remote sensing

MDPI

Article

Enhancing Soil Moisture Prediction in Drought-Prone Agricultural Regions Using Remote Sensing and Machine Learning Approaches

Xizhuoma Zha [1,2], Shaofeng Jia [2], Yan Han [2], Wenbin Zhu [2] and Aifeng Lv [2,3,*]

[1] College of Geographical Sciences, Qinghai Normal University, Xining 810000, China; 202147341024@stu.qhnu.edu.cn
[2] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100000, China; jiasf@igsnrr.ac.cn (S.J.); hanyan78@igsnrr.ac.cn (Y.H.); zhuwb@igsnrr.ac.cn (W.Z.)
[3] University of Chinese Academy of Sciences, Beijing 100000, China
* Correspondence: lvaf@igsnrr.ac.cn

Abstract: The North China Plain is a crucial agricultural region in China, but irregular precipitation patterns have led to significant water shortages. To address this, analyzing the high-resolution dynamics of root-zone soil moisture transport is essential for optimizing irrigation strategies and improving water resource efficiency. The Richards equation is a robust model for describing soil moisture transport dynamics across multiple soil layers, yet its application at large spatial scales is hindered by its sensitivity to boundary conditions and model parameters. This study introduces a novel approach that, for the first time, employs a continuous time series of near-surface soil moisture as the upper boundary condition in the Richards equation to estimate high-resolution root-zone soil moisture in the North China Plain, thus enabling its large-scale application. Singular spectrum analysis (SSA) was first applied to reconstruct site-specific time series, filling in missing and singular values. Leveraging observational data from 617 monitoring sites across the North China Plain and multiple spatial covariates, we developed a machine learning model to estimate near-surface soil moisture at a 1 km resolution. This high-resolution, continuous near-surface soil moisture series then served as the upper boundary condition for the Richards equation, facilitating the estimation of root-zone soil moisture across the region. The results indicated that the machine learning model achieved a correlation coefficient (R) of 0.92 for estimating spatial near-surface soil moisture. Analysis of spatial covariates showed that atmospheric forcing factors, particularly temperature and evaporation, had the most substantial impact on model performance, followed by static factors such as latitude, longitude, and soil texture. With a continuous time series of near-surface soil moisture, the Richards equation method accurately predicted multi-layer soil moisture and demonstrated its applicability for large-scale spatial use. The model yielded R values of 0.97, 0.78, 0.618, and 0.43, with RMSEs of 0.024, 0.06, 0.08, and 0.11, respectively, for soil layers at depths of 10 cm, 20 cm, 40 cm, and 100 cm across the North China Plain.

Keywords: North China Plain; Richards equation; near-surface soil moisture; root-zone soil moisture

check for updates

Academic Editor: Dusan Gleich

Received: 12 November 2024
Revised: 24 December 2024
Accepted: 31 December 2024
Published: 7 January 2025

Citation: Zha, X.; Jia, S.; Han, Y.; Zhu, W.; Lv, A. Enhancing Soil Moisture Prediction in Drought-Prone Agricultural Regions Using Remote Sensing and Machine Learning Approaches. *Remote Sens.* **2025**, *17*, 181. https://doi.org/10.3390/rs17020181

Copyright: © 2025 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

Root-zone soil moisture (RZSM) is a key environmental variable in agricultural and climate-related studies, serving as a primary driver of agricultural productivity and an

indicator of crop water stress [1]. RZSM significantly influences vegetation growth and resilience and is essential for various applications, including land surface modeling, numerical weather prediction, agricultural water management, flood forecasting, and drought assessment [2,3]. Furthermore, RZSM is instrumental in quantifying soil carbon fluxes [4], making its dynamic variability crucial for understanding and modeling biogeochemical cycles within the Earth system.

Monitoring RZSM can be achieved through in situ sensors installed along soil profiles. However, the labor-intensive nature of sensor installation and potential disturbance to soil properties limit the feasibility of precise RZSM monitoring over large spatial scales [5]. Remote sensing has thus emerged as the most effective approach for large-scale soil moisture (SM) monitoring. Current microwave remote sensing missions, such as Soil Moisture Active Passive (SMAP), Soil Moisture and Ocean Salinity (SMOS), and the Advanced Microwave Scanning Radiometer 2 (AMSR2), provide reliable SM measurements but are currently limited to near-surface soil moisture (<5 cm depth) with relatively coarse spatial and temporal resolutions. For example, SMAP Level 3 data has a spatial resolution of 9 km with a temporal resolution of 1–3 days, while SMOS offers a spatial resolution of 25 km with daily updates. Thus, methods for achieving high spatiotemporal resolution RZSM monitoring are still evolving. Nonetheless, advancements in remote sensing techniques have enabled the improved estimation of large-scale RZSM from near-surface soil moisture data [6].

RZSM estimation methods generally include data-driven approaches (such as time series analysis (TSA)) and hydrological models. TSA methods, including exponential filtering [7,8], cumulative distribution functions [9,10], and transfer functions [11], use near-surface soil moisture data to establish functional relationships with RZSM. However, high accuracy in new regions requires extensive parameter calibration. The Richards equation (RE), which is based on the physical principles governing water flow in unsaturated porous media, provides a theoretical framework linking RZSM states to environmental variables, such as meteorological conditions, vegetation growth, and soil properties [12,13]. Typically solved through numerical methods, the RE's complexity and sensitivity to multiple parameters present challenges. Research to address these issues generally takes two approaches: simplifying model structure to improve stability and reduce computational demands, or optimizing regional soil hydraulic parameters to enhance model accuracy. However, traditional studies often use atmospheric conditions as the upper boundary to simulate multi-layer soil moisture, with the RE employing meteorological data through the Penman equation to calculate SM flux. The complexity of these meteorological boundary conditions has limited the RE's large-scale applicability, particularly in regions with sparse meteorological data.

Continuous monitoring of near-surface soil moisture (SSM) reflects variations in surface hydrological processes, such as evaporation [14,15], precipitation, runoff [16], and irrigation [17]. SSM represents the net SM resulting from diverse water fluxes, including infiltration and evapotranspiration. Studies have shown that using SSM to estimate RZSM flux with the RE can yield high retrieval accuracy; however, previous research has focused on flux estimation rather than direct RZSM quantification. Consequently, large-scale, high-resolution RZSM estimation using SSM remains underexplored.

The North China Plain (NCP) is a crucial agricultural region in China, providing approximately 30% of the nation's wheat and 20% of its maize [18,19]. As China's largest agricultural production region, the NCP is characterized by high vegetation coverage and plant diversity, making it highly responsive to climate change [20,21]. Notably, the NCP is among the areas most affected by global climate change [22]. Since the 1980s, a significant greening trend in vegetation cover has been observed, leading to increased evapotran-

spiration. Combined with uneven rainfall distribution, these factors have exacerbated drought conditions, adversely impacting crop growth and yield. Recent national policies focused on food security and water efficiency have spurred extensive research on water resources and agricultural water-use efficiency in the region [23–27]. Findings consistently reveal alarming rates of water resource depletion, with research focusing on SM variability, drought risk assessment, agricultural water use, and optimized agricultural spatial layouts [28–31]. For example, the soil water deficit index (SWDI) and the soil moisture condition index (SMCI) have been developed as tools for monitoring extreme droughts [32]. More recently, the global land data assimilation system (GLDAS) has been utilized to assess the spatiotemporal patterns of multi-layer water availability, along with identifying the key natural and anthropogenic drivers of soil moisture anomalies [33]. Although RZSM is highly sensitive to climate variability, limited research has specifically addressed RZSM estimation within the NCP [34].

Estimating RZSM provides a theoretical foundation for optimized irrigation strategies. Yet, conventional estimates, often constrained by coarse spatial and temporal resolutions, are insufficient for effective drought monitoring and agricultural water management in the NCP [35–38]. To obtain high-resolution root-zone soil moisture (RZSM), exploring effective methods for acquiring high-resolution surface soil moisture (SSM) is essential. Traditional statistical downscaling methods establish statistical relationships between high-resolution auxiliary variables and satellite soil moisture data, providing a straightforward and practical approach. However, these methods often overlook dynamic changes within the system [39–41]. In contrast, multi-source data fusion techniques integrate remote sensing data and ground-based observations to enhance downscaling accuracy. These methods leverage the complementary advantages of different data sources to improve precision; however, discrepancies in temporal and spatial resolution among the data sources and higher computational costs present notable challenges [42–44]. Machine learning (ML), a powerful tool in data mining, can construct nonlinear mappings based on high-resolution auxiliary variables, resulting in high accuracy. Its computational capabilities have made it a widely adopted approach across various fields. Moreover, ML has increasingly been applied to estimate high-resolution spatial soil moisture, underscoring its growing significance in soil moisture estimation [45–49]. In this study, we first apply singular spectrum analysis to reconstruct site-specific time series and address gaps in observed surface time series data. We then employ a random forest regression model to estimate high spatiotemporal resolution SSM. Finally, the Richards equation, incorporating ML-estimated SSM as the upper boundary condition, is applied to predict RZSM, representing a novel application of the Richards equation across large spatial scales. This approach addresses the following questions: (1) Can ML models trained on ground-based observations accurately predict high-resolution SSM? (2) Can an RE-based approach using continuous SSM as the upper boundary support large-scale hydrological modeling with complex nonlinearities? (3) What are the spatial and temporal distribution patterns of RZSM across the NCP?

2. Study Area

The North China Plain (NCP), the second-largest plain in China, lies between latitudes 32° to 43°N and longitudes 113° to 123°E, covering an area of approximately 400,000 square kilometers. It spans five provinces—Hebei, Shandong, Henan, Anhui, and Jiangsu—as well as the municipalities of Beijing and Tianjin. This alluvial plain, characterized by its low-lying terrain with an average elevation below 50 m, showcases varied climatic conditions. According to the Köppen climate classification, northern parts such as Beijing and Tianjin experience a hot summer continental climate, while Hebei has a hot, semi-arid climate. Moving south, Henan and Shandong exhibit a temperate monsoon climate,

whereas Anhui and Jiangsu are classified within the humid subtropical climate zone. The NCP undergoes significant seasonal temperature and precipitation fluctuations, with annual mean temperatures ranging from 8 to 15 °C. Despite abundant solar radiation and thermal resources, precipitation ranges from 500 to 900 mm annually, while potential evapotranspiration (PET) consistently exceeds yearly rainfall, resulting in high irrigation demands for agriculture.

The fertile soils of the NCP, primarily loam and silt loam, benefit from sediment deposits from the Yellow, Hai, and Luan Rivers, making the region ideal for intensive agriculture. The NCP is a vital agricultural zone in China, contributing substantially to the nation's output of grains, cotton, and oilseed crops. It encompasses 27.9% of China's cultivated land, with significant areas dedicated to both dryland and paddy fields. The northern and western areas of the NCP feature hilly terrain with forest and grassland vegetation, while the central and eastern regions comprise extensive plains used predominantly for winter wheat and summer maize cultivation in a double-cropping system. However, agricultural productivity in the NCP is frequently challenged by substantial temporal variability in rainfall, leading to recurrent droughts, especially in spring, early summer, and autumn, making it one of China's most drought-prone regions. Most vegetation in the NCP has a root depth of less than one meter, which served as the depth limit for the calculations in this study. Figure 1 illustrates the NCP's regional zoning and vegetation distribution.

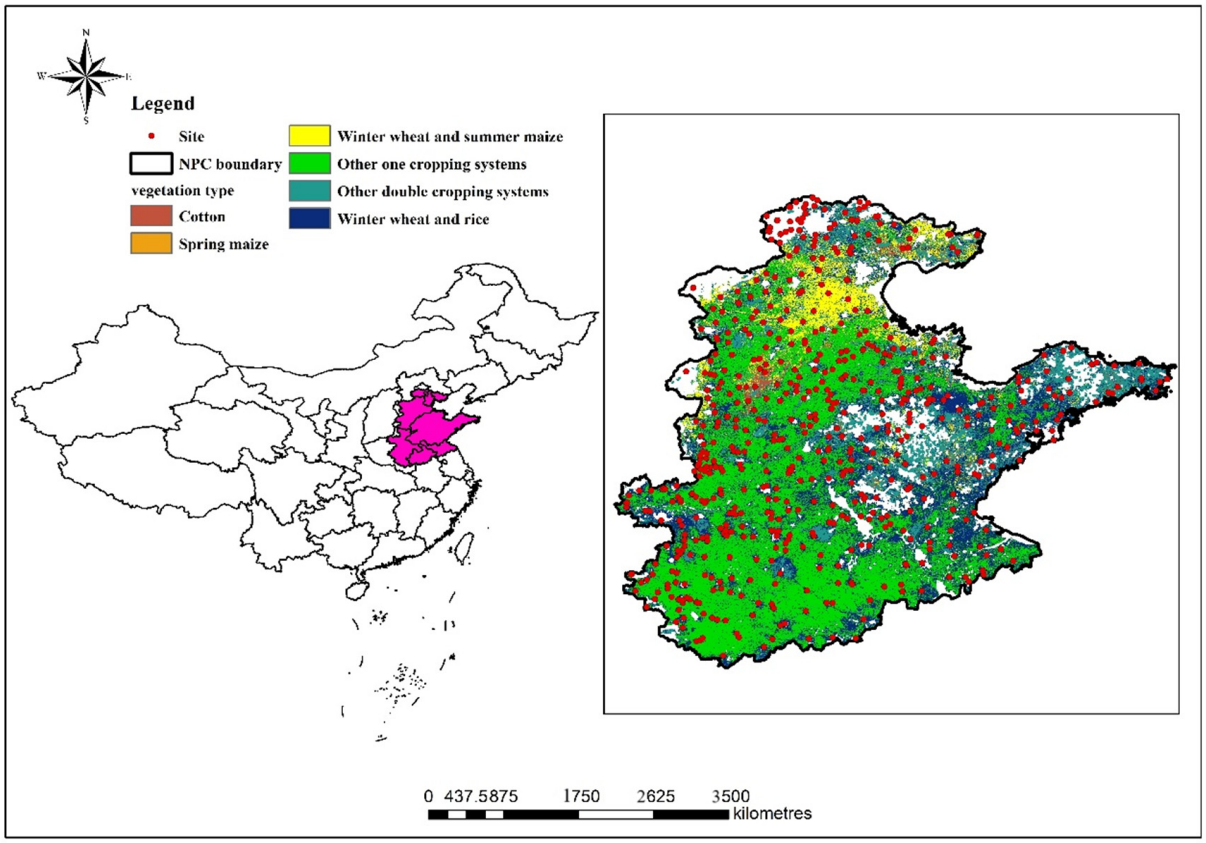


Figure 1. Distribution of vegetation types and administrative divisions in the NPC.

3. Datasets and Methods

The near-surface soil moisture in continuous time series reflects the cumulative effects of atmospheric processes. In other words, near-surface soil moisture represents the net result after atmospheric influences. To implement the calculation of the Richards equation with near-surface soil moisture as the upper boundary condition, we first employ singular

spectrum analysis (SSA) to process outliers and missing values in the monitoring data. Subsequently, the near-surface soil moisture at the site is considered the target variable. In contrast, other meteorological data, soil texture data, terrain data, and 1 km satellite data related to soil moisture serve as auxiliary variables. A random forest model is then utilized to estimate the spatial distribution of near-surface soil moisture at 1 km resolution. Finally, the estimated 1 km near-surface soil moisture is used as the upper boundary condition in the Richards equation to compute the variations in soil moisture profiles across the North China Plain. The workflow for each computational step in this study is illustrated in Figure 2.

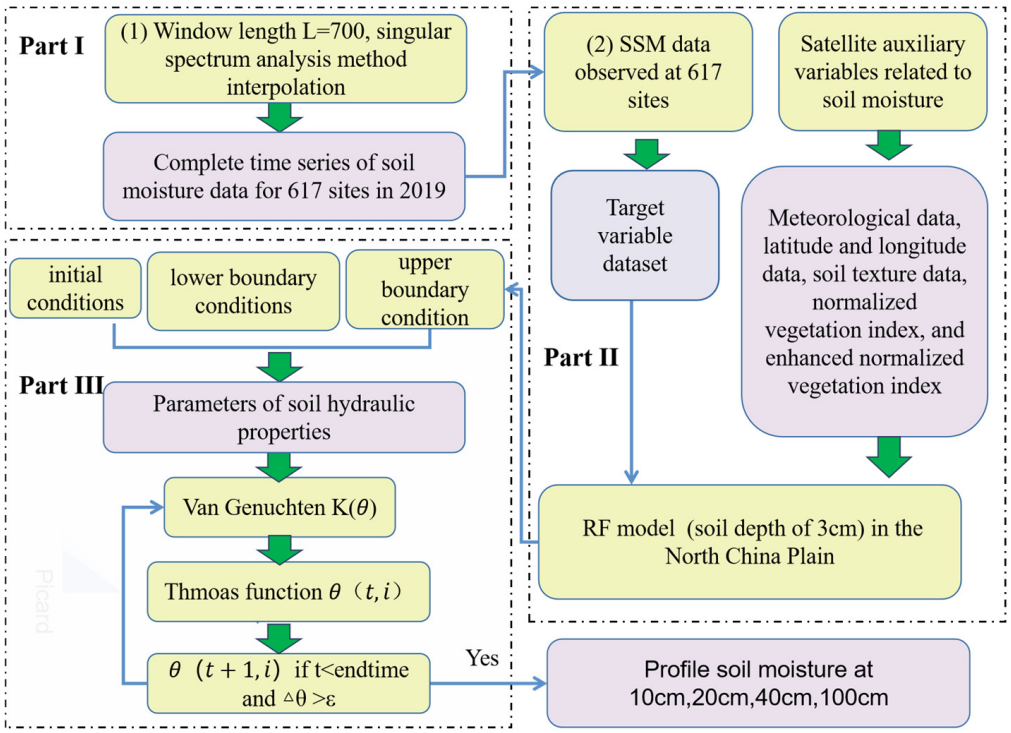


Figure 2. Flowchart for estimating soil moisture at different spatial layers in the North China Plain based on the Richards equation.

3.1. Datasets

Soil texture, meteorological conditions, topography, and vegetation cover collectively influence SM's spatial and temporal dynamics across the landscape [50]. Consequently, this study's random forest model for training and predicting regional SSM incorporates a comprehensive dataset that includes variables such as elevation, terrain features, atmospheric conditions, soil texture, soil hydraulic properties, and station-monitored data. These datasets are summarized below to outline their respective contributions and characteristics in predicting SSM dynamics across the study region.

3.1.1. ERA5_Land Reanalysis Data

ERA5-Land, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), represents the fifth generation of global atmospheric reanalysis, featuring high spatiotemporal resolution at 0.1° hourly intervals. ECMWF provides a coherent perspective on terrestrial variable evolution from 1950 to the present, with monthly aggregates available from hourly products [51]. Notably, the reanalysis process for ERA5-Land incorporates a vast array of historical observational data within its modeling and assimilation framework. Specifically, the number of assimilated observations increased significantly, from approximately 750,000 per day in 1979 to around 21 million per day by July 2018,

enhancing the accuracy of reanalysis and the reliability of near-real-time forecasts. The total evaporation variable represents the cumulative water vapor flux from the Earth's surface to the atmosphere, encompassing plant transpiration. This variable is accumulated from each forecast step's start to end. Similarly, the total precipitation variable provides the cumulative liquid water deposition at the surface, including rainfall and snowfall, and is accumulated over the forecast interval. The temperature refers to the air temperature measured 2 m above the ground, highly relevant to human activities, agriculture, and ecological research (https://cds.climate.copernicus.eu/, accessed on 17 July 2024). In this study, to ensure compatibility with the spatial resolution of auxiliary satellite datasets, we resampled the original datasets for total precipitation, total evaporation, and 2 m air temperature from their native 0.1° spatial resolution to 1 km.

3.1.2. MODIS Data

Research highlights the critical role of the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI), derived from moderate-resolution imaging spectroradiometer (MODIS) data, in assessing SM levels. Integrating NDVI and other vegetation parameters extracted from MODIS with Sentinel-1A radar data can effectively estimate soil moisture, and NDVI is important in distinguishing moisture conditions across diverse climates [52,53]. This study utilized NDVI and EVI as auxiliary variables to estimate high-resolution soil moisture (SSM). The NDVI and EVI datasets were derived from the MODIS 13A2 product generated by the MODIS sensors aboard NASA's Terra and Aqua satellites. This dataset provides global-scale NDVI and EVI measurements updated every 16 days, with a spatial resolution of 1 km. To obtain the daily NDVI and EVI values, we employed the Savitzky–Golay filter on the Google Earth Engine (GEE) platform. This approach facilitated temporal interpolation and smoothing of the data, ensuring continuity and reducing noise for more accurate analysis.

Furthermore, there exists a relationship between surface temperature and SM [54]. To accurately estimate soil saturation moisture (SSM), this study incorporates land surface temperature (LST) data derived from MODIS 11 A1, which is sourced from the USGS Earth Explorer portal https://daac.ornl.gov/, accessed on 17 July 2024. Additionally, the calculation of the diurnal temperature range (LST_Diff) is included in the analysis. The MODIS sensors on the Terra satellite generate the MODIS 11 A1 dataset, which is crucial for global surface temperature monitoring, providing daily surface temperature data with a spatial resolution of 1 km.

3.1.3. Soil Texture Datasets

The Harmonized World Soil Database (HWSD), developed jointly by the Food and Agriculture Organization (FAO) and the International Soil Reference and Information Centre (ISRIC), is a high-resolution (1 km) global soil dataset [55]. This extensive database provides detailed information on soil types, properties, textures, layer depths, and moisture levels, integrating and standardizing data from numerous ground measurements, remote sensing sources, soil classification systems, and a unified nomenclature framework. The HWSD data are available at FAO's Soil Data Hub (https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/, accessed on 17 July 2024). For this study, key soil texture parameters utilized include sand, silt, and clay components, measured in g/kg, along with soil porosity. Specifically, soil sand refers to particles between 0.05 mm and 2 mm in diameter, silt denotes particles between 0.002 mm and 0.05 mm, and clay particles measure less than 0.002 mm in diameter. These soil texture parameters provide foundational insights into the physical composition of soil, which is critical for modeling soil water dynamics and related ecological processes.

3.1.4. Elevation

The spatial distribution data for elevation in China (DEM) originates from the Shuttle Radar Topography Mission (SRTM), conducted by the U.S. Endeavour space shuttle. This study utilizes data resampled from the latest SRTM V4.1 version, achieving a spatial resolution of 1 km. The data are projected using the WGS84 ellipsoid reference system https://www.gscloud.cn/sources/, accessed on 17 July 2024, providing precise elevation information essential for analyzing spatial variability in environmental and ecological studies. Detailed auxiliary spatial data information is provided in Table 1.

Table 1. Data description.

Dataset	Variable	Spatial Resolution
ERA-5	temperature	0.1°
	Total_evaporation	0.1°
	Total_precipitation	0.1°
MODIS 13A2	NDVI	1 km
	EVI	1 km
HWSD	Clay	1 km
	Omc	1 km
	sand	1 km
	silt	1 km
MODIS 11A1	LST	1 km
DEM	elevation	1 km
	TI	1 km

3.1.5. Soil Hydraulic Properties Parameter Dataset

Soil hydraulic parameters are essential inputs for the Richards model, which simulates spatial SM dynamics, with parameter accuracy directly impacting the precision of model predictions. To estimate regional RZSM, this study utilizes soil texture data from a dataset published by the Land-Atmosphere Interaction Research Group at Sun Yat-sen University (http://globalchange.bnu.edu.cn/research/soil3, accessed on 17 July 2024). This dataset provides soil hydraulic properties across seven soil layers in China (depths of 4.5 cm, 9.1 cm, 16.6 cm, 28.9 cm, 49.3 cm, 82.9 cm, and 138.3 cm), calculated using five soil pedotransfer functions to derive parameters for the van Genuchten model and the Mualem function. These pedotransfer functions use soil particle size distribution, bulk density, and organic matter as inputs. The resulting hydraulic parameters include saturated hydraulic conductivity, residual water content, saturated water content, shape parameter (n), and air entry value (ALPHA). This dataset is presented at a 1 km × 1 km resolution. In the RZSM estimation for the NCP, simulations are conducted across five soil depth layers: 0–9.1 cm, 9.1–16.6 cm, 16.6–28.9 cm, 28.9–49.3 cm, and 49.3–100 cm.

3.1.6. In Situ SM Data

The soil volumetric water content data for validating regional RZSM in the study area were obtained from 617 monitoring stations operated by the China Meteorological Administration (CMA). This dataset covers daily SM readings from 1 January 2019 to 31 December 2019 across five soil depths: 10 cm, 20 cm, 40 cm, and 100 cm. For this study, we focused on stations providing continuous SM records over 365 days. The spatial distribution of these stations, shown in Figure 1, indicates a relatively even spread across the NCP.

In our analysis, SM at 10 cm depth is considered near-surface soil moisture, 20–40 cm as shallow SM, and 40–100 cm as RZSM. This depth-based categorization accurately assesses soil moisture dynamics across the soil profile.

3.2. Methods

3.2.1. Singular Spectrum Analysis

The singular spectrum analysis (SSA) method is powerful for decomposing time series into components ranked by variance intensity [56]. Widely employed to extract (nonlinear) trends and leading oscillatory modes or to denoise time signals [57], SSA's main advantage lies in its nonparametric nature and adaptability to data characteristics, making it extensively applicable, particularly in meteorology and oceanography [58–60]. The SSA process involves dividing a time series into lagged segments, analyzing correlations among these segments and reconstructing the time series based on these correlations. SSA can not only analyze and reconstruct complete time series, but also fill gaps in data by extrapolating from sample characteristics [61–63].

The singular spectrum analysis (SAA) is performed in four main steps. First, a window length L (with $1 < L < K/2$) is chosen, and the SSM $\delta_N = (\delta_1, \delta_2, \ldots, \delta_N)^T$ from various monitoring sites are arranged in a lagged sequence. This forms a trajectory matrix of length L:

$$D_{L\times K} = [d_1, d_2, \ldots, d_k] \quad (1)$$

where N represents the length of the data, and $K = N - L + 1$, $d_j = (\delta_j, \delta_{j+1}, \ldots, \delta_{j+L-1})^T$, where L is typically set to the length of a single period. In this study, the value of L is chosen as 700. Subsequently, the trajectory matrix is subjected to singular value decomposition (SVD) as follows:

$$D = \sum_{i=1}^{L} \sqrt{\lambda_i} U_i V_i = \sum_{i=1}^{L} D_i \quad (2)$$

$\sqrt{\lambda_i}$ represents the singular values in the equation, U_i and V_i are the left and right singular vector matrices, respectively, and D_i is referred to as the elementary matrix of the trajectory matrix. The indices of the elementary matrix D_i are divided into m disjoint subsets, denoted as $\{I_1, I_2, \ldots, I_m\}$. For any positive integer $q \epsilon [1, m]$, the grouped matrix is expressed as follows:

$$D_{Iq} = \sum_{i \epsilon Iq} D_i \quad (3)$$

Upon rewriting the above equation in grouped format, we obtain $D = D_{I1} + D_{I2} + \ldots + D_{Im}$, where D_{Iq} corresponds to the feature components associated with the time series. The contribution rate of these components is positively correlated with the singular values. Finally, the diagonal averaging method transforms the matrix D_{Iq} into a sequence of length N, yielding reconstructed time series data.

$$\delta_i = \sum_{q=1}^{m} y_{q'}, i = 1, 2, \ldots, N \quad (4)$$

where m represents the number of groups resulting from partitioning the elementary matrix D_i, $y_q = \sum_{k=1}^{N} y_{q,k}$.

Soil moisture data are influenced by atmospheric conditions and soil texture and exhibit high nonlinearity and nonstationarity. SSA is an effective method for analyzing complex, nonlinear, and nonstationary data. In this study, missing data points at monitoring sites primarily occur at the beginning and end of 2019. To ensure accurate data representation, we reconstructed the time series by integrating data from both 2019 and 2020.

3.2.2. Machine Learning

This study developed a regression model between station soil monitoring data and various auxiliary datasets to estimate high-resolution SSM. The random forest (RF) approach, an ensemble method, grows a collection of decision trees and averages their outputs to produce the final result. The "random" aspect refers to the use of bootstrapped subsamples from the original dataset for training and the random selection of predictors at each decision node, which greatly reduces the risk of overfitting [64–66]. Additionally, RF enables analysis of the relative importance of each predictor variable. In this study, we used Shapley Additive Explanations (SHAP) values to interpret the significance of covariates to the model.

The RF model employed in this study was configured with specific hyperparameters to optimize predictive performance and computational efficiency. The model comprises 20 decision trees. Each tree was configured with a minimum leaf population of 2, allowing the model to remain sensitive to finer details in the dataset. Additionally, the maximum depth of the trees was set to 30, providing a balance between capturing meaningful patterns and avoiding overfitting. These parameter choices were optimized through a grid search combined with three-fold cross-validation to minimize the root mean square error (RMSE). Training and validation were split, with 75% of the data used for model training and 25% for validation. The selection of these hyperparameters was informed by cross-validation experiments, with the correlation coefficient (R) serving as the primary metric for optimization. For the RF model, we integrated spatial covariate data with SSM observations from 617 monitoring stations, achieving SM predictions at a 1 km resolution. Covariate data encompass meteorological, vegetation, soil, topographical, and hydrological features. The meteorological data included daily precipitation (Preci), air temperature (Tair), evapotranspiration (Evapo), and land surface temperature (LST), represented both as a daily mean (LST_DAILY) and as its diurnal difference (LST_Diff). Vegetation indices, such as the EVI and the NDVI. Soil properties included proportions of clay, sand, and silt, soil porosity, and organic matter content (omc), reflecting the physical and chemical characteristics of the soil. Topographical variables, including elevation and the topographic index (TI), were utilized alongside geographic coordinates (latitude and longitude) to capture spatial variability. Before the model training, all input features were preprocessed to address missing values and normalized to ensure consistency across variables. These input variables were selected based on their established influence on soil moisture dynamics, enhancing the robustness and relevance of the model predictions.

3.2.3. Construction of the Richards' Equation with SSM as an Upper Boundary Condition

The Richards equation (RE) governs SM movement dynamics, which describe the spatial and temporal behavior of soil water. Derived from the unsaturated form of Darcy's law and the principle of mass conservation, the Richards equation provides a foundational framework for modeling soil water transport processes in unsaturated zones:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[D\frac{\partial \theta}{\partial z}\right] - \frac{\partial K(\theta)}{\partial z} - S(z,t) \quad (5)$$

where $S(z,t)$ is a sink term representing water uptake by roots expressed as volume of water per unit volume of soil per unit time, K is unsaturated hydraulic conductivity, and D

is diffusivity. These parameters are calculated using the van Genuchten method, with the unsaturated hydraulic conductivity formulated as follows:

$$\begin{cases} se = \frac{\theta-\theta_r}{\theta_s-\theta_r} = \left[\frac{1}{1+|ah|^n}\right]^m \\ K(\theta) = K_S se^{0.5}\left[1-\left(1-se^{\frac{1}{m}}\right)^m\right]^2 \end{cases} \quad (6)$$

$$m = 1 - 1/n \quad (7)$$

In the equation, θ_s represents the saturated water content, and θ_r is the residual water content. The parameters a, n, and m are soil texture-dependent parameters determined by the soil's texture and structure. Based on the soil texture composition at each monitoring site, the soil within the modeled depth is assumed to be homogeneous.

Solving the Richards equation (RE) requires specific initial and boundary conditions. This study uses the dynamic variation of near-surface SM over time as the upper boundary condition to estimate deeper SM levels. This approach, which applies SSM as the upper boundary in the RE framework, is referred to as the SSMRE model. The precise formulation of the upper boundary condition is provided in Equation (8):

$$\theta(Z_0, t) = \theta_0(t) \qquad Z_0 = 3 \text{ cm}, t > 0 \quad (8)$$

In the equation, Z_0 represents the depth of the near-surface soil layer, t denotes time, and θ_0 is the SM at the near-surface level. Given that vegetation roots in the NCP are generally distributed within a 1 m depth [67], and groundwater in the region is typically found at greater depths, a free drainage condition is applied as the lower boundary in this study. The specific formulation of the lower boundary condition is presented as follows:

$$\frac{\partial\theta}{\partial z} = 0 \qquad Z = 100 \text{ cm}, \ t > 0 \quad (9)$$

The initial conditions for SM are specifically expressed in Equation (10) to establish the baseline for the model's calculations.

$$\theta(Z_i, 0) = \theta_i(0) \qquad Z_i = 1, \ 2, \ 3, \ \ldots, t = 0 \quad (10)$$

In the footnotes, Z_i represents various soil depths (ranging from 1 to 50 cm), and θ_i denotes the SM for each corresponding soil layer.

3.2.4. Vegetation Water Uptake Modeling

The vegetation water uptake model is critical for understanding vegetation–atmosphere interactions, particularly in ecological and climate research. This model aims to quantify the uptake of SM by vegetation. In this study, we use the FD model to calculate root water uptake rates [68]. Within the FD model, the potential root water uptake S_p is estimated as a function of the potential transpiration T_p and the normalized root density distribution β, as shown in Equations (11) and (12):

$$S(z, t) = \alpha S_p(z, t) \quad (11)$$

$$S_p(z, t) = \beta_z T_p \quad (12)$$

The stress response function $\alpha[-]$ depends on dynamic soil state variables, such as soil water potential or moisture content. This function encompasses multiple stress factors, including soil hypoxia, total soil water potential, and atmospheric evaporative demand. These are critical in modeling the complex interactions between plant stress and

environmental conditions. Typically, this function is represented as a piecewise linear relationship between pressure head and potential transpiration rate. Specifically, the stress induced by matric potential is commonly modeled using a piecewise linear reduction function. The specific formula is shown below:

$$\alpha(\psi) = f(x) = \begin{cases} 0 & , \ \psi \le \psi_w \\ \frac{\psi - \psi_w}{\psi_{amc} - \psi_w}, & \psi_w < \psi < \psi_{amc} \\ 1, & \psi_{amc} < \psi < \psi_{fc} \\ \frac{\psi - \psi_a}{\psi_{fc} - \psi_a}, & \psi_{fc} < \psi < \psi_a \\ 0 & , \ \psi_a \le \psi \end{cases} \quad (13)$$

The $\alpha(\psi)[-]$ function describes the reduction in plant water uptake due to soil matric potential, a key factor contributing to SM stress. The shape of this function is determined by a series of specific matric potential thresholds, ranging from ψ_w to ψ_a, each corresponding to distinct physiological limits within the plant–soil system (Figure 3). When the soil matric potential exceeds a certain threshold, the plant's water uptake is limited due to insufficient oxygen. This indicates that the soil is excessively saturated, and hypoxic conditions inhibit root water uptake. As the matric potential decreases to ψ_a, oxygen availability reaches an optimal level, enabling transpiration to operate at peak efficiency. However, as the matric potential further drops below ψ_{amc}, the declining energy state of SM begins to restrict water uptake, marking the onset of water stress. Finally, when matric potential falls below ψ_w, SM reaches the wilting point, at which vegetation ceases water uptake entirely, indicating severe moisture stress where plants can no longer sustain transpiration.

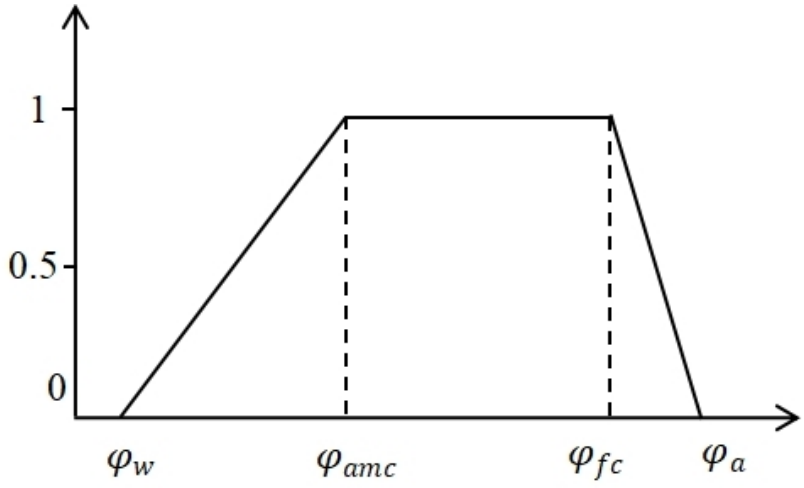


Figure 3. Definition sketch of $\alpha(\psi)$.

3.2.5. Calibration and Validation of the SSMRE Model

To assess the accuracy of the SSMRE model over large spatial scales, this study conducted a comparative analysis of the model's simulated results against observed SM measurements from monitoring stations across the NCP [69]. The R and root mean square error (RMSE) were used as metrics to validate the model, providing quantitative measures of the consistency between the SSMRE model's simulations and the observed SM dynamics. The calculation methods for R and RMSE are detailed in Equations (14) and (15); the RMSE of the model's predictions was calculated as m^3/m^3, consistent with the unit of the soil moisture variable:

$$R = E[(\theta_m - E[\theta_m](\theta_s - E[\theta_s]))](\sigma_m \sigma_s) \quad (14)$$

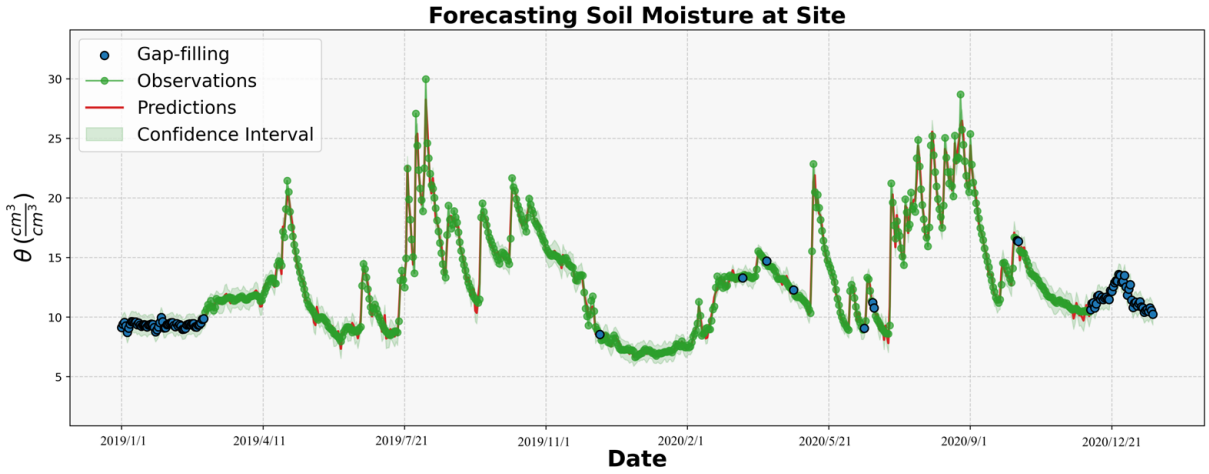
$$\mathrm{MSE} = \sqrt{\mathrm{E}\left[(\theta_s(i) - \theta_m(i))^2\right]} \quad (15)$$

In the equations, θ_s and θ_m represent the observed SM at the station and the SM simulated by the RE model, respectively. Within the time series, m denotes the time range, and i indicates specific time points. This notation facilitates precise comparison between observed and simulated SM data across the temporal domain.

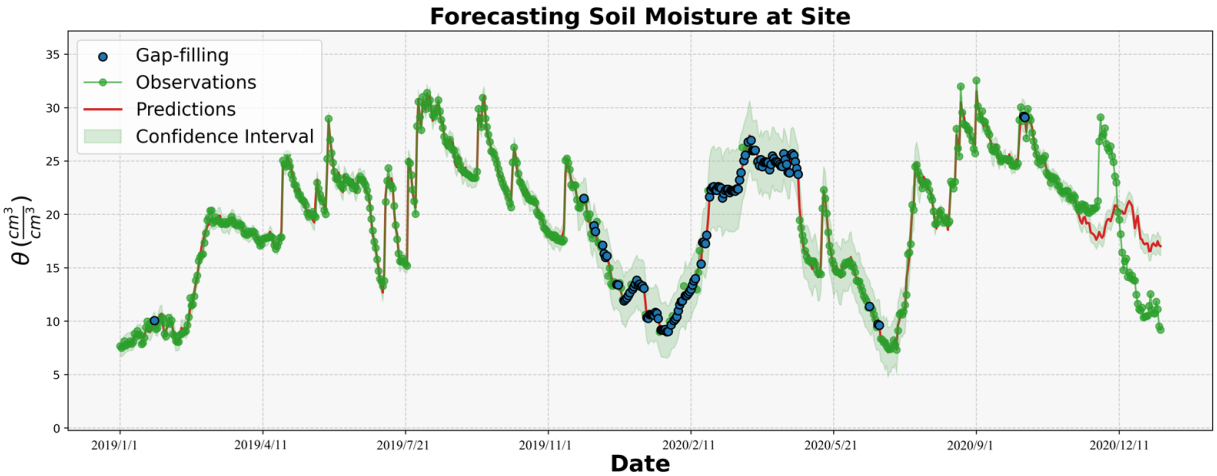
4. Results

4.1. Estimation of Site-Specific Soil Moisture Using SSA

The sample length for the SSA calculation was set to N = 24 months, with a focus on key stations exhibiting substantial data gaps at the start and end of 2019. Figure 4 displays the SSA results for SSM at these stations. A comparison between the SSA-reconstructed time series and the observed values at these stations shows a high level of accuracy, as indicated by the R of 0.98 and RMSE of 0.276. This high correlation demonstrates SSA's strong predictive capability for data distribution across the time series, enabling it to effectively fill in missing values across various stations.

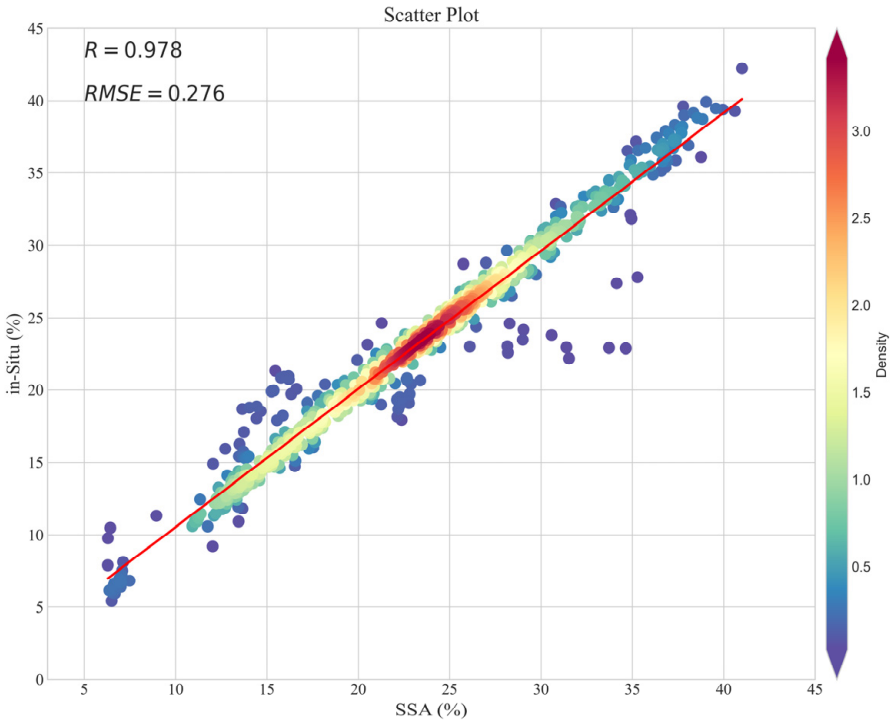


(**a**)



(**b**)

Figure 4. *Cont.*



(**c**)

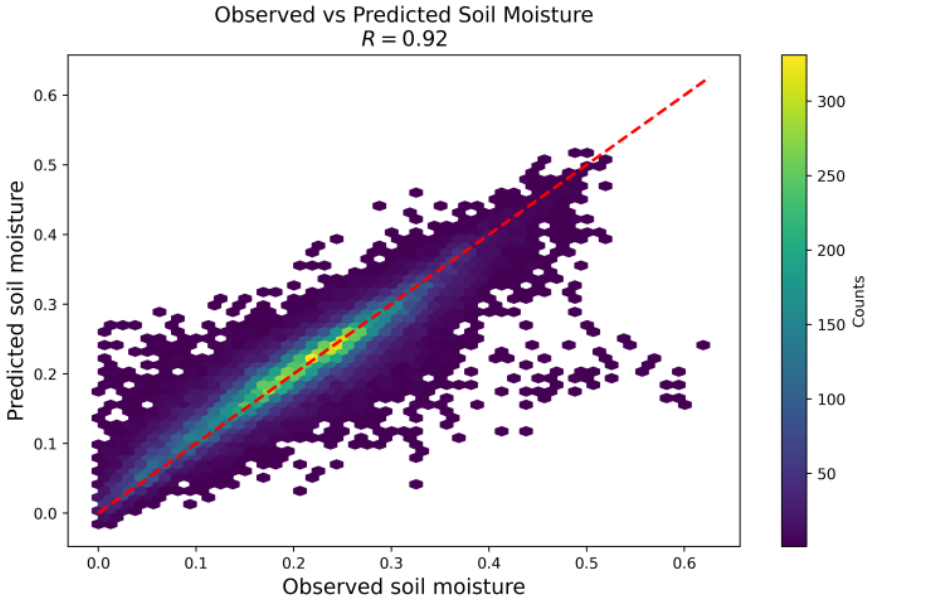
Figure 4. (**a**) Comparison between SSA-predicted time series and observed trends; (**b**) time series comparison for data imputation at the start and end of the year; (**c**) scatter plot of SSA results.

Trend analysis of the projections reveals that SM remains relatively stable during the drought-prone winter and early spring months, with a notable increase from March onward, aligning well with the observed seasonal patterns. The SSA method effectively captures SM trends mid-year (Figure 4a) and during the year's beginning and end (Figure 4b). Despite the dynamic nature of near-surface soil moisture, stations with extended data gaps (exceeding five months) at the start of 2019 showed minimal variation within the reconstructed 60-day series, displaying a nearly flat trend that diverges from expected seasonal SM changes. Consequently, these stations were excluded from the machine learning and RZSM validation processes.

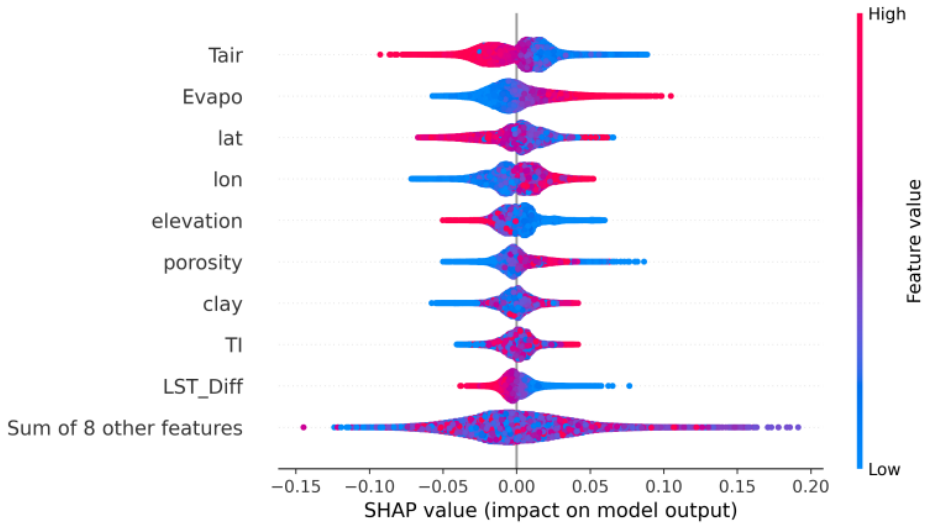
4.2. High-Resolution SSM Predictions Using the RF Method

To predict high-resolution SSM for 2019, an RF algorithm was employed, utilizing 75% of the data for training and 25% for validation. Shapley Additive Explanations (SHAP) were applied to interpret each feature's contribution to the model's predictions, highlighting the positive and negative impacts of individual features on the model output. Figure 5 presents the RF model's validation results for SSM predictions alongside the SHAP-based feature importance analysis.

As illustrated in Figure 5a, the RF model achieved a high R of 0.92, confirming the accuracy of its SSM predictions. According to the SHAP analysis, air temperature (Tair) exerted the most substantial influence on the predictions, with SHAP values ranging from 0.1 (positive impact) to −0.1 (negative effect). The feature importance ranked in descending order includes evaporation (Evapo), geographical coordinates (latitude and longitude), elevation, soil porosity, clay content, and diurnal land surface temperature difference (LSTDiff). Among these, LSTDiff had the smallest impact on the RF model, with SHAP values of 0.07 (positive) and −0.05 (negative).



(**a**)



(**b**)

Figure 5. (**a**) Scatter plot for RF validation; (**b**) SHAP value interpretation of RF model outputs. Note: In the figure, Tair represents the air temperature 2 m above the surface of land, ocean, or inland water bodies; LST_diff denotes the daily difference in land surface temperature.

This analysis reveals that air temperature plays a critical role in shaping the model's SSM predictions, while diurnal temperature variation has a relatively minor effect. These findings enhance our understanding of how different environmental variables contribute to SSM predictions within the RF model.

Figure 6 compares the spatially weighted average of SSM predicted by the RF model with site-observed SSM values. The RF model's predictions generally align well with the overall trend of SSM fluctuations recorded at monitoring stations. However, closer analysis of localized SSM patterns reveals some discrepancies. In June and October, when soil moisture abruptly decreased to 0.16 and 0.18, respectively, the RF result was overestimated, with a similar overestimation observed in November. Conversely, when soil moisture suddenly increased to 0.3 in August, the RF was underestimated. Following spatial weighting, the highest recorded soil moisture in the North China Plain (NCP) in 2019 was 0.3, while the

lowest was 0.159. A review of the maximum soil moisture values indicates that the North China Plain experienced relatively dry conditions in 2019.

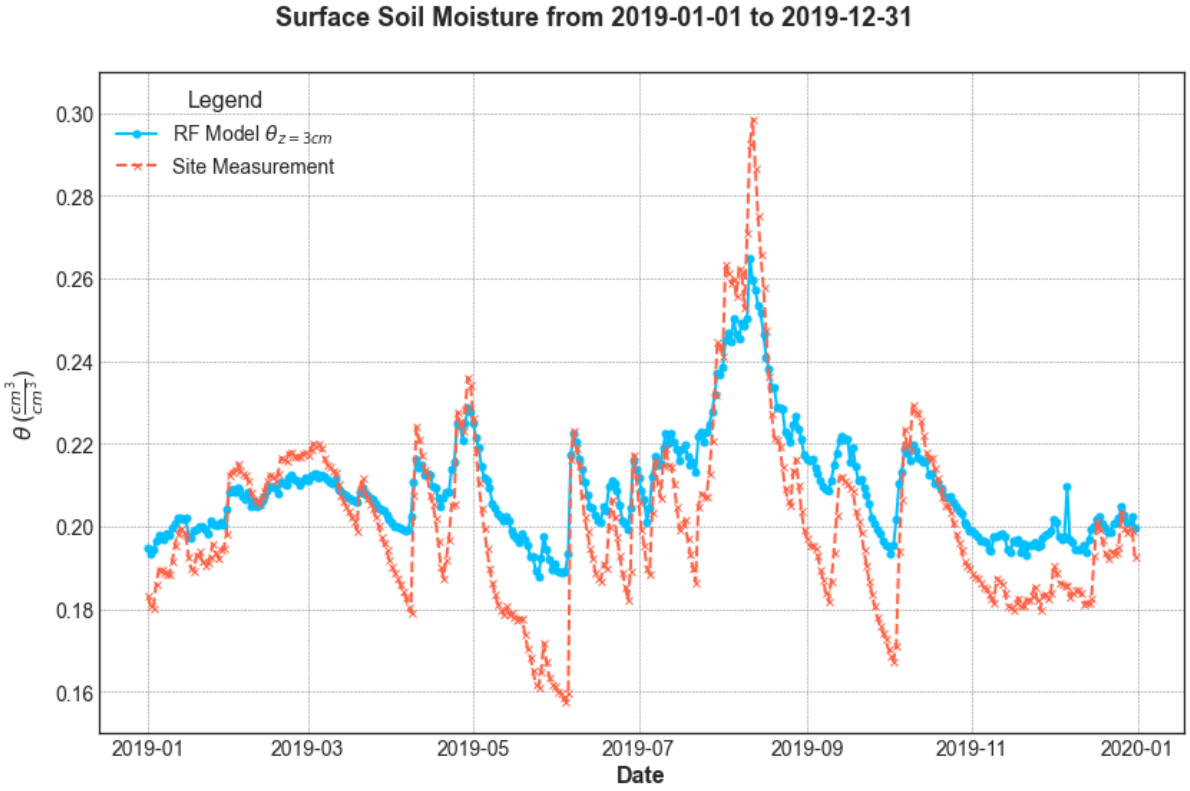


Figure 6. Comparison between RF model calculations and site observations.

Figure 7 illustrates the temporal trends of the spatially weighted average of random forest (RF) estimations alongside the R and RMSE values for site verification. A comparison of the RMSE and R values over time reveals that the R value consistently exceeds 0.75, while the RMSE remains below 0.07, indicating the relatively high overall accuracy of the RF model. However, a more detailed analysis indicates that RF performance during the summer (June, July, August) and autumn (September, October, November) is generally lower than during spring and winter. Specifically, the overall R values in spring and winter exceed 0.8, with R values at many sites in spring reaching 0.9. The RMSE values for spring and winter are below 0.05, which is lower than those recorded for the other two seasons, suggesting that the estimation accuracy of RF is superior in spring and winter. Month-by-month analysis reveals that lower accuracy is predominantly observed in June and October, with R and RMSE values of 0.78 and 0.065 in June and 0.75 and 0.068 in October, respectively.

Figure 8 illustrates the spatial distribution of the time-averaged R values and RMSE values. The overall maximum R value was 0.96, while the minimum is 0.8. Analyzing the spatial distribution of R values reveals that high values are predominantly located in the northern regions, whereas lower values are sparsely distributed in the central area. A more significant number of sites exhibit overall R values exceeding 0.9, suggesting that the spatial estimation accuracy of the random forest (RF) model is high. The maximum RMSE recorded is 0.12, with a minimum of 0.02. From a spatial perspective, the maximum RMSE occurs in the relatively humid region southeast of the North China Plain (NCP). Given that only one station reports an RMSE of 0.12 situated in this humid area, we consider this point an anomaly. Conversely, RMSE values below 0.02 are evenly distributed across the study

area. Overall, except for a few outliers, the spatial estimation accuracy of the RF model remains high.

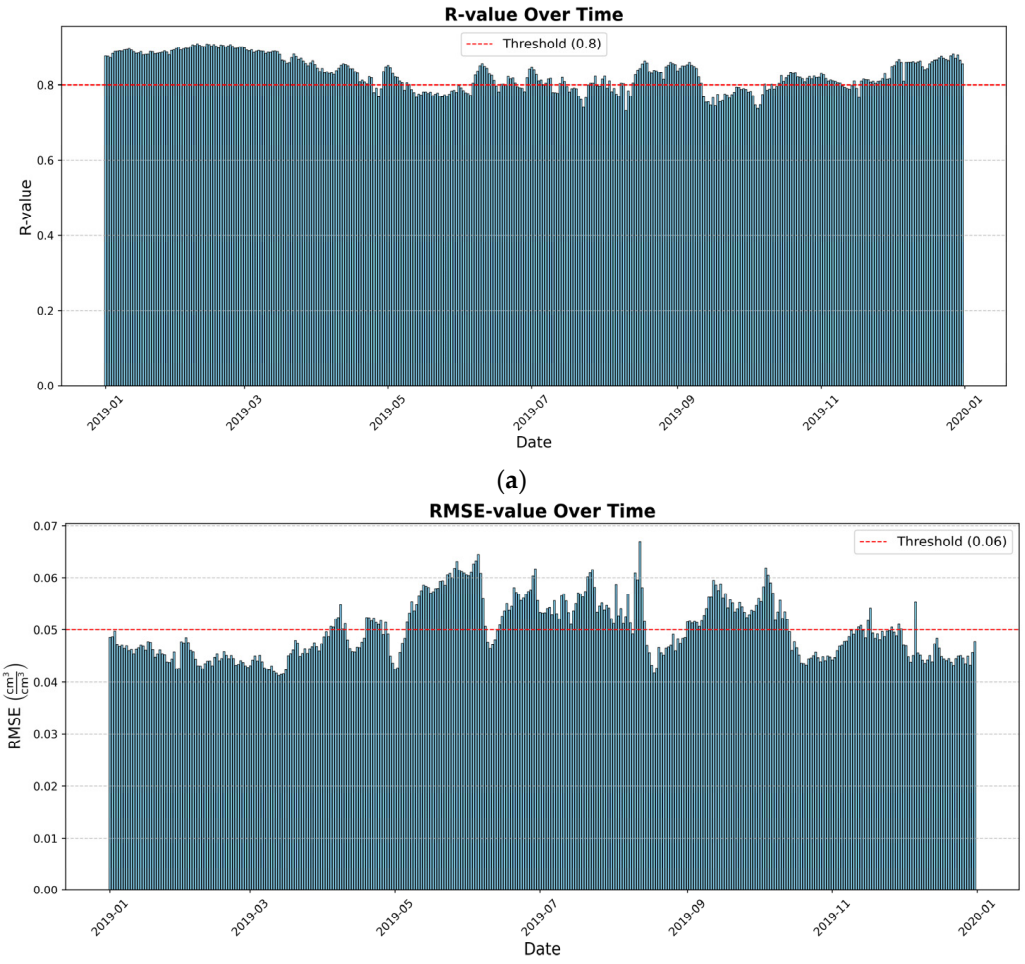


Figure 7. (**a**) Histogram of R value over time series; (**b**) Histogram of RMSE value over time series.

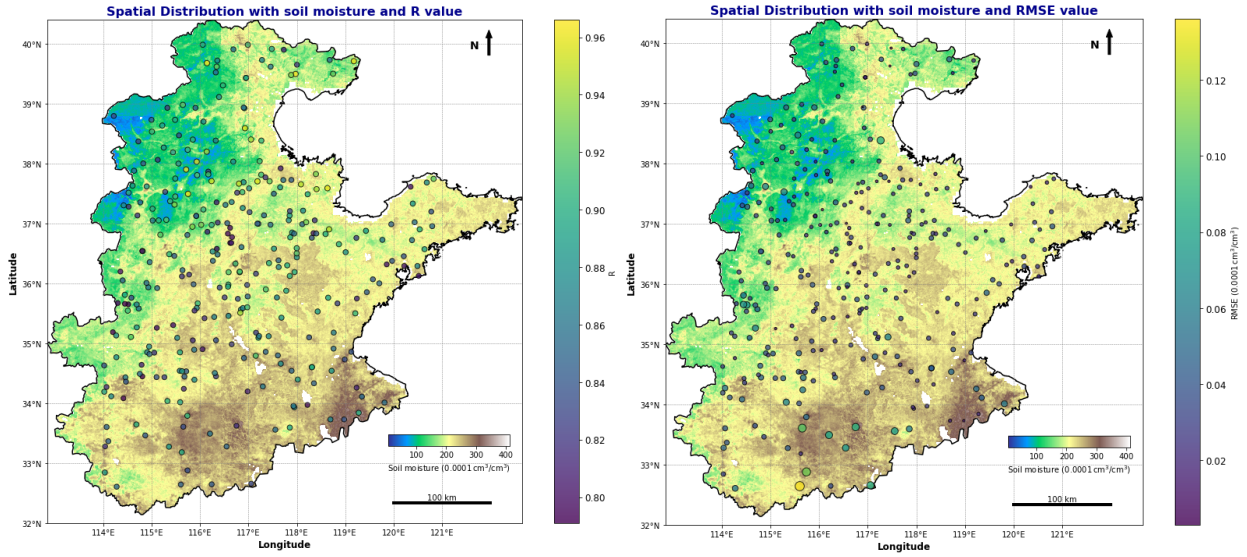
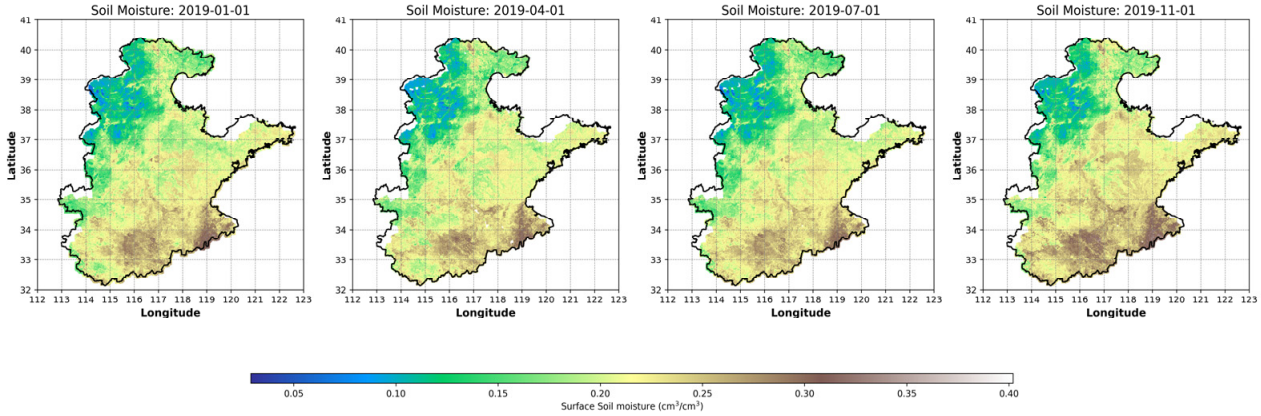


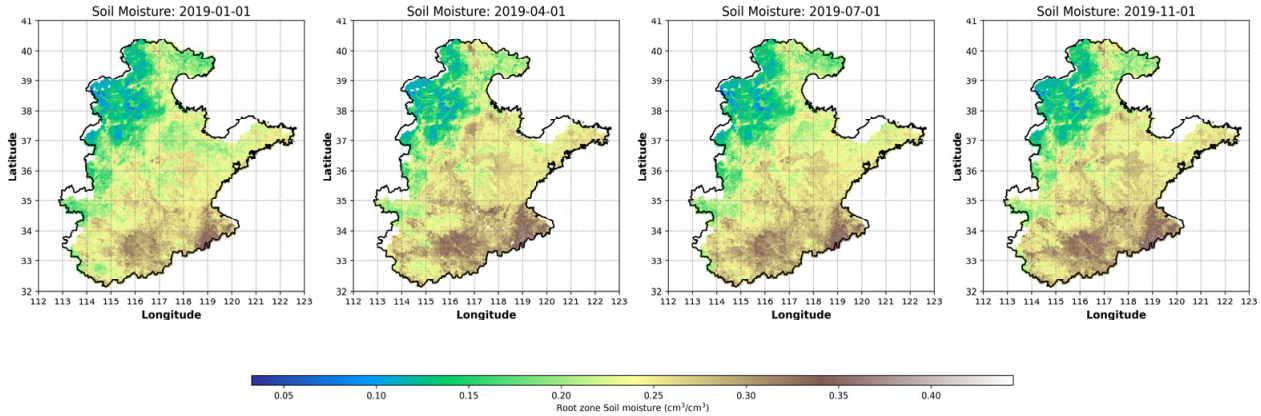
Figure 8. Spatial distribution of R values and RMSE values.

4.3. Estimation of RZSM in the NCP

Figure 9 illustrates the spatial distribution of SSM at 10 cm depth and RZSM at 100 cm depth, as calculated by the SSMRE model. An analysis of the spatial distribution of soil moisture across the four seasons—January, April, July, and November—revealed that the overall soil moisture was lowest on 1 July and highest on 1 November. Northern parts of the NCP tend to be drier than southern regions, with the driest conditions observed in northwestern areas such as Hebei Province, Beijing, and Tianjin. In contrast, southeastern regions, including Shandong, Anhui, and Jiangsu provinces, are comparatively more humid, with the highest SM levels found in northwest Jiangsu and northern Anhui.



(**a**)



(**b**)

Figure 9. (**a**) Spatial distribution of SSM; (**b**) Spatial distribution of RZSM.

A comparison of SSM and RZSM vertical profiles across months indicates that RZSM is generally higher than SSM due to reduced climatic sensitivity in the root zone. The spatial distribution analysis further shows that southern regions exhibit notably higher SSM and RZSM levels in July, while northern areas display the opposite pattern. A temperate monsoon climate with distinct seasonal variations in precipitation characterizes the NCP. The summer months are the wet season, with concentrated rainfall, while spring and winter are dry seasons with relatively low precipitation. Although July is typically rainy and humid, with expected increases in SSM and RZSM across the study area, localized analysis reveals that SSM and RZSM values in northern regions are actually lower during this month, with minimum values reaching 0.02. This finding underscores a regional disparity where, during the rainy season, the already humid southern areas become wetter, while arid conditions in northern regions are intensified.

Figure 10 presents a time series comparison between simulated SM at various depths, as modeled by the SSMRE model, and the observed measurements. Based on a comparison between the spatially weighted average simulations of each soil layer and the corresponding observed data, the overall SSMRE model accurately captured the soil moisture change trend, with results broadly consistent with those observed at the measurement sites. Vertical comparison of soil moisture across different layers revealed that as soil depth increased, the moisture content gradually stabilized, with the simulation accuracy being higher for the surface layer compared to the deeper layers. Moreover, a comparison of simulated and observed values for different periods indicated that the simulated surface soil moisture was lower than the observed values during June and October.

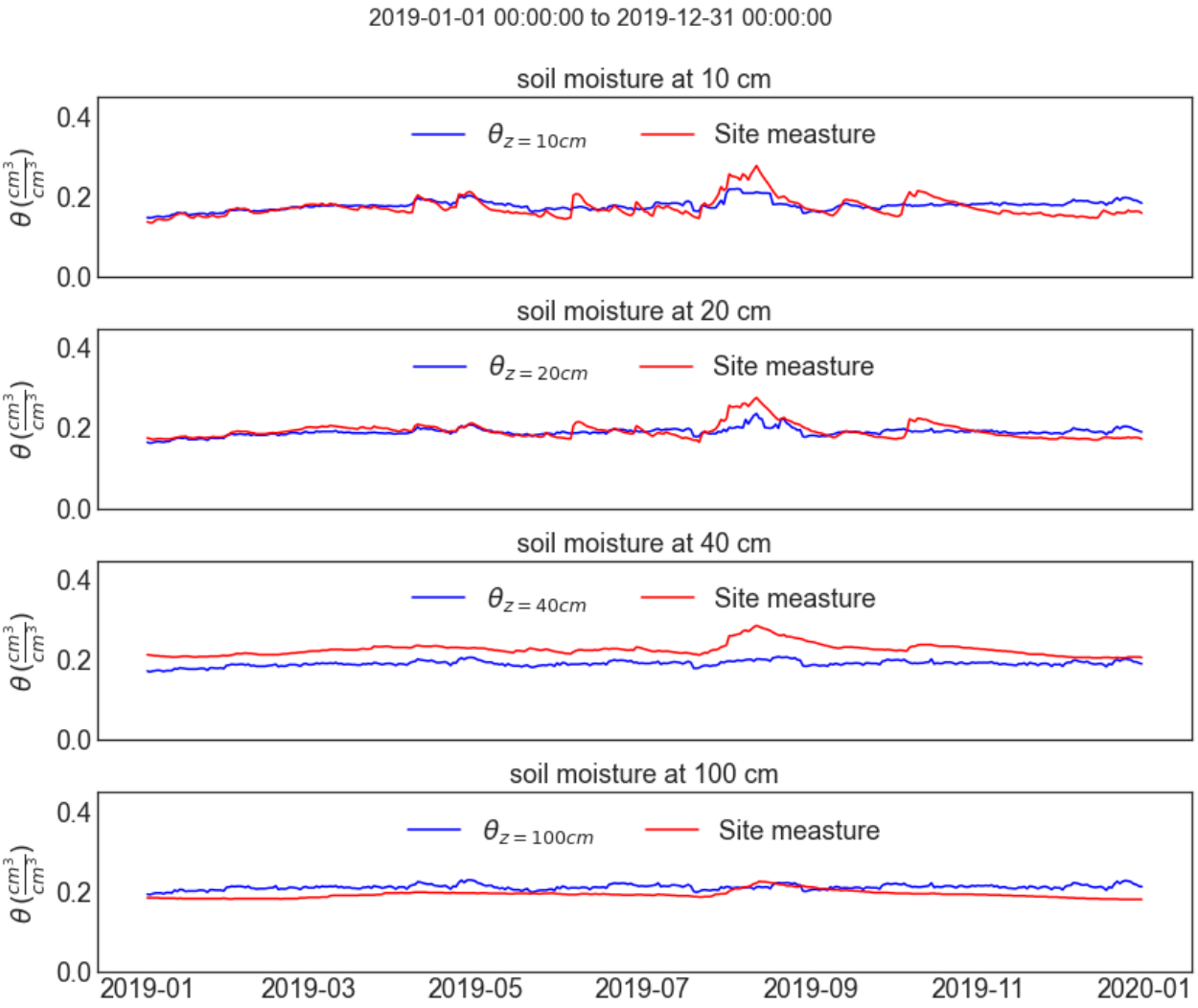


Figure 10. Comparison of soil moisture in each layer between SSMRE model simulation and site measured values.

Figure 11 presents the temporal variations in the RMSE values for surface and root-zone soil moisture. An analysis of the RMSE values for surface soil moisture reveals that the lowest RMSE, averaging 0.04, occurred between January and May. In contrast, the highest RMSE was observed in August, reaching 1.3, followed by June and October, with RMSE values of 0.079 and 0.072, respectively. For the root-zone soil moisture at a depth of 40 cm, the RMSE values were generally higher than those for the surface layer, with an average RMSE of 0.072 during the January–May period. The peak RMSE for the root zone was recorded in October at 0.9. Overall, higher RMSE values were observed in June,

August, and October, consistent with the findings from the RF verification. This suggests that the accuracy of surface soil moisture simulations plays a significant role in influencing the overall accuracy of soil moisture estimates.

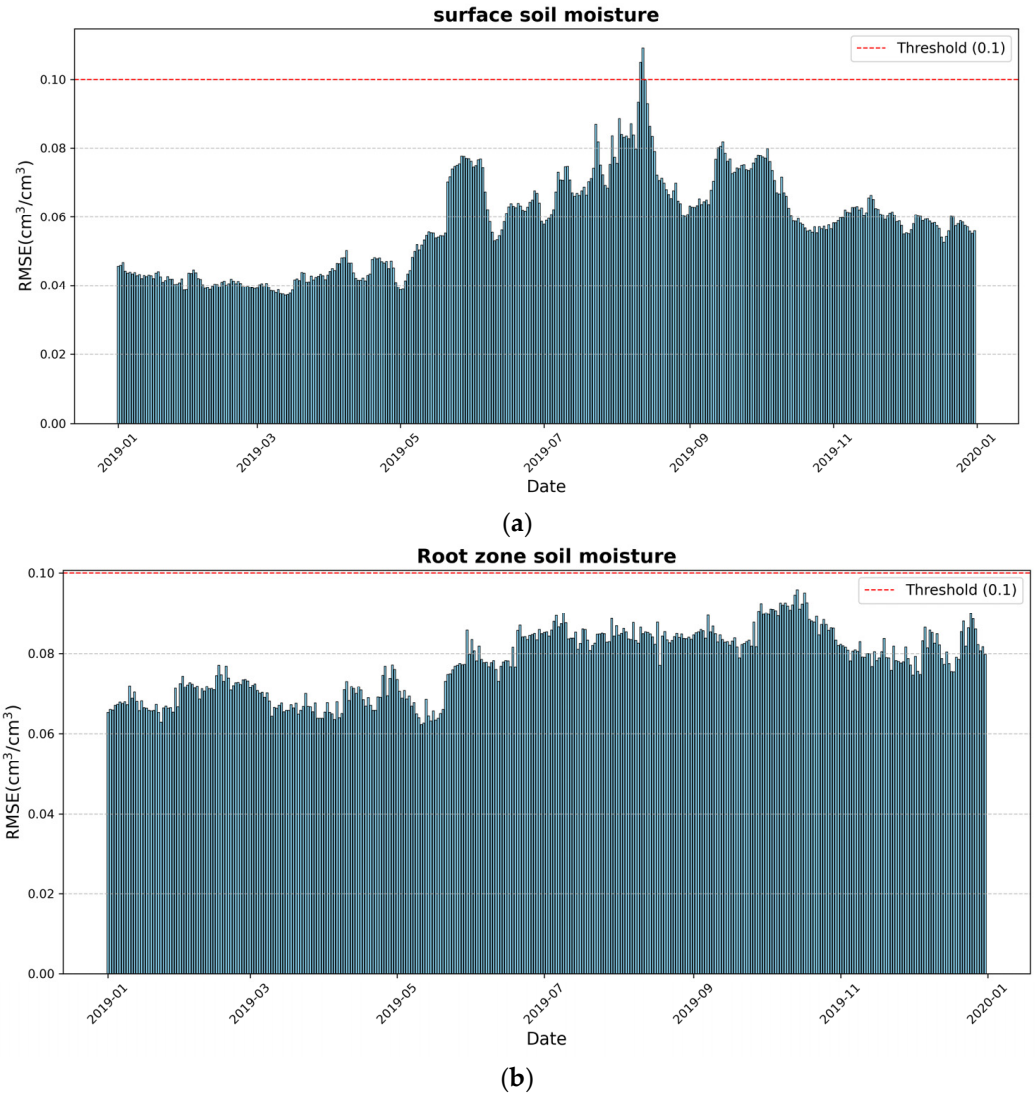


Figure 11. (**a**) SSM RMSE value; (**b**) Root layer soil moisture RMSE value changing trend.

Due to the limited number of monitoring stations, only 57 stations capable of measuring soil moisture at a depth of 100 cm were used to validate the soil moisture across various layers. Figure 12 shows the R and root RMSE values at each layer, based on these stations. A comparative analysis reveals that the R values gradually decrease as depth increases, while the RMSE values increase. The near-surface layer exhibits the highest accuracy, with R values greater than 0.85 at most stations and an RMSE of 0.04. R values for most stations are above 0.7 in the surface layer. The RMSE values in this layer range from 0.05 to 0.07, with outliers ranging from 0.1 to 0.11. Notably, the outlier stations are the same as those with low R values; both are concentrated in the D7806 region. For the root zone at a 40 cm depth, most stations show R values between 0.6 and 0.7, with RMSE values ranging from 0.8 to 0.9. At a depth of 100 cm, the R values for most stations are between 0.4 and 0.5, although some exceed 0.5. The RMSE values for this layer range from 0.1 to 0.13.



(**a**)



(**b**)

Figure 12. (**a**) R value of each layer verified at different sites; (**b**) RMSE value in each layer verified at different sites.

Figure 13 presents box plots of the R values and RMSE values for site verification across different soil layers. The analysis of these values at depths of 10 cm, 20 cm, 40 cm, and 100 cm reveals the following trends: the R values are 0.9, 0.77, 0.62, and 0.44, respectively, while the corresponding RMSE values are 0.05, 0.06, 0.09, and 0.11. This pattern demonstrates that as soil depth increases, the R values progressively decrease, while the RMSE values increase. These findings suggest that the accuracy and predictive performance of the SSMRE model diminish with increasing soil depth, indicating reduced model reliability for deeper soil layers.

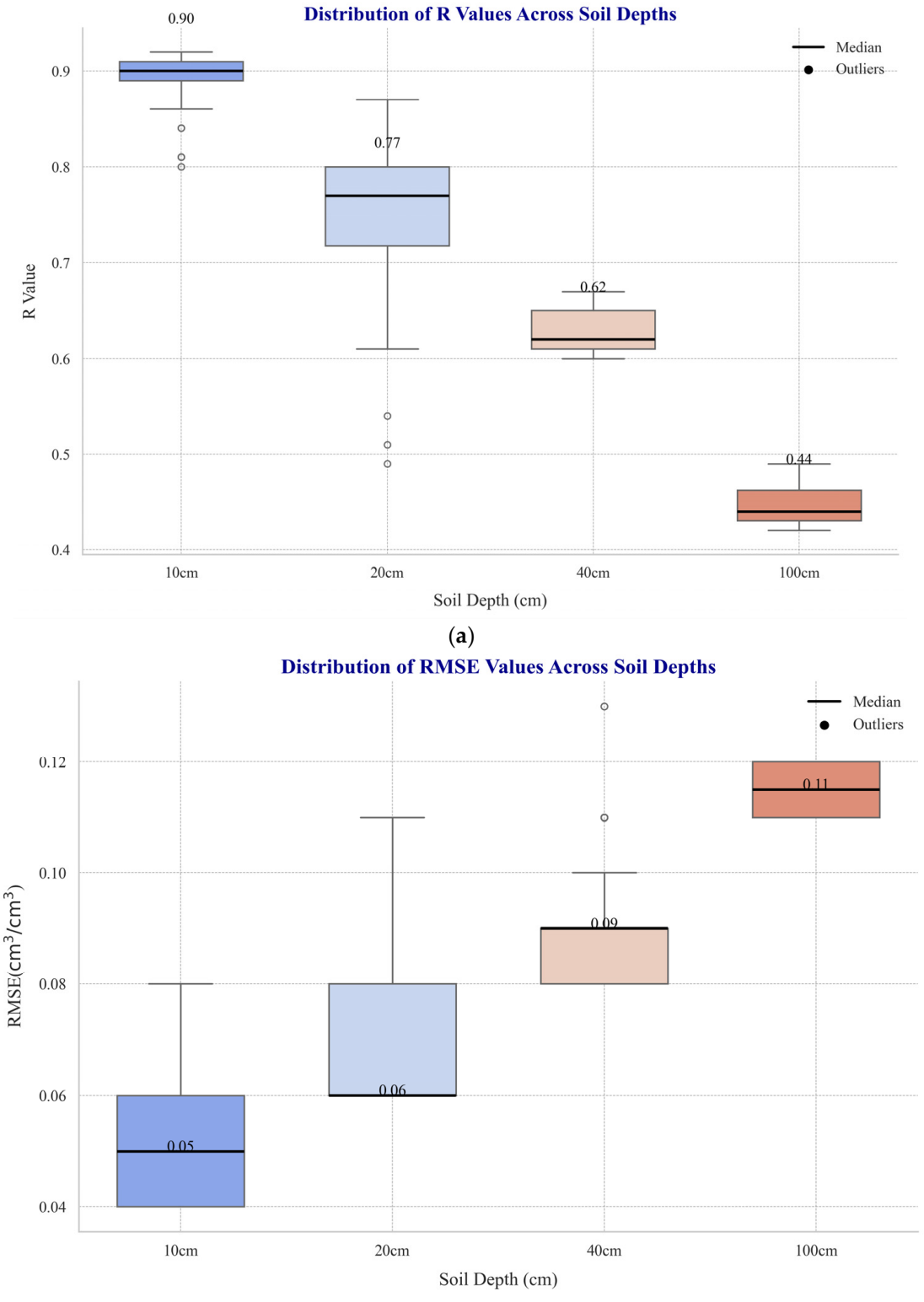


Figure 13. (**a**) The R values for the multi-layer soil profile were calculated using the Richards equation, with high-resolution SSM as the upper boundary condition; (**b**) The RMSE values for the multi-layer soil profile.

5. Discussion

SSM reflects the cumulative effect of multiple environmental factors, including evapotranspiration, rainfall, temperature, and irrigation, and can serve as a valuable indica-

tor of deep-layer SM dynamics. The Richards partial differential equation, commonly used to model SM dynamics, captures the nonlinear changes in soil moisture; however, many recent numerical models of this equation typically use moisture flux or meteorological variables as upper boundary conditions to estimate RZSM. While incorporating atmospheric boundary conditions enhances model accuracy, it also significantly increases computational complexity.

This study offers two primary advantages over models using atmospheric boundary conditions: (1) By utilizing SSM as an input for estimating deep soil moisture, the SSMRE method simplifies the calculation of complex upper boundary conditions, and (2) Temporal variations in SSM effectively mirror deep-layer SM patterns, validating the feasibility of estimating root-zone moisture using the Richards equation. The results of SSMRE modeling demonstrate strong performance, showing that SSM can reliably estimate root-zone moisture. This highlights the robustness of the SSMRE method in simplifying SM dynamics modeling while maintaining accuracy.

5.1. Performance and Covariate Importance Analysis of the RF Model

The spatial distribution and temporal trends of SSM are influenced by both surface factors (e.g., soil texture and topography) and atmospheric forcing factors (e.g., temperature, evaporation, and precipitation) [70,71]. Under drought conditions, the rate of SSM depletion varies based on climate characteristics, soil type and depth, vegetation cover, and the severity of drought [72,73]. Overall, the RF model demonstrates high accuracy in predicting surface SSM; however, areas with low SM exhibit greater model bias and uncertainty, with RF predictions typically exceeding observed values, a finding consistent with Du et al. (2024) [74]. Feature importance analysis reveals that dynamic forcing factors, such as air temperature (Tair) and evaporation, are primary drivers of SSM variability, while static factors, including latitude, longitude, and soil texture, have a secondary influence. The LST_diff exhibits the lowest contribution to predicting soil moisture dynamics. Tair directly influences soil moisture by driving surface evaporation; higher Tair accelerates evaporation, reducing soil moisture. Furthermore, Tair is closely associated with other meteorological variables, such as precipitation and evaporation, which enables it to indirectly capture the dynamics of soil moisture. These factors collectively account for the higher SHAP importance scores assigned to air temperature. By comparison, while LST_diff can reflect soil moisture conditions through surface temperature variations, such changes are limited to short-term fluctuations and confined to daily cycles. The effects of LST_diff are primarily observed as diurnal variations in evaporation rates, with little to no significant impact on soil moisture dynamics [75–77]. This distinction between dynamic and static variables emphasizes the predominant role of climate variability in driving SSM changes, with environmental properties providing an essential context for spatial distribution patterns.

Additionally, a comparison of spatial soil moisture trends between the monitoring sites and the RF simulations revealed certain discrepancies in June, August, and October. The histograms of R values and RMSE values over time indicated that, compared to the other months, the estimation accuracy of the RF model was lower during these three months. This reduction in accuracy is attributed to the abrupt changes in soil moisture that occurred during these periods. While the random forest (RF) model can efficiently capture complex nonlinear behaviors, it was trained using soil moisture data from only one year. As a result, when faced with sudden fluctuations that exceed the range or patterns present in the training data, the model's performance is limited. To enhance the simulation accuracy of RF, future studies should consider expanding the training dataset to include instances of sudden drought events and extreme fluctuations in soil moisture.

The spatial distribution of R values and RMSE shows that the estimation accuracy is higher in the arid northern regions compared to the more humid southern regions. Analyzing the underlying reasons reveals that soil moisture in the north experiences slower, more gradual changes, with relatively stable moisture patterns over time. In contrast, humid areas are subject to more complex climatic variability, where many factors, including precipitation, evaporation, and plant growth, influence soil moisture. These factors lead to more frequent and abrupt fluctuations in soil moisture levels. Consequently, soil moisture dynamics in humid regions are more volatile and heterogeneous, presenting a challenge for the random forest (RF) model to capture these sudden changes effectively. As a result, the model's accuracy is lower in humid areas, where the variability in soil moisture is more pronounced [78,79].

5.2. Performance of the SSMRE Model

While the SSMRE method performs well overall, its effectiveness diminishes with increasing depth in large-scale spatial applications, as indicated by the progressive decrease in R values and the corresponding increase in RMSE values with depth. This trend is consistent with the findings reported in other studies [80]. Soil hydraulic properties play a critical role in the SSMRE model, significantly affecting its predictive accuracy [81–84]. Validation of SM across different layers in the NCP reveals high precision for each layer simulated by the SSMRE model, with R values reaching up to 0.9 (e.g., Figure 10). However, when considering spatially averaged accuracy, a gradual decrease in R values with increasing depth is observed (Figure 13).

We employed the Monte Carlo method to perform sensitivity analysis (SSM) on the soil hydraulic parameters and boundary conditions to analyze the specific reasons for the model performance variation with increasing soil depth. Figure 14 illustrates the combined effects of the sensitivity parameters on the model outputs for each soil layer. The sensitivity of the model increases with soil depth. The root-zone soil moisture (RZSM) sensitivity index reaches −0.2. A detailed examination of the effects of different parameters on the soil moisture outputs at various depths, shown in Figure 15, reveals that SSM has the most significant impact on the model. The sensitivity of SSM for the 10 cm, 20 cm, 40 cm, and 100 cm layers are 1.00, 0.99, 0.96, and 0.49, respectively. The lower boundary condition (gbc) has the most significant influence on the RZSM, followed by the SSM. Among the soil hydraulic parameters, the residual water content (θ_r), shape factor (α), and saturated hydraulic conductivity (K_s) have sensitivity indices of 0.4, 0.28, and 0.17, respectively. Additionally, it was observed that the influence of soil hydraulic parameters on the model increases as the soil depth increases. Previous studies have demonstrated that uncertainty in soil hydraulic parameters significantly diminishes the performance of hydrological models, particularly impacting the simulation of RZSM [85]. In this study, the SSMRE model uses soil hydraulic parameters from the dataset, which is currently the only high-resolution, multi-layer spatial dataset of soil hydraulic properties available in China. However, due to limited observational data, the accuracy of this dataset remains unquantified. Additionally, the dataset highly depends on Pedotransfer Functions (PTFs), which estimate soil hydraulic parameters based on known soil properties. If the input soil characteristics are inaccurate, the derived hydraulic parameters are also likely to be imprecise [86]. Consequently, this study suggests that soil texture parameters likely explain the high accuracy observed in certain regions, despite the overall decline in accuracy with depth. Future research should focus on pixel-by-pixel optimization of SSMRE model parameters to improve accuracy in estimating deep soil moisture.

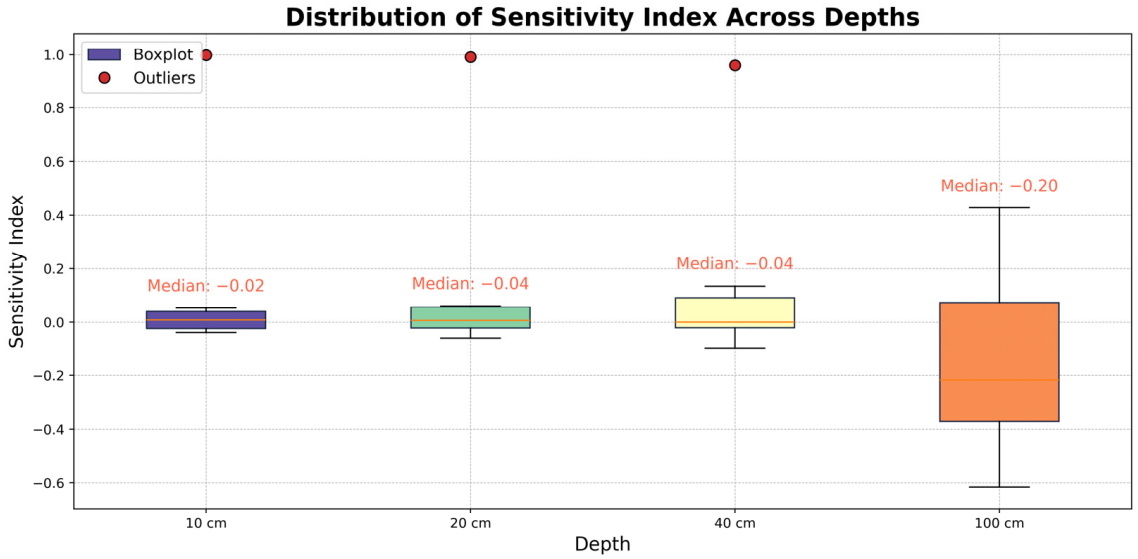


Figure 14. Box plots of parameter sensitivity analyses of soil moisture model outputs by layer.

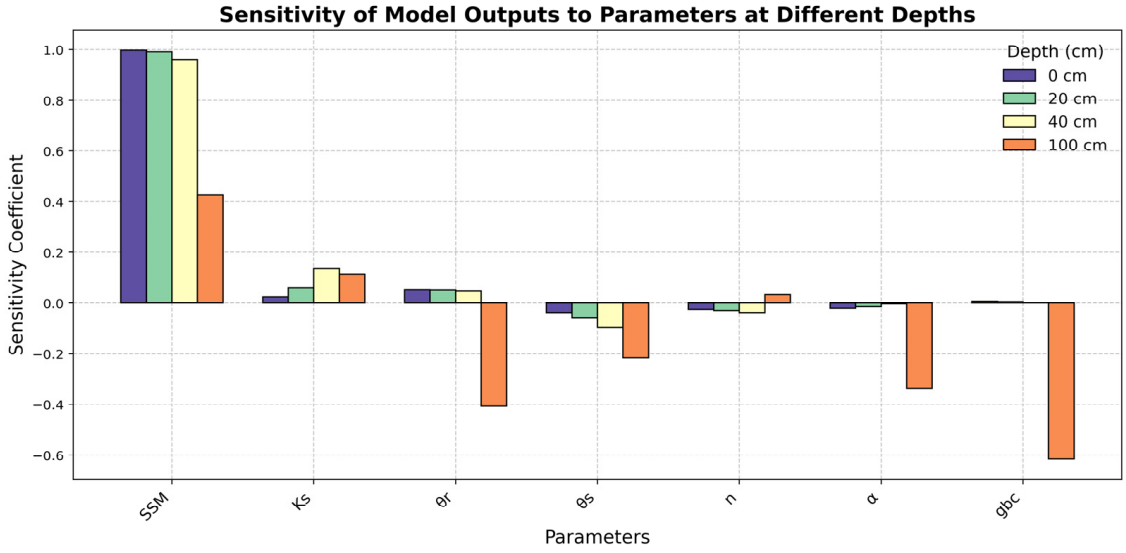


Figure 15. Histogram of the sensitivity analysis of different parameters.

Additionally, the R and RMSE values trends for both surface and root layers, as verified over the time series, are generally consistent with the results obtained from the RF model. Specifically, when the RF-estimated surface R value is higher, and RMSE values are lower, the accuracy of the SSMRE model also tends to be relatively high. Conversely, when the RF-estimated R value is lower and RMSE values are higher, the accuracy of the SSMRE model is also reduced. The same pattern is observed in the root layer estimates. These results highlight the critical role of SSM in determining the overall performance of models such as SSMRE. These models rely on accurate boundary conditions to estimate deeper soil moisture levels. Therefore, improving the accuracy of surface soil moisture estimates could substantially reduce the model's overall error.

5.3. Distribution Patterns of SM Across Different Layers in the NCP

As a key agricultural region, the NCP contributes significantly to China's agricultural output, comprising 20.4% of the nation's arable land and producing over 35% of its total grain. It is a crucial area for winter wheat production, accounting for 55% of the national wheat acreage and 60% of the total wheat yield [87,88]. Therefore, studying SM dynamics

in the NCP is essential for food security [89]. However, the region faces limited rainfall and frequent droughts, with research by Yuyin Chen et al. (2024) highlighting that the NCP has experienced three severe droughts in the past 17 years, including one in 2019 [90].

Time series SM data from Figure 14 reveal relatively low moisture levels across soil layers in June, the primary crop growth period, especially in deeper layers. In 2019, the NCP's spring and summer seasons experienced drought, with significantly reduced rainfall [91,92]. This low moisture is attributed to high temperatures, increased evaporation, and reduced rainfall during the summer. Given the spatial distribution of RZSM, optimizing crop layouts in the NCP is critical. For drier northwestern areas, planting drought-resistant crops and reducing corn acreage would reduce agricultural water consumption, conserve groundwater for subsequent wheat planting, and alleviate water scarcity [93]. Additionally, winter wheat cultivation depletes groundwater, exacerbated by insufficient natural precipitation during the wheat growing season. Reducing winter wheat acreage and implementing seasonal fallow periods could ease pressure on water resources, supporting sustainable groundwater management and national water security objectives [94,95].

6. Conclusions

SM is a critical variable in the hydrological cycle, essential for atmosphere–soil energy exchanges. It plays a pivotal role in agricultural drought monitoring, water resource management, and data assimilation for runoff predictions. This study first validated the performance of the SSMRE model using observational data from sites with identical vegetation cover but varying soil textures. To obtain high-resolution, large-scale RZSM estimates, we initially applied the SSA method to correct and fill missing values in site data with significant data gaps. Next, we constructed a machine learning model using processed site data and multiple spatial covariates associated with SSM to estimate SSM across the NCP at a 1 km resolution. Finally, using this 1 km SSM as the upper boundary condition in the Richards equation, we estimated large-scale RZSM, demonstrating the potential for integrating remote sensing data with the Richards equation to apply this approach at scale.

Overall, the RF model accurately estimated SSM at a 1 km resolution. However, time series analysis of SSM revealed an unusual trend: SSM values in the typically wet summer months (June and August) were similar to those in spring and winter, showing only minor variations. Additionally, RF predictions tended to overestimate SM during dry periods and underestimate it during wet periods. Based on other studies, we infer that a drought event affected the NCP in the summer of 2019. However, the RF model's performance in detecting drought anomalies, particularly flash droughts, remains to be evaluated. Feature importance analysis further revealed that dynamic variables, such as temperature and evaporation, exerted the most influence on RF predictions, followed by static factors, such as latitude, elevation, and soil texture.

In estimating RZSM using the SSMRE model, we observed a decline in accuracy with increasing soil depth, with SSM showing higher precision than deeper layers. This study did not perform spatial optimization of the model parameters; instead, it relied on China's only available soil texture dataset, although its accuracy has not been verified. Future research should focus on optimizing spatial soil texture parameters to improve the performance of the SSMRE model.

Spatial distribution analysis of the NCP revealed that the southern region is the wettest, while the northern region remains persistently dry year-round. During rainy seasons, the wetter southern areas become even more saturated, whereas the dry northern regions face exacerbated aridity. Time series analysis of SM layers further confirmed that the severe drought in 2019 led to uncharacteristically low SM in the typically rainy summer. Given the NCP's importance as China's largest agricultural zone, adjusting crop layout and selection

based on the spatial distribution of RZSM could enable more efficient water resource utilization across the region.

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

Funding: This study was financially supported by the China National Key R&D Program (No. 2021YFC3000201).

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: We sincerely thank the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, for providing the scientific research platform and data. We are also deeply grateful to Lü Aifeng of the institute for their valuable assistance in revising the manuscript.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Rigden, A.J.; Mueller, N.D.; Holbrook, N.M.; Pillai, N.; Huybers, P. Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting us maize yields. *Nat. Food* **2020**, *1*, 127–133. [CrossRef] [PubMed]
2. Tian, J.; Zhang, B.Q.; He, C.S.; Han, Z.; Bogena, H.R.; Huisman, J.A. Dynamic response patterns of profile soil moisture wetting events under different land covers in the Mountainous area of the Heihe River Watershed, Northwest China. *Agric. For. Meteorol.* **2019**, *271*, 225–239. [CrossRef]
3. Bolten, J.D.; Crow, W.T.; Zhan, X.; Jackson, T.J.; Reynolds, C.A. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *3*, 57–66. [CrossRef]
4. Kurc, S.A.; Small, E.E. Soil moisture variations and ecosystem-scale fluxes of water and carbon in semiarid grassland and shrubland. *Water Resour. Res.* **2007**, *43*. [CrossRef]
5. Nwogwu, N.A.; Okereke, N.; Ohanyere, S.; Chikwue, M.I. A Concise Review of Various Soil Moisture Measurement Techniques. In Proceedings of the 3rd South East Regional Conference of the Nigerian Institution of Agricultural Engineers, South East Region (NIAE-SER), Nsukka, Nigeria, 27–30 August 2018; pp. 613–624.
6. Ulaby, F.T.; Dubois, P.C.; Van Zyl, J. Radar mapping of surface soil moisture. *J. Hydrol.* **1996**, *184*, 57–84. [CrossRef]
7. Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ers scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207. [CrossRef]
8. Albergel, C.; Rüdiger, C.; Pellarin, T.; Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Piguet, B.; Martin, E. From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. *Hydrol. Earth Syst. Sci.* **2008**, *12*, 1323–1337. [CrossRef]
9. Gao, X.; Zhao, X.; Brocca, L.; Pan, D.; Wu, P. Testing of observation operators designed to estimate profile soil moisture from surface measurements. *Hydrol. Process.* **2019**, *33*, 575–584. [CrossRef]
10. Zhuang, R.; Zeng, Y.; Manfreda, S.; Su, Z. Quantifying long-term land surface and root zone soil moisture over tibetan plateau. *Remote Sens.* **2020**, *12*, 509. [CrossRef]
11. Pezij, M.; Augustijn, D.C.; Hendriks, D.M.; Hulscher, S.J. Applying transfer function-noise modelling to characterize soil moisture dynamics: A data-driven approach using remote sensing data. *Environ. Model. Softw.* **2020**, *131*, 104756. [CrossRef]
12. Cordova, J.R.; Bras, R.L. Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration. *Water Resour. Res.* **1981**, *17*, 93–106. [CrossRef]
13. Porporato, A.; Daly, E.; Rodriguez-Iturbe, I. Soil water balance and ecosystem response to climate change. *Am. Nat.* **2004**, *164*, 625–632. [CrossRef] [PubMed]
14. McColl, K.A.; Wang, W.; Peng, B.; Akbar, R.; Gianotti, D.J.S.; Lu, H.; Pan, M.; Entekhabi, D. Global characterization of surface soil moisture drydowns. *Geophys. Res. Lett.* **2017**, *44*, 3682–3690. [CrossRef]
15. Purdy, A.J.; Fisher, J.B.; Goulden, M.L.; Colliander, A.; Halverson, G.; Tu, K.; Famiglietti, J.S. SMAP soil moisture improves global evapotranspiration. *Remote Sens. Environ.* **2018**, *219*, 1–14. [CrossRef]

16. Almagro, A.; Oliveira, P.T.S.; Brocca, L. Assessment of bottom-up satellite rainfall products on estimating river discharge hydrologic signatures in Brazilian catchments. *J. Hydrol.* **2021**, *603*, 126897. [CrossRef]
17. Lawston, P.M.; Santanello, J.A., Jr.; Kumar, S.V. Irrigation signals detected from SMAP soil moisture retrievals. *Geophys. Res. Lett.* **2017**, *44*, 11860–11867. [CrossRef]
18. Cao, G.; Zheng, C.; Scanlon, B.R.; Liu, J.; Li, W. Use of flow modeling to assess sustainability of groundwater resources in the North China Plain. *Water Resour. Res.* **2013**, *49*, 159–175. [CrossRef]
19. Feng, T.; Shen, Y.; Chen, Q. Groundwater storage change driving factor analysis in North China using independent component decomposition. *J. Hydrol.* **2022**, *609*, 127708. [CrossRef]
20. Xiong, W.; Matthews, R.; Holman, I.; Lin, E.; Xu, Y. Modelling China's potential maize production at regional scale under climate change. *Clim. Chang.* **2007**, *85*, 433–451. [CrossRef]
21. Mo, X.; Liu, S.; Lin, Z.; Guo, R. Regional crop yield, water consumption and water use efficiency and their responses to climate change in the North China Plain. *Agric. Ecosyst. Environ.* **2009**, *134*, 67–78. [CrossRef]
22. Zhao, D.; Shen, J.; Lang, K.; Liu, Q.; Li, Q. Effects of irrigation and wide-precision planting on water use, radiation interception, and grain yield of winter wheat in the North China Plain. *Agric. Water Manag.* **2013**, *118*, 87–92.
23. Feng, W.; Zhong, M.; Lemoine, J.; Biancale, R.; Hsu, H.; Xia, J. Evaluation of groundwater depletion in North China using the gravity recovery and climate experiment (GRACE) data and ground-based measurements. *Water Resour. Res.* **2013**, *49*, 2110–2118. [CrossRef]
24. Huang, Z.; Pan, Y.; Gong, H.; Yeh, P.J.F.; Li, X.; Zhou, D.; Zhao, W. Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *Geophs. Res. Lett.* **2015**, *42*, 1791–1799. [CrossRef]
25. Pan, Y.; Zhang, C.; Gong, H.; Yeh, P.J.; Shen, Y.; Guo, Y.; Huang, Z.; Li, X. Detection of human-induced evapotranspiration using GRACE satellite observations in the Haihe River basin of China. *Geophs. Res. Lett.* **2017**, *44*, 190–199. [CrossRef]
26. Pang, Y.; Wu, B.; Cao, Y.; Jia, X. Spatiotemporal changes in terrestrial water storage in the Beijing-Tianjin sandstorm source region from GRACE satellites. *Int. Soil Water Conserv. Res.* **2020**, *8*, 295–307. [CrossRef]
27. Hao, H.; Liu, H.; Zhang, X. Terrestrial water storage variation in Hebei plain area of China, based on ground surface gravimetry. *Geod. Geodyn.* **2021**, *12*, 190–196. [CrossRef]
28. Wu, T.; Li, B.; Lian, L.; Zhu, Y.; Chen, Y. Assessment of the combined risk of drought and high-temperature heat wave events in the North China Plain during summer. *Remote Sens.* **2022**, *14*, 4588. [CrossRef]
29. Liu, J.; Si, Z.; Li, S.; Hamani, A.K.M.; Zhang, Y.; Wu, L.; Gao, Y.; Duan, A. Variations in water sources used by winter wheat across distinct rainfall years in the North China Plain. *J. Hydrol.* **2023**, *618*, 129186. [CrossRef]
30. Shu, J.; Bai, Y.; Chen, Q.; Weng, C.; Zhang, F. Dynamic simulation of the water-land-food nexus for the sustainable agricultural development in the North China Plain. *Sci. Total Environ.* **2024**, *912*, 168771. [CrossRef]
31. Li, X.; Ren, G.; You, Q.; Wang, S.; Zhang, W. Soil moisture continues declining in North China over the regional warming slowdown of the past 20 years. *J. Hydrometeorol.* **2021**, *22*, 3001–3015. [CrossRef]
32. Cao, M.; Chen, M.; Liu, J.; Liu, Y. Assessing the performance of satellite soil moisture on agricultural drought monitoring in the North China Plain. *Agric. Water Manag.* **2022**, *263*, 107450. [CrossRef]
33. Feng, T.; Shen, Y.; Wang, F.; Chen, Q.; Ji, K. Spatiotemporal variability and driving factors of the shallow soil moisture in North China during the past 31 years. *J. Hydrol.* **2023**, *619*, 129331. [CrossRef]
34. Berg, A.; Sheffield, J.; Milly, P.C.D. Divergent surface and total soil moisture projections under global warming. *Geophys. Res. Lett.* **2017**, *44*, 236–244. [CrossRef]
35. Yuan, W.; Zheng, Y.; Piao, S.; Ciais, P.; Lombardozzi, D.; Wang, Y.; Ryu, Y.; Chen, G.; Dong, W.; Hu, Z.; et al. Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Sci. Adv.* **2019**, *5*, eaax1396. [CrossRef]
36. Zhang, Y.; Parazoo, N.C.; Williams, A.P.; Zhou, S.; Gentine, P. Large and projected strengthening moisture limitation on end-of-season photosynthesis. *Proc. Natl. Acad. Sci.* **2020**, *117*, 9216–9222. [CrossRef]
37. Bablet, A.; Viallefont-Robinet, F.; Jacquemoud, S.; Fabre, S.; Briottet, X. Highresolution mapping of in-depth SM through a laboratory experiment coupling a spectroradiometer and two hyperspectral cameras. *Remote Sens. Environ.* **2020**, *236*, 111533. [CrossRef]
38. Dumedah, G.; Walker, J.P.; Merlin, O. Root-zone soil moisture estimation from assimilation of downscaled soil moisture and ocean salinity data. *Adv. Water Resour.* **2015**, *84*, 14–22. [CrossRef]
39. Merlin, O.; Chehbouni, A.; Boulet, G.; Kerr, Y. Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data. *J. Hydrometeorol.* **2006**, *7*, 1308–1322. [CrossRef]
40. Montaldo, N.; Albertson, J.D. Multi-scale assimilation of surface soil moisture data for robust root zone moisture predictions. *Adv. Water Resour.* **2003**, *26*, 33–44. [CrossRef]
41. Mascaro, G.; Vivoni, E.R.; Deidda, R. Soil moisture downscaling across climate regions and its emergent properties. *J. Geophys. Res. Atmos.* **2011**, *116*. [CrossRef]

42. Verma, K.; Katpatal, Y.B. Groundwater monitoring using GRACE and GLDAS data after downscaling within basaltic aquifer system. *Groundwater* **2020**, *58*, 143–151. [CrossRef] [PubMed]
43. Xu, Y.; Wang, L.; Ma, Z.; Li, B.; Bartels, R.; Liu, C.; Zhang, X.; Dong, J. Spatially explicit model for statistical downscaling of satellite passive microwave soil moisture. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 1182–1191. [CrossRef]
44. Cheng, M.; Zhong, L.; Ma, Y.; Zou, M.; Ge, N.; Wang, X.; Hu, Y. A study on the assessment of multi-source satellite soil moisture products and reanalysis data for the Tibetan Plateau. *Remote Sens.* **2019**, *11*, 1196. [CrossRef]
45. Park, S.; Park, S.; Im, J.; Rhee, J.; Shin, J.; Park, J.D. Downscaling GLDAS soil moisture data in East Asia through fusion of multi-sensors by optimizing modified regression trees. *Water* **2017**, *9*, 332. [CrossRef]
46. Han, L.; Wang, C.; Liu, Q.; Wang, G.; Yu, T.; Gu, X.; Zhang, Y. Soil moisture mapping based on multi-source fusion of optical, near-infrared, thermal infrared, and digital elevation model data via the bayesian maximum entropy framework. *Remote Sens.* **2020**, *12*, 3916. [CrossRef]
47. Abbaszadeh, P.; Moradkhani, H.; Zhan, X. Downscaling SMAP radiometer soil moisture over the CONUS using an ensemble learning method. *Water Resour. Res.* **2019**, *55*, 324–344. [CrossRef]
48. Fang, K.; Shen, C. Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive data integration kernel. *J. Hydrometeorol.* **2020**, *21*, 399–413. [CrossRef]
49. Liu, Y.; Jing, W.; Wang, Q.; Xia, X. Generating high-resolution soil moisture by using spatial downscaling techniques: A comparison of six machine learning algorithms. *Adv. Water Resour.* **2020**, *141*, 103601. [CrossRef]
50. Liu, Y.; Xia, X.; Yao, L.; Jing, W.; Zhou, C.; Huang, W.; 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]
51. Mao, H.; Kathuria, D.; Duffield, N.; Mohanty, B.P. Gap filling of high-resolution soil moisture for SMAP/Sentinel-1: A two-layer machine learning-based framework. *Water Resour. Res.* **2019**, *55*, 6986–7009. [CrossRef]
52. Muñoz-Sabater, J.; Dutra, E.; Agustí-Panareda, A.; Albergel, C.; Arduini, G.; Balsamo, G.; Boussetta, S.; Choulga, M.; Harrigan, S.; Hersbach, H.; et al. ERA5-Land: A state-of-theart global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **2021**, *13*, 4349–4383. [CrossRef]
53. Han, Y.; Bai, X.; Shao, W.; Wang, J. Retrieval of soil moisture by integrating Sentinel-1A and MODIS data over agricultural fields. *Water* **2020**, *12*, 1726. [CrossRef]
54. Mohseni, F.; Mokhtarzade, M. A new soil moisture index driven from an adapted long-term temperature-vegetation scatter plot using MODIS data. *J. Hydrol.* **2020**, *581*, 124420. [CrossRef]
55. Yuan, L.; Li, L.; Zhang, T.; Chen, L.; Zhao, J.; Hu, S.; Liu, W. Soil moisture estimation for the Chinese Loess Plateau using MODIS-derived ATI and TVDI. *Remote Sens.* **2020**, *12*, 3040. [CrossRef]
56. Trok, J.T.; Davenport, F.V.; Barnes, E.A.; Diffenbaugh, N.S. Using Machine Learning with Partial Dependence Analysis to Investigate Coupling Between Soil Moisture and Near-Surface Temperature. *J. Geophys. Res. Atmos.* **2023**, *128*, e2022JD038365. [CrossRef]
57. Wieder, W.R.; Boehnert, J.; Bonan, G.B.; Langseth, M. *Regridded Harmonized World Soil Database v1.2. Data Set*; Oak Ridge National Laboratory Distributed Active Archive Center: Oak Ridge, TN, USA, 2014. Available online: http://daac.ornl.gov (accessed on 2 October 2024).
58. Vautard, R.; Yiou, P.; Ghil, M. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Phys. D Nonlinear Phenom.* **1992**, *58*, 95–126. [CrossRef]
59. Hassani, H. Singular spectrum analysis: Methodology and comparison. *J. Data Sci.* **2007**, *5*, 20. [CrossRef]
60. Bond, G.; Kromer, B.; Beer, J.; Muscheler, R.; Evans, M.N.; Showers, W.; Hoffmann, S.; Lotti-Bond, R.; Hajdas, I.; Bonani, G. Persistent solar influence on north Atlantic climate during the Holocene. *Science* **2001**, *294*, 2130–2136. [CrossRef]
61. Ghil, M.; Allen, M.R.; Dettinger, M.D.; Ide, K.; Kondrashov, D.; Mann, M.E.; Robertson, A.W.; Saunders, A.; Tian, Y.; Varadi, F.; et al. Advanced spectral methods for climatic time series. *Rev. Geophys.* **2002**, *40*, 3-1–3-42. [CrossRef]
62. Schlesinger, M.E.; Ramankutty, N. An oscillation in the global climate system of period 65–70 years. *Nature* **1994**, *367*, 723–726. [CrossRef]
63. Kondrashov, D.; Ghil, M. Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Process. Geophys.* **2006**, *13*, 151–159. [CrossRef]
64. Schoellhamer, D.H. Singular spectrum analysis for time series with missing data. *Geophys. Res. Lett.* **2001**, *28*, 3187–3190. [CrossRef]
65. Yi, S.; Sneeuw, N. Filling the data gaps within GRACE missions using singular spectrum analysis. *J. Geophys. Res. Solid Earth* **2021**, *126*, e2020JB021227. [CrossRef]
66. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
67. Zhang, Y.Q.; Chiew, F.H.S.; Li, M.; Post, D. Predicting runoff signatures using regression and hydrological modeling approaches. *Water Resour. Res.* **2018**, *54*, 7859–7878. [CrossRef]

68. Li, M.; Zhang, Y.; Wallace, J.; Campbell, E. Estimating annual runoff in response to forest change: A statistical method based on random forest. *J. Hydrol.* **2020**, *589*, 125168. [CrossRef]
69. Yang, Y.; Donohue, R.J.; McVicar, T.R. Global estimation of effective plant rooting depth: Implications for hydrological modeling. *Water Resour. Res.* **2016**, *52*, 8260–8276. [CrossRef]
70. Feddes, R.A.; Kowalik, P.J.; Zaradny, H. *Simulation of Field Water Use and Crop Yield*; Centre for Agricultural Publishing and Documentation: Wageningen, The Netherlands, 1978.
71. Brocca, L.; Melone, F.; Moramarco, T.; Morbidelli, R. Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resour. Res.* **2010**, *46*. [CrossRef]
72. Entin, J.K.; Robock, A.; Vinnikov, K.Y.; Hollinger, S.E.; Liu, S.; Namkhai, A. Temporal and spatial scales of observed soil moisture variations in the extratropics. *J. Geophys. Res.* **2000**, *105*, 11865–11877. [CrossRef]
73. McColl, K.A.; Alemohammad, S.H.; Akbar, R.; Konings, A.G.; Yueh, S.; Entekhabi, D. The global distribution and dynamics of surface soil moisture. *Nat. Geosci.* **2017**, *10*, 100–104. [CrossRef]
74. Vicente-Serrano, S.M.; Quiring, S.M.; Peña-Gallardo, M.; Yuan, S.; Domínguez-Castro, F. A review of environmental droughts: Increased risk under global warming? *Earth-Sci. Rev.* **2020**, *201*, 102953. [CrossRef]
75. Du, J.; Kimball, J.S.; Jencso, K.; Hoylman, Z.; Brust, C.; Ketchum, D.; Xu, Y.; Lu, H.; Sheffield, J. Machine-learning based multi-layer soil moisture forecasts—An application case study of the Montana 2017 flash drought. *Water Resour. Res.* **2024**, *60*, e2023WR036973. [CrossRef]
76. Schmugge, T.; Blanchard, B.; Anderson, A.; Wang, J. Soil moisture sensing with aircraft observations of the diurnal range of surface temperature 1. *JAWRA J. Am. Water Resour. Assoc.* **1978**, *14*, 169–178. [CrossRef]
77. Schlosser, C.A.; Milly, P.C.D. A model-based investigation of soil moisture predictability and associated climate predictability. *J. Hydrometeorol.* **2002**, *3*, 483–501. [CrossRef]
78. Ford, T.W.; Quiring, S.M. In situ soil moisture coupled with extreme temperatures: A study based on the Oklahoma Mesonet. *Geophys. Res. Lett.* **2014**, *41*, 4727–4734. [CrossRef]
79. Kellogg, W.W.; Zhao, Z.C. Sensitivity of soil moisture to doubling of carbon dioxide in climate model experiments. Part I: North America. *J. Clim.* **1988**, *1*, 348–366. [CrossRef]
80. Walker, J.P.; Houser, P.R. A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *J. Geophys. Res. Atmos.* **2001**, *106*, 11761–11774. [CrossRef]
81. He, J.; Hantush, M.M.; Kalin, L.; Rezaeianzadeh, M.; Isik, S. A two-layer numerical model of soil moisture dynamics: Model development. *J. Hydrol.* **2021**, *602*, 126797. [CrossRef]
82. Inforsato, L.; de Jong van Lier, Q.; Durner, W. A novel fitting procedure for soil hydraulic properties with improved parameter uncertainty assessment. In *EGU General Assembly Conference Abstracts, Proceedings of the EGU General Assembly 2022, Vienna, Austria & Online, 23–27 May 2022*; EGU: Munich, Germany, 2022; p. EGU22-8760.
83. da Silva, A.J.P.; Pinheiro, E.A.R.; de Jong van Lier, Q. Determination of soil hydraulic properties and its implications for mechanistic simulations and irrigation management. *Irrig. Sci.* **2020**, *38*, 223–234. [CrossRef]
84. Dastranj, M.; Duus Børgesen, C.; Vangsø Iversen, B. Analysis of Soil Hydraulic Parameters effects on soil water modelling based on Danish soil water monitoring systems. In *EGU General Assembly Conference Abstracts, Proceedings of the EGU General Assembly 2024, Vienna, Austria & Online, 14–19 April 2024*; EGU: Munich, Germany, 2024; p. 10446.
85. Wöhling, T.; Mietrach, R. Richard's equation revisited-the challenge to reconstruct non-equilibrium field retention data with soil hydraulic models. In *EGU General Assembly Conference Abstracts, Proceedings of the EGU General Assembly 2022, Vienna, Austria & Online, 23–27 May 2022*; EGU: Munich, Germany, 2022; p. EGU22-7865.
86. Baroni, G.; Facchi, A.; Gandolfi, C.; Ortuani, B.; Horeschi, D.; Van Dam, J.C. Uncertainty in the determination of soil hydraulic parameters and its influence on the performance of two hydrological models of different complexity. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 251–270. [CrossRef]
87. Dai, Y.; Shangguan, W.; Duan, Q.; Liu, B.; Fu, S.; Niu, G. Development of a China dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling. *J. Hydrometeorol.* **2013**, *14*, 869–887. [CrossRef]
88. Zhao, H.; Zheng, Y.; Zhang, Y.; Li, T. Evaluating the effects of surface O_3 on three main food crops across China during 2015–2018. *Environ. Pollut.* **2020**, *258*, 113794. [CrossRef] [PubMed]
89. Yin, J.; Yan, D.; Yang, Z.; Zhe, Y.; Yong, Y.; Cheng, Z. Projection of extreme precipitation in the context of climate change in Huang-Huai-Hai region, China. *J. Earth Syst. Sci.* **2016**, *125*, 417–429. [CrossRef]
90. Zhang, S.; Yang, S.; Wang, J.; Wu, X.; Malak, H.; Tehseen, J.; Zhang, J. BAI Yun. Integrating a novel irrigation approximation method with a process-based remote sensing model to estimate multi-years winter wheat yield over the North China Plain. *J. Integr. Agric.* **2023**, *22*, 2865–2881. [CrossRef]
91. Chen, Y.; Zhang, Y.; Tian, J.; Tang, Z.; Wang, L.; Yang, X. Understanding the Propagation of Meteorological Drought to Groundwater Drought: A Case Study of the North China Plain. *Water* **2024**, *16*, 501. [CrossRef]

92. Zhang, L.; Chu, Q.Q.; Jiang, Y.L.; Chen, F.; Lei, Y.D. Impacts of climate change on drought risk of winter wheat in the North China Plain. *J. Integr. Agric.* **2021**, *20*, 2601–2612. [CrossRef]
93. Wan, W.; Liu, Z.; Li, K.; Wang, G.; Wu, H.; Wang, Q. Drought monitoring of the maize planting areas in Northeast and North China Plain. *Agric. Water Manag.* **2021**, *245*, 106636. [CrossRef]
94. Ren, D.; Yang, Y.; Yang, Y.; Richards, K.; Zhou, X. Land-Water-Food Nexus and indications of crop adjustment for water shortage solution. *Sci. Total Environ.* **2018**, *626*, 11–21. [CrossRef]
95. Huang, F.; Yang, X.; Fang, Y.; Wang, S.; Kang, S. Exploring water-and-land-adapted spatial layout of crop planting in North China. *Strateg. Study Chin. Acad. Eng.* **2022**, *24*, 89–96. [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.