



Article

Optimized Soil Moisture Mapping Strategies on the Tibetan Plateau Using Downscaled and Interpolated Maps as Mutual Covariates

Mo Zhang ^{1,2}, Yong Ge ^{1,2,3,4,*} and Jianghao Wang ^{1,2}

¹ State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; zhangmo@lreis.ac.cn (M.Z.); wangjh@lreis.ac.cn (J.W.)

² College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100101, China

³ Key Laboratory of Poyang Lake Wetland and Watershed Research Ministry of Education, Jiangxi Normal University, Nanchang 330022, China

⁴ Key Laboratory of Intelligent Monitoring and Comprehensive Management of Watershed Ecology, Nanchang 330022, China

* Correspondence: gey@lreis.ac.cn

Abstract: Accurate high-resolution soil moisture maps are crucial for a better understanding of hydrological processes and energy cycles. Mapping strategies such as downscaling and interpolation have been developed to obtain high-resolution soil moisture maps from multi-source inputs. However, research on the optimization performance of integrating downscaling and interpolation, especially through the use of mutual covariates, remains unclear. In this study, we compared four methods—two standalone methods based on downscaling and interpolation strategies and two combined methods that utilize soil moisture maps as mutual covariates within each strategy—in a case study of daily soil moisture mapping at a 1 km resolution in the Tibetan Plateau. We assessed mapping performance in terms of prediction accuracy and differences in spatial coverage. The results indicated that introducing interpolated soil moisture maps into the downscaling strategy significantly improved prediction accuracy (RMSE: -5.94% , correlation coefficient: $+14.02\%$) but was limited to localized spatial coverage (6.9% of grid cells) near in situ sites. Conversely, integrating downscaled soil moisture maps into the interpolation strategy resulted in only modest gains in prediction accuracy (RMSE: -1.07% , correlation coefficient: $+1.04\%$), yet facilitated broader spatial coverage (40.4% of grid cells). This study highlights the critical differences between downscaling and interpolation strategies in terms of accuracy improvement and spatial coverage, providing a reference for optimizing soil moisture mapping over large areas.

Keywords: spatial prediction; machine learning; random forest; variable importance; digital soil mapping



Citation: Zhang, M.; Ge, Y.; Wang, J. Optimized Soil Moisture Mapping Strategies on the Tibetan Plateau Using Downscaled and Interpolated Maps as Mutual Covariates. *Remote Sens.* **2024**, *16*, 3939. <https://doi.org/10.3390/rs16213939>

Academic Editors: Junyu He, George Christakos and Jiaping Wu

Received: 4 September 2024

Revised: 18 October 2024

Accepted: 21 October 2024

Published: 23 October 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

Today, more research focuses on the processes, mechanisms, and interactions between soil moisture and natural factors such as vegetation, precipitation, and atmosphere [1–3]. These studies are typically modeled at a fine spatiotemporal scale. Therefore, more accurate soil moisture maps with high-resolution is crucial for a better understanding of hydrological processes and energy cycles [4–6].

The heterogeneity of soil moisture makes it highly variable and nonlinear in both time and space [7]. Soil moisture dynamics are influenced by many factors such as temperature [8], precipitation [9], vegetation [10], land cover [11], and soil properties [12]. For example, precipitation and temperature control the water cycle and evapotranspiration, which are primary drivers of long-term soil moisture variability [9,13]. Additionally, vegetation morphology is directly related to evapotranspiration and influences the feedback of evapotranspiration on soil moisture [10]. The latest advances in soil mapping

provide an unprecedented opportunity but also pose challenges due to inappropriate uses or inaccurate estimates of these products [14,15].

Various mapping strategies, including downscaling and interpolation, have been developed to obtain high-resolution soil moisture products by utilizing multi-source soil moisture inputs with different spatial supports. The downscaling strategy enhances the resolution of soil moisture maps by integrating coarse-resolution soil moisture satellite data with high-resolution auxiliary data [16–18]. It includes microwave data-fusion-based techniques [19], statistical models [20,21], and data assimilation techniques [22]. For example, Zheng et al. [23] constructed a global daily 1 km resolution soil moisture dataset using a downscaling method based on multi-source optical remote sensing data. However, satellite-based soil moisture data, such as those from the Soil Moisture Active Passive (SMAP) mission, offer wide coverage but often have resolutions exceeding tens of kilometers, which may not provide accurate spatial estimates of soil moisture at finer scales [24].

The interpolation strategy estimates soil moisture at unobserved locations by utilizing available site-scale measurements, thereby producing a continuous spatial distribution of soil moisture [25–27]. It involves incorporating covariates following the “SCORPAN” factors of digital soil mapping [28]. This strategy includes machine learning and geostatistical methods [29,30]. For example, Li et al. [31] presented a 1 km resolution long-term soil moisture dataset derived through the random-forest-based interpolation method trained by in situ observations in China. Despite in situ soil moisture data providing more accurate measurements, their limited spatial support leads to issues of insufficient spatial representativeness in sparsely observed areas [32].

With the increasing popularity of machine learning in spatial predictive modeling [33,34], a growing number of studies have used it to integrate multiple covariates and model their nonlinear relationships for predicting soil moisture through both downscaling and interpolation mapping strategies [18]. However, few studies have effectively combined downscaling and interpolation strategies to fully utilize the strengths of each method in soil moisture mapping, particularly in terms of integrating their results as mutual covariates. In addition, it is still necessary to assess the importance of various covariates in different strategies to ensure the appropriate selection of covariates and accurate estimation of soil moisture.

In this study, we assessed the performance of machine-learning-based downscaling and interpolation strategies for soil moisture mapping over the Tibetan Plateau, both with and without the inclusion of mutual covariates. Using random forest, we compared four methods: the downscaling method, the interpolation method, and two methods that combine soil moisture maps as mutual covariates with each strategy. These prediction maps may differ in accuracy and show variations across broader or more localized areas. Therefore, the mapping performance was evaluated based on prediction accuracy and differences in spatial coverage. To examine the contribution of different covariates in downscaling and interpolation strategies, we introduced 13 covariates across four categories—climate, vegetation, topography, and soil factors—to evaluate their influence on modeling. The objectives of this work are (i) to evaluate whether incorporating mutual covariates improves prediction accuracy compared to methods that do not use mutual covariates across different strategies; (ii) to analyze differences in spatial coverage between the soil moisture maps and determine whether these differences are concentrated in specific small areas or spread across larger regions; and (iii) to investigate the contributions of various types of covariates within different strategies.

2. Materials and Methods

2.1. Study Area

The Tibetan Plateau (TP), located in southwestern China and covering about 2.5 million km² (26.5°–40°N, 73°–105°E), is the largest and highest plateau in China. Its average elevation surpasses 4000 m, with the terrain rising higher in the west and sloping lower towards the east (Figure 1a). Known as the “roof of the world”, the TP is charac-

terized by extensive permafrost and glaciers [35,36]. It has significant variation in annual precipitation, ranging from approximately 1500 mm to less than 100 mm, mainly concentrated from June to September [37–39]. The dominant vegetation includes meadows and grasslands in the central and eastern regions, forests and shrubs in the southeast, and deserts in the northwest, with meadows and grasslands comprising about 70% of the total vegetation cover [40,41].

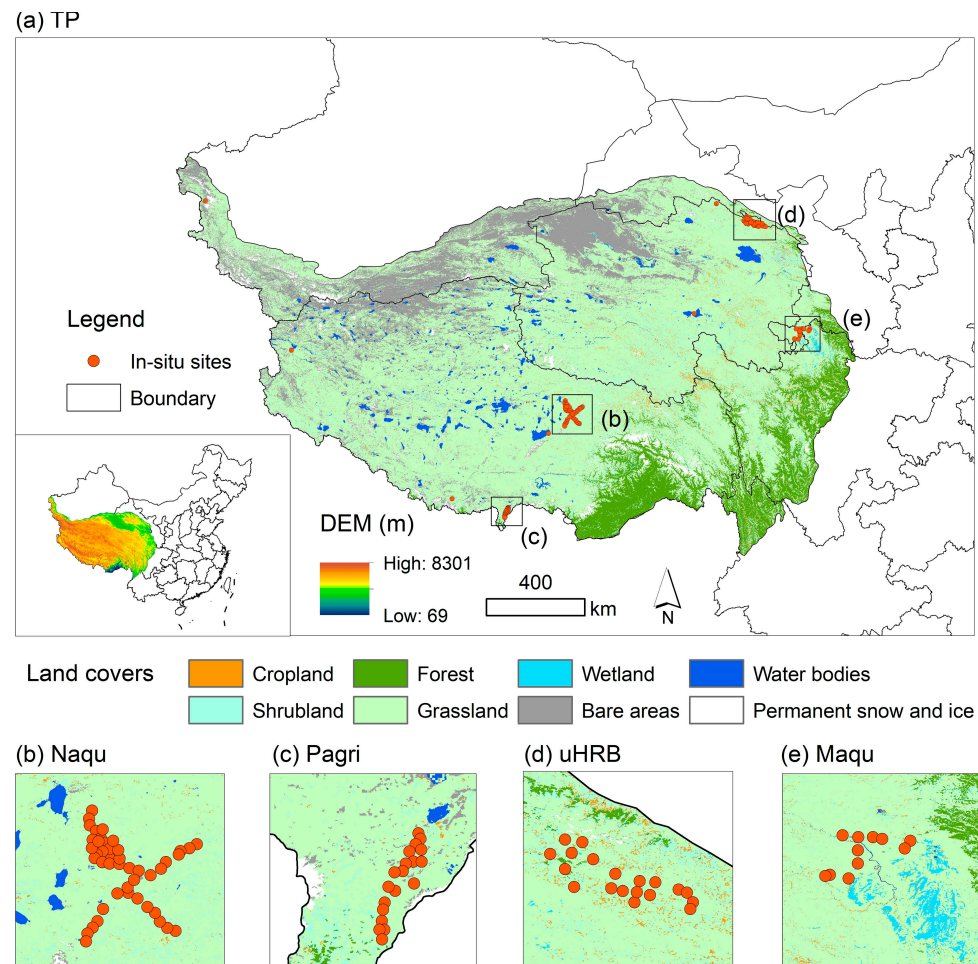


Figure 1. The location, land cover types, and elevation of the TP (a). Land covers and sensor locations for four networks: (b) Naqu, (c) Pagri, (d) uHRB, and (e) Maqu.

2.2. Data

2.2.1. In Situ Data

The in situ soil moisture observations originate from the Tibet-Obs network, which monitors soil temperature and moisture on the Tibetan Plateau (Tibet-Obs), and from the wireless sensor network in the upper reaches of the Heihe River Basin (uHRB), Qinghai Province, China [42,43]. Tibet-Obs includes stations at Naqu, Pagri, and Maqu (Figure 1b,c,e) [44,45], while uHRB is strategically located in the Babao River basin of the upper Heihe River (Figure 1d), and has been optimized for data collection [46]. Table 1 details these networks. Additionally, we utilized the most recently released soil moisture dataset from the long-term, quality-assured land–atmosphere interactions project [47], which includes sporadic data outside the four main networks, with a total of 9 samples. We processed data from the unfrozen period between June and September 2016. After ensuring data availability and high quality, records from 117 sites were retained for modeling and analysis.

Table 1. Information of in situ soil moisture observation networks.

Networks Name	Naqu	Pagri	Maqu	uHRB
Number of stations	53	19	11	25
Time frequency	15 min	15 min	15 min	1 h
Latitude (N)	31.0°–31.9°	27.7°–28.2°	33.6°–34.0°	37.9°–38.3°
Longitude (E)	91.7°–92.5°	89.1°–89.3°	101.8°–102.6°	100.2°–101°

2.2.2. Satellite Data

The soil moisture satellite products used in this study were sourced from the Soil Moisture Active Passive (SMAP) mission, managed by NASA [48]. The L-band radiometer of SMAP provides daily surface soil moisture estimates for the topsoil layer (top ~5 cm) at a spatial resolution of approximately 36 km [48,49]. Specifically, we used data from the SMAP L3 Radiometer Global Daily 36 km descending product by NASA, Pasadena, California, USA. This dataset includes advanced calibration methods for brightness temperature and a sophisticated land surface model, significantly enhancing the accuracy of soil moisture estimation in the TP [40,50]. To align with in situ observations and ensure comprehensive data coverage, SMAP collected daily data throughout the thawing period from June to September 2016.

2.2.3. Auxiliary Data

The auxiliary data (i.e., environmental covariates) for this study were obtained from the National Tibetan Plateau Data Center (<https://data.tpdc.ac.cn/>, accessed on 22 October 2020). These covariates include 13 covariates in four categories: climate, vegetation, terrain, and soil physical and chemical properties. Climate and vegetation factors vary over time, whereas the others remain constant (Table 2). To maintain consistency across our analyses, we resampled all covariates to raster datasets at a resolution of 1000 m based on the bilinear interpolation, ensuring that information was uniformly extracted for all sites.

Table 2. Information of environmental covariates in our study.

Type	Covariate ¹	Original Resolution		Data Source
Climate	Rainfall	1000 m	1 days	Jiang et al. [51]
	LST_D	1000 m	1 day	Tang et al. [52]
	LST_N	1000 m	1 day	
Vegetation	NDVI	250 m	30 days	Gao et al. [53]
Terrain	DEM	30 m	–	https://earthdata.nasa.gov (accessed on 22 October 2020)
	Aspect	30 m	–	
	Slope	30 m	–	
Soil physical and chemical properties	Sand	250 m	–	Liu et al. [54]
	Silt	250 m	–	
	Clay	250 m	–	
	pH	250 m	–	
	SOC	250 m	–	
	BD	250 m	–	

¹ Abbreviations: LST_D and LST_N, land surface temperature during day and night; NDVI, normalized difference vegetation index; DEM, digital elevation model; SOC, soil organic carbon content; BD, soil bulk density.

2.3. Prediction Models and Processes

We introduced four methods: the downscaling method (DM), the interpolation method (IM), and two combined downscaling–interpolation methods that integrate soil moisture maps as mutual covariates—one based on the downscaling framework (CB1) and the other based on the interpolation framework (CB2) (Figure 2). Random forest was used to model and predict soil moisture for each method (RF_{DM}, RF_{IM}, RF_{CB1}, and RF_{CB2}). Additionally, we used boosted regression trees instead of RF to perform the same steps in order to

validate the robustness of the modeling results. The results presented in the main text are based on the RF predictions; the results generated by boosted regression trees are provided in the Supplementary Materials. First, all covariates listed in Table 2 were utilized for the DM and IM to generate 1 km soil moisture maps, as indicated by the gray boxes in Figure 2. Next, these two fine-scale soil moisture maps were incorporated as additional covariates into the optimized downscaling and interpolation frameworks, respectively, to update the combination of covariates. Based on the updated covariates, the two combined methods within the downscaling and interpolation frameworks produced optimized soil moisture prediction results. This allowed for a comparative analysis between the DM and CB1 in the downscaling framework, and the IM and CB2 in the interpolation framework.

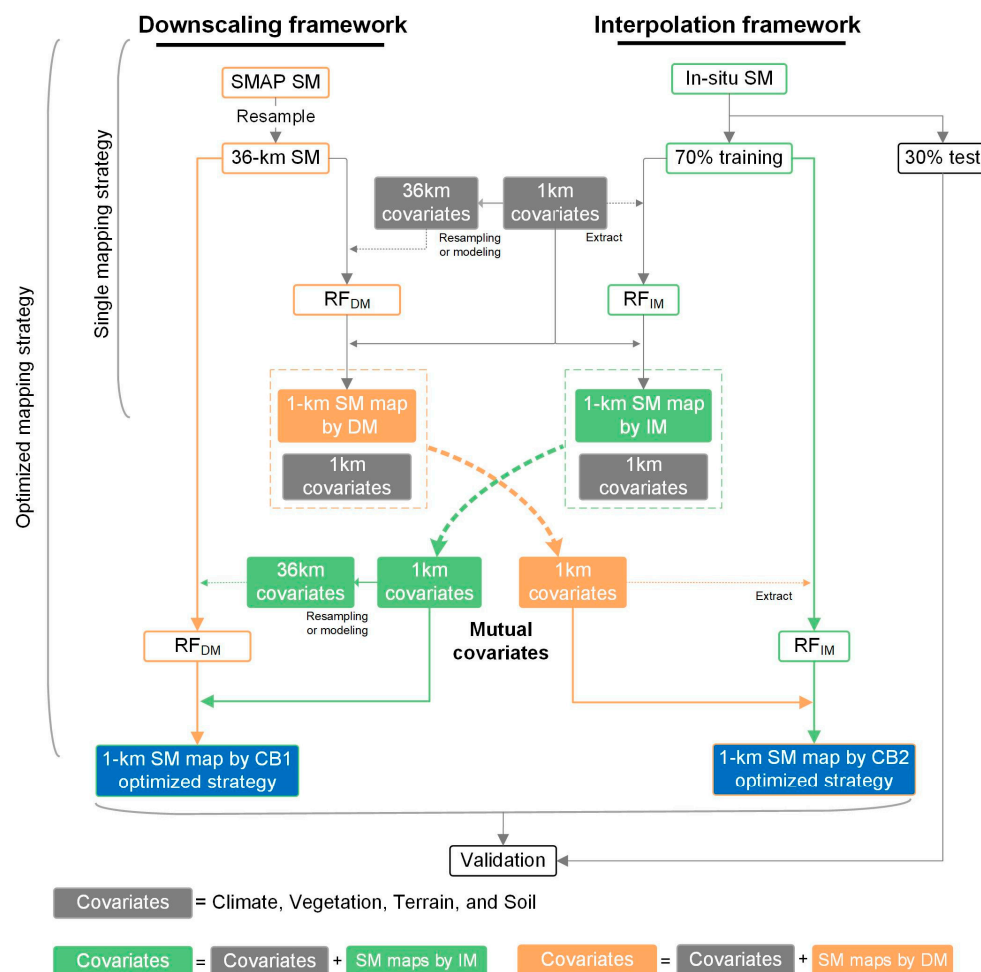


Figure 2. Flowchart of the downscaling, interpolation, and combined downscaling–interpolation methods used in this study. Color coding: orange for downscaling methods, green for interpolation methods, and blue for combined methods. Abbreviations: SM, soil moisture; RF, random forest; DM, downscaling method; IM, interpolation method; CB1, combined method incorporating interpolated SM as covariate within a downscaling framework; CB2, combined method including downscaled SM as covariate within an interpolation framework.

2.3.1. Machine Learning Models

Random forest (RF) was used in this study to predict soil moisture based on four distinct methods. RF is a machine learning algorithm that combines the bagging method with random variable selection, using a series of “trees” to form a “forest” [55]. RF adopts bootstrap sampling for each tree, with the out-of-bag error serving as an indicator of model performance [56]. It is used to predict the dependent variable from covariates with its training dataset consisting of paired observations of the dependent and independent

variables. Depending on the type of dependent variables (continuous or categorical), RF can handle both regression and classification problems. In this work, we used RF for regression, as soil moisture is a continuous variable.

The advantages of RF include its ability to handle covariates without requiring normalization. By aggregating predictions from multiple trees, RF mitigates the risk of overfitting compared to individual decision trees. In RF, variable selection is achieved by randomly selecting subsets of covariates at each node split. The split at each node of a tree is determined by the covariate from the subset that achieves the largest reduction of the loss function. Adjustable parameters in RF include the number of trees (*ntree*), the minimum number of observations per leaf node (*nodesize*), and randomly selected covariates for each node split (*mtry*). Specifically, *ntree* influences the robustness of RF by averaging predictions across trees; *nodesize* ensures each leaf node has enough samples for reliable predictions; *mtry* determines the number of features considered at each split, promoting diversity among trees. The importance function generates rankings of environmental covariates based on their significance for soil moisture when modeling. Variable importance was assessed with the permutation method, which assesses the importance by randomly permuting feature values and observing the resulting decrease in model accuracy. The RF algorithm is available within the “caret” R package (version 6.0-80) [57].

To demonstrate the robustness of the results in this study, we also used the boosted regression trees (BRT) model [58,59] to perform the same modeling process as RF for different strategies and compared it with the RF model. Detailed information and results regarding the BRT model can be found in Supplementary Materials.

2.3.2. Downscaling Method

The downscaling method (DM) utilizes remote sensing satellite soil moisture data as inputs, employing RF (denoted as RF_{DM}) to model the relationship between environmental covariates and soil moisture at the satellite scale. Subsequently, this coarse-scale relationship is combined with fine-scale environmental covariates to predict fine-scale soil moisture. Therefore, the DM follows a structured three-step process:

Step 1: Pre-processing of soil moisture satellite products and obtaining coarse- and fine-scale covariates. We first employed the bilinear interpolation to resample the daily SMAP soil moisture products into regular 36 km × 36 km grids. Next, all covariates were also resampled from their original spatial resolutions to both 1 km and 36 km using the same bilinear interpolation method, thereby producing fine- and coarse-scale covariates, respectively.

Step 2: Establishment of a RF regression model at the coarse scale in RF downscaling framework. The RF_{DM} was used to establish relationships between daily coarse-scale soil moisture SMAP products and covariates at a spatial resolution of 36 km. We established separate regression models for soil moisture and covariates for each day. In selecting covariates, we considered that surface soil moisture for a given day may depend not only on the time-varying covariates of that day but also on covariate values from previous days, reflecting the delayed effect of these covariate changes on soil moisture [60]. Therefore, we included average rainfall and land surface temperature from the previous 14 days.

Step 3: Prediction of soil moisture at the fine scale. After training models at 36 km resolution using coarse soil moisture satellite products and covariates, this relationship derived from RF_{DM} was combined with 1 km covariates to obtain the prediction of downscaled soil moisture at 1 km resolution (Figure 2). This involved predicting 1 km soil moisture using the RF_{DM} regression model from step 2 and 1 km covariates from step 1 within an area-to-area prediction framework.

2.3.3. Interpolation Method

The interpolation method (IM), differing from the DM, utilizes in situ data rather than remote sensing data. It employs the same covariates as those used in the downscaling framework based on the “SCORPAN” factors of digital soil mapping [28]. RF was also applied to the IM to predict soil moisture (denoted as RF_{IM}). Therefore, the primary

distinction between the IM and DM lies in the type of soil moisture data used during model training. The process of predicting soil moisture using the IM consists of two main steps:

Step 1: Extraction of covariate information from in situ sites and establishment of an RF regression model at the site scale. Initially, to use in situ soil moisture for the interpolation of soil moisture, we first aggregated these in situ measurements to daily averages to align with dynamic covariates with the highest temporal resolution, such as rainfall and LST. Next, we extracted environmental information from covariates at 1 km resolution at the measurement locations and dates from in situ sites (Table 1). Using this data, RF_{IM} was employed to model a regression that describes the relationship between soil moisture and covariates at the site scale.

Step 2: Generation of interpolated soil moisture maps at 1 km grids. We predicted the site-scale relationship into grid scale to generate spatial interpolation of soil moisture. The relationships derived from the site-scale RF_{IM} were applied across 1 km grids to create soil moisture interpolated maps by combining with fine-scale covariates. Thus, the predictions of 1 km soil moisture using the IM require the RF_{IM} regression relationship and 1 km covariates (Figure 2) within a point-to-area prediction framework.

2.3.4. Combined Downscaling–Interpolation Methods

In this study, leveraging the established downscaling and interpolation frameworks, the two combined methods, CB1 and CB2, integrated soil moisture maps as mutual covariates. Specifically, CB1 incorporated soil moisture maps generated by the IM into the downscaling framework, whereas CB2 integrated DM soil moisture maps into the interpolation framework (Figure 2). These approaches allowed each combined method to leverage the strengths of both single methods, enabling more comprehensive modeling of soil moisture using data from multiple scales. The specific steps of the two combined methods (CB1 and CB2) are as follows:

Step 1: Generation of fine-scale soil moisture maps at 1 km resolution using single mapping strategies for both downscaling and interpolation. Initially, 1 km soil moisture maps were generated using the DM and IM based on the downscaling and interpolation frameworks, respectively. In addition, the optimization strategy in the downscaling framework (CB1) requires integrating soil moisture maps at both coarse and fine scales for modeling and prediction. This involves generating both 1 km and 36 km interpolated soil moisture maps. The latter can be obtained by either combining observed in situ soil moisture with coarse-scale covariates or resampling the 1 km soil moisture maps (Figure 2). The methods for the single downscaling and interpolation strategies are detailed in Sections 2.3.2 and 2.3.3.

Step 2: Updating the combination of environmental covariates used in different frameworks. Based on the selected environmental covariates in this study (Table 2), different 1 km soil moisture maps were added to form updated combinations of covariates for downscaling and interpolation strategies. Specifically, for the CB1 strategy in the downscaling framework, the updated covariates consisted of all previous covariates along with the soil moisture maps generated by the IM; for the CB2 strategy in the interpolation framework, the updated covariates comprised all previous covariates in conjunction with the soil moisture maps generated by the DM.

Step 3: Generation of new soil moisture predictions combining with the updated combinations of covariates within different frameworks. Based on the updated combinations of covariates, the two combined strategies, CB1 and CB2, repeated the key steps of the DM and IM under the downscaling and interpolation frameworks to produce optimized soil moisture maps.

2.4. Validation and Evaluation

2.4.1. Validation Method

To ensure comparability across all models (DM, IM, CB1, and CB2), an independent dataset validation method was employed, even though the DM does not directly use in situ data during the modeling process. The soil moisture in situ data were randomly divided into two independent datasets: a training set (60%) and a test set (40%). To maintain data

integrity, data from each location were explicitly assigned to either the training or the test set. These sets were then used to train the models and evaluate their performance, respectively. This validation process was repeated 30 times to ensure robust results.

2.4.2. Validation Indicators

For validation indicators, we used the mean error (ME, m^3/m^3), root mean square error (RMSE, m^3/m^3), and the correlation coefficient (R, dimensionless) to evaluate model performance. This direct validation method, which compares in situ observations with 1 km prediction results, was also used to reflect the mapping performance. Scatter density plots were applied to evaluate the overall effectiveness of the predictions, comparing observed outcomes to predicted results. A closer alignment of the linear fitting slope to the 1:1 line indicates superior model performance. Furthermore, we monitored the daily prediction performance of the four methods—DM, IM, CB1, and CB2—and analyzed the daily improvements achieved by the combined methods over the single methods. The equations for ME, RMSE, and R are as follows:

$$ME = \frac{1}{n} \sum_{i=1}^n (O_i - P_i), \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2}, \quad (2)$$

$$R = \frac{\sum_{i=1}^n ((O_i - \bar{O}) \times (P_i - \bar{P}))}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \times \sum_{i=1}^n (P_i - \bar{P})^2}}, \quad (3)$$

where O_i and P_i are the observed and predicted values of soil moisture at site i ; $\bar{O}$ and $\bar{P}$ represent the mean of observed and predicted values of soil moisture, respectively; n denotes the number of in situ sites. ME and RMSE evaluate the bias and accuracy of the model, respectively. An ME value close to 0 and a lower RMSE value indicate better model performance. R values approaching 1 signify high consistency between predicted and observed values.

To evaluate the spatial differences between soil moisture maps generated from the combined and single methods across the TP, we first set the 15% and 85% percentile of the differences between prediction maps as the significance threshold, which was that the prediction differences of these locations were significant. Subsequently, we quantified the percentage of grid cells that showed significant differences out of the total number of grid cells across the TP. In addition, to reveal the impact of in situ observation on spatial differences, a buffer zone of 100 km around the in situ sites was established, which accounted for 14.3% of the total area of the TP. We calculated the percentage of significantly different grid cells within the buffer zone relative to the total number of significantly different grids:

$$P_{total} = \frac{N_{sig}}{N_{total}} \times 100\%, \quad (4)$$

$$P_{site} = \frac{N_{site}}{N_{sig}} \times 100\%, \quad (5)$$

where P_{total} is the percentage of all significantly different grid cells in the TP, N_{sig} denotes the number of grid cells with significant numerical differences, N_{total} represents the total number of grid cells in the TP, and P_{site} and N_{site} indicate the percentage and the number of significantly different grid cells within the 100 km buffer zones, respectively. All data preprocessing, model training, and validation processes were conducted using R software (version 4.3.0) [61].

3. Results

3.1. Accuracy Evaluation of Single and Combined Models

The independent dataset validations revealed that both the DM and IM produced nearly unbiased predictions overall. However, the IM demonstrated superior performance, with its regression slope closer to 1, a lower RMSE, and a higher R compared to the DM (Figure 3a,c). In comparison, the combined methods—CB1 and CB2—which utilized soil moisture maps as mutual covariates, outperformed the single models. Specifically, the improvement offered by CB1 relative to the DM was more substantial, with a decrease in RMSE by 5.94% and an increase in R by 14.02%. Conversely, the enhancements from CB2 compared to the IM were more modest, with RMSE reducing by 1.07% and R increasing by 1.04% (Figure 3b,d). The validation results from the BRT model also reached similar conclusions to RF, further highlighting that integrating soil moisture maps into the downscaling and interpolation framework led to varying degrees of improvement in overall prediction accuracy (Figure S1).

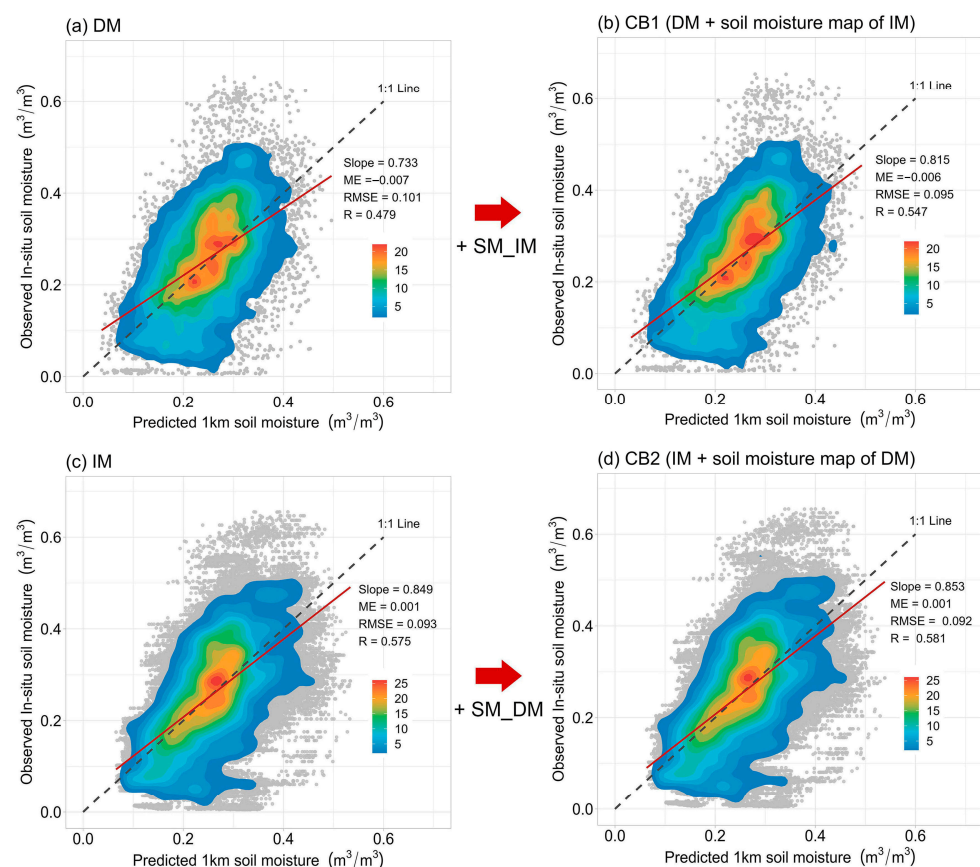


Figure 3. Scatter density plot of soil moisture based on RF comparing the overall accuracy of observation and prediction of (a) DM, (b) CB1, (c) IM, and (d) CB2. Red solid lines represent the linear regression fitting for the observed and predicted values, while black dashed line indicates the 1:1 line. The gray dots represent the collection of predicted values and corresponding observed values from the test set across all available time periods and sites in the independent dataset validation. The captions SM_IM and SM_DM refer to the soil moisture maps generated by the IM and DM.

Figure 4 demonstrates the differences in daily performance between the combined and single methods over the available time periods using RF. The optimization pattern observed aligns with the overall evaluation—integrating interpolated soil moisture maps into the DM framework (CB1) led to more substantial daily improvements compared to the standalone DM (Figure 4a–d). Specifically, CB1 demonstrated daily enhancements compared to the DM, with average improvements of 5.9% in RMSE and 16.7% (Figure 4e). However, the

enhancements from incorporating downscaled soil moisture maps into the IM framework (CB2) were less consistent. The number of days in which RMSE and R worsened accounted for 29–33% of the available time periods, leading to a decrease in average improvement (Figure 4f). Furthermore, in the comparison of R and BRT, the improvement in R based on BRT was stronger than that of RF for certain days (Figure S2f), highlighting the differences between RF and BRT modeling. However, the predictions from BRT also revealed a similar pattern to that of RF, where the improvement in daily prediction accuracy of CB1 relative to the DM was more pronounced than that of CB2 relative to the IM (Figure S2).

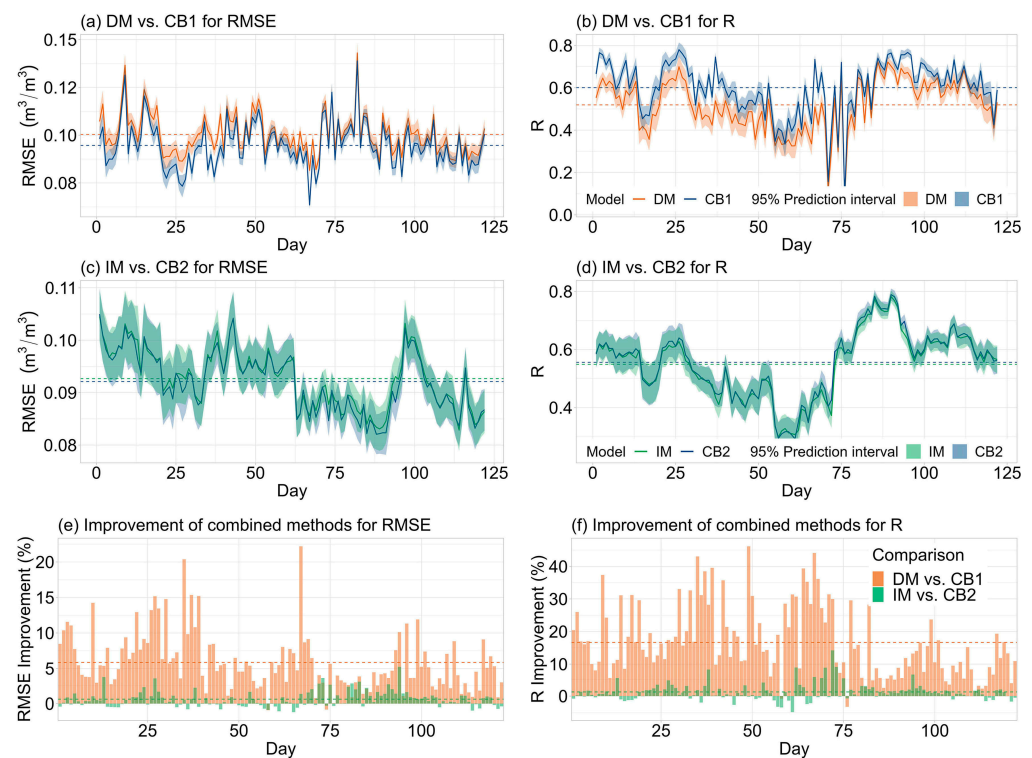


Figure 4. Daily prediction performance and differences for the single methods (DM and IM) and the combined methods (CB1 and CB2) using RF. (a,b): daily comparisons of RMSE and R between DM and CB1; (c,d): daily comparisons of RMSE and R between IM and CB2; (e,f): percentage improvement in RMSE and R of CB1 and CB2 relative to their respective single method (DM and IM). Day 1 corresponds to 1 June 2016. An average value over the available time periods is represented by a horizontal dashed line, and shaded areas represent the 95% prediction intervals.

3.2. Spatial Coverage Differences and Improvement Patterns

To reveal the differences in spatial coverage between the optimized strategies and the single strategies, we conducted difference mapping within the downscaling and interpolation frameworks. The analysis of spatial coverage differences between the combined and single methods revealed distinct patterns of improvement and limitations. The spatial improvement of CB1 over the DM was notably confined, primarily around in situ data sites. Specifically, on Day 67, which saw the greatest improvement in RMSE for CB1, the predictions of CB1 closely mirrored those of the DM in the northwestern parts of the TP. This similarity is attributed to the lack of in situ data in these regions, highlighting the dependence of CB1 on local data for enhancing predictions. Only 24.6% of the grid cells were significantly different; however, the grid cells around the in situ sites contributed 17.7% (Figure 5a).

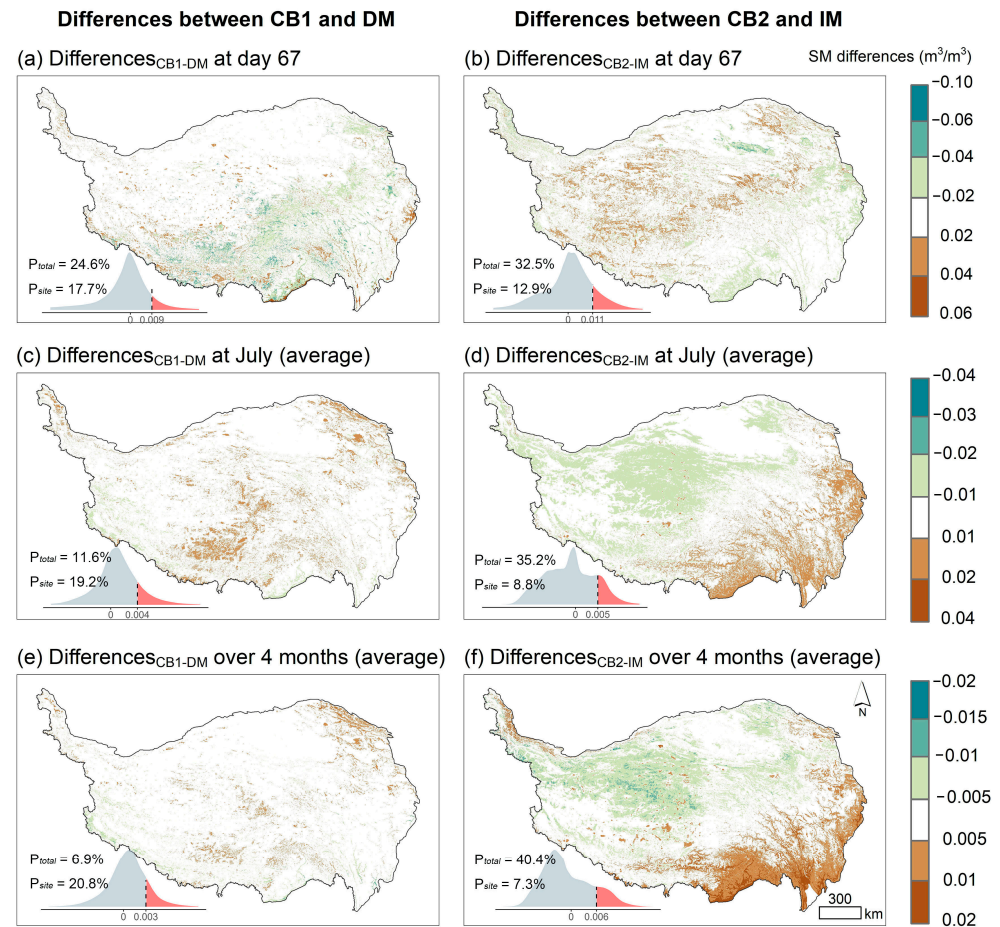


Figure 5. Difference mapping for spatial coverage in 1 km downscaled or interpolated soil moisture maps generated using single (DM and IM) and combined (CB1 and CB2) methods across different temporal resolutions based on RF: (a,b) daily scales; (c,d) monthly scales; (e,f) four-monthly scales. Soil moisture at the monthly scale was assessed by averaging values over the specified periods. The distribution of spatial differences is plotted in the lower left corner of each graph, with the red area representing the 85% percentile.

In contrast, the improvements by CB2 over the IM were more widely distributed in the TP, with 32.5% grid cells showing significant differences (Figure 5b). However, only 12.9% were close to the in situ sites. This widespread enhancement underscores the effectiveness of CB2 in leveraging downscaled data to refine interpolation across diverse geographical areas. As the analysis time scale extended, the magnitude of spatial coverage differences between methods decreased, and these differences for CB1 were increasingly concentrated in in situ sites (Figure 5c,e). The average difference over four months showed that only 6.9% of grid cells were significant, of which 20.8% came from the buffer zone around the in situ observations, suggesting that the benefits of this method are most pronounced where dense data are available. On the contrary, the optimization achieved by CB2 remained extensive across the TP; about 40.4% of the area of the TP had benefited from this improvement and was not affected by in situ observations (7.3%), indicating the broad applicability of the remote sensing data (Figure 5d,f). The spatial differences based on different frameworks were also clearly evident in the results generated by the BRT model (Figure S3).

3.3. Covariate Contributions in Downscaling and Interpolation

To elucidate the differing impacts of various covariates in the downscaling and interpolation frameworks, a variable importance analysis revealed that climate and vegetation covariates, such as daily rainfall and LST, were crucial for both downscaling and interpola-

tion. These dynamic covariates directly influence the variability of soil moisture, making them central to both modeling approaches.

For terrain and soil variables, the downscaling framework highlighted that terrain variables (e.g., elevation, slope) were deemed more significant than soil properties (Figure 6a). This suggests that in areas where downscaling is applied, the physical geography plays a crucial role in determining soil moisture variability. Conversely, in the interpolation framework, soil properties (such as BD and SOC) held higher importance than terrain variables. This indicates that local soil characteristics are more influential in predicting soil moisture at specific sites where interpolation is utilized (Figure 6b). Additionally, the downscaled and interpolated soil moisture maps indicated different importance in the two combined methods.

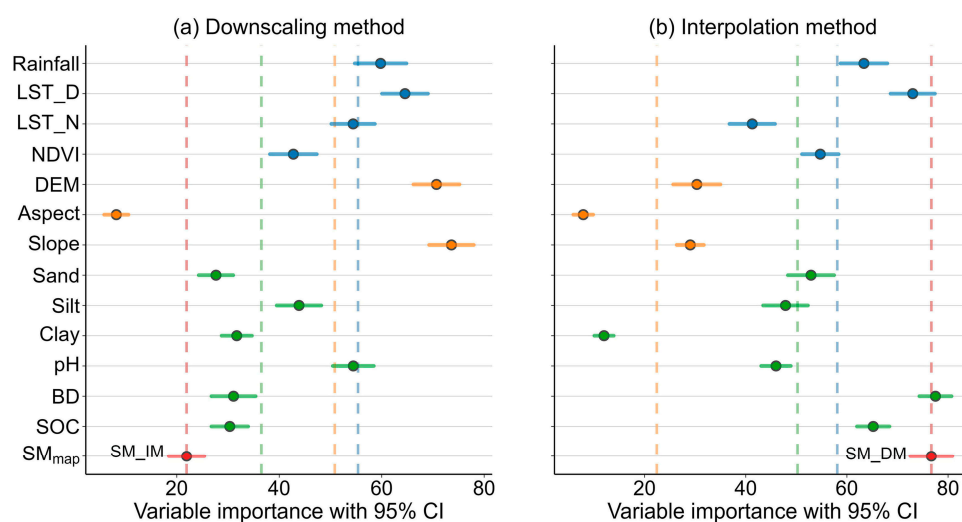


Figure 6. The variable importance within downscaling and interpolation frameworks for different covariates including soil moisture maps. SM_{map} refers to the soil moisture map produced by the DM or IM. Points indicate the daily average importance, and short lines represent the 95% confidence intervals (CIs). The mean importance of different covariate types is highlighted using vertical dashed lines in various colors.

4. Discussion and Conclusions

4.1. Factors Influencing Differences in Soil Moisture Predictions

Consistent with prior research, our results confirm that incorporating additional or diverse covariates typically enhances the performance of machine learning models [62–65]. However, we emphasize that the integration of soil moisture maps as mutual covariates into the downscaling and interpolation frameworks (i.e., the optimized strategies) resulted in notable differences in both prediction accuracy and spatial coverage compared to the single strategies.

The different characteristics of the soil moisture data used by the two frameworks lead to variations in the prediction results. The in situ observation data utilized by the IM are typically concentrated within small geographic networks for ease of monitoring and maintenance [66,67], resulting in lower sampling density but higher local accuracy. In contrast, the DM relies on remote sensing data, providing broader spatial coverage and greater environmental adaptability, albeit with lower observational accuracy [68]. Therefore, in CB1, integrating high-accuracy interpolated soil moisture maps into the downscaling framework significantly enhanced prediction accuracy. However, due to the limitations of localized observations from in situ stations, this optimized strategy could only achieve notable coverage differences within a small area of the TP. For CB2, since the downscaled soil moisture maps depend on low-precision remote sensing data, the improvement in prediction accuracy was limited when integrated into the interpolation framework, as the interpolation algorithm placed more emphasis on the spatial structure and correlation of

local data. Nevertheless, the broad coverage of remote sensing data allowed the optimized strategy to generate coverage differences over a larger area of the TP.

The importance of different types of covariates also highlighted the differences. First, the topographical features of the TP are complex, and the influence of terrain on soil moisture is particularly pronounced in this unique cold plateau region [69,70]. Factors such as slope, aspect, and altitude all affect soil moisture distribution. For instance, in areas with steeper slopes, moisture is more prone to runoff [71,72]; aspect influences the amount of solar radiation received, which in turn affects the evaporation of soil moisture [73]. The DM, which is based on remote sensing data, is capable of capturing more of the topographical features of the TP, which is crucial for spatial coverage differences. Therefore, the downscaling framework's reliance on the complex terrain information of the TP indicated that terrain factors play a more significant role than soil factors (Figure 6a), consistent with recent research findings [74,75]. By further integrating the downscaled soil moisture maps into the interpolation framework, CB1 revealed differences from the single strategy by reflecting more detailed terrain features.

Second, the terrain information extracted from the locations of the in situ stations may be limited due to the relatively homogeneous terrain features of these local networks. However, these stations contain richer soil information, leading to a higher importance of soil attributes (including soil moisture) (Figure 6b). The vegetation types in the TP are relatively uniform (Figure 1), which restricts the variability of soil characteristics across the entire region. Therefore, the in situ stations are sufficiently representative of the soil characteristics in the TP. Additionally, the relatively high importance of soil attributes in the interpolation framework is crucial for enhancing local prediction accuracy. From the perspective of soil influence mechanisms, more detailed soil information can improve prediction accuracy based on site-scale validation. For example, different soil particle sizes have varying impacts on moisture retention and movement concerning soil physical properties [76,77]. Fine-textured soils (e.g., clay) can retain moisture better, while coarse-textured soils (e.g., sand) drain more quickly [78]. Furthermore, incorporating more detailed soil chemical properties, such as pH and SOC, can provide a more comprehensive reflection of soil structure and water-holding capacity [79,80]. Using the IM soil moisture map as a covariate in the downscaling framework is beneficial for predicting soil moisture behavior, thereby improving the accuracy of local predictions.

An intriguing aspect of our analysis was the variable importance disparities between different soil moisture maps. Although interpolated soil moisture maps significantly improved accuracy, their importance was relatively minor within the DM framework, likely because the DM prioritizes broader spatial patterns captured through remote sensing [81–83]. In contrast, the importance of downscaled soil moisture maps was comparable to climate factors, underscoring the value of integrating downscaled data into the IM to enhance regional-scale understanding [84–86].

4.2. Applicability and Limitations

The range of applicability of the study is limited to statistical machine learning models, particularly tree-based machine learning. In this study, we utilized two tree-based models, RF and BRT, to generate prediction results for different strategies, revealing the conclusions through a focused analysis of RF results. In addition, the comparison of predictions from the two tree-based models demonstrated that the conclusions are consistent across both models, highlighting the robustness of these results. However, modeling methods that integrate more geographical information and geoscience knowledge may yield different outcomes. Therefore, we emphasize that the results of this study are robust within the context of tree-based machine learning, underscoring the mapping differences between single and optimized methods under various strategies. Future research should consider more advanced methods to enhance the extensibility of the study. For example, physics-informed machine learning can incorporate geographic and physical knowledge to provide more interpretable results [87,88]. Additionally, geostatistical extension methods, like

the Bayesian maximum entropy method, can integrate different mapping strategies by combining hard and soft data [89,90].

We also recognize that our study has certain limitations. First, our study is limited to modeling soil moisture in the top layer of 0–5 cm. As the depth increases, the deeper soil moisture maps (e.g., root-zone soil moisture) generated by different strategies might affect our conclusions, especially when considering soil moisture processes and mechanisms (e.g., vertical infiltration) [91,92].

Second, our case study in the Tibetan Plateau cannot demonstrate the generalizability of our conclusions to other regions or even globally. Future research needs to test multiple datasets in various locations to provide more robust evidence. Therefore, we recommend that researchers carefully consider the spatial and temporal scales and depth of soil moisture, as well as the prediction models and strategies to produce the most accurate soil moisture maps in future studies. Currently, for optimizing soil moisture mapping over large areas, if there is a wealth of available station data within the study area, then we recommend integrating interpolated soil moisture maps as covariates into the downscaling framework (i.e., the CB1 method).

Finally, from the perspective of data input, the integration of downscaling and interpolation merges information from multiple scales, potentially mitigating errors associated with scale effects. From the validation perspective, scale effects may still exist. In this study, site-based validation metrics such as RMSE and R were used to represent the performance of spatial predictions, where in situ data were employed to validate 1 km soil moisture predictions. This validation approach, known as direct validation, is one of the main methods currently used for downscaled soil moisture validation [93–95]. However, this leads to a scale mismatch between in situ and downscaled soil moisture [94,96]. A potential solution is to strengthen the upscaling of in situ measurements in validation studies and consider various indirect validation methods, such as cross-validation based on multiple predictions (or products) [97] and evaluating the consistency between the downscaled results and environmental covariates [98]. These approaches can systematically and accurately assess the errors and uncertainties in soil moisture spatial predictions, thus improving their reliability [99,100].

4.3. Conclusions

This paper evaluates the performance of machine-learning-based downscaling and interpolation strategies, as well as their optimized strategies using the downscaled and interpolated soil moisture maps as mutual covariates, applying an example of the Tibetan Plateau (TP) daily soil moisture mapping at a 1 km resolution. We found that incorporating interpolated soil moisture maps within the downscaling framework significantly enhanced its accuracy but was restricted to narrow spatial coverage linked to in situ observations. Conversely, incorporating downscaled soil moisture maps into the interpolation framework showed a modest enhancement in prediction accuracy, but this improvement extended across the entire TP. Our results indicated the critical differences in the improvement of prediction accuracy and spatial coverage when using downscaled and interpolated soil moisture maps as mutual covariates. Moreover, terrain and soil factors showed different variable importance within the downscaling and interpolation frameworks, implying the need for the targeted selection of appropriate covariates for different methods.

Our findings underscore the roles that mutual covariates play in modeling, which helps in tailoring appropriate modeling approaches to better capture the spatial and temporal variability of soil moisture. Moving forward, the challenge lies in balancing the optimization of soil moisture prediction concerning both prediction accuracy and spatial coverage. This study provides a reference for optimizing soil moisture mapping at a large scales and highlights the critical differences between methods in improving prediction accuracy and expanding spatial coverage.

Supplementary Materials: The following support information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs16213939/s1>, Supplementary Material Section S1: The principle of boosted regression trees (BRT) model. Supplementary Material Section S2: The differences in overall prediction accuracy (Figure S1), daily prediction accuracy (Figure S2), and spatial coverage (Figure S3) between the optimized strategies obtained and the single strategies using the BRT model. References [101,102] are cited in Supplementary Materials.

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

Funding: This study was supported by the National Natural Science Foundation of China (No. 42230110 and 42222110).

Data Availability Statement: The original data presented in the study are openly available, the in situ observation of soil moisture and environmental covariates were obtained from the National Tibetan Plateau/Third Pole Environment Data Center: <https://data.tpdc.ac.cn/> (accessed on 22 October 2020). The SMAP soil moisture satellite products are available at: <https://smap.jpl.nasa.gov/> (accessed on 22 October 2020).

Acknowledgments: We acknowledge the National Tibetan Plateau/Third Pole Environment Data Center for providing soil moisture data.

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

References

1. Ford, T.W.; Steiner, J.; Mason, B.; Quiring, S.M. Observation-Driven Characterization of Soil Moisture-Precipitation Interactions in the Central United States. *J. Geophys. Res.-Atmos.* **2023**, *128*, e2022JD037934. [\[CrossRef\]](#)
2. Wang, C.; Fu, B.J.; Zhang, L.; Xu, Z.H. Soil moisture-plant interactions: An ecohydrological review. *J. Soils Sediments* **2019**, *19*, 1–9. [\[CrossRef\]](#)
3. Hsu, H.; Dirmeyer, P.A.; Seo, E. Exploring the Mechanisms of the Soil Moisture-Air Temperature Hypersensitive Coupling Regime. *Water Resour. Res.* **2024**, *60*, e2023WR036490. [\[CrossRef\]](#)
4. Li, M.; Foster, E.J.; Le, P.V.V.; Yan, Q.; Stumpf, A.; Hou, T.; Papanicolaou, A.N.; Wacha, K.M.; Wilson, C.G.; Wang, J.K.; et al. A new dynamic wetness index (DWI) predicts soil moisture persistence and correlates with key indicators of surface soil geochemistry. *Geoderma* **2020**, *368*, 17. [\[CrossRef\]](#)
5. Luo, Z.T.; Niu, J.Z.; He, S.Q.; Zhang, L.; Chen, X.W.; Tan, B.; Wang, D.; Berndtsson, R. Linking roots, preferential flow, and soil moisture redistribution in deciduous and coniferous forest soils. *J. Soils Sediments* **2023**, *23*, 1524–1538. [\[CrossRef\]](#)
6. Lee, E.; Kim, S. Spatiotemporal soil moisture response and controlling factors along a hillslope. *J. Hydrol.* **2022**, *605*, 127382. [\[CrossRef\]](#)
7. Dari, J.; Morbidelli, R.; Saltalippi, C.; Massari, C.; Brocca, L. Spatial-temporal variability of soil moisture: Addressing the monitoring at the catchment scale. *J. Hydrol.* **2019**, *570*, 436–444. [\[CrossRef\]](#)
8. Peng, C.; Zeng, J.; Chen, K.S.; Ma, H.; Bi, H. Spatiotemporal Patterns and Influencing Factors Of Soil Moisture At A Global Scale. In Proceedings of the IGARSS 2023—2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023; pp. 3174–3177.
9. 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. [\[CrossRef\]](#)
10. Wei, X.; Li, Q.; Zhang, M.; Giles-Hansen, K.; Liu, W.; Fan, H.; Wang, Y.; Zhou, G.; Piao, S.; Liu, S. Vegetation cover—Another dominant factor in determining global water resources in forested regions. *Glob. Chang. Biol.* **2017**, *24*, 786–795. [\[CrossRef\]](#)
11. Vereecken, H.; Amelung, W.; Bauke, S.L.; Bogaen, H.; Brüggemann, N.; Montzka, C.; Vanderborght, J.; Bechtold, M.; Blöschl, G.; Carminati, A.; et al. Soil hydrology in the Earth system. *Nat. Rev. Earth Environ.* **2022**, *3*, 573–587. [\[CrossRef\]](#)
12. Peng, C.C.; Zeng, J.Y.; Chen, K.S.; Li, Z.; Ma, H.L.; Zhang, X.; Shi, P.F.; Wang, T.T.; Yi, L.; Bi, H.Y. Global spatiotemporal trend of satellite-based soil moisture and its influencing factors in the early 21st century. *Remote Sens. Environ.* **2023**, *291*, 13. [\[CrossRef\]](#)
13. Cheng, S.; Huang, J. Enhanced soil moisture drying in transitional regions under a warming climate. *J. Geophys. Res. Atmos.* **2016**, *121*, 2542–2555. [\[CrossRef\]](#)
14. Li, C.X.; Yu, G.; Wang, J.L.; Horton, D.E. Toward Improved Regional Hydrological Model Performance Using State-Of-The-Science Data-Informed Soil Parameters. *Water Resour. Res.* **2023**, *59*, e2023WR034431. [\[CrossRef\]](#)
15. Sabaghy, S.; Walker, J.P.; Renzullo, L.J.; Akbar, R.; Chan, S.; Chaubell, J.; Das, N.; Dunbar, R.S.; Entekhabi, D.; Gevaert, A.; et al. Comprehensive analysis of alternative downscaled soil moisture products. *Remote Sens. Environ.* **2020**, *239*, 111586. [\[CrossRef\]](#)

16. Wu, T.J.; Yang, C.F.; Luo, J.C.; Dong, W.; Zhou, Y.N.; Yang, Y.P.; Zhao, W.; Xi, J.B.; Wang, C.P. Land Geoparcels-Based Spatial Downscaling for the Microwave Remotely Sensed Soil Moisture Product. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [\[CrossRef\]](#)
17. Chen, Q.Q.; Tang, X.W.; Li, B.; Tang, Z.Y.; Miao, F.; Song, G.L.; Yang, L.; Wang, H.; Zeng, Q.Y. Spatial Downscaling of Soil Moisture Based on Fusion Methods in Complex Terrains. *Remote Sens.* **2023**, *15*, 4451. [\[CrossRef\]](#)
18. Senanayake, I.P.; Arachchilage, K.; Yeo, I.Y.; Khaki, M.; Han, S.C.; Dahlhaus, P.G. Spatial Downscaling of Satellite-Based Soil Moisture Products Using Machine Learning Techniques: A Review. *Remote Sens.* **2024**, *16*, 2067. [\[CrossRef\]](#)
19. Wang, Q.M.; Ji, P.; Atkinson, P.M. Fusion of Surface Soil Moisture Data for Spatial Downscaling of Daily Satellite Precipitation Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 1053–1065. [\[CrossRef\]](#)
20. Im, J.; Park, S.; Rhee, J.; Baik, J.; Choi, M. Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environ. Earth Sci.* **2016**, *75*, 1120. [\[CrossRef\]](#)
21. Liu, Y.X.Y.; Xia, X.L.; Yao, L.; Jing, W.L.; Zhou, C.H.; Huang, W.M.; Li, Y.; Yang, J. Downscaling Satellite Retrieved Soil Moisture Using Regression Tree-Based Machine Learning Algorithms Over Southwest France. *Earth Space Sci.* **2020**, *7*, e2020EA001267. [\[CrossRef\]](#)
22. Dandridge, C.; Fang, B.; Lakshmi, V. Downscaling of SMAP Soil Moisture in the Lower Mekong River Basin. *Water* **2020**, *12*, 56. [\[CrossRef\]](#)
23. Zheng, C.; Jia, L.; Zhao, T. A 21-year dataset (2000–2020) of gap-free global daily surface soil moisture at 1-km grid resolution. *Sci. Data* **2023**, *10*, 139. [\[CrossRef\]](#) [\[PubMed\]](#)
24. 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. [\[CrossRef\]](#)
25. Usowicz, B.; Lipiec, J.; Lukowski, M.; Slominski, J. Improvement of Spatial Interpolation of Precipitation Distribution Using Cokriging Incorporating Rain-Gauge and Satellite (SMOS) Soil Moisture Data. *Remote Sens.* **2021**, *13*, 1039. [\[CrossRef\]](#)
26. Yan, P.; Lin, K.R.; Wang, Y.R.; Zheng, Y.; Gao, X.; Tu, X.J.; Bai, C.M. Spatial interpolation of red bed soil moisture in Nanxiong basin, South China. *J. Contam. Hydrol.* **2021**, *242*, 103860. [\[CrossRef\]](#)
27. Zeyliger, A.; Chinilin, A.; Ermolaeva, O. Spatial Interpolation of Gravimetric Soil Moisture Using EM38-mk Induction and Ensemble Machine Learning (Case Study from Dry Steppe Zone in Volgograd Region). *Sensors* **2022**, *22*, 6153. [\[CrossRef\]](#)
28. McBratney, A.B.; Santos, M.L.M.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [\[CrossRef\]](#)
29. Tunçay, T. Comparison Quality of Interpolation Methods to Estimate Spatial Distribution of Soil Moisture Content. *Commun. Soil Sci. Plant Anal.* **2021**, *52*, 353–374. [\[CrossRef\]](#)
30. Carranza, C.; Nolet, C.; Peziz, M.; van der Ploeg, M. Root zone soil moisture estimation with Random Forest. *J. Hydrol.* **2021**, *593*, 125840. [\[CrossRef\]](#)
31. Li, X.W.; Hua, G.W.; Cheng, T.C.E.; Choi, T.M. What Does Cross-Industry-Production Bring Under COVID-19? A Multi-Methodological Study. *IEEE Trans. Eng. Manag.* **2022**, *15*, 1230–1244. [\[CrossRef\]](#)
32. Jiang, H.R.; Zheng, G.H.; Yi, Y.H.; Chen, D.L.; Zhang, W.J.; Yang, K.; Miller, C.E. Progress and Challenges in Studying Regional Permafrost in the Tibetan Plateau Using Satellite Remote Sensing and Models. *Front. Earth Sci.* **2020**, *8*, 560403. [\[CrossRef\]](#)
33. Wang, X.W.; Yang, Y.Z.; Lv, J.L.; He, H.L. Past, present and future of the applications of machine learning in soil science and hydrology. *Soil Water Res.* **2023**, *18*, 67–80. [\[CrossRef\]](#)
34. Wadoux, A.; Minasny, B.; McBratney, A.B. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Sci. Rev.* **2020**, *210*, 103359. [\[CrossRef\]](#)
35. Yang, K.; Ye, B.; Zhou, D.; Wu, B.; Foken, T.; Qin, J.; Zhou, Z. Response of hydrological cycle to recent climate changes in the Tibetan Plateau. *Clim. Chang.* **2011**, *109*, 517–534. [\[CrossRef\]](#)
36. Cheng, G.; Wu, T. Responses of permafrost to climate change and their environmental significance, Qinghai-Tibet Plateau. *J. Geophys. Res. Earth Surf.* **2007**, *112*, F02S03. [\[CrossRef\]](#)
37. Ma, Z.; Shi, Z.; Zhou, Y.; Xu, J.; Yu, W.; Yang, Y. A spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai-Tibet Plateau with the effects of systematic anomalies removed. *Remote Sens. Environ.* **2017**, *200*, 378–395. [\[CrossRef\]](#)
38. Wang, C.; Gao, Q.; Yu, M. Quantifying Trends of Land Change in Qinghai-Tibet Plateau during 2001–2015. *Remote Sens.* **2019**, *11*, 2435. [\[CrossRef\]](#)
39. Chen, H.; Zhu, Q.; Peng, C.; Wu, N.; Wang, Y.; Fang, X.; Gao, Y.; Zhu, D.; Yang, G.; Tian, J.; et al. The impacts of climate change and human activities on biogeochemical cycles on the Qinghai-Tibetan Plateau. *Glob. Chang. Biol.* **2013**, *19*, 2940–2955. [\[CrossRef\]](#)
40. Shangguan, Y.; Min, X.; Shi, Z. Inter-comparison and integration of different soil moisture downscaling methods over the Qinghai-Tibet Plateau. *J. Hydrol.* **2023**, *617*, 129014. [\[CrossRef\]](#)
41. Piao, S.; Cui, M.; Chen, A.; Wang, X.; Ciais, P.; Liu, J.; Tang, Y. Altitude and temperature dependence of change in the spring vegetation green-up date from 1982 to 2006 in the Qinghai-Xizang Plateau. *Agric. For. Meteorol.* **2011**, *151*, 1599–1608. [\[CrossRef\]](#)
42. Jin, R.; Li, X.; Yan, B.; Li, X.; Luo, W.; Ma, M.; Guo, J.; Kang, J.; Zhu, Z.; Zhao, S. A Nested Ecohydrological Wireless Sensor Network for Capturing the Surface Heterogeneity in the Midstream Areas of the Heihe River Basin, China. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 2015–2019. [\[CrossRef\]](#)
43. Kang, J.; Jin, R.; Li, X.; Ma, C.; Qin, J.; Zhang, Y. High spatio-temporal resolution mapping of soil moisture by integrating wireless sensor network observations and MODIS apparent thermal inertia in the Babao River Basin, China. *Remote Sens. Environ.* **2017**, *191*, 232–245. [\[CrossRef\]](#)

44. Chen, Y.; Yang, K.; Qin, J.; Cui, Q.; Lu, H.; La, Z.; Han, M.; Tang, W. Evaluation of SMAP, SMOS, and AMSR2 soil moisture retrievals against observations from two networks on the Tibetan Plateau. *J. Geophys. Res. Atmos.* **2017**, *122*, 5780–5792. [\[CrossRef\]](#)
45. Yang, K.; Qin, J.; Zhao, L.; Chen, Y.; Tang, W.; Han, M.; Zhu, L.; Chen, Z.; Lv, N.; Ding, B.; et al. A Multi-Scale Soil Moisture and Freeze-Thaw Monitoring Network on the Third Pole. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1907–1916. [\[CrossRef\]](#)
46. Ge, Y.; Wang, J.; Heuvelink, G.; Jin, R.; Wang, J. Sampling design optimization of a wireless sensor network for monitoring ecohydrological processes in the Babao River basin, China. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 92–110. [\[CrossRef\]](#)
47. Ma, Y.; Xie, Z.; Chen, Y.; Liu, S.; Che, T.; Xu, Z.; Shang, L.; He, X.; Meng, X.; Ma, W.; et al. Dataset of spatially extensive long-term quality-assured land–atmosphere interactions over the Tibetan Plateau. *Earth Syst. Sci. Data* **2024**, *16*, 3017–3043. [\[CrossRef\]](#)
48. Entekhabi, D.; Njoku, E.G.; Neill, P.E.O.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; et al. The Soil Moisture Active Passive (SMAP) Mission. *Proc. IEEE* **2010**, *98*, 704–716. [\[CrossRef\]](#)
49. Das, N.N.; Entekhabi, D.; Dunbar, R.S.; Chaubell, M.J.; Colliander, A.; Yueh, S.; Jagdhuber, T.; Chen, F.; Crow, W.; O'Neill, P.E.; et al. The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. *Remote Sens. Environ.* **2019**, *233*, 111380. [\[CrossRef\]](#)
50. Qu, Y.; Zhu, Z.; Montzka, C.; Chai, L.; Liu, S.; Ge, Y.; Liu, J.; Lu, Z.; He, X.; Zheng, J.; et al. Inter-comparison of several soil moisture downscaling methods over the Qinghai-Tibet Plateau, China. *J. Hydrol.* **2021**, *592*, 125616. [\[CrossRef\]](#)
51. Jiang, Y.; Yang, K.; Qi, Y.; Zhou, X.; He, J.; Lu, H.; Li, X.; Chen, Y.; Li, X.; Zhou, B.; et al. TPHiPr: A long-term (1979–2020) high-accuracy precipitation dataset (1/30°, daily) for the Third Pole region based on high-resolution atmospheric modeling and dense observations. *Earth Syst. Sci. Data* **2023**, *15*, 621–638. [\[CrossRef\]](#)
52. Tang, W.; Zhou, J.; Ma, J.; Wang, Z.; Ding, L.; Zhang, X.; Zhang, X. TRIMS LST: A daily 1 km all-weather land surface temperature dataset for China's landmass and surrounding areas (2000–2022). *Earth Syst. Sci. Data* **2024**, *16*, 387–419. [\[CrossRef\]](#)
53. Gao, J.; Zhang, H.; Zhang, W.; Chen, X.; Shen, W.; Xiao, T.; Zhang, Y. *China Regional 250 m Normalized Difference Vegetation Index Data Set (2000–2023)*; National Tibetan Plateau Data Center, Ed.; National Tibetan Plateau Data Center: Beijing, China, 2024. [\[CrossRef\]](#)
54. Liu, F.; Wu, H.; Zhao, Y.; Li, D.; Yang, J.-L.; Song, X.; Shi, Z.; Zhu, A.X.; Zhang, G.-L. Mapping high resolution National Soil Information Grids of China. *Sci. Bull.* **2022**, *67*, 328–340. [\[CrossRef\]](#)
55. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
56. Liaw, A.; Wiener, M. Classification and Regression by RandomForest. *Forest* **2001**, *23*, 18–22.
57. Max, K. *Caret: Classification and Regression Training*, R package version 6.0-80; R Development Core Team: Vienna, Austria, 2018.
58. Elith, J.; Leathwick, J.; Hastie, T. A Working Guide to Boosted Regression Trees. *J. Anim. Ecol.* **2008**, *77*, 802–813. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Ridgeway, G. *Gbm: Generalized Boosted Regression Models*, R package version 2.2.2; R Development Core Team: Vienna, Austria, 2024.
60. Heuvelink, G.B.M.; Angelini, M.E.; Poggio, L.; Bai, Z.; Batjes, N.H.; van den Bosch, R.; Bossio, D.; Estella, S.; Lehmann, J.; Olmedo, G.F.; et al. Machine learning in space and time for modelling soil organic carbon change. *Eur. J. Soil Sci.* **2021**, *72*, 1607–1623. [\[CrossRef\]](#)
61. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
62. Minasny, B.; McBratney, A.B. Digital soil mapping: A brief history and some lessons. *Geoderma* **2016**, *264*, 301–311. [\[CrossRef\]](#)
63. Hengl, T.; Heuvelink, G.B.M.; Kempen, B.; Leenaars, J.G.B.; Walsh, M.G.; Shepherd, K.D.; Sila, A.; MacMillan, R.A.; Mendes de Jesus, J.; Tamene, L.; et al. Mapping Soil Properties of Africa at 250 m Resolution: Random Forests Significantly Improve Current Predictions. *PLoS ONE* **2015**, *10*, e0125814. [\[CrossRef\]](#)
64. Nussbaum, M.; Spiess, K.; Baltensweiler, A.; Grob, U.; Keller, A.; Greiner, L.; Schaepman, M.E.; Papritz, A. Evaluation of digital soil mapping approaches with large sets of environmental covariates. *Soil* **2018**, *4*, 1–22. [\[CrossRef\]](#)
65. Cheng, Z.; Wang, J.; Zhu, K.; Ge, Y.; Zhou, C. Evaluating spatial statistical and machine learning models in urban dynamic population mapping. *Trans. Urban Data Sci. Technol.* **2022**, *1*, 37–55. [\[CrossRef\]](#)
66. Malone, B.; Biggins, D.; Sharman, C.; Searle, R.; Glover, M.; Brown, S. An experiential account with recommendations for the design, installation, operation and maintenance of a farm-scale soil moisture sensing and mapping system. *Soil Res.* **2024**, *62*, 1–17. [\[CrossRef\]](#)
67. Zhang, P.; Zheng, D.; van der Velde, R.; Wen, J.; Ma, Y.; Zeng, Y.; Wang, X.; Wang, Z.; Chen, J.; Su, Z. A dataset of 10-year regional-scale soil moisture and soil temperature measurements at multiple depths on the Tibetan Plateau. *Earth Syst. Sci. Data* **2022**, *14*, 5513–5542. [\[CrossRef\]](#)
68. Zhan, X.W. Accuracy issues associated with satellite remote sensing soil moisture data and their assimilation. In Proceedings of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Shanghai, China, 25–27 June 2008; pp. 213–220.
69. Cao, W.; Sheng, Y.; Wu, J.C.; Peng, E.X. Differential response to rainfall of soil moisture infiltration in permafrost and seasonally frozen ground in Kangqiong small basin on the Qinghai-Tibet Plateau. *Hydrol. Sci. J. J. Des. Sci. Hydrol.* **2021**, *66*, 525–543. [\[CrossRef\]](#)
70. Cao, W.; Sheng, Y.; Wu, J.C.; Li, J. Spatial variability and its main controlling factors of the permafrost soil-moisture on the northern-slope of Bayan Har Mountains in Qinghai-Tibet Plateau. *J. Mt. Sci.* **2017**, *14*, 2406–2419. [\[CrossRef\]](#)

71. He, Z.M.; Jia, G.D.; Liu, Z.Q.; Zhang, Z.Y.; Yu, X.X.; Hao, P.Q. Field studies on the influence of rainfall intensity, vegetation cover and slope length on soil moisture infiltration on typical watersheds of the Loess Plateau, China. *Hydrol. Process.* **2020**, *34*, 4904–4919. [\[CrossRef\]](#)
72. Mo, M.H.; Liu, Z.; Yang, J.; Song, Y.J.; Tu, A.G.; Liao, K.T.; Zhang, J. Water and sediment runoff and soil moisture response to grass cover in sloping citrus land, Southern China. *Soil Water Res.* **2019**, *14*, 10–21. [\[CrossRef\]](#)
73. Liu, H.Y.; He, S.Y.; Anenkhonov, O.A.; Hu, G.Z.; Sandanov, D.V.; Badmaeva, N.K. Topography-Controlled Soil Water Content and the Coexistence of Forest and Steppe in Northern China. *Phys. Geogr.* **2012**, *33*, 561–573. [\[CrossRef\]](#)
74. Deng, Q.H.; Yang, J.J.; Zhang, L.P.; Sun, Z.Z.; Sun, G.Z.; Chen, Q.; Dou, F.K. Analysis of Seasonal Driving Factors and Inversion Model Optimization of Soil Moisture in the Qinghai Tibet Plateau Based on Machine Learning. *Water* **2023**, *15*, 2859. [\[CrossRef\]](#)
75. Wang, J.P.; Wu, X.D.; Tang, R.Q.; Ma, D.J.; Zeng, Q.C.; Xiao, Q.; Wen, J.G. The first assessment of coarse-pixel soil moisture products within the multi-scale validation framework over Qinghai-Tibet Plateau. *J. Hydrol.* **2022**, *613*, 128454. [\[CrossRef\]](#)
76. Sudnitsyn, I.I. Effect of the size of elementary soil particles on the soil moisture characteristic curve. *Eurasian Soil Sci.* **2015**, *48*, 735–741. [\[CrossRef\]](#)
77. Guber, A.K.; Rawls, W.J.; Shein, E.V.; Pachepsky, Y.A. Effect of soil aggregate size distribution on water retention. *Soil Sci.* **2003**, *168*, 223–233. [\[CrossRef\]](#)
78. Shwetha, P.; Varija, K. Soil water retention curve from saturated hydraulic conductivity for sandy loam and loamy sand textured soils. In Proceedings of the International Conference on Water Resources, Coastal and Ocean Engineering (ICWRCOE), Natl Inst Technol Karnataka, Mangaluru, India, 11–14 March 2015; pp. 1142–1149.
79. Qiu, D.X.; Xu, R.R.; Wu, C.X.; Mu, X.M.; Zhao, G.J.; Gao, P. Vegetation restoration improves soil hydrological properties by regulating soil physicochemical properties in the Loess Plateau, China. *J. Hydrol.* **2022**, *609*, 127730. [\[CrossRef\]](#)
80. Vu, E.; Schaumann, G.E.; Buchmann, C. The contribution of microbial activity to soil-water interactions and soil microstructural stability of a silty loam soil under moisture dynamics. *Geoderma* **2022**, *417*, 115822. [\[CrossRef\]](#)
81. Cheng, T.; Hong, S.Y.; Huang, B.S.; Qiu, J.; Zhao, B.K.; Tan, C. Passive Microwave Remote Sensing Soil Moisture Data in Agricultural Drought Monitoring: Application in Northeastern China. *Water* **2021**, *13*, 2777. [\[CrossRef\]](#)
82. Zheng, X.M.; Zhao, K.; Ding, Y.L.; Jiang, T.; Zhang, S.Y.; Jin, M.J. The spatiotemporal patterns of surface soil moisture in Northeast China based on remote sensing products. *J. Water Clim. Chang.* **2016**, *7*, 708–720. [\[CrossRef\]](#)
83. Kang, J.; Jin, R.; Li, X. Regression Kriging-Based Upscaling of Soil Moisture Measurements From a Wireless Sensor Network and Multiresource Remote Sensing Information Over Heterogeneous Cropland. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 92–96. [\[CrossRef\]](#)
84. Wu, D.D.; Wang, T.J.; Di, C.L.; Wang, L.C.; Chen, X. Investigation of controls on the regional soil moisture spatiotemporal patterns across different climate zones. *Sci. Total Environ.* **2020**, *726*, 138214. [\[CrossRef\]](#)
85. Diek, S.; Chabrillat, S.; Nocita, M.; Schaepman, M.E.; de Jong, R. Minimizing soil moisture variations in multi-temporal airborne imaging spectrometer data for digital soil mapping. *Geoderma* **2019**, *337*, 607–621. [\[CrossRef\]](#)
86. Burns, T.T.; Berg, A.A.; Cockburn, J.; Tetlock, E. Regional scale spatial and temporal variability of soil moisture in a prairie region. *Hydrol. Process.* **2016**, *30*, 3639–3649. [\[CrossRef\]](#)
87. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [\[CrossRef\]](#)
88. Li, L.; Dai, Y.J.; Wei, Z.W.; Wei, S.G.; Wei, N.; Zhang, Y.G.; Li, Q.L.; Li, X.X. Enhancing Deep Learning Soil Moisture Forecasting Models by Integrating Physics-based Models. *Adv. Atmos. Sci.* **2024**, *41*, 1326–1341. [\[CrossRef\]](#)
89. Zhu, Z.Z.; Bo, Y.C.; Sun, T.T.; Zhang, X.R.; Sun, M.; Shen, A.J.; Zhang, Y.S.; Tang, J.; Cao, M.F.; Wang, C.Y. A downscaling-and-fusion framework for generating spatio-temporally complete and fine resolution remotely sensed surface soil moisture. *Agric. For. Meteorol.* **2024**, *352*, 110044. [\[CrossRef\]](#)
90. He, J.; Yin, J.; Wu, J.; Christakos, G. Accurate carbon storage estimation for the salt marsh ecosystem based on Bayesian maximum entropy approach—A case study for the *Spartina alterniflora* ecosystem. *J. Environ. Manag.* **2024**, *354*, 120278. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Liao, Y.; Dong, L.B.; Li, A.; Lv, W.W.; Wu, J.Z.; Zhang, H.L.; Bai, R.H.; Liu, Y.L.; Li, J.W.; Shangguan, Z.P.; et al. Soil physicochemical properties and crusts regulate the soil infiltration capacity after land-use conversions from farmlands in semiarid areas. *J. Hydrol.* **2023**, *626*, 130283. [\[CrossRef\]](#)
92. Liu, Y.; Cui, Z.; Huang, Z.; López-Vicente, M.; Wu, G.L. Influence of soil moisture and plant roots on the soil infiltration capacity at different stages in arid grasslands of China. *Catena* **2019**, *182*, 104147. [\[CrossRef\]](#)
93. Liu, J.; Chai, L.; Lu, Z.; Liu, S.; Qu, Y.; Geng, D.; Song, Y.; Guan, Y.; Guo, Z.; Wang, J.; et al. Evaluation of SMAP, SMOS-IC, FY3B, JAXA, and LPRM Soil Moisture Products over the Qinghai-Tibet Plateau and Its Surrounding Areas. *Remote Sens.* **2019**, *11*, 792. [\[CrossRef\]](#)
94. Peng, J.; Loew, A.; Merlin, O.; Verhoest, N.E.C. A review of spatial downscaling of satellite remotely sensed soil moisture. *Rev. Geophys.* **2017**, *55*, 341–366. [\[CrossRef\]](#)
95. Colliander, A.; Jackson, T.J.; Bindlish, R.; Chan, S.; Das, N.; Kim, S.B.; Cosh, M.H.; Dunbar, R.S.; Dang, L.; Pashaian, L.; et al. Validation of SMAP surface soil moisture products with core validation sites. *Remote Sens. Environ.* **2017**, *191*, 215–231. [\[CrossRef\]](#)

96. Crow, W.T.; Berg, A.A.; Cosh, M.H.; Loew, A.; Mohanty, B.P.; Panciera, R.; de Rosnay, P.; Ryu, D.; Walker, J.P. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Rev. Geophys.* **2012**, *50*, RG2002. [[CrossRef](#)]
97. Xu, T.; Guo, Z.; Xia, Y.; Ferreira, V.G.; Liu, S.; Wang, K.; Yao, Y.; Zhang, X.; Zhao, C. Evaluation of twelve evapotranspiration products from machine learning, remote sensing and land surface models over conterminous United States. *J. Hydrol.* **2019**, *578*, 124105. [[CrossRef](#)]
98. Sun, Y.; Huang, S.; Ma, J.; Li, J.; Li, X.; Wang, H.; Chen, S.; Zang, W. Preliminary Evaluation of the SMAP Radiometer Soil Moisture Product over China Using In Situ Data. *Remote Sens.* **2017**, *9*, 292. [[CrossRef](#)]
99. Liu, J.; Chai, L.; Dong, J.; Zheng, D.; Wigneron, J.P.; Liu, S.; Zhou, J.; Xu, T.; Yang, S.; Song, Y.; et al. Uncertainty analysis of eleven multisource soil moisture products in the third pole environment based on the three-corned hat method. *Remote Sens. Environ.* **2021**, *255*, 112225. [[CrossRef](#)]
100. Li, X.; Liu, S.; Ding, J.; Song, L.; Xu, T.; Ma, Y.; Xu, Z.; Yang, X.; Zhang, Y.; Wang, J. A Framework for Quantifying the Uncertainty in Upscaling Evapotranspiration From Homogeneous to Heterogeneous Underlying Surface. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 4413624. [[CrossRef](#)]
101. Groemping, U. Relative importance for linear regression in r: The package relaimpo. *J. Stat. Softw.* **2006**, *17*, 1–27. [[CrossRef](#)]
102. Rätsch, G.; Onoda, T.; Müller, K.R. Soft margins for adaboost. *Mach. Learn.* **2001**, *42*, 287–320. [[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.