J.; et al. A Roadmap for High-Resolution Satellite Soil Moisture Applications—Confronting Product Characteristics with User Requirements. *Remote Sens. Environ.* **2021**, *252*, 112162. [\[CrossRef\]](#)
12. Gruber, A.; De Lannoy, G.; Albergel, C.; Al-Yaari, A.; Brocca, L.; Calvet, J.-C.; Colliander, A.; Cosh, M.; Crow, W.; Dorigo, W.; et al. Validation Practices for Satellite Soil Moisture Retrievals: What Are (the) Errors? *Remote Sens. Environ.* **2020**, *244*, 111806. [\[CrossRef\]](#)
13. Pulvirenti, L.; Pierdicca, N.; Teuling, A.J. On the Potential of Sentinel-1 for Sub-Field Scale Soil Moisture Monitoring. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *120*, 103342.
14. Yinglan, A.; Guoqiang, W.; Peng, H.; Xiaoying, L.; Baolin, X.; Qingqing, F. Root-Zone Soil Moisture Estimation Based on Remote Sensing Data and Deep Learning. *Environ. Res.* **2022**, *212*, 113278.
15. Virnodkar, S.S.; Pachghare, V.K.; Patil, V.C.; Jha, S.K. Remote Sensing and Machine Learning for Crop Water Stress Determination in Various Crops: A Critical Review. *Precis. Agric.* **2020**, *21*, 1121–1155. [\[CrossRef\]](#)
16. García-Tejero, I.F.; Rubio, A.E.; Viñuela, I.; Hernández, A.; Gutiérrez-Gordillo, S.; Rodríguez-Pleguezuelo, C.R.; Durán-Zuazo, V.H. Thermal Imaging at Plant Level to Assess the Crop-Water Status in Almond Trees (Cv. Guara) under Deficit Irrigation Strategies. *Agric. Water Manag.* **2018**, *208*, 176–186. [\[CrossRef\]](#)
17. María, A.; Claudia, N.; Ángel, C.-B.M.; Miguel, A.L.; Jesús, Á.-M. Evaluation of Soil Moisture Estimation Techniques Based on Sentinel-1 Observations over Wheat Fields. *Agric. Water Manag.* **2023**, *287*, 108422.
18. Zhang, Z.; Bian, J.; Han, W.; Fu, Q.P.; Chen, S.B.; Cui, T. Cotton moisture stress diagnosis based on canopy temperature characteristics calculated from UAV thermal infrared image. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 77–84. [\[CrossRef\]](#)
19. Yu, Z.; Wenting, H.; Huihui, Z.; Xiaotao, N.; Guomin, S. Evaluating Soil Moisture Content under Maize Coverage Using UAV Multimodal Data by Machine Learning Algorithms. *J. Hydrol.* **2023**, *617*, 129086.
20. Hong, Q.; Sun, H.; Chen, Y.S. Comparisons and Classification System of Typical Remote Sensing Indexes for Agricultural Drought. *Trans. Chin. Soc. Agric. Eng.* **2012**, *28*, 147–154.
21. Guide County Local Records Compilation Committee. *Guide Yearbook (2021)*; Sanqin Press: Xi'an, China, 2021; Volume 12.
22. Hua, Q.C. Changes of Live—top Biomass, Plant Diversity and Soil Factors at Sunny and Shady Slope on Alpine Kobresia Meadow. *J. Grassl. Forage Sci.* **2024**, *4*, 22–25.
23. Lu, S.L.; Liu, S.W.; Wu, Z.L.; Ho, T.N.; Zhou, L.H.; Huang, R.F.; Pan, J.T.; Editorial Committee of Flora of China, Chinese Academy of Sciences. *Flora of China*; Science Press: Beijing, China, 1993; Volume 4, pp. 156–161.
24. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [\[CrossRef\]](#)
25. Yi, Q.X. Remote estimation of cotton LAI using Sentinel-2 multispectral data. *Trans. Chin. Soc. Agric. Eng.* **2019**, *35*, 189–197.
26. Yuxiang, Z.; Jiheng, H.; Dasa, G.; Haixu, B.; Yuyun, F.; Yipu, W.; Rui, L. Simulation of Isoprene Emission with Satellite Microwave Emissivity Difference Vegetation Index as Water Stress Factor in Southeastern China during 2008. *Remote. Sens.* **2022**, *14*, 1740.
27. Zhijun, Z.; Shengbo, C.; Tiangang, Y.; Eric, C.; Nicolas, L.; Jordan, G.; Michael, H.; Wenhan, Q.; Lisai, C.; Jian, L.; et al. Using the Negative Soil Adjustment Factor of Soil Adjusted Vegetation Index (SAVI) to Resist Saturation Effects and Estimate Leaf Area Index (LAI) in Dense Vegetation Areas. *Sensors* **2021**, *21*, 2115. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Qin, J.; Ma, M.; Shi, J.; Ma, S.; Wu, B.; Su, X.; Su, X. The Time-Lag Effect of Climate Factors on the Forest Enhanced Vegetation Index for Subtropical Humid Areas in China. *Int. J. Environ. Res. Public Health* **2023**, *20*, 799. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Allbed, A.; Kumar, L.; Aldakheel, Y.Y. Assessing Soil Salinity Using Soil Salinity and Vegetation Indices Derived from IKONOS High-Spatial Resolution Imageries: Applications in a Date Palm Dominated Region. *Geoderma* **2014**, *231*, 1–8. [\[CrossRef\]](#)
30. Liu, H.Q.; Huete, A. A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 457–465. [\[CrossRef\]](#)
31. Khan, N.M.; Rastoskuev, V.V.; Sato, Y.; Shiozawa, S. Assessment of Hydrosaline Land Degradation by Using a Simple Approach of Remote Sensing Indicators. *Agric. Water Manag.* **2005**, *77*, 96–109. [\[CrossRef\]](#)
32. Liu, H.J.; Wang, X.; Zhang, X.K.; Zhang, X.L.; Jin, H.N.; Dou, X. High spectral prediction model for soil moisture in songnen plain. *Chin. J. Soil Sci.* **2018**, *49*, 38–44.
33. Sobrino, J.A.; Jiménez-Muñoz, J.C.; Soria, G.; Romaguera, M.; Guanter, L.; Moreno, J.F.; Plaza, A.J.; Martínez, P. Land Surface Emissivity Retrieval from Different VNIR and TIR Sensors. *IEEE Trans Geosci. Remote Sens.* **2008**, *46*, 316–327. [\[CrossRef\]](#)
34. Yang, L.P.; Ren, J.; Wang, Y.; Zhang, J.; Wang, T.; Li, K. Soil Salinity Estimation Model in Juyanze Based on Multi-source Remote Sensing Data. *Tran. Chin. Soc. Agric. Mach.* **2022**, *53*, 226–235.

35. Krzeminska, D.; Bloem, E.; Starkloff, T.; Stolte, J. Combining FDR and ERT for monitoring soil moisture and temperature patterns in undulating terrain in south-eastern Norway. *Catena* **2022**, *212*, 106100. [[CrossRef](#)]
36. Wang, S.N.; Li, R.P.; Wu, Y.J.; Zhao, S.X.; Wang, X.Q. Multi-model comprehensive inversion of surface soil moisture based on model averaging method. *Trans. Chin. Soc. Agric. Eng.* **2022**, *38*, 87–94.
37. Sohrabinia, M.; Rack, W.; Zawar-Reza, P. Errata: Soil Moisture Derived Using Two Apparent Thermal Inertia Functions over Canterbury, New Zealand. *Remote Sens.* **2014**, *8*, 083624. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.