

Article

Multi-Source Remote Sensing for Soil Texture Mapping in an Arid Oasis of the Central Section of the Northern Tianshan Slope

Xin Ma ^{1,2}, Hongqi Wu ^{1,2,*}, Yanmin Fan ^{1,2,*}, Pengfei Wang ^{1,2}, Wenxian Ma ^{1,2}, Haobo Zhang ^{1,2} and Ruihong Zhong ^{1,2}

- ¹ College of Resources and Environment, Xinjiang Agricultural University, Urumqi 830052, China; 320243739@stu.xjau.edu.cn (X.M.); 320243749@stu.xjau.edu.cn (P.W.); 320243738@stu.xjau.edu.cn (W.M.); 320243759@stu.xjau.edu.cn (H.Z.); 320243762@stu.xjau.edu.cn (R.Z.)
² Xinjiang Engineering Technology Research Center of Soil Big Data, Urumqi 830052, China
* Correspondence: whq@xjau.edu.cn (H.W.); fanyym@xjau.edu.cn (Y.F.)

Abstract

Soil texture is essential for soil surveys, cropland quality assessment, and soil and water management. However, complex landforms and spatial heterogeneity challenge prediction in arid oases. Single-source data and conventional single-pixel models may inadequately capture spatial variation. This study combined field-collected soil sample data with multi-year remote sensing observations. A predictor system centered on multi-year bare-soil composites (BSCs) integrated additional remote sensing and terrain information. Following isometric log-ratio (ILR) transformation, random forest (RF), XGBoost, and Multi-Modal Vision Transformer (MMVT) were compared across spatial windows. Models were evaluated through repeated random partitions and five-fold spatial cross-validation. Grouped permutation importance and the dissimilarity index (DI) assessed modality contributions and spatial applicability, respectively. MMVT achieved the best overall performance using a 10×10 -pixel window (approximately 300×300 m). Across repeated random partitions, mean R^2 values were 0.586 ± 0.027 , 0.519 ± 0.029 , and 0.706 ± 0.026 . These corresponded to sand, silt, and clay, respectively. Corresponding RMSE values were 4.695%, 3.391%, and 3.335%, and RPIQ values were 2.107, 2.152, and 2.324. Five-fold spatial cross-validation yielded mean R^2 values of 0.550, 0.504, and 0.657, respectively. BSCs showed relatively high normalized grouped permutation importance in all three models. Its importance values were 62.68%, 61.75%, and 50.39% for RF, XGBoost, and MMVT, respectively. Mapping showed higher sand contents along the northern desert margin and greater clay enrichment in the western oasis plain. Approximately 85% of the valid mapped area fell within the area of applicability (AOA). Higher DI values occurred mainly in some urban built-up areas, low-mountain river valleys, and hilly transition zones. The findings suggest that combining multi-year bare-soil composites with spatial context at appropriate scales can enhance soil texture prediction. They may provide useful reference information for regional soil surveys, soil type identification, and land suitability analysis.

Academic Editor: Gian Franco Capra

Received: 17 September 2026

Revised: 19 October 2026

Accepted: 20 October 2026

Published: 22 October 2026

Copyright: © 2026 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/).

Keywords: soil texture; Vision Transformer; bare-soil composite; digital soil mapping; arid oasis

1. Introduction

Soil texture is determined by the relative contents of sand, silt, and clay. It is an important physical property that affects soil pore structure, soil and water conservation, and cultivation suitability [1,2]. As results from China's Third National Soil Survey become available, there is a growing need for continuous and detailed soil texture data. These data can improve soil resource maps, support soil type identification, and help cropland quality assessment [3]. Soil texture is especially important in arid oasis regions. These regions have low rainfall, strong evaporation, and a high dependence on irrigation for crop production [4]. Soil texture patterns in these regions are jointly influenced by piedmont alluvial proluvial processes, aeolian deposition, landform position, and oasis land use, resulting in strong spatial heterogeneity among different landform units. However, traditional field surveys are costly and have limited spatial coverage. They are therefore difficult to use for detailed mapping of local soil variation [5]. Existing soil products also have limits in spatial resolution, model transferability, and class representation. As a result, they may not show local transition features within oasis areas well [6,7]. Therefore, making full use of field observations and multi-year remote sensing data to improve the continuous spatial representation of soil texture remains an important task for regional soil surveys and land resource assessment.

Remote sensing has been widely used in digital soil mapping (DSM), and many studies have shown its value with different methods and datasets [8,9]. Satellite data can provide information on soil color, moisture, and shortwave-infrared response. However, in arid regions, oasis cropland, desert bare land, and built-up areas are spatially interspersed, making it difficult for single-date imagery to stably represent the spectral characteristics of soil itself. Long-term bare-soil compositing provides a way to reduce this limitation. By selecting exposed-soil observations from multiple dates, it can increase effective bare-soil coverage and reduce the influence of conditions associated with individual image acquisitions [10,11]. This approach has already been used for mapping soil organic carbon, soil texture, and other soil properties [12,13]. Other environmental data can also provide useful information. Radar backscatter can reflect surface structure and moisture. Land surface temperature can describe thermal differences. Terrain data can represent landscape position and depositional conditions [14,15]. Soil texture is also affected by parent material, climate, hydrology, topography, and long-term soil formation processes. Other soil properties may also provide related information [16,17]. However, these data sources often differ in regional coverage, mapping scale, temporal consistency, and independent availability, which limits their direct use for representing fine-scale spatial variation within oasis areas. Therefore, developing a stable regional predictor system that considers both the environmental relevance and spatial scale of variables while utilizing complementary information from multiple data sources remains an important research need.

Random Forest (RF) and eXtreme Gradient Boosting (XGBoost) can describe nonlinear relationships between soil properties and high-dimensional environmental variables. Both methods have been widely used for soil prediction [18,19]. However, these models usually use the environmental values of the pixel that contains each soil sample. They do not directly use spatial information from nearby field boundaries, bare-soil patches, or landform transitions [20]. Convolutional Neural Networks (CNNs) can extract local spatial features from image blocks. This provides another way to use neighborhood information for soil prediction [21,22]. Vision Transformers use self-attention to build relationships among different spatial positions. They can also be used to learn relationships among different data modalities [23]. Previous studies have applied Multi-Modal Vision Transformer (MMVT) to agricultural soil texture mapping. These studies showed that a suitable image extent can improve the prediction of sand, silt, and clay [24–26]. However, the useful spatial context depends on landform structure, field size, and environmental

heterogeneity. The oasis region along the central section of the northern Tianshan slope contains piedmont alluvial proluvial fans, oasis plains, desert margins, and mixed urban rural surfaces. The suitable spatial context for this region therefore needs to be tested separately. Another issue is the closed nature of soil particle-size data. Sand, silt, and clay always sum to 100%. Many studies model these fractions separately and then restore the closure by normalization or residual calculation. Log-ratio methods, including additive, centered, and isometric log-ratio transformations, have been used to deal with this problem. The isometric log-ratio transformation (ILR) represents a three-part composition with non-redundant orthogonal coordinates. It can therefore provide suitable response variables for both machine learning and deep learning models [27,28]. A comparison between single-pixel models and spatial-context models should also keep the compositional relationship among the three particle-size fractions.

Based on field soil survey data and remote sensing observations, this study integrates multi-year bare-soil composites, multi-source environmental information, and spatial-context modeling to improve soil texture prediction and mapping in complex arid oasis landscapes. The objectives are to: (1) construct a multi-source variable system centered on multi-year bare-soil composites and evaluate the contribution of different data sources; (2) compare RF, XGBoost, and MMVT models with different spatial windows within a compositional data framework, determine an appropriate spatial-context scale, and assess regional model applicability; and (3) generate continuous maps of sand, silt, and clay contents and soil texture classes using the optimal model, and characterize their spatial patterns. The study aims to improve soil texture prediction and spatial representation in arid oasis regions and provide more accurate spatial information for regional soil resource surveys, soil type identification, and land suitability assessment.

2. Materials and Methods

2.1. Study Area

The study area is located in the central section of the northern Tianshan slope in Xinjiang. It extends from 43°20′–45°00′ N, 85°17′–91°32′ E, with a delineated mapping area of approximately 22,900 km². The area includes Urumqi, Changji Hui Autonomous Prefecture, and Wujiaqu. It is a typical piedmont oasis–desert transition zone. The landform changes gradually from the Tianshan Mountains in the south to the Junggar Basin in the north. The study area includes mountain areas, piedmont alluvial proluvial fans, oasis plains, deserts, and Gobi areas. These landform units differ greatly in relief and depositional conditions [29]. The region has a typical temperate continental arid to semiarid climate. Mean annual precipitation in the oasis plain is about 190–215 mm, while potential evaporation exceeds 1200 mm. Agricultural and ecological water mainly comes from mountain precipitation, snow and glacier meltwater, and irrigation systems. Long-term alluvial proluvial deposition, aeolian processes, and irrigation have jointly affected soil formation in this region. These processes have resulted in strong spatial variation in soil texture. Figure 1 shows the study area and the distribution of the sampling sites.

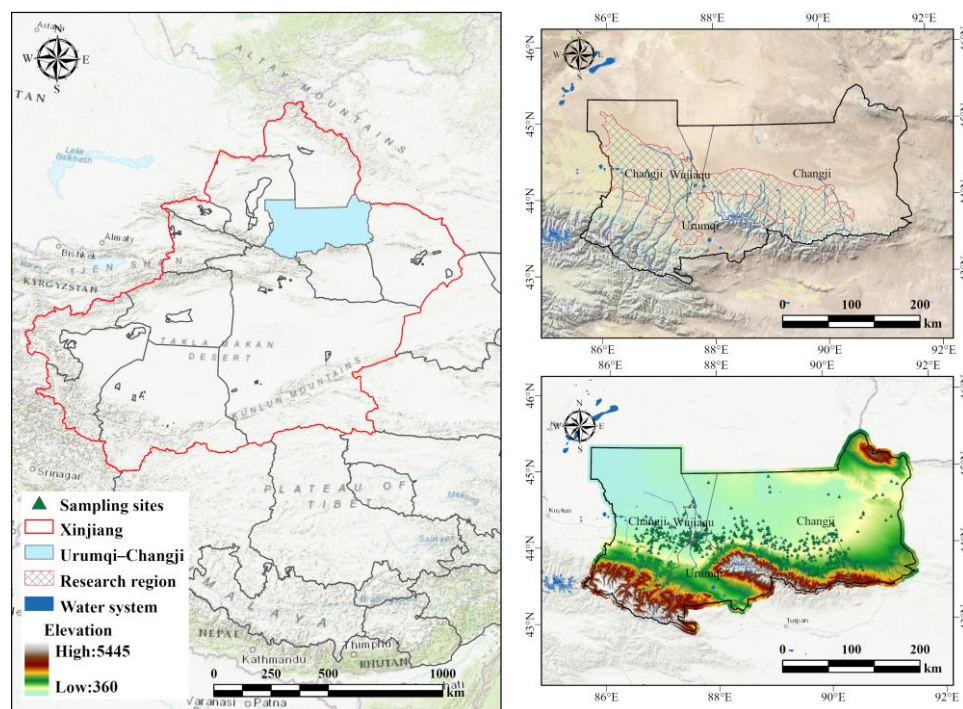


Figure 1. Study area and soil sampling sites.

2.2. Study Workflow

Figure 2 shows the main workflow of this study. We used the collected soil data and multi-source remote sensing data from 2015 to 2025 to build six environmental data modalities. We used the ILR transformation for the soil particle-size data. We then compared the soil texture prediction performance of RF, XGBoost, and MMVT. We selected the best model to predict sand, silt, and clay across the study area. We also used the predictions to produce a soil texture map based on the International Society of Soil Science (ISSS) classification. Finally, we analyzed the contribution of each data modality and the regional spatial pattern.

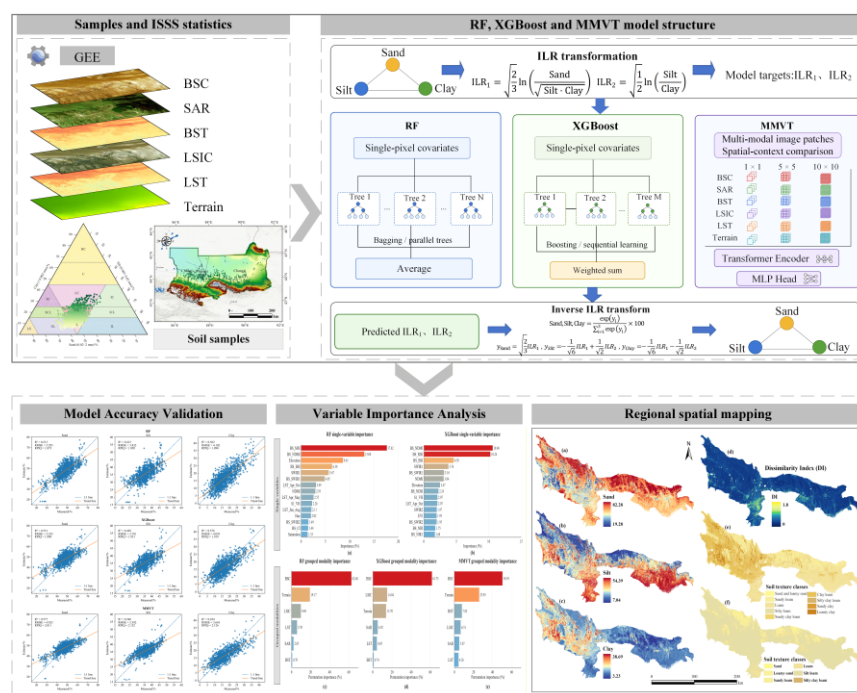


Figure 2. Workflow of the study.

2.3. Data Sources and Processing

2.3.1. Soil Sample Data

Soil samples were collected from 2023 to 2025. The sampling sites covered the main landform units and land-use types in the study area. Each sampling site included a soil sample from the 0–30 cm layer. We selected the sampling sites by considering soil type, landform position, land use, vegetation cover, and accessibility. We also avoided roads, shelterbelts, drainage ditches, and areas with clear human disturbance. This helped the samples better represent the surrounding pixel or field. For the 0–30 cm samples used in this study, we removed surface litter, plant roots, and other foreign materials. We then mixed the soil thoroughly and retained a representative sample using the quartering method. Soil particle-size distribution was measured using the pipette method following the Chinese standard HJ 1068–2019 [30]. Sand (0.02–2 mm), silt (0.002–0.02 mm), and clay (<0.002 mm) were defined according to the ISSS classification system. All fractions were expressed as mass percentages [31]. To minimize closely spaced duplicate sampling within local areas and reduce potential spatial dependence between neighboring samples, a spatial screening procedure with a minimum inter-sample distance of 2 km was applied.

2.3.2. Multi-Source Remote Sensing and Environmental Variables

We used multi-source remote sensing data from 2015 to 2025 and terrain information to construct six groups of environmental variables: multispectral bare-soil composites (BSCs), radar variables (SAR), bare-soil temperature composites (BSTs), long-term spectral index composites (LSICs), temporal land surface temperature composites (LSTs), and terrain variables (Table 1). Given the relative temporal stability of soil texture in the absence of substantial erosion, deposition, or disturbance, these multi-year observations were used to characterize environmental conditions associated with soil samples collected during 2023–2025. Multi-year compositing also increased valid bare-soil observations and reduced short-term surface variability. These modalities provide complementary information on soil texture variation [26,32,33]. All variables were processed in Google Earth Engine (GEE) [34], resampled, and aligned to a common 30 m grid. The calculation formulas are provided in Supplementary Table S1, and the processing steps for each modality are described below.

(1) Multispectral Bare-Soil Composite

Given strong seasonal and land-use changes in surface conditions, multi-year imagery was used to identify bare-soil pixels and construct a long-term composite, reducing the effects of vegetation, crop residues, soil moisture, and varying observation conditions. The BSC was produced from the Harmonized Landsat and Sentinel-2 (HLS) dataset [35]. A long image record was screened to retain spectral observations acquired when soil was bare or sparsely covered [10,12,13]. Based on regional crop phenology and the availability of exposed-soil observations, mid-March to late May and early September to late October were selected as observation windows. These periods increased the number of usable bare-soil pixels while avoiding peak vegetation cover. Thresholds were selected through visual interpretation of representative surface conditions and sensitivity testing. Bare soil was identified jointly with the normalized difference vegetation index (NDVI), normalized burn ratio 2 (NBR2), normalized difference snow index (NDSI), and bare soil index (BSI) [10,12,36]. The adopted thresholds were $NDVI < 0.25$, $NBR2 < 0.175$, $NDSI < 0.10$, and $BSI > 0.05$. Valid observations from all years were summarized by their median to create the multispectral composite bands, after which indices related to soil moisture, surface exposure, and shortwave-infrared response were calculated [37].

(2) Radar Variables

SAR backscatter responds to surface structure, roughness, and moisture and can supplement soil-surface information that optical imagery does not fully capture [14,32,33]. We used a bare-soil Sentinel-1 composite together with annual PALSAR/PALSAR-2 mosaics [38,39]. Each Sentinel-1 acquisition was matched to an HLS observation acquired within five days. When more than one SAR acquisition occurred within a matching window, those acquisitions were averaged first. The bare-soil mask derived from the corresponding HLS observation was then used to remove non-bare-soil pixels, and pixelwise means of the valid VV and VH observations were calculated over 2015–2025. PALSAR/PALSAR-2 data came from the 25 m global annual SAR mosaics distributed by the Japan Aerospace Exploration Agency (JAXA). Their L-band HH and HV backscatter supplied complementary annual-scale radar information.

(3) Bare-Soil Temperature Composite

The BST was derived from the Landsat surface temperature (ST) band [40], providing information on variations in soil moisture status, surface exposure, and thermal behavior [41]. Because Landsat multispectral and ST observations are acquired nearly simultaneously, a bare-soil mask was generated independently for each screened multispectral scene and transferred to the corresponding ST band. Non-bare-soil pixels were thereby removed before the valid observations were combined. The median across all years was then calculated to represent the thermal background of exposed soil.

(4) Long-Term Spectral Index Composite

The LSIC used HLS observations from 2015–2025 to characterize the long-term surface spectral background, vegetation development, and moisture response [42]. Following preprocessing, the median of all valid observations was calculated for each pixel. The modality contains the original reflectance bands together with NDVI, NDMI, NDWI, EVI, GCVI, and NMDI. These layers describe persistent surface cover, vegetation vigor, and moisture conditions and therefore provide indirect information on soil-vegetation relationships and the broader environmental setting [33,43].

(5) Temporal Land-Surface-Temperature Composites

Temporal LST variables were built from the Landsat surface temperature (ST) band [44]. These variables describe the long-term surface thermal conditions of the study area and their seasonal differences [41]. We considered seasonal changes in agricultural activity and surface cover. Based on these changes, we produced LST composites for April–May, June–August, September–October, and the whole April–October period. These composites represent the long-term thermal conditions during different stages of the main agricultural season. We did not include winter because snow and cloud cover greatly reduced the number of valid observations.

(6) Terrain Variables

Topography influences soil formation, material redistribution, and water movement, so terrain information was included as an environmental covariate. Elevation was obtained from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) [45]. Slope, aspect, and the topographic wetness index (TWI) were derived from the DEM. TWI represents relative wetness or potential water availability within a landscape and combines slope with upslope contributing area [46].

HLS L30 and S30 share blue, green, red, near-infrared, and two shortwave-infrared bands. L30 B5 and S30 B8A were both named NIR, while S30 B8 was named NIR1. RE1, RE2, and RE3 correspond to S30 B5, B6, and B7, respectively. These red-edge bands were available only from Sentinel-2. HLS applies spectral response adjustments and observation geometry normalization to the shared bands. Both products were processed using the same seasonal selection and quality control criteria. Valid observations were used to produce multi-year median composites for each band. The shared bands combined observations from L30 and S30. In contrast, RE1, RE2, RE3, and NIR1 used only S30 observations.

Table 1. Multi-source environmental variables used for soil texture prediction.

Variable Category	Data Source	Period	Spatial Resolution	Variables
Multispectral bare-soil composite	HLS L30/S30	2015–2025	30 m	B, G, R, RE1, RE2, RE3, NIR1, NIR, SWIR1, SWIR2, NDVI, MSAVI2, BSI, NBR2, NDSI, NDMI, NDMI2, NMDI, MSI, MSI2, BI, CI, SI, Saturation, Hue, Value
Radar variables	Sentinel-1	2015–2025	10 m	S1_VV, S1_VH
	PALSAR	2015–2023	25 m	PALSAR_HH, PALSAR_HV
Bare-soil temperature composite	Landsat ST	2015–2025	30 m	BST
Long-term spectral index composite	HLS L30/S30	2015–2025	30 m	B, G, R, RE1, RE2, RE3, NIR1, NIR, SWIR1, SWIR2, NDVI, NDMI, NDWI, EVI, GCVI, NMDI
Temporal land-surface-temperature composites	Landsat ST	2015–2025	30 m	LST_Apr_May, LST_Jun_Aug, LST_Sep_Oct, LST_Apr_Oct
Terrain variables	DEM	—	30 m	Elevation, Slope, Aspect, TWI

Notes: RE1, RE2, RE3, and NIR1 were derived only from HLS S30. NIR corresponds to L30 B5 and S30 B8A. Some spectral indices have the same names in the BSC and LSIC modalities. In this study, BSC variables were labeled with the prefix “BS_”, while the corresponding LSIC variables retained their original names.

2.4. Compositional Data Transformation

Because sand, silt, and clay sum to 100%, they form closed compositional data. Modeling the fractions independently can introduce spurious correlation and yield predictions that do not satisfy closure [47,48]. We therefore transformed the ternary composition into two orthogonal real-valued coordinates using the ILR transformation and recovered the fractions by inverse transformation after prediction [28]. If zero values were present, multiplicative replacement would be applied by substituting each zero with half the detection limit and proportionally rescaling the nonzero fractions to preserve compositional closure and their relative ratios. Let sand, silt, and clay be denoted by x_s , x_{si} , and x_c , respectively. An orthonormal basis contrasting sand with the combined silt-clay component and silt with clay was used; the ILR coordinates are:

$$z_1 = \sqrt{\frac{2}{3}} \ln \left(\frac{x_s}{\sqrt{x_{si}x_c}} \right) \quad (1)$$

$$z_2 = \sqrt{\frac{1}{2}} \ln \left(\frac{x_{si}}{x_c} \right) \quad (2)$$

Here, z_1 and z_2 are the two transformed real-valued coordinates. They served as model responses, and the particle-size fractions were recovered after prediction by the inverse transformation. Define:

$$y_s = \sqrt{\frac{2}{3}} z_1 \quad (3)$$

$$y_{si} = -\frac{1}{\sqrt{6}} z_1 + \frac{1}{\sqrt{2}} z_2 \quad (4)$$

$$y_c = -\frac{1}{\sqrt{6}}z_1 - \frac{1}{\sqrt{2}}z_2 \quad (5)$$

The inverse transformation is:

$$x_i = \frac{\exp(y_i)}{\exp(y_s) + \exp(y_{si}) + \exp(y_c)} \times 100 \quad (6)$$

where i denotes sand, silt, or clay. This procedure guarantees that the predicted particle-size fractions sum to 100%.

2.5. Soil Texture Prediction and Mapping Models

To compare different modeling approaches for spatial soil texture prediction, we used RF, XGBoost, and MMVT. RF models nonlinear relationships between environmental variables and soil texture by combining multiple decision trees [18]. The RF model used 300 trees, a maximum depth of 20, and a maximum feature ratio of 0.7. The minimum samples required for splitting a node and forming a leaf were 20 and 5, respectively. XGBoost uses gradient boosting and regularization to improve predictions and model complex interactions among high-dimensional variables [19]. It used 300 trees, a maximum depth of 6, and a learning rate of 0.05. The feature and sample subsampling ratios were 0.5 and 0.85, respectively.

MMVT extracted neighborhood image patches centered on each sampling site from six environmental modalities. Each modality was converted into feature tokens through spatial partitioning, feature embedding, and positional encoding. These tokens were jointly processed by the Transformer encoder to learn spatial and cross-modal relationships. The model used an embedding dimension of 64, four Transformer encoder layers, and four attention heads. The MLP hidden dimension was 256, and the dropout rate was 0.10. Training used the AdamW optimizer and mean squared error (MSE) loss. The initial learning rate was 0.001, with a weight decay of 1×10^{-5} . The batch size was 32, and the maximum number of training epochs was 120. These settings were selected to balance prediction accuracy and computational efficiency. Multi-head self-attention captured long-range dependencies and interactions among modalities and spatial positions [20,23,49]. This allowed the model to represent spatial continuity in bare-soil patches, field structures, and landform transitions.

All input data were resampled to a common spatial resolution of 30 m. MMVT used 12 input window sizes, ranging from 1×1 to 12×12 pixels. For 1×1 to 3×3 windows, each modality generated one token. Windows of 4×4 pixels or larger used 2×2 -pixel patches (Table 2). The model predicted two isometric log-ratio (ILR) coordinates. These were inverse-transformed into sand, silt, and clay percentages that sum to 100%. The resulting fractions were used for spatial soil texture mapping. Detailed model parameters are provided in Supplementary Table S2.

Table 2. Input-window sizes and patch settings for the MMVT model.

Image Size	Spatial Extent	Total Number of Patches
1×1	30×30 m	6
2×2	60×60 m	6
3×3	90×90 m	6
4×4	120×120 m	24
5×5	150×150 m	54
6×6	180×180 m	54
7×7	210×210 m	96

8 × 8	240 × 240 m	96
9 × 9	270 × 270 m	150
10 × 10	300 × 300 m	150
11 × 11	330 × 330 m	216
12 × 12	360 × 360 m	216

Note: For 1 × 1–3 × 3 windows, each modality was encoded as a single token. For windows of 4 × 4 pixels or larger, tokens were generated using 2 × 2-pixel spatial blocks. For odd-sized windows, the edge patches overlapped with adjacent patches by one pixel to ensure complete spatial coverage without padding or pixel omission.

2.6. Model Evaluation, Variable Importance, Spatial Applicability

(1) Model Performance Evaluation

Model performance was evaluated using 50 repeated random splits and five-fold spatial cross-validation. For each random split, samples were divided into training and external test sets at a 7:3 ratio. Within the training set, 15% of samples were reserved for internal validation. This subset was used for hyperparameter tuning, early stopping, and window selection. The external test set assessed predictive performance after model settings were finalized.

Five-fold spatial cross-validation was also used to assess spatial generalization [50]. Based on the spatial autocorrelation ranges estimated from empirical semivariograms of the observed sand, silt, and clay fractions [51], the study area was divided into regular spatial blocks of 30 × 30 km (Figure S1). These blocks were assigned to five folds using a greedy sample-balancing algorithm to ensure relatively balanced sample sizes across folds. RF, XGBoost, and MMVT used the same spatial folds. The mean and standard deviation of performance metrics were calculated across the five folds. All out-of-fold (OOF) predictions were then combined to calculate overall performance metrics.

Model performance was evaluated with the coefficient of determination (R^2), root mean square error (RMSE), and ratio of performance to interquartile range (RPIQ). R^2 measures the proportion of observed variation represented by the predictions, RMSE gives the absolute magnitude of error, and RPIQ divides the interquartile range (IQR) by RMSE to relate prediction error to the spread of the observations [52]. The metrics were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

$$\text{RPIQ} = \frac{\text{IQR}}{\text{RMSE}} \quad (9)$$

where y_i is the measured value for sample i , $\hat{y}_i$ is its prediction, $\bar{y}$ is the mean of the measured values, n is the number of evaluation samples, and IQR is the difference between the 75th and 25th percentiles. Better performance is indicated by larger R^2 and RPIQ values and a smaller RMSE.

(2) Variable Importance Analysis

To evaluate the contributions of individual environmental variables and different data modalities to soil texture prediction, individual-variable importance and grouped modality permutation importance analyses were conducted [53]. Model-specific variable importance was calculated separately for RF and XGBoost to evaluate the contribution of individual environmental variables. Grouped permutation was further performed for the

six environmental data modalities while keeping the model structure and parameters unchanged [54]. For RF and XGBoost, all environmental variables within the selected modality were treated as a group and jointly permuted among validation samples using the same random permutation index; for MMVT, the complete image block corresponding to the selected modality was exchanged among validation samples using the same random permutation index, while all other modalities remained unchanged. Each modality was permuted 50 times. Using RMSE as the error metric, the performance change caused by permuting modality m in repetition r was expressed as:

$$\Delta \text{RMSE}_{m,r} = \text{RMSE}_{m,r}^{\text{perm}} - \text{RMSE}_{\text{base}} \quad (10)$$

where $\text{RMSE}_{\text{base}}$ is the baseline RMSE for the unaltered validation set and $\text{RMSE}_{m,r}^{\text{perm}}$ is the RMSE obtained after modality m was shuffled in permutation run r . The importance of modality m was defined as the mean change across the 50 permutations:

$$I_m = \frac{1}{R} \sum_{r=1}^R \Delta \text{RMSE}_{m,r} \quad (11)$$

To facilitate comparison among modalities, importance was converted to a percentage:

$$P_m = \frac{I_m}{\sum_{m=1}^M I_m} \times 100 \quad (12)$$

where $R = 50$, $M = 6$, and P_m is the normalized importance of modality m . Grouped permutation importance therefore measures the loss in predictive performance when the information carried by an entire modality is disrupted.

(3) Spatial Applicability Assessment

In addition, the Dissimilarity Index (DI) and Area of Applicability (AOA) were used to evaluate the spatial applicability of the final MMVT model [55]. The DI was calculated using the same six modalities used for model development. For each training sample and prediction location, the same 10×10 -pixel neighborhood used by MMVT was extracted, and the neighborhood mean of each environmental variable was used to construct the environmental feature vector. All variables were standardized using the mean and standard deviation of the training samples. The six modalities were weighted according to their normalized grouped permutation importance derived from MMVT, while variables within each modality equally shared the corresponding modality weight. Distances between prediction locations and training samples were calculated in the weighted environmental feature space and used to derive the DI. The AOA threshold was determined using the 50 repeated 7:3 data partitions. For each soil sample, DI values obtained when the sample was assigned to the held-out test set were collected and averaged to obtain a single DI value. Based on the distribution of these sample-level DI values, unusually high values exceeding $(Q_3 + 1.5 \times \text{IQR})$ were excluded, and the maximum remaining DI was used as the AOA threshold.

3. Results

3.1. Descriptive Statistics of Soil Samples

A total of 1051 soil sampling locations were used for model development and assessment. Table 3 summarizes the particle-size fractions, and Figure 3 shows their distributions. Sand was the largest fraction, with a mean of 47.78%, a median of 47.82%, and a range of 16.63–69.08%. Silt averaged 35.53%, had a median of 35.61%, and ranged from 18.10–49.65%. Clay was lower overall, with a mean of 16.70%, a median of 15.66%, and a range of 1.30–39.52%.

Figure 4 places the samples in the ISSS soil texture triangle. Clay loam, loam, sandy loam, and loamy clay accounted for 37.96%, 33.21%, 11.89%, and 10.18% of the samples, respectively, together representing 93.24%. Other classes were uncommon. The sampled topsoils were therefore dominated by loam and clay loam, with relatively few extremely sandy or clayey observations. The classification criteria for the 12 ISSS soil texture classes are provided in Supplementary Table S3.

Table 3. Descriptive statistics of particle-size fractions in the soil samples (%).

Soil Particle-Size Fraction	Max	Min	Mean	Median	SD	CV
Sand	69.08	16.63	47.78	47.82	7.57	15.85
Silt	49.65	18.10	35.53	35.61	5.00	14.08
Clay	39.52	1.30	16.70	15.66	6.20	37.15

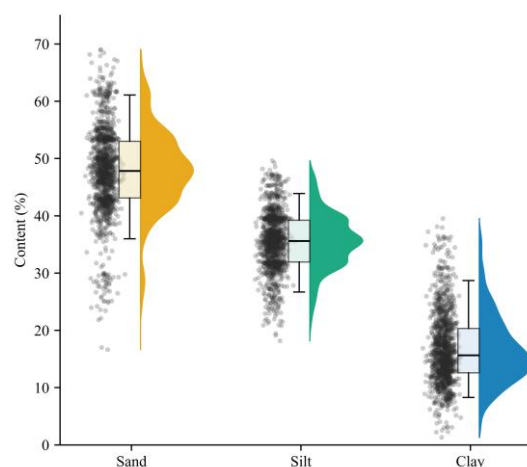


Figure 3. Distributions of sand, silt, and clay contents.

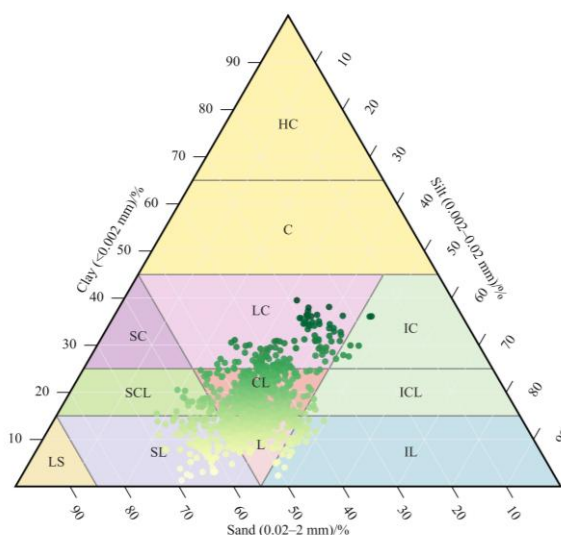


Figure 4. Distribution of samples in the ISSS soil texture triangle.

3.2. Composite Patterns of Environmental Variables

Figure 5 shows the multispectral bare-soil composite, Sentinel-1 bare-soil SAR composite, bare-soil temperature composite, and temporal land surface temperature composite. All four products cover most of the study area. They also show clear regional differences and local variation.

Several similar patterns can be seen among the four products. Permanently bare land in the central-western part of the study area and along the northern desert margin appears light yellow to brown in the BSC image. The Sentinel-1 SAR image shows a relatively continuous green background in these areas, with some bright and dark patches. BST and LST are also relatively high in the same areas. In the central croplands and around towns, the BSC image is generally darker. The SAR pattern is more fragmented, and more medium- and low-temperature patches appear in the BST and LST images.

The four products also show some clear differences. BSC shows the color differences among fields, bare land, and the oasis margin more clearly. Sentinel-1 SAR shows a similar general pattern, but some small local differences are less clear. BST and LST differ more strongly. LST includes all valid observations from April to October, including the warm period from June to August. In contrast, BST only uses bare-soil observations from spring and autumn. For this reason, BST is generally cooler and shows more scattered local patches. LST has a more continuous regional pattern. Higher values mainly occur in the western and northern parts, while lower values occur along the southern piedmont and in some eastern areas. Overall, these products show different surface features and provide useful information for soil texture prediction.

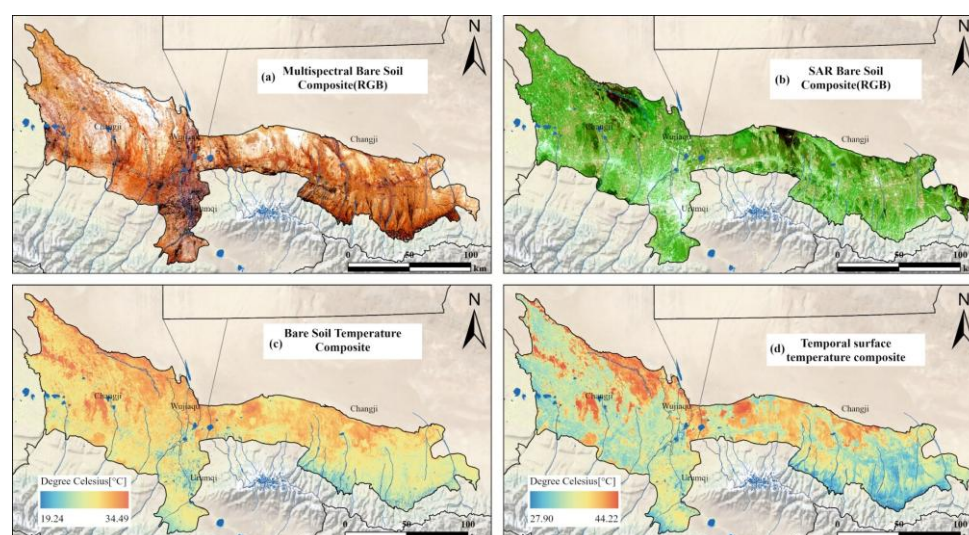


Figure 5. Environmental composite images used in this study. (a) Multispectral bare-soil composite shown as RGB (R/G/B); (b) Sentinel-1 bare-soil SAR composite shown as RGB (VH/VV/VH + VV); (c) bare-soil temperature composite; (d) April–October land surface temperature composite (LST_Apr_Oct).

3.3. Model Accuracy Assessment

Model performance was assessed through internal validation, repeated external testing, and five-fold spatial cross-validation. Internal validation showed that MMVT performance generally increased with window size before declining at larger windows (Figure 6). The 10×10 window showed relatively high predictive accuracy across all three particle-size fractions. This configuration was selected for subsequent model comparisons and spatial mapping.

Table 4 presents the results of 50 repeated external tests for three models using 1×1 and 10×10 windows. At 1×1 , XGBoost generally showed higher predictive accuracy than RF and MMVT. At 10×10 , MMVT achieved mean R^2 values of 0.586 ± 0.027 , 0.519 ± 0.029 , and 0.706 ± 0.026 . These values correspond to sand, silt, and clay, respectively. The corresponding RMSE values were $4.695 \pm 0.118\%$, $3.391 \pm 0.119\%$, and $3.335 \pm 0.116\%$. RPIQ values were 2.107 ± 0.124 , 2.152 ± 0.125 , and 2.324 ± 0.122 , respectively. Compared with RF

at 10×10 , mean R^2 increased by 0.073, 0.118, and 0.131. Performance distributions across 50 external tests for both window sizes are shown in Supplementary Figure S2.

Spatial cross-validation generally produced lower R^2 values and greater variation across folds than repeated external testing (Table 5). Table 5 compares the same three models and two window sizes. At 1×1 , XGBoost generally maintained higher predictive accuracy. At 10×10 , MMVT achieved mean R^2 values of 0.550 ± 0.077 , 0.504 ± 0.090 , and 0.657 ± 0.092 . The corresponding RMSE values were $4.948 \pm 0.569\%$, $3.410 \pm 0.469\%$, and $3.557 \pm 0.592\%$. RPIQ values were 2.009 ± 0.212 , 1.980 ± 0.226 , and 2.119 ± 0.336 , respectively. Compared with repeated external testing, MMVT mean R^2 decreased by 0.036, 0.015, and 0.049, respectively. However, its mean R^2 remained higher than those of RF and XGBoost at 10×10 .

Figure 7 compares the three models using pooled spatial out-of-fold (OOF) predictions at 10×10 . MMVT achieved R^2 values of 0.577, 0.540, and 0.654 for sand, silt, and clay, respectively. The corresponding RMSE values were 4.923%, 3.392%, and 3.645%. RPIQ values were 2.011, 2.152, and 2.126, respectively. All three models showed some overestimation at low values and underestimation at high values. Additional analyses focused on the spatial OOF predictions from MMVT at 10×10 . Additional analyses included ILR predictions for all three models (Figure S3), MMVT residuals (Figure S4), and classification confusion matrices (Figure S5). Additionally, Aitchison distances were calculated from the pooled five-fold spatial out-of-fold (OOF) predictions of RF, XGBoost, and MMVT using a 10×10 window to evaluate the overall prediction errors of soil particle-size compositions (Supplementary Table S4).

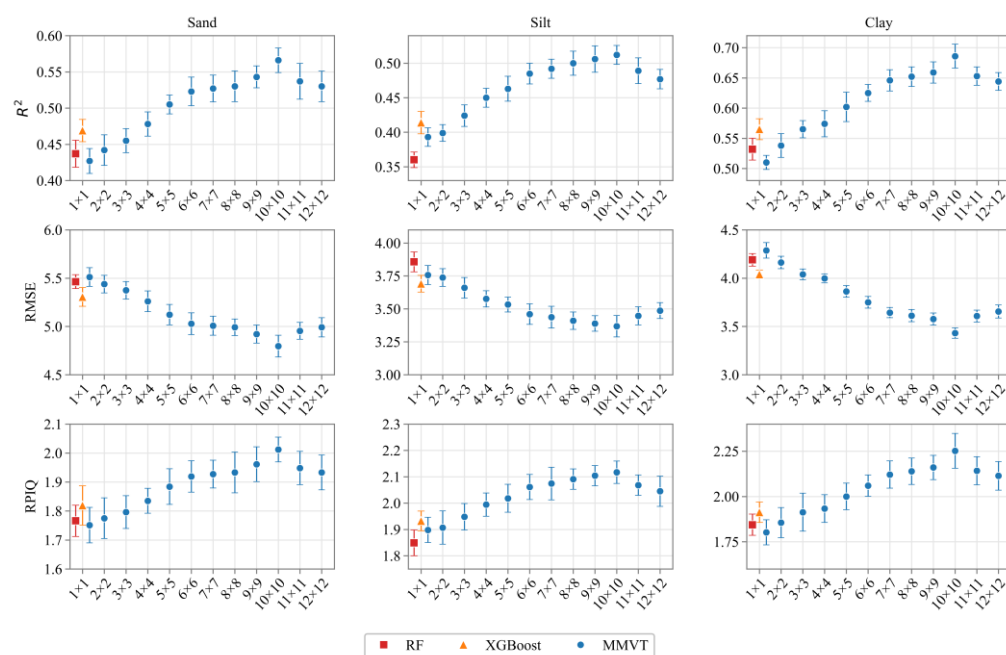


Figure 6. Effects of input window size on MMVT performance during repeated internal validation. Results are shown for sand, silt, and clay contents across window sizes from 1×1 to 12×12 pixels. Points represent mean performance across 50 repetitions, and error bars indicate 95% confidence intervals.

Table 4. Model performance across 50 repeated external tests with 1×1 and 10×10 windows (mean $\pm$ SD).

Model	Metric	1×1			10×10		
		Sand	Silt	Clay	Sand	Silt	Clay
RF	R ²	0.448 $\pm$ 0.043	0.367 $\pm$ 0.048	0.544 $\pm$ 0.045	0.513 $\pm$ 0.034	0.401 $\pm$ 0.041	0.575 $\pm$ 0.029
	RMSE	5.384 $\pm$ 0.153	3.828 $\pm$ 0.163	4.110 $\pm$ 0.158	5.065 $\pm$ 0.146	3.744 $\pm$ 0.139	3.978 $\pm$ 0.145
	RPIQ	1.837 $\pm$ 0.147	1.907 $\pm$ 0.157	1.886 $\pm$ 0.150	1.953 $\pm$ 0.127	1.950 $\pm$ 0.127	1.948 $\pm$ 0.127
XGBoost	R ²	0.483 $\pm$ 0.038	0.429 $\pm$ 0.043	0.582 $\pm$ 0.040	0.527 $\pm$ 0.034	0.451 $\pm$ 0.038	0.604 $\pm$ 0.027
	RMSE	5.219 $\pm$ 0.145	3.662 $\pm$ 0.154	3.957 $\pm$ 0.150	5.002 $\pm$ 0.139	3.598 $\pm$ 0.131	3.854 $\pm$ 0.134
	RPIQ	1.895 $\pm$ 0.142	1.993 $\pm$ 0.149	1.959 $\pm$ 0.144	1.977 $\pm$ 0.129	2.029 $\pm$ 0.132	2.011 $\pm$ 0.131
MMVT	R ²	0.441 $\pm$ 0.048	0.407 $\pm$ 0.049	0.524 $\pm$ 0.048	0.586 $\pm$ 0.027	0.519 $\pm$ 0.029	0.706 $\pm$ 0.026
	RMSE	5.433 $\pm$ 0.165	3.729 $\pm$ 0.168	4.213 $\pm$ 0.168	4.695 $\pm$ 0.118	3.391 $\pm$ 0.119	3.335 $\pm$ 0.116
	RPIQ	1.820 $\pm$ 0.157	1.958 $\pm$ 0.158	1.839 $\pm$ 0.158	2.107 $\pm$ 0.124	2.152 $\pm$ 0.125	2.324 $\pm$ 0.122

Note: RF and XGBoost use central-pixel values (1×1) or neighborhood means (10×10). MMVT uses image patches.

Table 5. Model performance under five-fold spatial cross-validation with 1×1 and 10×10 windows (mean $\pm$ SD).

Model	Metric	1×1			10×10		
		Sand	Silt	Clay	Sand	Silt	Clay
RF	R ²	0.421 $\pm$ 0.078	0.335 $\pm$ 0.089	0.475 $\pm$ 0.088	0.497 $\pm$ 0.076	0.376 $\pm$ 0.091	0.521 $\pm$ 0.093
	RMSE	5.551 $\pm$ 0.554	3.839 $\pm$ 0.435	4.289 $\pm$ 0.523	5.208 $\pm$ 0.552	3.767 $\pm$ 0.440	4.122 $\pm$ 0.553
	RPIQ	1.789 $\pm$ 0.163	1.756 $\pm$ 0.186	1.754 $\pm$ 0.203	1.911 $\pm$ 0.185	1.795 $\pm$ 0.196	1.830 $\pm$ 0.232
XGBoost	R ²	0.460 $\pm$ 0.072	0.392 $\pm$ 0.097	0.514 $\pm$ 0.081	0.500 $\pm$ 0.081	0.403 $\pm$ 0.097	0.538 $\pm$ 0.083
	RMSE	5.383 $\pm$ 0.528	3.730 $\pm$ 0.458	4.145 $\pm$ 0.501	5.202 $\pm$ 0.578	3.708 $\pm$ 0.456	4.057 $\pm$ 0.513
	RPIQ	1.846 $\pm$ 0.166	1.810 $\pm$ 0.207	1.817 $\pm$ 0.208	1.911 $\pm$ 0.196	1.822 $\pm$ 0.208	1.858 $\pm$ 0.222
MMVT	R ²	0.412 $\pm$ 0.086	0.384 $\pm$ 0.087	0.447 $\pm$ 0.093	0.550 $\pm$ 0.077	0.504 $\pm$ 0.090	0.657 $\pm$ 0.092
	RMSE	5.604 $\pm$ 0.584	3.748 $\pm$ 0.430	4.394 $\pm$ 0.539	4.948 $\pm$ 0.569	3.410 $\pm$ 0.469	3.557 $\pm$ 0.592
	RPIQ	1.775 $\pm$ 0.170	1.800 $\pm$ 0.193	1.714 $\pm$ 0.198	2.009 $\pm$ 0.212	1.980 $\pm$ 0.226	2.119 $\pm$ 0.336

Note: Input settings are consistent with Table 4.

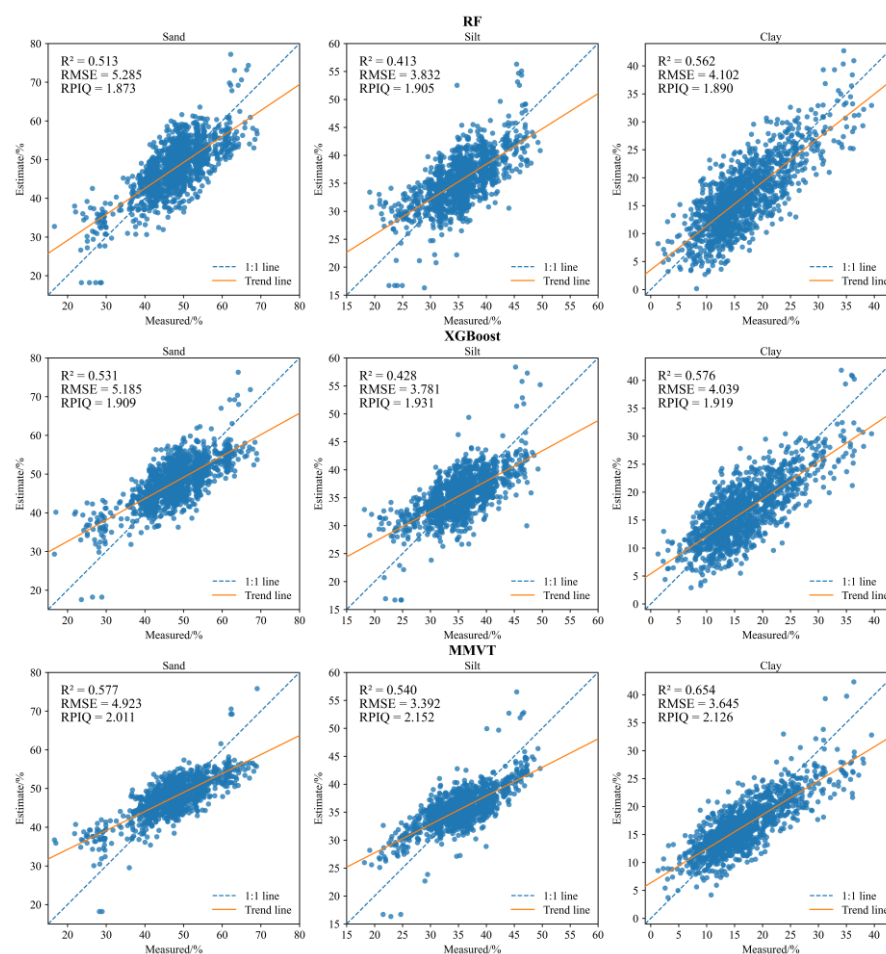


Figure 7. Observed versus predicted soil particle-size fractions for RF, XGBoost, and MMVT under five-fold spatial cross-validation. All models used a 10×10 -pixel neighborhood. Predictions were pooled from the five spatial test folds.

3.4. Variable Importance Analysis

To assess the contributions of different environmental variables and the six data modalities to soil texture prediction, single-variable importance was calculated for RF and XGBoost. Grouped modality permutation importance was further calculated for RF, XGBoost, and MMVT using the same method. The results are shown in Figure 8. Only the top 15 variables ranked by single-variable importance are presented for RF and XGBoost. Overall, BSC had the highest contribution in all three models. However, the models differed in their dependence on the other environmental information.

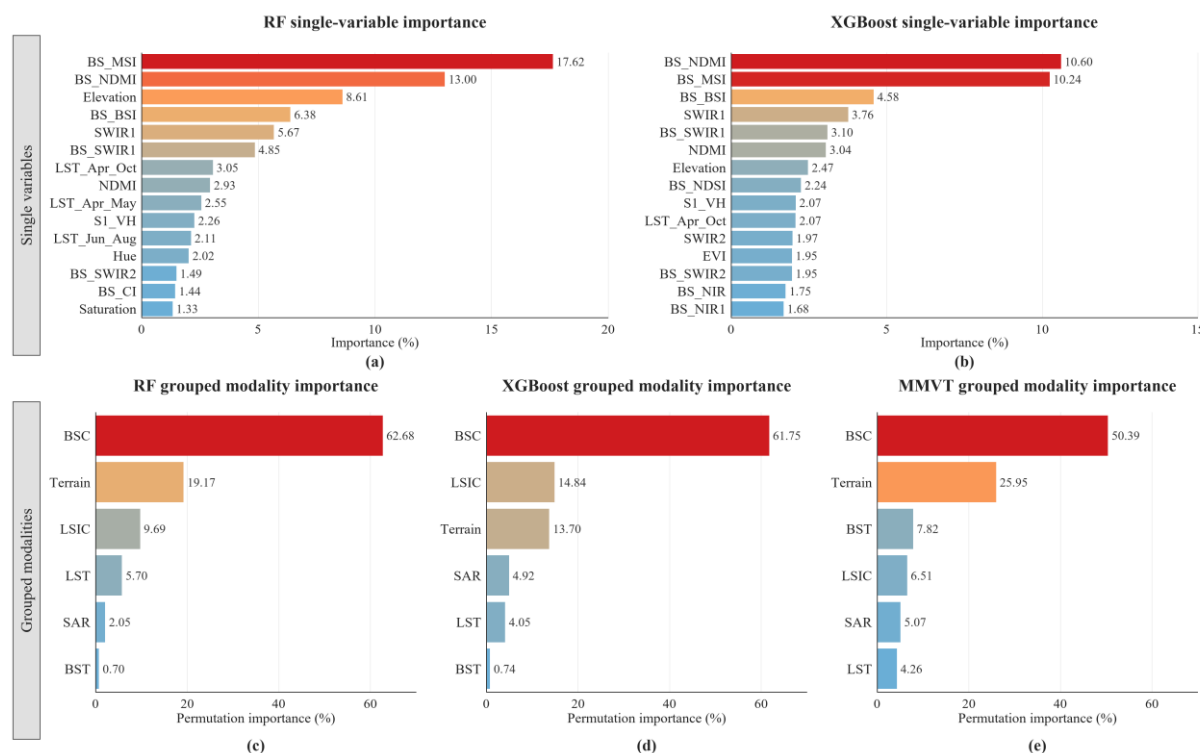


Figure 8. Variable and grouped modality importance for RF, XGBoost, and MMVT: (a) RF; (b) XGBoost; (c) RF grouped modality; (d) XGBoost grouped modality; (e) MMVT grouped modality. Top 15 variables are shown in (a,b). Some spectral indices have the same names in the BSC and LSIC modalities. In this study, BSC variables were labeled with the prefix “BS_”, while the corresponding LSIC variables retained their original names.

For RF, BS_MSI and BS_NDMI were the two most important variables, with importance values of 17.62% and 13.00%, respectively. Elevation, BS_BSI, SWIR1, and BS_SWIR1 followed, with values of 8.61%, 6.38%, 5.67%, and 4.85%. LST_Apr_Oct, NDMI, LST_Apr_May, and S1_VH were also among the top ten variables. For XGBoost, BS_NDMI and BS_MSI ranked first and second, with importance values of 10.60% and 10.24%. They were followed by BS_BSI, SWIR1, BS_SWIR1, NDMI, Elevation, BS_NDSI, S1_VH, and LST_Apr_Oct. Most of the highly ranked variables in both models belonged to BSC. Variables from LSIC, Terrain, SAR, and LST were also included.

The grouped modality permutation results further showed that BSC had the highest contribution in all three models. In RF, the normalized importance of BSC was 62.68%. Terrain and LSIC followed, with values of 19.17% and 9.69%, respectively. In XGBoost, the contribution of BSC was 61.75%. LSIC and Terrain contributed 14.84% and 13.70%, respectively. BSC also had the highest contribution in MMVT, with a normalized importance of 50.39%. This value was lower than those of RF and XGBoost. Terrain ranked second in MMVT, with a contribution of 25.95%. The contributions of BST, LSIC, SAR, and LST were 7.82%, 6.51%, 5.07%, and 4.26%, respectively. Compared with RF and XGBoost, the modality contributions of MMVT were more dispersed. MMVT also showed greater dependence on Terrain, BST, and SAR.

Overall, BSC exhibited the highest normalized grouped permutation importance across all three models. Its relative importance was higher in RF and XGBoost than in MMVT. Terrain showed a higher relative importance in MMVT. These results reflect model sensitivity to modality perturbations rather than independent contributions from each information source.

3.5. Spatial Mapping of Soil Texture

Using the MMVT model with a 10×10 window, continuous maps of sand, silt, and clay contents were generated. The Dissimilarity Index (DI) was also calculated to evaluate model applicability across the study area. The continuous predictions were further converted into ISSS soil texture classes. Classification results from the China High-Resolution National Soil Information Grids Basic Attribute Dataset (2010–2018) were used only as a reference for broad spatial patterns (Figure 9) [6].

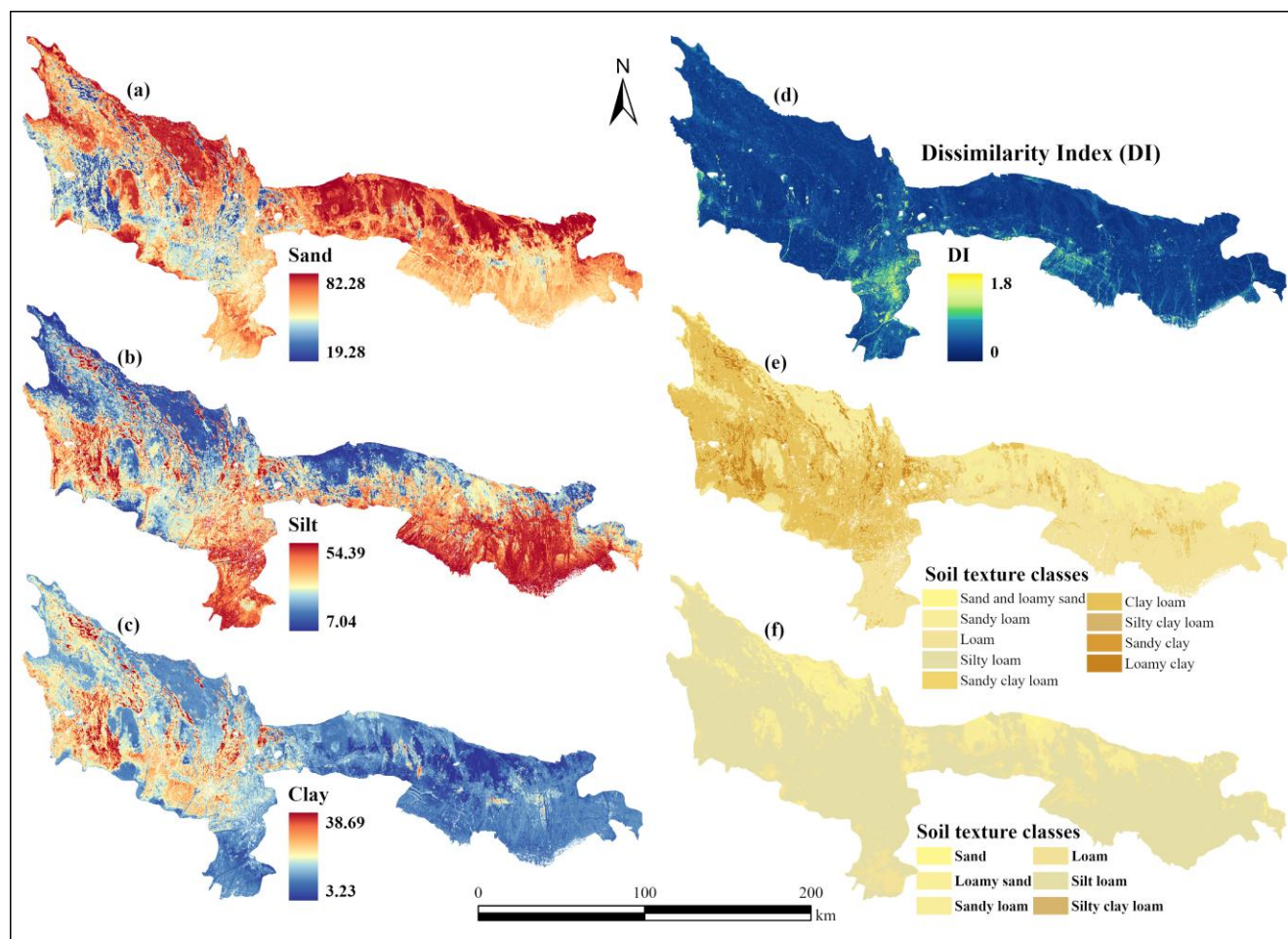


Figure 9. Continuous soil texture maps, DI, and classification comparison: (a) sand content; (b) silt content; (c) clay content; (d) dissimilarity index (DI); (e) ISSS texture classes derived in this study; and (f) classification results from an existing soil data product, used only for comparison of broad spatial patterns.

The predictions showed a coherent regional pattern. Sand content was high along the northern desert margin, forming a relatively continuous sandy belt. In the southeastern piedmont hills and foothill croplands, sand and silt occurred in similar proportions and clay was generally low. The western oasis plain was comparatively enriched in clay. Fragmented cropland in the center and transitions between landforms were more heterogeneous, with the three fractions interspersed in small patches.

Most of the study area had relatively low DI values, indicating environmental conditions similar to those represented by the training samples. Higher DI values occurred mainly in some urban areas, low-mountain river valleys, and hilly transition zones. With an AOA threshold of 0.524, approximately 85% (19,400 km²) of the study area was within the model AOA, whereas about 15% exceeded the threshold. The high-sand areas in the

north generally remained within low-to-moderate DI ranges and below the AOA threshold. For the 10×10 MMVT model, pooled spatial OOF predictions showed higher R^2 and lower RMSE inside the AOA (Figure S6).

The ISSS classification was dominated by loam (34.28%), sandy loam (23.16%), clay loam (18.56%), and loamy clay (9.14%). Their corresponding sample proportions were 33.21%, 11.89%, 37.96%, and 10.18%, respectively. The largest differences occurred in sandy loam (+11.27 percentage points) and clay loam (−19.40 percentage points). The existing soil data product showed a broadly similar spatial pattern. Because the two classification systems differ in particle-size limits and class definitions, only qualitative spatial comparison was made. Detailed class distributions are provided in Supplementary Table S5.

3.6. Comparison of Mapping Results in Representative Subregions

Four representative subregions (3×3 km) were selected to compare RF, XGBoost, and MMVT mapping results (Figure 10). These included a desert–oasis transition zone, cropland, mountainous hills, and a river valley. The three models showed broadly similar spatial patterns but differed in particle-size proportions, local patch structures, and spatial continuity.

In the desert–oasis transition zone (subregion 1), all models showed a clear spatial transition. Sand-rich soils occurred in the west, while relatively silt-rich soils appeared in the east. Mean sand contents were 67.86% for RF, 69.11% for XGBoost, and 60.74% for MMVT. RF and XGBoost showed more fragmented pixel-scale variations, whereas MMVT exhibited more gradual spatial transitions. In cropland (subregion 2), all models distinguished the central area from surrounding fields. RF and XGBoost showed greater within-field variability. MMVT produced more continuous patches while retaining differences between adjacent fields.

In the mountainous and hilly area (subregion 3), the models differed in local patch shapes and spatial arrangements. RF and XGBoost showed frequent pixel-scale changes, whereas MMVT displayed more continuous patches. In the river valley (subregion 4), all models indicated relatively high sand contents near the river channel. They also showed soil texture differences on either side of the valley. MMVT showed more concentrated spatial patches and less localized variation than RF and XGBoost.

Quantitative results showed lower standard deviations and neighboring-pixel absolute gradients (NAGs) for MMVT across all fractions and subregions. For sand in subregion 3, MMVT had an NAG of 2.369 percentage points (pp). The corresponding values were 2.878 pp for RF and 3.796 pp for XGBoost. These results indicate smaller local spatial variations in the MMVT maps. Detailed statistics are provided in Supplementary Table S6.

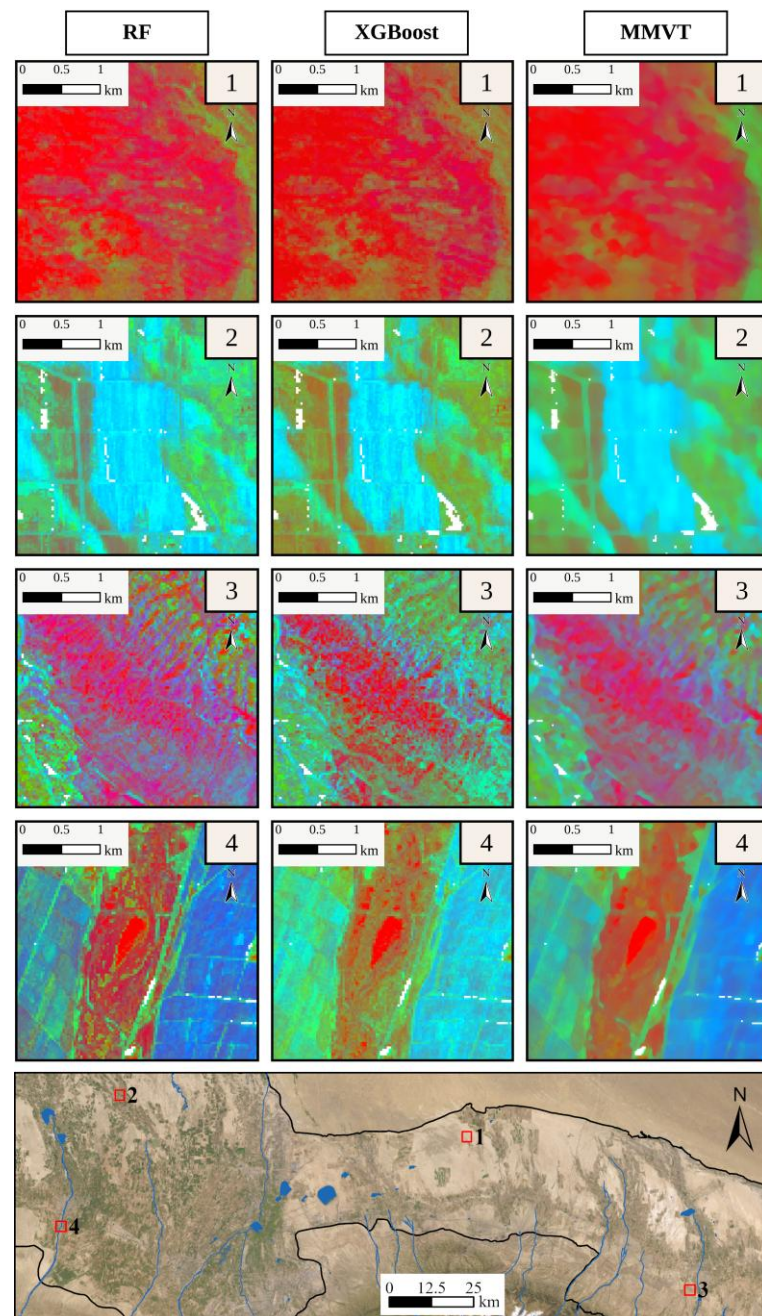


Figure 10. Sand–silt–clay RGB maps produced by RF, XGBoost, and MMVT in four representative subregions. The locator map shows the spatial distribution of the four subregions. Subregion 1 is a desert–oasis transition zone, subregion 2 is cropland, subregion 3 is a mountainous and hilly area, and subregion 4 is a river valley. All maps use the same RGB color stretch for consistent comparison.

4. Discussion

4.1. Contributions of Multimodal Environmental Variables to Soil Texture Prediction

The six data modalities describe soil texture from different aspects. They include bare-soil spectral response, radar backscatter, bare-soil temperature, long-term surface spectral conditions, long-term surface temperature, and terrain. This follows the general idea of digital soil mapping, where environmental variables are used to describe soil–environment relationships [56,57]. The six modalities are based on different types of observations, so they do not provide the same information. Their main role is to add different information for soil texture prediction rather than simply increase the number of variables [58,59].

The grouped permutation results showed that BSC had the highest relative importance among the six modalities in all three models. Its normalized modality importance was 62.68% for RF, 61.75% for XGBoost, and 50.39% for MMVT. This result was generally consistent with previous studies. Silvero et al. found that multitemporal bare-soil composites increased the number of usable bare-soil observations and improved the prediction of surface soil properties compared with single-date images [60]. Møller et al. also found that optical and radar bare-soil composites were useful for predicting mineral soil texture in Denmark. However, radar data made a larger contribution in their study [61]. These results show that the importance of each data type can differ among regions and processing methods. In the central section of the northern Tianshan slope, low rainfall and strong evaporation create seasonal periods of soil exposure. Surface moisture and tillage conditions, however, can vary between observation dates. Using bare-soil observations from several years reduces the effects of vegetation and short-term surface conditions. It also retains more stable spectral information related to soil color, mineral composition, and particle-size differences [62,63].

Terrain and LSIC also contributed to soil texture prediction, whereas SAR, BST, and LST showed lower grouped permutation importance. BSC and LSIC shared some HLS-derived spectral information but represented different surface conditions. BSC focused on bare-soil spectral responses, whereas LSIC described long-term land-surface characteristics. Their spectral overlap may influence the relative importance assigned to these modalities, as reported for correlated predictors [64]. Nevertheless, BSC consistently showed the highest grouped permutation importance across all three models, highlighting its relevance for soil texture prediction.

4.2. Model Performance and Differences Among Particle-Size Fractions

Spatial cross-validation produced lower prediction accuracy than repeated random testing, reflecting the greater difficulty of predicting geographically separated samples. Random splitting may retain spatial similarities between training and test samples, whereas spatial validation provides a more demanding evaluation. Nevertheless, MMVT maintained relatively high prediction accuracy under spatial cross-validation, particularly with the 10×10 window. This suggests that incorporating neighborhood information can help improve soil texture prediction accuracy.

Prediction accuracy varied among soil particle-size fractions and study regions. Clay showed relatively strong predictive performance in the present study. In contrast, Swain et al. [65] reported relatively high accuracy for sand in the western catchment of Chilika Lagoon, India. Their R^2 values were 0.62, 0.54, and 0.54 for sand, silt, and clay, respectively. Wittstruck et al. [26] also reported relatively high accuracy for sand across German agricultural soils. Their corresponding R^2 values were 0.74, 0.73, and 0.52. These differences may reflect regional variations in soil-forming conditions and the spatial distribution of particle-size fractions. In the northern Tianshan piedmont–oasis system, alluvial–proluvial transport and sediment sorting influence the redistribution of fine particles [17,29]. Clay enrichment in parts of the oasis plain may produce relatively continuous spatial gradients associated with landform position. These gradients may be captured by terrain variables and bare-soil spectral characteristics. Meanwhile, aeolian inputs and mixed sediment sources may increase local variations in sand and silt. Such regional characteristics may partly explain the relatively strong predictive performance for clay in this study.

Prediction performance also depends on the spatial extent and environmental complexity of the study area. A national 90 m soil texture product for China reported overall R^2 values of 0.43–0.50 [66]. This finding provides a broader context for the challenges of predicting soil texture across diverse environmental conditions. The relatively focused

piedmont–oasis setting of this study may provide clearer environmental gradients for soil texture prediction. Such gradients could partly contribute to the relatively strong predictive performance for clay under spatial cross-validation.

4.3. Spatial Differentiation Patterns of Soil Texture

Loam (34.28%), sandy loam (23.16%), clay loam (18.56%), and loamy clay (9.14%) dominated the mapped soil texture classes. Their spatial distribution generally followed the mountain–oasis–desert gradient along the northern Tianshan slope. Sand-rich soils along the northern desert margin may reflect aeolian inputs from the nearby Gurbantunggut Desert [59,67]. In the southeastern piedmont hills, sand and silt occurred in similar proportions, suggesting mixed sedimentary influences. Hillslope erosion, gully runoff, piedmont flooding, and windblown dust may jointly contribute to these deposits [29,67,68]. In contrast, clay was relatively enriched in the western oasis plain.

Regional differences may also be associated with the north–south width of the plains and river discharge. The western plain has a broader transition from the mountain front to the northern lowlands than the eastern plain. The Manas River in the western sector has greater annual runoff than rivers in the eastern part of the study area [69]. As rivers leave the mountains, declining slopes and transport energy promote sediment sorting [70]. Coarser particles generally settle earlier, while finer particles can travel farther downstream. The broader western plain and greater river discharge may favor the transport and accumulation of fine sediments. These conditions may partly explain the relatively higher clay content in the western oasis.

Differences between sample and mapped class proportions were particularly evident for sandy loam (+11.27 pp) and clay loam (−19.40 pp). The original sampling campaign considered site accessibility and spatial differences in soil properties across the study area. Field access was limited in some remote mountainous and grassland areas, contributing to variations in sampling density. Previous digital soil mapping studies have also reported the influence of accessibility on sample distributions [71]. Furthermore, small prediction differences near ISSS classification boundaries may alter the mapped proportions of individual texture classes.

Approximately 85% of the mapped area fell within the AOA, while the remaining 15% was outside. For clay, spatial OOF R^2 was 0.658 inside the AOA and 0.530 outside. Predictions within the AOA mainly represent interpolation under environmental conditions covered by the training samples. Higher DI values in urban areas, river valleys, and ecological transition zones indicate greater extrapolation potential. The spatial distribution of sampling sites may partly contribute to the higher DI values in urban areas. Urban impervious surfaces also have distinct spectral and thermal characteristics compared with natural soil surfaces [72]. These differences may further increase their environmental dissimilarity to the sampled locations. The AOA analysis therefore provides additional information for interpreting the spatial applicability of soil texture predictions.

4.4. Limitations and Future Directions

This study integrated field measurements, multi-year bare-soil composites, and environmental variables for regional soil texture prediction. Comparisons among RF, XGBoost, and MMVT addressed predictive accuracy, spatial context, modality contributions, and model applicability. Nevertheless, several limitations warrant consideration. To begin with, optical and thermal observations mainly characterize near-surface conditions, whereas measured particle-size fractions represent the integrated 0–30 cm soil layer. In areas with pronounced vertical textural stratification, surface observations may not adequately represent subsurface fractions, potentially contributing to local prediction errors [73]. In addition, the selected spatial window and bare-soil compositing strategy reflect

the sampling distribution, landscape characteristics, crop phenology, and soil exposure conditions of this region. Their transferability, particularly that of seasonal observation windows and spectral thresholds, requires further evaluation in other environments [74]. Furthermore, satellite and terrain variables were prioritized to ensure consistent spatial coverage. Other potentially informative factors, including parent material, climate, hydrology, organic matter, and salinity, were excluded because available datasets differed in spatial scale or temporal coverage. Their omission may limit the representation of soil-forming processes and local variability. Finally, this study focused on matched-scale comparisons among three covariate-driven models, without incorporating an ordinary kriging baseline [75]. Direct quantitative comparisons with existing soil products were also not conducted because of differences in particle-size classification standards. These limitations restrict the assessment of model performance relative to conventional interpolation methods and existing soil products. Future studies should incorporate geostatistical benchmarks, harmonized soil-product comparisons, depth-specific measurements, additional environmental variables, independent cross-regional validation, and pixel-level uncertainty assessment. Overall, integrating multi-year bare-soil observations and neighborhood information shows potential for regional soil texture prediction and may support soil resource surveys and management in arid oasis environments.

5. Conclusions

This study integrated 1051 soil samples and multi-year remote sensing observations to map soil texture in the central northern Tianshan slope. RF, XGBoost, and MMVT were evaluated through repeated external testing and five-fold spatial cross-validation. The main conclusions are as follows:

(1) Multi-source environmental variables provided useful information for soil texture prediction. BSC showed substantial normalized grouped permutation importance, reaching 62.68%, 61.75%, and 50.39% for RF, XGBoost, and MMVT, respectively. These findings highlight the value of multi-year bare-soil observations for regional soil texture mapping.

(2) Incorporating neighborhood information contributed to soil texture prediction. MMVT showed relatively strong performance using a 10×10 -pixel window (300×300 m). Across 50 repeated external tests, mean R^2 values were 0.586, 0.519, and 0.706 for sand, silt, and clay, respectively. Under five-fold spatial cross-validation, the corresponding mean R^2 values were 0.550, 0.504, and 0.657. Despite lower R^2 values under spatial validation, MMVT maintained useful predictive performance for all three fractions in this study.

(3) The results revealed distinct regional patterns of soil texture. Sand was relatively abundant along the northern desert margin, while clay was enriched in the western oasis plain. Loam, sandy loam, clay loam, and loamy clay accounted for 34.28%, 23.16%, 18.56%, and 9.14%, respectively. Approximately 85% of the mapped area fell within the AOA. Higher DI values occurred mainly in urban areas, river valleys, and ecological transition zones, indicating greater environmental dissimilarity in these locations.

In conclusion, this study highlights the potential of integrating multi-year bare-soil observations, multi-source environmental information, and neighborhood context for soil texture prediction. The findings may provide useful reference information for soil resource surveys and agricultural management in the central section of the northern Tianshan slope.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/agronomy16212079/s1>. Table S1. Spectral indices and their formulas used in this study; Table S2. Hyperparameter and training configurations of RF, XGBoost, and MMVT; Table S3. International soil texture classification; Table S4. Aitchison distance of soil

texture predictions under five-fold spatial cross-validation using a 10×10 window; Table S5. Comparison of observed and mapped proportions across the 12 ISSS soil texture classes; Table S6. Comparison of soil texture spatial variability and neighboring pixel gradients among RF, XGBoost, and MMVT models across four representative subregions; Figure S1. Empirical semivariograms and spatial autocorrelation ranges of soil particle-size fractions; Figure S2. Distributions of RF, XGBoost, and MMVT performance with 1×1 and 10×10 windows across 50 repeated random holdout tests; Figure S3. Observed versus predicted ILR coordinates for RF, XGBoost, and MMVT using 10×10 windows; Figure S4. Residual analysis of MMVT spatial out-of-fold predictions for sand, silt, and clay; Figure S5. Confusion matrices of ISSS soil texture classes predicted by RF, XGBoost, and MMVT using 10×10 windows; Figure S6. Comparison of MMVT prediction performance inside and outside the area of applicability.

Author Contributions: Conceptualization, X.M. and H.W.; methodology, Y.F.; software, X.M.; validation, X.M., Y.F. and W.M.; formal analysis, X.M.; investigation, W.M., H.Z., P.W. and R.Z.; resources, H.W. and Y.F.; data curation, X.M., W.M., H.Z., P.W. and R.Z.; writing—original draft preparation, X.M.; writing—review and editing, H.W. and Y.F.; visualization, X.M.; supervision, H.W. and Y.F.; project administration, H.W. and Y.F.; funding acquisition, H.W. and Y.F. All authors have read and agreed to the published version of the manuscript.

Funding: Integration and Demonstration of Technical Models for Soil Improvement and Capacity Enhancement in Northwest Saline Soil Area 2023YFD1901503.

Data Availability Statement: The data presented in this study are available on request from the corresponding authors due to privacy reasons.

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

References

1. Vereecken, H.; Amelung, W.; Bauke, S.L.; Bogaen, H.; Brüggemann, N.; Montzka, C.; Vanderborght, J. Soil Hydrology in the Earth System. *Nat. Rev. Earth Environ.* **2022**, *3*, 573–587. <https://doi.org/10.1038/s43017-022-00324-6>.
2. Mgohele, R.N.; Massawe, B.H.J.; Shitindi, M.J.; Sanga, H.G.; Omar, M.M. Prediction of Soil Texture Using Remote Sensing Data: A Systematic Review. *Front. Remote Sens.* **2024**, *5*, 1461537. <https://doi.org/10.3389/frsen.2024.1461537>.
3. State Council of the People's Republic of China. Notice of the State Council on Conducting the Third National Soil Survey. In *Bulletin of the State Council of the People's Republic of China*; State Council of the People's Republic of China: Beijing, China, 2022; pp. 29–31.
4. Zhao, J.; Wu, H.; Gu, H.; Fan, Y.; Zhao, Z.; Wang, P.; Li, C. Climate Surpasses Soil Texture in Driving Soil Salinization Alleviation in Arid Xinjiang. *Remote Sens.* **2025**, *17*, 3812. <https://doi.org/10.3390/rs17233812>.
5. Wadoux, A.M.J.-C.; Minasny, B.; McBratney, A.B. Machine Learning for Digital Soil Mapping: Applications, Challenges and Suggested Solutions. *Earth-Sci. Rev.* **2020**, *210*, 103359. <https://doi.org/10.1016/j.earscirev.2020.103359>.
6. Liu, F.; Wu, H.; Zhao, Y.; Li, D.; Yang, J.; Song, X.; Shi, Z.; Zhu, A.-X.; Zhang, G. Mapping High Resolution National Soil Information Grids of China. *Sci. Bull.* **2022**, *67*, 328–340. <https://doi.org/10.1016/j.scib.2021.10.013>.
7. Poggio, L.; de Sousa, L.M.; Batjes, N.H.; Heuvelink, G.B.M.; Kempen, B.; Ribeiro, E.; Rossiter, D. SoilGrids 2.0: Producing Soil Information for the Globe with Quantified Spatial Uncertainty. *SOIL* **2021**, *7*, 217–240. <https://doi.org/10.5194/soil-7-217-2021>.
8. Tziolas, N.; Tsakiridis, N.; Chabrilat, S.; Dematté, J.A.; Ben-Dor, E.; Gholizadeh, A.; Zalidis, G.; Van Wesemael, B. Earth Observation Data-Driven Cropland Soil Monitoring: A Review. *Remote Sens.* **2021**, *13*, 4439. <https://doi.org/10.3390/rs13214439>.
9. Mirzaeitalarposhti, R.; Shafizadeh-Moghadam, H.; Taghizadeh-Mehrjardi, R.; Demyan, M.S. Digital Soil Texture Mapping and Spatial Transferability of Machine Learning Models Using Sentinel-1, Sentinel-2, and Terrain-Derived Covariates. *Remote Sens.* **2022**, *14*, 5909. <https://doi.org/10.3390/rs14235909>.
10. Zepp, S.; Heiden, U.; Bachmann, M.; Möller, M.; Wiesmeier, M.; van Wesemael, B. Optimized Bare Soil Compositing for Soil Organic Carbon Prediction of Topsoil Croplands in Bavaria Using Landsat. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 287–302. <https://doi.org/10.1016/j.isprsjprs.2023.06.003>.

11. Castaldi, F.; Koparan, M.H.; Wetterlind, J.; Żydelis, R.; Vinci, I.; Savaş, A.Ö.; Kıvrak, C.; Tunçay, T.; Volungevičius, J.; Obber, S. Assessing the Capability of Sentinel-2 Time-Series to Estimate Soil Organic Carbon and Clay Content at Local Scale in Croplands. *ISPRS J. Photogramm. Remote Sens.* **2023**, *199*, 40–60. <https://doi.org/10.1016/j.isprsjprs.2023.03.016>.
12. Dvorakova, K.; Heiden, U.; Pepers, K.; Staats, G.; van Os, G.; van Wesemael, B. Improving Soil Organic Carbon Predictions from a Sentinel-2 Soil Composite by Assessing Surface Conditions and Uncertainties. *Geoderma* **2023**, *429*, 116128. <https://doi.org/10.1016/j.geoderma.2022.116128>.
13. Rogge, D.; Bauer, A.; Zeidler, J.; Mueller, A.; Esch, T.; Heiden, U. Building an Exposed Soil Composite Processor (SCMaP) for Mapping Spatial and Temporal Characteristics of Soils with Landsat Imagery (1984–2014). *Remote Sens. Environ.* **2018**, *205*, 1–17. <https://doi.org/10.1016/j.rse.2017.11.004>.
14. Bousbih, S.; Zribi, M.; Pelletier, C.; Gorraeb, A.; Lili-Chabaane, Z.; Baghdadi, N.; Aissa, N.B.; Mougnot, B. Soil Texture Estimation Using Radar and Optical Data from Sentinel-1 and Sentinel-2. *Remote Sens.* **2019**, *11*, 1520. <https://doi.org/10.3390/rs11131520>.
15. He, W.; Xiao, Z.; Lu, Q.; Wei, L.; Liu, X. Digital Mapping of Soil Particle Size Fractions in the Loess Plateau, China, Using Environmental Variables and Multivariate Random Forest. *Remote Sens.* **2024**, *16*, 785. <https://doi.org/10.3390/rs16050785>.
16. McBratney, A.B.; Mendonça Santos, M.L.; Minasny, B. On Digital Soil Mapping. *Geoderma* **2003**, *117*, 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4).
17. Gebauer, A.; Sakhaee, A.; Don, A.; Poggio, M.; Ließ, M. Topsoil Texture Regionalization for Agricultural Soils in Germany—An Iterative Approach to Advance Model Interpretation. *Front. Soil Sci.* **2022**, *1*, 770326. <https://doi.org/10.3389/fsoil.2021.770326>.
18. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/10.1023/A:1010933404324>.
19. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016*; Association for Computing Machinery: New York, NY, USA; pp. 785–794. <https://doi.org/10.1145/2939672.2939785>.
20. Padarian, J.; Minasny, B.; McBratney, A.B. Using Deep Learning for Digital Soil Mapping. *Soil* **2019**, *5*, 79–89. <https://doi.org/10.5194/soil-5-79-2019>.
21. Tziolas, N.; Tsakiridis, N.; Heiden, U.; van Wesemael, B. Soil Organic Carbon Mapping Utilizing Convolutional Neural Networks and Earth Observation Data: A Case Study in Bavaria, Germany. *Geoderma* **2024**, *444*, 116867. <https://doi.org/10.1016/j.geoderma.2024.116867>.
22. Ließ, M.; Sakhaee, A. Deep Learning with a Multi-Task Convolutional Neural Network to Generate a National-Scale 3D Soil Data Product: The Particle Size Distribution of the German Agricultural Soil Landscape. *Agriculture* **2024**, *14*, 1230. <https://doi.org/10.3390/agriculture14081230>.
23. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16 × 16 Words: Transformers for Image Recognition at Scale. In *Proceedings of the International Conference on Learning Representations, Vienna, Austria, 4 May 2021*. Available online: <https://openreview.net/forum?id=YicbFdNTTy> (accessed on 23 July 2026).
24. Kakhani, N.; Rangzan, M.; Jamali, A.; Attarchi, S.; Alavipanah, S.K.; Mommert, M.; Tziolas, N.; Scholten, T. SSL-SoilNet: A Hybrid Transformer-Based Framework with Self-Supervised Learning for Large-Scale Soil Organic Carbon Prediction. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–15. <https://doi.org/10.1109/TGRS.2024.3446042>.
25. Tresson, P.; Dumont, M.; Jaeger, M.; Borne, F.; Boivin, S.; Marie-Louise, L.; François, J.; Boukcim, H.; Goëau, H. Self-Supervised Learning of Vision Transformers for Digital Soil Mapping Using Visual Data. *Geoderma* **2024**, *450*, 117056. <https://doi.org/10.1016/j.geoderma.2024.117056>.
26. Wittstruck, L.; Waske, B.; Jarmer, T. Multi-Modal Vision Transformer for High-Resolution Soil Texture Prediction of German Agricultural Soils Using Remote Sensing Imagery. *Remote Sens. Environ.* **2025**, *331*, 114985. <https://doi.org/10.1016/j.rse.2025.114985>.
27. Zhang, M.; Shi, W.; Ma, Y.; Ge, Y. Mapping Soil Particle-Size Fractions Based on Compositional Balances. *Catena* **2024**, *234*, 107643. <https://doi.org/10.1016/j.catena.2023.107643>.
28. Egozcue, J.J.; Pawłowsky-Glahn, V.; Mateu-Figueras, G.; Barceló-Vidal, C. Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* **2003**, *35*, 279–300. <https://doi.org/10.1023/A:1023818214614>.
29. Yao, J.; Chen, Y.; Guan, X.; Zhao, Y.; Chen, J.; Mao, W. Recent Climate and Hydrological Changes in a Mountain–Basin System in Xinjiang, China. *Earth-Sci. Rev.* **2022**, *226*, 103957. <https://doi.org/10.1016/j.earscirev.2022.103957>.
30. HJ 1068-2019; Soil—Determination of Particle Size Distribution-Pipette Method and Hydrometer Method. China Environmental Publishing Group: Beijing, China, 2019.

31. Juo, A.S.R.; Franzluebbers, K. Soil Physics. In *Tropical Soils: Properties and Management for Sustainable Agriculture*; Oxford University Press: New York, NY, USA, 2003; pp. 47–70, ISBN 978-0-19-511598-7. <https://doi.org/10.1093/oso/9780195115987.003.0007>.
32. Zhou, T.; Geng, Y.; Lv, W.; Xiao, S.; Zhang, P.; Xu, X.; Chen, J.; Wu, Z.; Pan, J.; Si, B.; et al. Effects of Optical and Radar Satellite Observations within Google Earth Engine on Soil Organic Carbon Prediction Models in Spain. *J. Environ. Manag.* **2023**, *338*, 117810. <https://doi.org/10.1016/j.jenvman.2023.117810>.
33. Shafizadeh-Moghadam, H.; Minaei, F.; Talebi-Khiyavi, H.; Xu, T.; Homaei, M. Synergetic Use of Multi-Temporal Sentinel-1, Sentinel-2, NDVI, and Topographic Factors for Estimating Soil Organic Carbon. *Catena* **2022**, *212*, 106077. <https://doi.org/10.1016/j.catena.2022.106077>.
34. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
35. Ju, J.; Zhou, Q.; Freitag, B.; Roy, D.P.; Zhang, H.K.; Sridhar, M.; Mandel, J.; Arab, S.; Schmidt, G.; Crawford, C.J.; et al. The Harmonized Landsat and Sentinel-2 Version 2.0 Surface Reflectance Dataset. *Remote Sens. Environ.* **2025**, *324*, 114723. <https://doi.org/10.1016/j.rse.2025.114723>.
36. Urbina-Salazar, D.; Vaudour, E.; Richer-de-Forges, A.C.; Chen, S.; Martelet, G.; Baghdadi, N.; Arrouays, D. Sentinel-2 and Sentinel-1 Bare Soil Temporal Mosaics of 6-Year Periods for Soil Organic Carbon Content Mapping in Central France. *Remote Sens.* **2023**, *15*, 2410. <https://doi.org/10.3390/rs15092410>.
37. Gholizadeh, A.; Borůvka, L.; Saberioon, M.; Vašát, R. Soil Organic Carbon and Texture Retrieving and Mapping Using Proximal, Airborne and Sentinel-2 Spectral Imaging. *Remote Sens. Environ.* **2018**, *218*, 89–103. <https://doi.org/10.1016/j.rse.2018.09.015>.
38. Miranda, N.; Torres, R.; Geudtner, D.; Pinheiro, M.; Potin, P.; Gratadour, J.-B.; O’Connell, A.; Bibby, D.; Navas-Traver, I.; Cossu, M. Sentinel-1 First Generation Status, Past and Future. In *Proceedings of the IGARSS 2023—IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023*; IEEE: New York, NY, USA, 2023; pp. 4560–4563. <https://doi.org/10.1109/IGARSS52108.2023.10281990>.
39. Tadono, T.; Ohki, M.; Motohka, T.; Rosenqvist, A.; Sobue, S. Advanced Land Observing Satellite-2 “DAICHI-2” (ALOS-2): Overview and Operational Status. *IEEE Geosci. Remote Sens. Mag.* **2025**, *13*, 41–48. <https://doi.org/10.1109/MGRS.2025.3535720>.
40. Crawford, C.J.; Roy, D.P.; Arab, S.; Barnes, C.; Vermote, E.; Hulley, G.; Gerace, A.; Choate, M.; Engebretson, C.; Micijevic, E.; et al. The 50-Year Landsat Collection 2 Archive. *Sci. Remote Sens.* **2023**, *8*, 100103. <https://doi.org/10.1016/j.srs.2023.100103>.
41. Wang, D.C.; Zhang, G.L.; Zhao, M.S.; Pan, X.Z.; Zhao, Y.G.; Li, D.C.; Macmillan, B. Retrieval and Mapping of Soil Texture Based on Land Surface Diurnal Temperature Range Data from MODIS. *PLoS ONE* **2015**, *10*, e0129977. <https://doi.org/10.1371/journal.pone.0129977>.
42. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 Surface Reflectance Data Set. *Remote Sens. Environ.* **2018**, *219*, 145–161. <https://doi.org/10.1016/j.rse.2018.09.002>.
43. Maynard, J.J.; Levi, M.R. Hyper-Temporal Remote Sensing for Digital Soil Mapping: Characterizing Soil-Vegetation Response to Climatic Variability. *Geoderma* **2017**, *285*, 94–109. <https://doi.org/10.1016/j.geoderma.2016.09.024>.
44. U.S. Geological Survey. *Landsat 8–9 Collection 2 Level-2 Science Product Guide, Version 6.0*; U.S. Geological Survey: Sioux Falls, SD, USA, 2024.
45. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, RG2004. <https://doi.org/10.1029/2005RG000183>.
46. Beven, K.J.; Kirkby, M.J. A Physically Based, Variable Contributing Area Model of Basin Hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. <https://doi.org/10.1080/02626667909491834>.
47. Aitchison, J. The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1982**, *44*, 139–160. <https://doi.org/10.1111/j.2517-6161.1982.tb01195.x>.
48. Amirian-Chakan, A.; Minasny, B.; Taghizadeh-Mehrjardi, R.; Akbarifazli, R.; Darvishpasand, Z.; Khordehbin, S. Some Practical Aspects of Predicting Texture Data in Digital Soil Mapping. *Soil Tillage Res.* **2019**, *194*, 104289. <https://doi.org/10.1016/j.still.2019.06.006>.
49. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In *Proceedings of the Advances in Neural Information Processing Systems*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30, pp. 5998–6008. Available online: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (accessed on 28 July 2026).

50. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography* **2017**, *40*, 913–929. <https://doi.org/10.1111/ecog.02881>.
51. Valavi, R.; Elith, J.; Lahoz-Monfort, J.J.; Guillera-Arroita, G. blockCV: An r Package for Generating Spatially or Environmentally Separated Folds for k-Fold Cross-Validation of Species Distribution Models. *Methods Ecol. Evol.* **2019**, *10*, 225–232. <https://doi.org/10.1111/2041-210X.13107>.
52. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.-M.; McBratney, A. Critical Review of Chemometric Indicators Commonly Used for Assessing the Quality of the Prediction of Soil Attributes by NIR Spectroscopy. *TrAC Trends Anal. Chem.* **2010**, *29*, 1073–1081. <https://doi.org/10.1016/j.trac.2010.05.006>.
53. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinform.* **2007**, *8*, 25. <https://doi.org/10.1186/1471-2105-8-25>.
54. Fisher, A.; Rudin, C.; Dominici, F. All Models Are Wrong, but Many Are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously. *J. Mach. Learn. Res.* **2019**, *20*, 177.
55. Meyer, H.; Pebesma, E. Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models. *Methods Ecol. Evol.* **2021**, *12*, 1620–1633. <https://doi.org/10.1111/2041-210X.13650>.
56. Broeg, T.; Don, A.; Gocht, A.; Scholten, T.; Taghizadeh-Mehrjardi, R.; Erasmi, S. Using Local Ensemble Models and Landsat Bare Soil Composites for Large-Scale Soil Organic Carbon Maps in Cropland. *Geoderma* **2024**, *444*, 116850. <https://doi.org/10.1016/j.geoderma.2024.116850>.
57. Azizi, K.; Garosi, Y.; Ayoubi, S.; Tajik, S. Integration of Sentinel-1/2 and Topographic Attributes to Predict the Spatial Distribution of Soil Texture Fractions in Some Agricultural Soils of Western Iran. *Soil Tillage Res.* **2023**, *229*, 105681. <https://doi.org/10.1016/j.still.2023.105681>.
58. Shokri, N.; Hassani, A.; Sahimi, M. Multi-Scale Soil Salinization Dynamics from Global to Pore Scale: A Review. *Rev. Geophys.* **2024**, *62*, e2023RG000804. <https://doi.org/10.1029/2023RG000804>.
59. Cao, Q.; Li, J.; Wang, G.; Wang, D.; Xin, Z.; Xiao, H. On the Spatial Variability and Influencing Factors of Soil Organic Carbon and Total Nitrogen Stocks in a Desert Oasis Ecotone of Northwestern China. *Catena* **2021**, *206*, 105533. <https://doi.org/10.1016/j.catena.2021.105533>.
60. Silvero, N.E.Q.; Demattê, J.A.M.; Amorim, M.T.A.; dos Santos, N.V.; Rizzo, R.; Safanelli, J.L.; Poppiel, R.R.; Mendes, W.d.S.; Bonfatti, B.R. Soil Variability and Quantification Based on Sentinel-2 and Landsat-8 Bare Soil Images: A Comparison. *Remote Sens. Environ.* **2021**, *252*, 112117. <https://doi.org/10.1016/j.rse.2020.112117>.
61. Møller, A.B.; Nyborg, L.; Grogan, K.; Druce, D.; Svane, S.F.; Greve, M.B.; Gutierrez, S.; Styczen, M.; Greve, M.H.; Knudsen, L.; et al. High-Resolution 3D Soil Texture Mapping in Denmark Using Satellite Time Series and Bare Soil Composites. In *Advances in Agronomy*; Sparks, D.L., Ed.; Academic Press: Cambridge, MA, USA, 2025; Volume 194, pp. 133–186. <https://doi.org/10.1016/bs.agron.2025.07.001>.
62. Guo, L.; Sun, X.; Fu, P. Mapping Soil Organic Carbon Stock by Hyperspectral and Time-Series Multispectral Remote Sensing Images in Low-Relief Agricultural Areas. *Geoderma* **2021**, *398*, 115118. <https://doi.org/10.1016/j.geoderma.2021.115118>.
63. Demattê, J.A.M.; Safanelli, J.L.; Poppiel, R.R.; Rizzo, R.; Silvero, N.E.Q.; Mendes, W.d.S.; Bonfatti, B.R.; Dotto, A.C.; Salazar, D.F.U. Bare Earth's Surface Spectra as a Proxy for Soil Resource Monitoring. *Sci. Rep.* **2020**, *10*, 4461. <https://doi.org/10.1038/s41598-020-61408-1>.
64. Kaneko, H. Cross-Validated Permutation Feature Importance Considering Correlation between Features. *Anal. Sci. Adv.* **2022**, *3*, 278–287. <https://doi.org/10.1002/ansa.202200018>.
65. Swain, S.R.; Chakraborty, P.; Panigrahi, N.; Vasava, H.B.; Reddy, N.N.; Roy, S.; Majeed, I.; Das, B.S. Estimation of Soil Texture Using Sentinel-2 Multispectral Imaging Data: An Ensemble Modeling Approach. *Soil Tillage Res.* **2021**, *213*, 105134. <https://doi.org/10.1016/j.still.2021.105134>.
66. Liu, F.; Zhang, G.-L.; Song, X.; Li, D.; Zhao, Y.; Yang, J.; Wu, H.; Yang, F. High-Resolution and Three-Dimensional Mapping of Soil Texture of China. *Geoderma* **2020**, *361*, 114061. <https://doi.org/10.1016/j.geoderma.2019.114061>.
67. Wang, J.; Tian, Q.; Yan, A.; Liu, X.; Jiao, Z.; Bai, Z. Discriminating Arid Ecosystems and Tracking Desertification Dynamics: Evidence From Grain-Size End-Member in the Manas River Basin, Northern Tianshan Mountains. *Land Degrad. Dev.* **2026**, *37*, 10174–10186. <https://doi.org/10.1002/ldr.70762>.
68. Ventra, D.; Clarke, L.E. Geology and Geomorphology of Alluvial and Fluvial Fans: Current Progress and Research Perspectives. *Geol. Soc. Lond. Spec. Publ.* **2018**, *440*, 1–21. <https://doi.org/10.1144/SP440.16>.

69. Zhou, L.; Chen, F.; He, C.; Wu, F.; Zhang, R. Spatiotemporal Dynamics and Attribution of Meteorological-Hydrological Drought Propagation in the Central Section of the Northern Tianshan Mountains. *Geomat. Nat. Hazards Risk* **2026**, *17*, 2649445. <https://doi.org/10.1080/19475705.2026.2649445>.
70. Latrubesse, E.M. Large Rivers, Megafans and Other Quaternary Avulsive Fluvial Systems: A Potential “Who’s Who” in the Geological Record. *Earth-Sci. Rev.* **2015**, *146*, 1–30. <https://doi.org/10.1016/j.earscirev.2015.03.004>.
71. Rossiter, D.G.; Poggio, L.; Beaudette, D.; Libohova, Z. How Well Does Digital Soil Mapping Represent Soil Geography? An Investigation from the USA. *Soil* **2022**, *8*, 559–586. <https://doi.org/10.5194/soil-8-559-2022>.
72. Su, S.; Tian, J.; Dong, X.; Tian, Q.; Wang, N.; Xi, Y. An Impervious Surface Spectral Index on Multispectral Imagery Using Visible and Near-Infrared Bands. *Remote Sens.* **2022**, *14*, 3391. <https://doi.org/10.3390/rs14143391>.
73. Laborczi, A.; Szatmári, G.; Kaposi, A.D.; Pásztor, L. Comparison of Soil Texture Maps Synthetized from Standard Depth Layers with Directly Compiled Products. *Geoderma* **2019**, *352*, 360–372. <https://doi.org/10.1016/j.geoderma.2018.01.020>.
74. Xue, J.; Zhang, X.; Huang, Y.; Chen, S.; Dai, L.; Chen, X.; Yu, Q.; Ye, S.; Shi, Z. A Two-Dimensional Bare Soil Separation Framework Using Multi-Temporal Sentinel-2 Images across China. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *134*, 104181. <https://doi.org/10.1016/j.jag.2024.104181>.
75. Oliver, M.A.; Webster, R. A Tutorial Guide to Geostatistics: Computing and Modelling Variograms and Kriging. *Catena* **2014**, *113*, 56–69. <https://doi.org/10.1016/j.catena.2013.09.006>.

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.