


Article

High-Resolution Mapping of Cropland Soil Organic Carbon in Northern China

Rui Wang ^{1,*}, Wenbo Du ¹, Ping Li ², Zelong Yao ² and Huiwen Tian ³ ¹ Shanxi Farmland Quality Monitoring and Protection Center, Taiyuan 030001, China² College of Resources and Environment, Shanxi Agricultural University, Taigu 030800, China³ College of Urban and Rural Construction, Shanxi Agricultural University, Taigu 030800, China

* Correspondence: wangruiisxy@126.com

Abstract: Mapping the high-precision spatiotemporal dynamics of soil organic carbon (SOC) in croplands is crucial for enhancing soil fertility and carbon sequestration and ensuring food security. We conducted field surveys and collected 1121 soil samples from cropland in Changzhi, northern China, in 2010 and 2020. Random Forest (RF) models combined with 19 environmental covariates were used to map the topsoil (0–20 cm) SOC in 2010 and 2020, and uncertainty maps were used to calculate the dynamic changes in cropland SOC between 2010 and 2020. Finally, RF and Structural Equation Modeling (SEM) were employed to explore the effects of climate, vegetation, topography, soil properties, and agricultural management on SOC variation in croplands. Compared to the prediction model using only natural variables (RF_C), the model incorporating agricultural management (RF_A) significantly improved the simulation accuracy of SOC. The coefficient of determination (R^2) increased from 0.77 to 0.85, while the Root Mean Square Error (RMSE) decreased from 1.74 to 1.53 g kg⁻¹, and the Mean Absolute Error (MAE) was reduced from 1.10 to 0.94 g kg⁻¹. The uncertainty in our predictions was low, with an average value of only 0.39–0.66 g kg⁻¹. From 2010 to 2020, SOC in the Changzhi croplands exhibited an overall increasing trend, with an average increase of 1.57 g kg⁻¹. Climate change, agricultural management, and soil properties strongly influence SOC variation. Mean annual precipitation (MAP), drainage condition (DC), and net primary productivity (NPP) were the primary drivers of SOC variability. Our findings highlight the effectiveness of agricultural management for predicting SOC in croplands. Overall, the study confirms that improved agricultural management has great potential to increase soil carbon stocks, which may contribute to sustainable agricultural development.

Keywords: soil organic carbon; soil science; digital soil mapping; random forest; agricultural management



Academic Editors: Audrius Kačergius and Eugenija Bakšienė

Received: 23 December 2024

Revised: 24 January 2025

Accepted: 28 January 2025

Published: 30 January 2025

Citation: Wang, R.; Du, W.; Li, P.; Yao, Z.; Tian, H. High-Resolution Mapping of Cropland Soil Organic Carbon in Northern China. *Agronomy* **2025**, *15*, 359. <https://doi.org/10.3390/agronomy15020359>

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

Soil Organic Carbon (SOC) is a crucial factor for maintaining the stability and function of terrestrial ecosystems and serves as a bridge between various natural components [1]. The total SOC stock (within 1 m depth) is estimated to be 2- to 3-fold higher than that of the terrestrial vegetation carbon pool and more than twice that of the atmospheric carbon pool [2]. Even small changes in SOC can exert significant effects on atmospheric CO₂ concentrations. The decomposition and transformation of organic carbon are accompanied by the release, utilization, and redistribution of life elements that influence plant growth on the surface. In agricultural ecosystems, SOC is commonly used as an indicator of soil fertility, agricultural production potential, and land degradation [3,4]. Therefore, mapping

SOC in croplands is important for regional ecosystem stability, improving agricultural management, increasing crop yields, and promoting sustainable agricultural development.

According to the soil formation theory, SOC can be considered a function of environmental factors such as climate, topography, and soil properties [5,6]. Digital Soil Mapping (DSM) has utilized these environmental covariates to create dynamic maps of soil organic carbon at regional, national, and global scales [7–11]. Stockmann et al. employed this method to develop a global surface soil organic carbon model, with land cover as the primary driver of SOC variation, resulting in global SOC change maps at 1 km resolution for the years 2001 and 2009 [7]. Similarly, Wang et al. and Liu et al. used legacy soil data and the CENTURY model to map soil organic carbon in China and its alpine regions, providing insights into SOC stock changes over the past three decades [8,9]. However, these studies overlooked the impact of human activities on SOC, particularly in agricultural areas where human activities are intensive. In fact, agricultural practices have a significant influence on the changes in SOC in cropland areas [12,13]. Specifically, once agricultural activities such as irrigation levels and fertilization practices interfere with the soil, SOC undergoes changes within a short time, leading to greater spatial heterogeneity of SOC in croplands [14,15]. Therefore, the integration of agricultural management information possesses the potential to reveal dynamic changes in cropland SOC and improve the accuracy of SOC predictions. However, due to the large workforce and time required to collect agricultural management data, most DSM studies overlook these factors, limiting the accuracy of the results and reducing their value for guiding agricultural production at the field scale. Although some researchers have attempted to incorporate agricultural management information, in addition to natural environmental covariates, into SOC mapping—where agricultural management information mainly includes indirect relationship data such as distances to roads, rivers, irrigation canals, and villages [16], or panel data from statistical yearbooks [17]—the coarse spatial resolution of these data may introduce biases into DSM products and limit predictive accuracy [18]. Finally, these studies have predominantly focused on the spatial variability of cropland SOC rather than its temporal changes [15,19]. This gap makes it challenging to accurately capture the dynamics of the SOC over time. Thus, it is possible to model dynamic changes in cropland SOC by combining direct measurements of SOC changes over two periods from regional-scale resampling activities and integrating regional agricultural management information. This approach can provide key information regarding how agricultural management practices and natural environmental variables influence the SOC of croplands.

Changzhi, located on the southeastern edge of the Loess Plateau in China, possesses extensive mountainous areas, leading to a fragmented and scattered cropland distribution. Therefore, spatiotemporal SOC data from croplands are urgently needed to support precise soil management and agricultural production. However, relevant information remains scarce. To fill these gaps, the main objectives of this study are:

- (1) To introduce agricultural management information and natural environmental variables into DSM, using Random Forest (RF) models to map the 30 m resolution SOC in cropland for 2010 and 2020;
- (2) To assess the prediction accuracy and uncertainty of the SOC maps, examine the effectiveness of agricultural management information in predicting SOC, and estimate the changes in cropland SOC stocks between the two periods;
- (3) To use RF and Structural Equation Modeling (SEM) to quantify the effects of agricultural management and natural environmental variables on SOC variability, identifying the key factors influencing SOC changes.

This study aimed to produce spatiotemporal SOC change maps for Changzhi, providing more effective guidance for regional soil quality improvement and agricultural

development and offering a basis for designing climate change, carbon neutrality, and soil security policies.

2. Materials and Methods

2.1. Study Area

Changzhi is located in the southeastern Loess Plateau (Figure 1a). The total area of the city is 13,955 km², with approximately 3715 km² of croplands. The elevation varies by 2100 m, and the topography is complex. Mountainous areas account for 50.7% of the total area, whereas hilly areas cover 33.4%, primarily in the northern part of the city. Plains areas accounted for 15.9%, primarily in the southern Shangdang Basin and intermontane plains. The annual average temperature in the city is 9 °C, with the hottest month averaging 28 °C, and the coldest month averaging −12.5 °C. The annual precipitation is 618.9 mm, and the annual evaporation is 1578.8 mm. Due to variations in climate and elevation, the region exhibits diverse soil types, comprising six major soil classes, primarily including Alfisols, Ochrepts and Lithosols. The main crops grown in this area are maize, winter wheat, and soybeans. Nitrogen fertilizer application is relatively high, averaging 0.88 kg N km² yr^{−1}. Irrigation methods in the croplands mainly include furrow irrigation, flood irrigation, and ditch irrigation, while drainage conditions and cropland shelterbelts vary with topography.

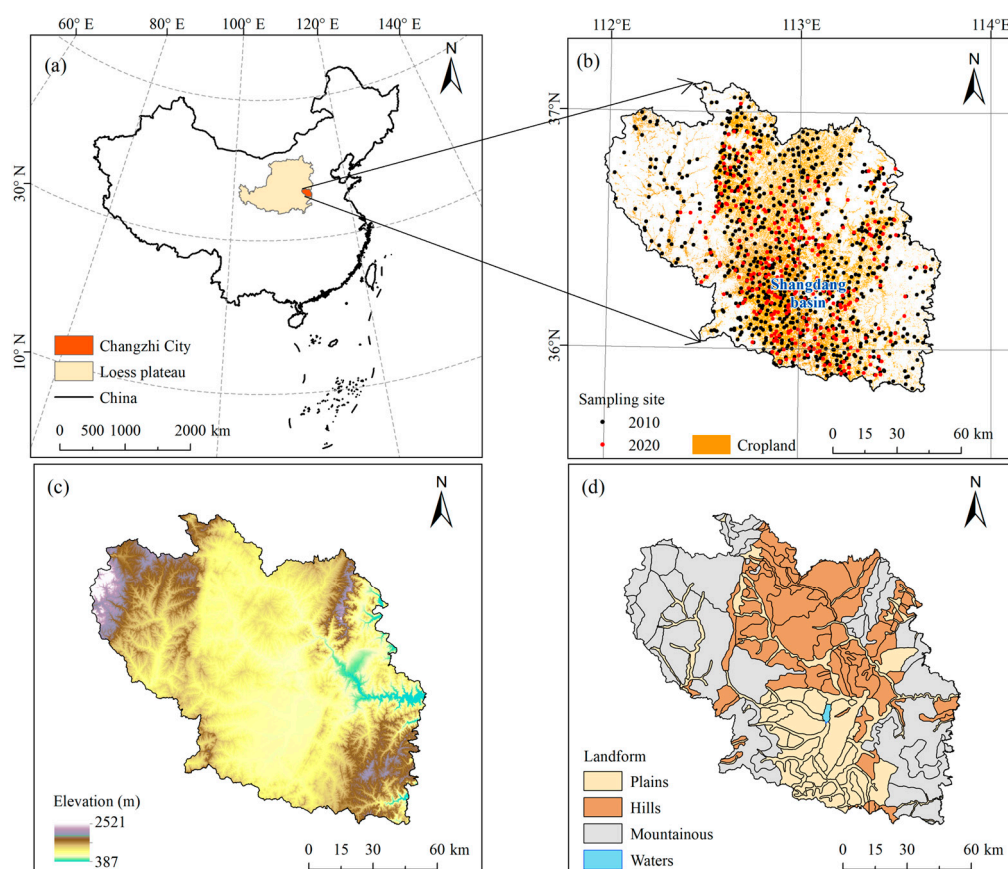


Figure 1. Study area (a), soil sampling sites (b), elevation (c) and landform (d).

2.2. Soil Sampling

In this study, soil samples were collected from croplands in Changzhi in July of 2010 and 2020, totaling 1121 samples. For the 2010 sampling, the sampling scheme was designed such that one sample point was established approximately every 5 km² of cropland, with random sampling conducted at each point, resulting in a total of 711 soil samples. In 2020, the sampling density was increased based on the 2010 sampling points, with one sample point placed approximately every 3 km² of cropland, yielding 410 new sample points. The distribution of the sampling points for both 2010 and 2020 is presented in Figure 1b. At each point, surface soil samples (0–20 cm) were collected, sealed, and returned to the laboratory. The exact coordinates of each soil-sampling point were recorded using a portable GPS system. After natural air-drying, the soil samples were sieved after removing plant roots and stones. SOC was determined using the potassium dichromate oxidation method [20].

2.3. Agricultural Management Survey

Agricultural management information for croplands in Changzhi was surveyed in 2010 and 2020. The amount of nitrogen fertilizer applied to the sampled plots was determined through interviews with landowners. All landowners received basic training to ensure data accuracy. Additionally, agricultural management information for the sampled plots, including irrigation and drainage conditions and cropland shelterbelts, was gathered through field surveys and discussions with local agricultural departments. Actual cropland vector data at a 1:5000 scale were collected. Using the spatial join tool in ArcGIS 10.8, the agricultural management information was assigned to the corresponding cropland polygons. For a small number of cropland polygons for which data could not be surveyed, information from neighboring polygons was used as a substitute. Finally, the agricultural management vector data containing nitrogen fertilizer application, irrigation conditions, drainage conditions, and cropland shelter rates were converted into raster data using ArcGIS 10.8. These raster data were reprojected to the EPSG:3857—WGS 84/World Mercator coordinate system, and the nearest-neighbor resampling method was applied to prepare a 30 m resolution raster dataset. All raster data were then masked to the cropland boundaries of Changzhi.

2.4. Selection of Environmental Covariates

Cropland soil organic carbon exhibits significant heterogeneity at both regional and global scales, with this variability being controlled by soil formation factors. Jenny proposed that environmental conditions govern soil development, asserting that soil is a natural body formed through the combined actions of climate, biology (vegetation), topography, parent material, and other soil properties [5]. Additionally, cropland SOC is influenced by agricultural activities, particularly irrigation and drainage conditions, shelterbelt coverage, and fertilizer application, all of which directly affect the soil environment. Therefore, we selected a total of 19 covariates for the Digital Soil Mapping (DSM) model, including both natural environmental variables and agricultural management variables. Table 1 provides the categories, resolution, and sources of all environmental covariates.

The topographic factors we selected were elevation (Ele), slope (Slope), aspect (Aspect) and topographic humidity index (TWI). These variables were derived using SAGA GIS software (<https://saga-gis.org/>, accessed on 16 August 2024) from the digital elevation model (DEM) obtained from the Shuttle Radar Topography Mission (SRTM) with a resolution of 30 m. The DEM data were accessed via the Geospatial Data Cloud platform (<https://gscloud.cn>, accessed on 11 August 2024).

Table 1. Composition of environmental covariates.

Variable Categories	Factor	Resolution	Year	Source	
Natural variables	Elevation(Ele)	30 m	-	USGS ASTGTM	
	Topography	Slope	30 m	-	Calculated from Elevation
		Aspect	30 m	-	Calculated from Elevation
		Topographic Wetness Index (TWI)	30 m	-	Calculated from Elevation
		Climate	Mean Annual Temperature (MAT)	1 km	2010, 2020
	Mean Annual Precipitation (MAP)		1 km	2010, 2020	
	Mean Annual Evapotranspiration (MAE)		1 km	2010, 2020	
	Vegetation	Net primary productivity (NPP)	250 m	2010, 2020	https://www.resdc.cn/ , accessed on 21 August 2024
		Normalized difference vegetation index (NDVI)	30 m	2010, 2020	USGS ASTGTM
	Soil property	Bulk density (BD)	250 m	-	SoilGrids250m 2.0
		Soil type (STP)	250 m	-	SoilGrids250m 2.0
		Sand	250 m	-	SoilGrids250m 2.0
		Silt	250 m	-	SoilGrids250m 2.0
		Clay	250 m	-	SoilGrids250m 2.0
Geology	Parent material (PM)	1:1,000,000	-	https://soil.geodata.cn , accessed on 16 August 2024	
Agricultural management	Application amount of nitrogen fertilizer (NF)	Vector data	2010, 2020	This study	
	Irrigation condition (IC)	Vector data	2010, 2020	This study	
	Drainage condition (DC)	Vector data	2010, 2020	This study	
	Cropland shelterbelt (CS)	Vector data	2010, 2020	This study	

Geology and soil property data, including parent material (PM), were sourced from the China Soil and Water Digital Database (<https://soil.geodata.cn>, accessed on 16 August 2024) managed by the National Earth System Science Data Center at a scale of 1:1,000,000. This dataset includes a 1:1,000,000 rock-type vector distribution map of China from the 1990s onward. Using ArcGIS 10.8 software, the vector layer was converted into a 30 m resolution raster map. Other soil properties such as soil type (STP) and soil texture (Clay, Silt, and Sand) were obtained from SoilGrids 250m v2.0 (<https://soilgrids.org/>, accessed on 21 August 2024), with 0–20 cm soil texture data derived using depth-weighted averaging functions [21].

The climate covariates, including mean annual temperature (MAT), mean annual precipitation (MAP) and mean annual evapotranspiration (MAE) for 2010 and 2020, were sourced from the National Tibetan Plateau Data Center (<https://data.tpdc.ac.cn/>, accessed on 21 August 2024). For vegetation data, the mean annual normalized difference vegetation index (NDVI) and mean annual net primary productivity (NPP) for 2010 and 2020 were selected. The NDVI data were generated from Landsat imagery available through the United States Geological Survey (USGS) website with a spatial resolution of 30 m. NPP data were obtained from the China Resource Environment and Scientific Data Center

(<http://www.resdc.cn>, accessed on 21 August 2024) with a spatial resolution of 250 m. These datasets are pre-existing. Agricultural management data were obtained as described in Section 2.3.

To address the differences in coordinate reference systems and resolutions, all datasets were resampled to 30 m × 30 m raster data using the nearest-neighbor resampling method in ArcGIS 10.8, and the data were projected onto the EPSG:3857—WGS 84/World Mercator coordinate system.

2.5. Random Forest

Random Forest (RF) is a machine learning technique that utilizes multiple decision trees in an ensemble approach. It is widely adopted in various fields, including Digital Soil Mapping (DSM), owing to its simplicity, interpretability, and robust performance [22,23]. RF is capable of handling both classification and regression tasks and processing high-dimensional datasets. Furthermore, it allows for the assessment of variable importance in predicting the target variable [24]. In this study, the RF model was employed to estimate the spatial distribution of Soil Organic Carbon (SOC) in the surface layer (20 cm) of cropland in Changzhi for the period from 2010 to 2020.

To validate the RF model's predictive capability (or robustness), SOC sample observations from two different years were randomly divided, with 80% used for training and 20% for testing (The SOC sample points for 2010 and 2020 were divided separately). The sample division was performed using the "Subset Features" tool in the "Geospatial Analysis" toolbox of ArcGIS 10.8 to ensure the non-bias and uniformity of the model sampling and point distribution.

2.6. Model Construction and Accuracy Evaluation

In this study, dynamic environmental covariate values (climate, vegetation, and agricultural management factors that vary by year) and static environmental covariate values (topography and soil properties that do not change over a considerable period) were extracted for each soil sample point corresponding to the sampling time. The Random Forest (RF) model was employed to establish the relationship between SOC sample data and environmental covariates. The testing set was then used to evaluate the model's predictive performance for SOC. Finally, the model's simulation accuracy was assessed using statistical indicators, including the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) [25,26]. A higher R^2 and lower MAE and RMSE values were considered indicative of superior model performance. The formulas for calculating MAE, RMSE, and R^2 are as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |o_i - p_i| \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (o_i - p_i)^2}{n}} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (p_i - o_i)^2}{\sum_{i=1}^n (o_i - \bar{o})^2} \quad (3)$$

where n represents the total number of soil samples, p_i denotes the observed value for the i -th sample, o_i is the predicted value of the i -th soil sample, and \bar{o} indicates the average of all the sample values.

Building the RF model requires setting several key parameters, including the number of trees (n_{tree}), the minimum number of terminal node (node size), and the number of covariates that are randomly selected at each tree (m_{try}) [27]. In this study, the n_{tree} , node

size, and mtry parameters were optimized through cross-validation to achieve the best estimation performance. We used 10-fold cross-validation to select the optimal combination of model parameters, with ntree set to 500 and mtry set to 5. Additionally, the RF model evaluates variable importance through two methods: Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG). In this study, MDA was chosen due to its generally higher reliability compared to MDG. For further details on these methods, please refer to [28,29]. The RF model was constructed, trained, and optimized using the 'Random Forest' package in R version 4.2.2 [30].

2.7. SOC Mapping and Its Uncertainty

In this study, the spatiotemporal substitution method (that is, the dynamic variables corresponding to each year were used to simulate the SOC of the corresponding year) was used to map the SOC distribution of the surface 20 cm depth of the cropland in Changzhi for 2010 and 2020. Specifically, the dynamic environmental covariate raster maps for 2010 and 2020 (dynamic environmental variables) were input into the trained RF model. The RF model was run 100 times annually, generating 100 SOC maps per year. The final SOC map was derived by averaging these outputs.

The uncertainty reflects the robustness of the model in terms of spatial predictions and the errors caused by random sampling. The uncertainty in SOC predictions was quantified by calculating the standard deviation across all SOC maps produced by the RF model for each period [31]. Finally, the spatiotemporal distribution maps of SOC and their associated prediction uncertainty maps were created using ArcGIS 10.8.

2.8. Partial Least Squares Structural Equation Modeling

Structural Equation Modeling (SEM) is highly suitable for evaluating the relationships between variables, as it presumes causal links among latent variables and decomposes correlations into direct and indirect effects within the model [32]. In this study, PLS-SEM was employed to assess the direct and indirect impacts of various variable categories on cropland SOC in Changzhi.

The initial hypothesis for the PLS-SEM model is as follows: agricultural management, topography, climate, vegetation, soil properties, and geology all have direct causal relationships with soil organic carbon formation. Agricultural management also exerts an indirect causal effect on SOC through its influence on soil properties or geology. Topography has an indirect causal effect on SOC through its impact on agricultural management. Climate influences SOC indirectly by affecting vegetation, soil properties, and geology. Vegetation also affects SOC indirectly through its impact on soil properties and geology. Based on this hypothesized mechanism, we incorporated the SOC data for both 2010 and 2020, along with the corresponding environmental covariates, into the constructed PLS-SEM model based on the hypothesized mechanisms. The 'piecewiseSEM' package in R was applied to evaluate the direct and indirect effects of variables on SOC. Model fit quality across iterations was assessed using the chi-squared/degrees of freedom ratio ($\chi^2/df \in [0, 3]$) and the Comparative Fit Index ($CFI > 0.9$) [33]. The optimal fit results were obtained, and all analyses were conducted using R version 4.2.2.

The workflow of this study is shown in Figure 2, which primarily consists of three parts: the collection of cropland soil organic carbon and environmental covariates, digital mapping and uncertainty analysis of soil organic carbon, and the analysis of the driving factors of soil organic carbon.

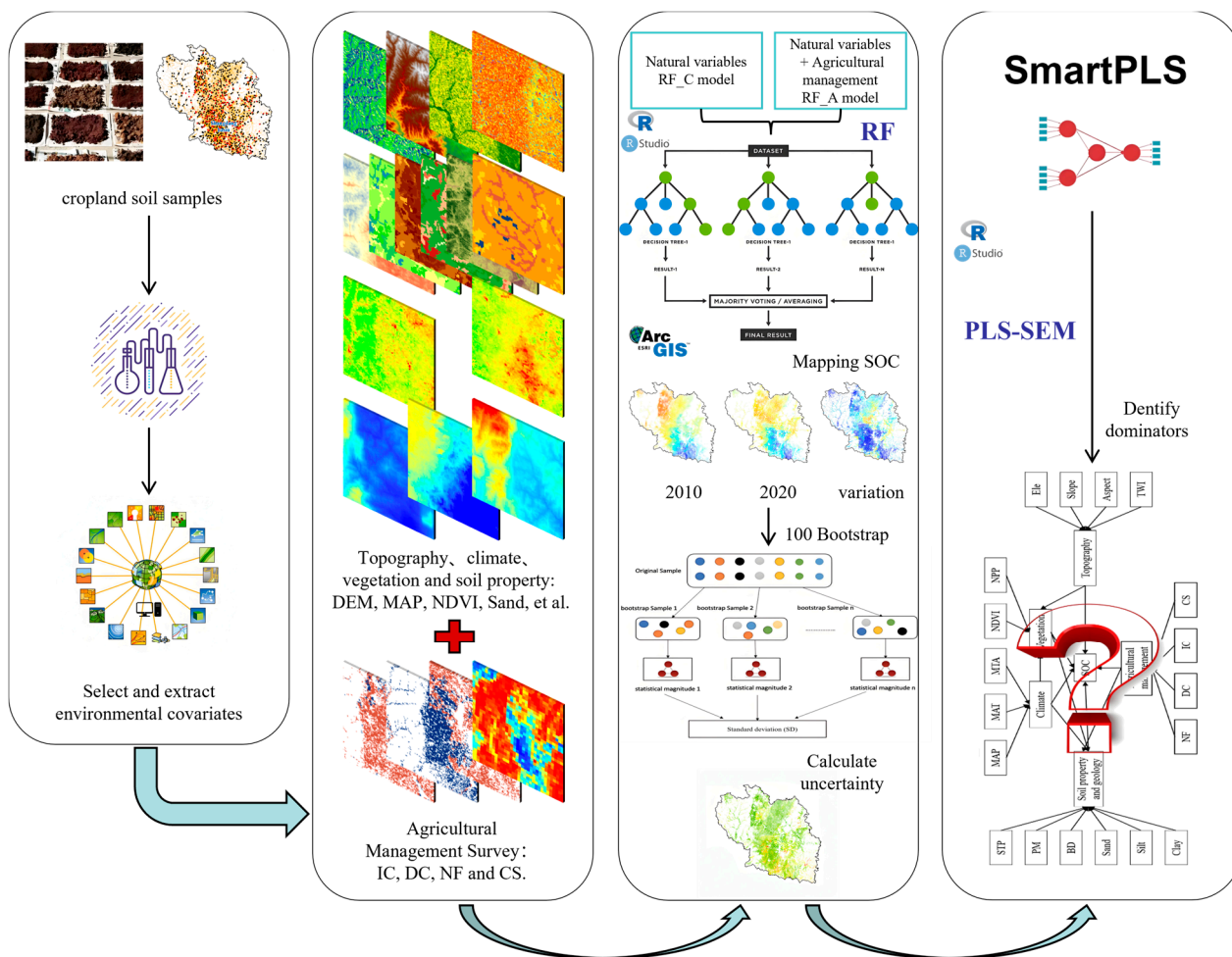


Figure 2. The workflow of this study.

3. Results

3.1. Descriptive Statistics

Table 2 presents the descriptive statistics of the SOC in Changzhi. The SOC range for the 1121 soil samples in Changzhi was from 2.03 g kg⁻¹ to 41.44 g kg⁻¹, with a mean value of 11.25 g kg⁻¹ and a standard deviation of 3.62 g kg⁻¹. The coefficient of variation (CV) was 32.22%. Based on the CV, the degree of variability was classified into three categories that included high variability (CV > 100%), moderate variability (10% < CV < 100%), and low variability (CV < 10%) [14]. From a temporal perspective, from 2010 to 2020 the average SOC value in Changzhi increased by 2.10 g kg⁻¹, whereas the CV increased from 23.10% to 42.69%, indicating a higher degree of SOC variability in 2020.

Table 2. Descriptive statistics of SOC.

	Sampling Point	Min (g kg ⁻¹)	Max (g kg ⁻¹)	Mean (g kg ⁻¹)	SD (g kg ⁻¹)	CV (%)
All	1121	2.03	41.44	11.25	3.62	32.22
2010	829	2.57	37.40	10.70	2.47	23.10
2020	292	2.03	41.44	12.80	5.46	42.69

3.2. Model Performance Evaluation

To evaluate the effectiveness of agricultural management factors for predicting cropland SOC, we constructed an RF model that excluded agricultural management factors. Here, RF_A represents the RF model using all 19 covariates, whereas RF_C represents

the RF model that uses only natural variables (i.e., excluding agricultural management information). Table 3 presents the performance evaluation results for the RF models with two different variable combinations. The results indicate that, compared to the model using natural variables (RF_C.), the inclusion of agricultural management variables (RF_A) improved the prediction accuracy, with an R^2 increase of 0.08 (0.77 vs. 0.85), a reduction in RMSE (1.74 g kg^{-1} vs. 1.53 g kg^{-1}), and a decrease in MAE (1.10 g kg^{-1} vs. 0.94 g kg^{-1}). Specifically, incorporating agricultural management variables resulted in a higher prediction performance for cropland SOC.

Table 3. Predicted values and accuracy evaluation of RF model under two different variable combinations.

Model	MAE (g kg^{-1})	RMSE (g kg^{-1})	R^2
RF_A	0.94	1.53	0.85
RF_C	1.10	1.74	0.77

3.3. Relative Importance of Variables

Previous studies have demonstrated that the importance of environmental variables varies depending on region and scale [34]. In this study, the relative importance of each variable in the RF model was determined, with their contributions adjusted to total 100%. In the RF_A model, mean annual precipitation (15.71%) emerged as the most influential factor for SOC variation, followed by IC (10.32%), MAT (8.20%), NPP (7.18%), clay (6.51%), and NF (6.09%). These variables accounted for 54.02% of the total relative importance, indicating that they were the primary environmental variables influencing variation in cropland SOC in the study area (Figure 3). Climate variables contributed the most (28.03%), followed by agricultural management (24.59%), soil properties and geology (20.98%), vegetation (13.28%), and topographic variables (13.12%) (Figure 3). Compared to the relative importance results of the RF_A model, in the RF_C model, the importance ranking of the other four categorical variables remains unchanged, except for the exclusion of the agricultural management variable. Regarding individual covariate factors, MAP still holds the highest relative importance, with NPP, MAT, NDVI, and Clay remaining as factors with relatively strong importance. Aside from MAE, the relative importance ranking of the individual factors does not change significantly.

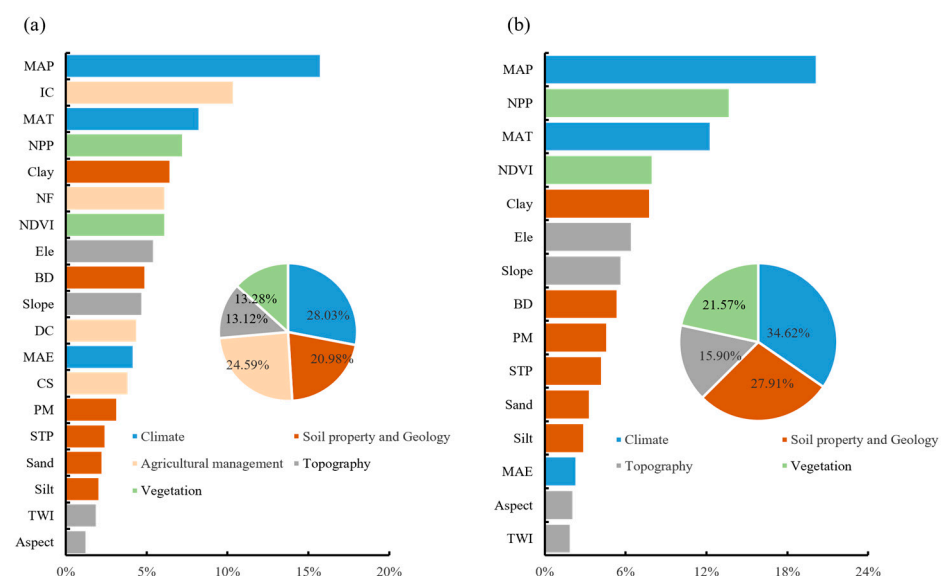


Figure 3. Ranking of the relative importance of environmental covariates in the RF_A model (a) and RF_C model (b). All abbreviations (MAP, IC, MAT, etc.) can be found in Table 1.

3.4. Spatiotemporal Distribution of Soil Organic Carbon

The spatial distributions of cropland SOC in Changzhi from 2010 to 2020, as predicted by both RF_A and RF_C models, showed similar patterns, with pronounced spatial variability in carbon content (Figure 4). The mean values predicted by both models were very close to the observed SOC values. We analyzed the spatiotemporal distribution map, generated using the optimal model (RF_A), was conducted (Figure 4a,b). Overall, surface SOC in the croplands increased from north to south, with a significant difference between the northern and southern regions. The southern Shangdang Basin has higher SOC, whereas the northern mountainous and hilly areas have lower SOC. In 2010, the surface SOC in the croplands was relatively low, with an average of 10.70 g kg^{-1} (Figure 4a). In 2020, the surface SOC in cropland generally increased to varying degrees, with an average of 12.27 g kg^{-1} (Figure 4b). According to estimates, the average organic carbon content in China's cropland surface soil ranges from 11.95 to 12.67 g kg^{-1} [35], and the SOC content in this study is within this range. In the northeastern black soil region of China, SOC reserves were the highest, at nearly double the national average. However, some studies reported SOC loss in the northeastern region due to agricultural practices [36]. Our study demonstrated an increasing trend in cropland SOC in the Loess Plateau (Changzhi) from 2010 to 2020, suggesting that this region, as one of China's most important soil carbon reservoirs, will play a more critical role in the future.

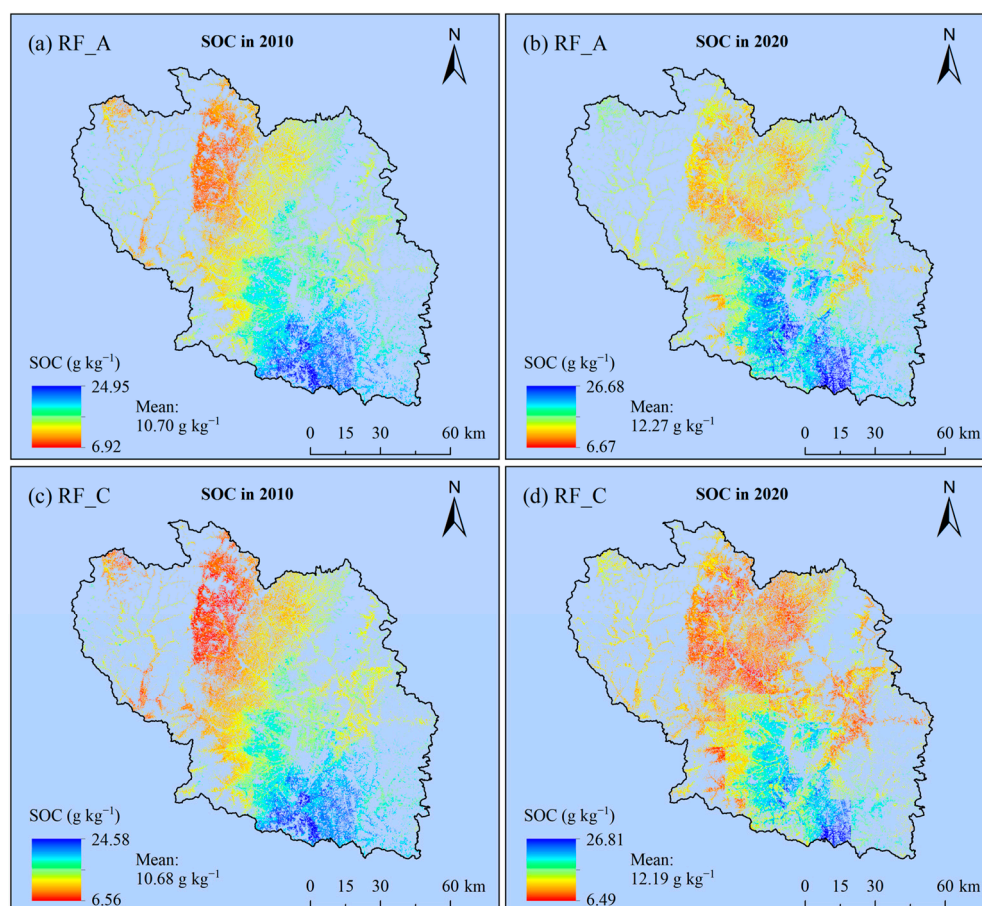


Figure 4. Spatial distribution of SOC in Changzhi for 2010 and 2020 as predicted by the RF_A (a,b) and RF_C (c,d) models. RF_A is the random forest model using all variables, and RF_C is the random forest model using only natural variables.

3.5. Uncertainty Analysis

Figure 5 indicates the uncertainty in SOC mapping for Changzhi in 2010 and 2020 based on the RF_A and RF_C models. Regions with low uncertainty indicated reliable predictions, where the SOC differences after 100 RF model simulations were small. However, areas with high uncertainty suggest relatively large variability in the prediction results after 100 simulations. From a spatiotemporal perspective, the spatial distribution of the uncertainty in Changzhi exhibited similar trends (Figure 5). The average uncertainty of the topsoil (0–20 cm) SOC in Changzhi ranges from 0.39 g kg^{-1} to 0.66 g kg^{-1} , with the uncertainty in most regions being less than 1.00 g kg^{-1} . This suggests that the SOC maps generated in this study are highly reliable. Generally, the density of soil sample points influences the uncertainty in SOC maps. In the RF_C model predictions (Figure 5c,d), the relationship between sample point density and uncertainty is notable. Larger uncertainties are observed in the sparse sampling regions of the east and west, particularly in 2010. In areas with higher sampling point density, except for parts of the northern plains (Shangdang Basin), uncertainty is relatively higher. This is primarily due to the higher SOC in the plains (Shangdang Basin), which may lead to larger errors in multiple prediction results. It is noteworthy that the RF_A model, which incorporates agricultural management information, does not exhibit this pattern. Compared to RF_C, RF_A shows lower uncertainty, both in 2010 and 2020. The main difference between the two models is that RF_A exhibits consistently lower uncertainty across the entire region, with a significant reduction in uncertainty in the southern Shangdang Basin. This can be attributed to improvements in agricultural management practices such as irrigation and drainage conditions, cropland shelterbelts, and fertilization in the low-altitude basin. These measures significantly enhanced the SOC in the region. Therefore, incorporating agricultural management practices into SOC prediction models can effectively reduce mapping uncertainty, particularly in low-altitude areas.

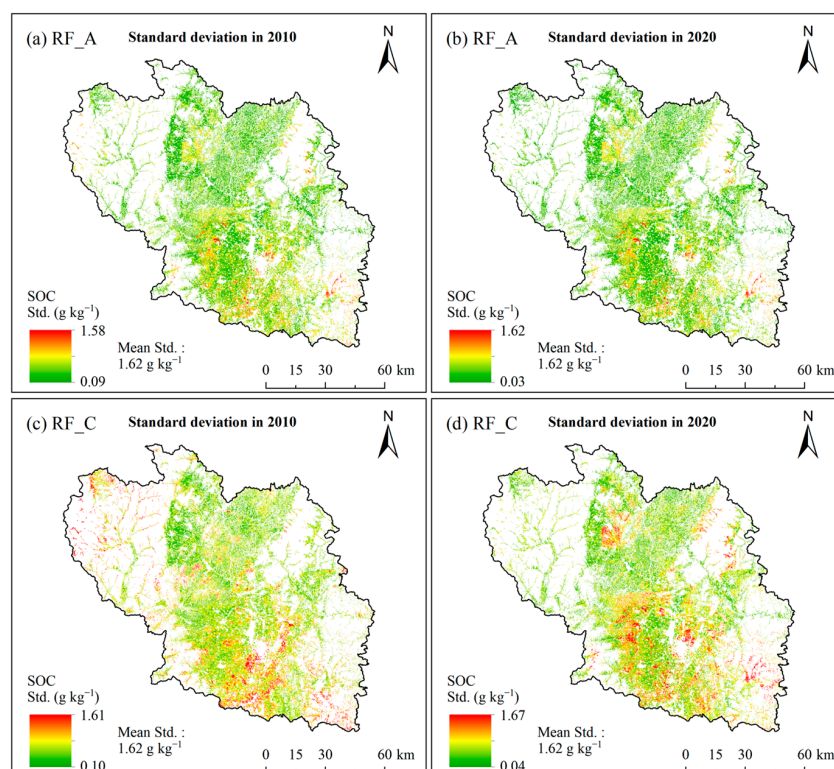


Figure 5. Uncertainty in the SOC mapping for Changzhi in 2010 and 2020 based on the RF_A (a,b) and RF_C (c,d) models. RF_A is the random forest model using all variables, while RF_C is the random forest model using only natural variables.

3.6. PLS-SEM of Soil Organic Carbon and Covariates

In addition to using the RF model to quantify the relative importance of covariates for predicting cropland SOC, we explored the complex causal relationships between the environmental covariates and SOC using a PLS-SEM model. Based on existing soil knowledge, the paths of the PLS-SEM model were adjusted, resulting in the best-fitting structural equation model ($\chi^2/df = 1.147$, CFI = 0.996) that explained 62.20% of the variance in SOC (Figure 6). The influence of different categories of covariates on SOC variation was ranked as follows: climate > agricultural management > soil properties and geology > vegetation > topography. This was consistent with ranking the importance of different categories of covariates in the RF model. Specifically, climatic factors (path coefficient = 0.669, $p < 0.001$), agricultural management (path coefficient = 0.461, $p < 0.01$), soil properties and geology (path coefficient = 0.339, $p < 0.01$), and vegetation factors (path coefficient = 0.214, $p < 0.05$) were significantly positively correlated with SOC, whereas topography (path coefficient = 0.187, $p < 0.05$) was significantly negatively correlated. This result indicates that SOC content is not controlled by a single environmental factor but by the combined effects of climate, soil, geology, vegetation, topography, and agricultural management. This further confirmed the effectiveness of agricultural management for improving the prediction of cropland SOC in Changzhi. It is important to note that the ranking of the relative importance of the factors in the RF model does not precisely match the ranking of the path coefficients in the PLS-SEM model due to differences in the model’s principles or evaluation perspectives. However, both models indicated that MAT, IC, and NPP were the key variables influencing cropland SOC.

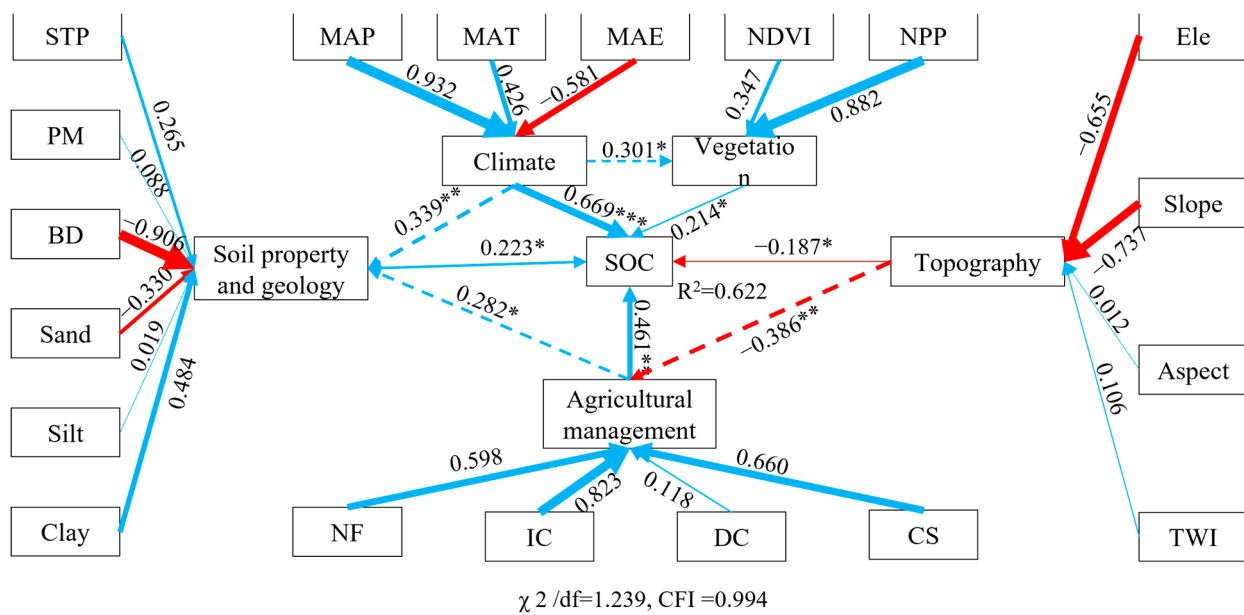


Figure 6. PLS-SEM path analysis results for the effects of climate, vegetation, topography, soil properties, and agricultural management on SOC. The names in the rectangles represent individual variables or categories. Rectangles denote variables or categories, with numbers in parentheses indicating loading scores. Positive and negative path coefficients or loadings are shown by blue and red lines, respectively. Solid lines represent direct effects, while dashed lines indicate indirect effects, with line widths proportional to path coefficients or loadings. Statistical significance is denoted by asterisks: *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

4. Discussion

4.1. Dynamics of Cropland SOC

We subtracted the 2010 SOC from the 2020 SOC (simulated using the optimal RF_A model) to map the dynamic changes in SOC in the Changzhi croplands from 2010 to 2020 (Figure 7). Over the ten years from 2010 to 2020, the SOC of croplands in Changzhi exhibited an overall increasing trend. The southern region exhibited a larger increase, the northern region exhibited a smaller increase, and the central region experienced a slight decrease in SOC. Specifically, 68.38% of the cropland area exhibited an increase in SOC, with the increase primarily concentrated between 1–2 g kg⁻¹ and >2 g kg⁻¹. The areas with the largest increase in surface SOC (>2 g kg⁻¹) were located in the southern Shangdang Basin and northern mountain plains. Although urban construction, greenhouse gas emissions, and industrial development in these areas may cause SOC loss in croplands, the Chinese government has implemented a series of basic high-standard farmland construction projects to ensure food security. These projects prioritize areas such as plains or basins that are relatively flat and involve measures such as reasonable fertilization, improved irrigation and drainage conditions, and the construction of cropland shelterbelts. These management practices have promoted SOC accumulation, compensating for losses caused by urban construction and other factors.

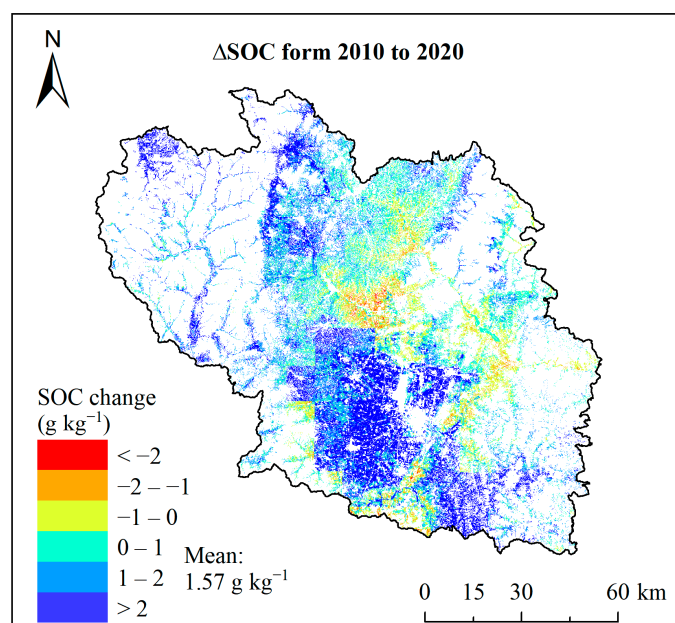


Figure 7. Spatial distribution of cropland SOC changes in Changzhi from 2010 to 2020.

Areas with a decrease in SOC accounted for 32.62% of the cropland area, with the largest decrease in surface SOC (−1 to −2 g kg⁻¹) occurring in the transitional zones between the central basin and hilly mountainous areas. Although the disturbance from urban construction in these regions is relatively small compared to that in the plains and basins, it still leads to SOC loss in croplands. The lack of significant improvements in agricultural management practices in these areas has hindered the effective accumulation of SOC, resulting in a net decline in the cropland SOC content. The SOC change was relatively small in the northern mountainous regions, at within 1 g kg⁻¹. These areas are less accessible and, apart from agricultural activities, are less disturbed by other human activities. Moreover, agricultural management practices in the mountains have remained relatively stable with fewer external environmental changes, thus contributing to the overall balance of cropland SOC in these regions from 2010 to 2020.

Overall, changes in cropland SOC in Changzhi from 2010 to 2020 were influenced by multiple factors. Although there were areas within the region where SOC loss occurred due to urbanization and industrialization, the agricultural sustainability measures implemented by the government compensated for these losses to some extent, particularly in the plains and basin areas. Concurrently, some transitional zones and mountainous areas continue to face the challenges of SOC depletion due to relatively lagging agricultural management. In the future, efforts should be made to strengthen agricultural management in these regions, particularly in key areas where SOC loss is significant, and more effective measures should be implemented to achieve sustainable growth of cropland SOC.

4.2. Effects of Different Environmental Covariates on SOC

The PLS-SEM results demonstrated that climate not only exerted a direct impact on SOC but also indirectly controlled SOC by influencing vegetation (path coefficient = 0.301, $p < 0.01$) and soil properties (path coefficient = 0.339, $p < 0.01$) (Figure 6). The average precipitation in the study area increased by 95.5 mm from 2010 to 2020. Increased MAP can reduce soil aeration, creating anaerobic conditions that decrease soil respiration and slow the rate of soil carbon oxidation, thereby promoting SOC [37]. However, the close relationship between MAP, MAE, and soil moisture enhances plant uptake of available water in areas with high soil moisture, increasing plant productivity and leading to more litter entering the soil, and this in turn increases SOC content [38]. MAT also exerts a strong effect on cropland SOC primarily due to the observation that temperature is an important condition for soil development and controls the decomposition of soil humus. Additionally, climate significantly affects soil microbial activity by regulating the rate at which microbes decompose plant residues entering the soil, thus altering the quantity and quality of soil carbon inputs [39,40]. Although climate change in Changzhi has been relatively stable, climate variables still play a significant role in SOC prediction, further emphasizing the importance of climate variables in SOC forecasting.

Vegetation factors directly affect cropland SOC (Figure 6), as vegetation is a key source of SOC, particularly in agricultural ecosystems where organic carbon is continuously input into the soil through plant residues, root exudates, and other pathways [41]. In SOC mapping studies, vegetation indices (such as NDVI and NPP) are often used as primary predictive indicators. Vegetation growth increases aboveground biomass and provides a carbon source to the soil through root systems, and this alters the soil microenvironment, enhances biotic and abiotic interactions in underground ecological processes, stimulates the carbon cycle, and increases soil carbon storage [42]. The RF model demonstrated that NPP was more important than NDVI (Figure 3), indicating that NPP better reflects the overall productivity of vegetation and reveals the influence of vegetation growth on soil carbon input. NPP directly reflects the net primary production of plants, thus capturing the total amount of fixed carbon during the plant-growing season. In croplands in northern China, straw return is a common agricultural practice in which leftover crop residues are directly returned to the soil after harvest. Straw is an important source of organic carbon in cropland soils, as it contains a large amount of organic carbon that accumulates during plant growth. NPP controls the amount of carbon fixed by vegetation and determines the carbon input from returned straw. Therefore, NPP is not only a vegetation index but also a critical controlling factor for soil carbon input. By contributing to the cropland carbon cycle, NPP possesses greater application value for predicting SOC stocks.

Agricultural management factors contributed the most to changes in cropland SOC (Figures 3 and 6) and indirectly affected SOC by influencing soil properties (path coefficient = 0.282, $p < 0.05$). This study demonstrated that IC is a key factor influencing changes in cropland SOC (Figures 3 and 6). Appropriate irrigation increases soil moisture,

improves crop growth conditions, and promotes microbial activity in soils with adequate moisture, aiding the decomposition and transformation of soil organic matter [43]. Irrigation also regulates soil temperature, particularly during dry or hot periods, by cooling the soil surface through evaporative cooling. Lower soil temperatures help slow organic carbon mineralization, thus reducing carbon dioxide emissions and stabilizing organic carbon. DC and CS are also important for predicting SOC in croplands (Figures 3 and 6), as they can influence SOC by affecting soil properties. The balance between drainage and irrigation conditions is crucial for maintaining a good soil structure, enhancing the stability of soil aggregates, and reducing the risk of erosion, thereby aiding organic carbon retention and preventing soil salinization [44]. Furthermore, CS improves SOC content and stability through various mechanisms, such as by improving microclimates, reducing erosion, enhancing biodiversity, improving soil structure, and increasing microbial activity. The importance of CS ranked second among the agricultural management factors after IC (Figure 3). Nitrogen fertilization replenishes the nitrogen lost during agricultural production, and the soil nitrogen content is one of the key factors in the soil carbon cycle. Nitrogen fertilizers promote crop growth and increase biomass, particularly root growth [45]. Consequently, residual roots and crop residues (such as leaves and stems) enter the soil, increasing the organic carbon input and promoting carbon accumulation. Additionally, the carbon-to-nitrogen (C/N) ratio influenced the decomposition rate and accumulation of organic matter. Nitrogen fertilization alters the C/N ratio, and appropriate nitrogen application can maintain the C/N ratio at levels favorable for organic carbon accumulation [46]. Therefore, the appropriate management of agricultural infrastructure is essential for enhancing soil carbon stocks and achieving carbon sink functions in cropland ecosystems.

Topography is a critical factor in soil formation, as it influences water content, temperature, and distribution of the parent materials. Terrain variables derived from digital elevation models (DEMs) are commonly used as key predictors in soil mapping and digital soil modeling. In this study, topography directly affected cropland SOC and indirectly controlled SOC by influencing agricultural management (path coefficient = 0.386; $p < 0.01$). The significant topographical variation in Changzhi leads to diverse agricultural practices across terrains, affecting carbon decomposition and soil transformation, and ultimately causing SOC variations. Additionally, this study demonstrated that elevation was the most important topographical factor for predicting cropland SOC ('Ele' in Figure 3). Elevation influences the vertical distribution of water and heat and plays a crucial role in microclimate development [47], affecting SOC decomposition and transformation [48]. The slope was almost as important as altitude (Figure 3). The slope typically affects the movement of solutes, water, sediment, and soil moisture, influencing soil development and the spatial distribution of soil properties. In previous soil property prediction studies, elevation has also been identified as the most effective topographical parameter, with slope, aspect, and TWI being key factors influencing the spatial distribution of SOC [4–50].

Increasing evidence highlights the pivotal role of soil conditions in controlling the stability of organic carbon, which directly influences its decomposition rate [51]. Variations in soil properties significantly affect both the quantity and quality of SOC by regulating its stability through diverse physical and chemical mechanisms. These properties are essential for the long-term storage of carbon in cropland soils. In the present study, clay and BD were the most influential soil properties on cropland SOC (Figures 3 and 6). Clay particles protect organic carbon by forming soil aggregates and adsorbing organic matter, thereby preventing rapid microbial decomposition. Higher clay content results in more tightly bound soil particles and a more stable aggregate structure, and this helps to physically protect the SOC [52]. BD reflects soil compaction, and an increase in bulk density typically indicates a decrease in soil porosity, tighter soil structure, and lower soil carbon stocks [53].

When microbial activity in the soil is low, the rate of organic carbon decomposition is reduced, and this slows carbon mineralization and release. Although the importance of soil texture parameters (e.g., sand, silt, ST, PM) for predicting cropland SOC was slightly lower in this study (Figure 3), these soil properties remain crucial for SOC accumulation and stability.

4.3. Limitations and Perspectives

Although the model validation and uncertainty results suggest that the SOC simulations in this study are reasonable and acceptable, certain limitations remain. One potential issue is the differing initial spatial resolutions of the input data. This study used multisource data to predict SOC distribution and dynamics, some of which were sourced from global-scale datasets. For instance, the SoilGrids 250m soil parameters, derived from a large global field survey (150,000 soil profiles), provide a higher spatial resolution compared to other global soil property datasets [54]. Although methods were employed to ensure dataset accuracy, the performance and precision of such high-resolution global datasets may be less reliable in the regional context of China, potentially introducing biases and contributing to uncertainty in the study results. Another limitation was that the predictive model did not include some factors potentially influencing SOC changes. For example, variables such as the amount of straw returned to the soil and potassium (K) fertilizer and phosphorus (P) fertilizer application were difficult to quantify accurately during the survey. It is difficult to combine comprehensive agricultural management information to improve the accuracy of the DSM. Additionally, the effects of dynamic variables such as climate, vegetation, and agricultural management on farm SOC may exhibit time lags. Therefore, establishing long-term monitoring stations to study their impact on SOC mapping is important.

5. Conclusions

In this study, we applied the Random Forest (RF) model to map SOC in Changzhi's croplands at a 30 m resolution for 2010 and 2020. The conclusions are as follows. (1) Incorporating agricultural management information into the RF model improves the spatiotemporal prediction accuracy of SOC for this study area. (2) The spatial distribution trend of SOC in Changzhi's cropland was consistent between 2010 and 2020, with the average SOC increasing from 10.70 g kg⁻¹ in 2010 to 12.27 g kg⁻¹ in 2020. (3) Climate variables, agricultural management, soil properties and geology were the major contributors to SOC modeling (73.60%), with MAP, IC, and NPP identified as the most important covariates affecting SOC changes. Our findings confirm that effective agricultural management can enhance cropland soil carbon stocks, which may contribute to sustainable agricultural development. Future research should focus on incorporating laboratory-measured soil properties and utilizing remote sensing technologies to identify additional agricultural management information for predicting or mapping regional cropland soil organic carbon. Similarly, we recommend further improvements in agricultural management practices in mountainous and hilly areas to enhance both regional cropland soil organic carbon and soil quality.

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

Funding: This study was supported by the Major State Basic Research Development Program (2021YFD1600301) and Henan University Postgraduate Talents Project (No: SYL20060133).

Data Availability Statement: The authors do not have permission to share data.

Acknowledgments: We appreciate the helpful comments from the editor and anonymous reviewers.

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

References

1. Qin, H.; Liu, Y.; Chen, C.; Chen, A.; Liang, Y.; Cornell, C.R.; Guo, X.; Bai, E.; Hou, H.; Wang, D.; et al. Differential contribution of microbial and plant-derived organic matter to soil organic carbon sequestration over two decades of natural revegetation and cropping. *Sci. Total Environ.* **2024**, *949*, 174960. [[CrossRef](#)] [[PubMed](#)]
2. Lal, R. Soil carbon sequestration impacts on global climate change and food security. *Science* **2004**, *304*, 1623–1627. [[CrossRef](#)] [[PubMed](#)]
3. de Alencar, G.V.; Gomes, L.C.; Barros, V.M.D.S.; Ortiz Escobar, M.E.; de Oliveira, T.S.; Mendonça, E.D.S. Organic farming improves soil carbon pools and aggregation of sandy soils in the Brazilian semi-arid region. *Soil Use Manag.* **2024**, *40*, e13097. [[CrossRef](#)]
4. Awoonor, J.K.; Amoako, E.E.; Dogbey, B.F.; Wiredu, I. Quantitative analysis of soil degradation in response to land use change in the Guinea savanna zone of Ghana. *Geoderma Reg.* **2024**, *37*, e00779. [[CrossRef](#)]
5. Jenny, H. *Factors of Soil Formation, A System of Quantitative Pedology*; McGrawHill: New York, NY, USA, 1941; pp. 1–20.
6. Zhu, H.; Xu, Z.; Jing, Y.; Bi, R.; Yang, W. Spatial variation and predictions of soil organic matter and total nitrogen based on vnir reflectance in a Basin of Chinese Loess Plateau. *J. Soil Sci. Plant Nutr.* **2018**, *18*, 1126–1141. [[CrossRef](#)]
7. Stoorvogel, J.J.; Bakkenes, M.; Temme, A.J.; Batjes, N.H.; ten Brink, B.J. S-world: A global soil map for environmental modelling. *Land Degrad. Dev.* **2017**, *28*, 22–33. [[CrossRef](#)]
8. Wang, S.; Xu, L.; Zhuang, Q.; He, N. Investigating the spatio-temporal variability of soil organic carbon stocks in different ecosystems of China. *Sci. Total Environ.* **2021**, *758*, 143644. [[CrossRef](#)]
9. Liu, X.; Zhao, Y.; Shi, X.; Wang, S.; Feng, X.; Yan, F. Spatio-temporal Changes and Associated Uncertainties of CENTURY-modelled SOC for Chinese Upland Soils, 1980–2010. *Chin. Geogr. Sci.* **2021**, *31*, 126–136. [[CrossRef](#)]
10. Azamat, S.; Ilgiz, A.; Ruslan, S.; Mirsayapov, R.; Ilyusya, G.; Tuktarova, I.; Belan, L. Assessing and mapping of soil organic carbon at multiple depths in the semi-arid Trans-Ural steppe zone. *Geoderma Reg.* **2024**, *38*, e00855. [[CrossRef](#)]
11. Szatmári, G.; Pirkó, B.; Koós, S.; Laborczi, A.; Bakacsi, Z.; Szabó, J.; Pásztor, L. Spatio-temporal assessment of topsoil organic carbon stock change in Hungary. *Soil Tillage Res.* **2019**, *195*, 104410. [[CrossRef](#)]
12. Zhou, T.; Geng, Y.; Chen, J.; Liu, M.; Haase, D.; Lausch, A. Mapping soil organic carbon content using multisource remote sensing variables in the Heihe River Basin in China. *Ecol. Indic.* **2020**, *114*, 106288. [[CrossRef](#)]
13. Zhou, T.; Geng, Y.; Chen, J.; Pan, J.; Haase, D.; Lausch, A. High-resolution digital mapping of soil organic carbon and soil total nitrogen using DEM derivatives, Sentinel-1 and Sentinel-2 data based on machine learning algorithms. *Sci. Total Environ.* **2020**, *729*, 138244. [[CrossRef](#)]
14. Tian, H.; Zhang, J.; Zheng, Y.; Shi, J.; Qin, J.; Ren, X.; Bi, R. Prediction of soil organic carbon in mining areas. *Catena* **2022**, *215*, 106311. [[CrossRef](#)]
15. Hu, B.; Xie, M.; Zhou, Y.; Chen, S.; Zhou, Y.; Ni, H.; Peng, J.; Ji, W.; Hong, Y.; Li, H.; et al. A high-resolution map of soil organic carbon in cropland of Southern China. *Catena* **2024**, *237*, 107813. [[CrossRef](#)]
16. Dong, W.; Wu, T.; Luo, J.; Sun, Y.; Xia, L. Land parcel-based digital soil mapping of soil nutrient properties in an alluvial-diluvia plain agricultural area in China. *Geoderma* **2019**, *340*, 234–248. [[CrossRef](#)]
17. Deng, X.; Chen, X.; Ma, W.; Ren, Z.; Zhang, M.; Grieneisen, M.L.; Long, W.; Ni, Z.; Zhan, Y.; Lv, X. Baseline map of organic carbon stock in farmland topsoil in East China. *Agric. Ecosyst. Environ.* **2018**, *254*, 213–223. [[CrossRef](#)]
18. Martin, M.P.; Wattenbach, M.; Smith, P.; Meersmans, J.; Jolivet, C.; Boulonne, L.; Arrouays, D. Spatial distribution of soil organic carbon stocks in France. *Biogeosciences* **2011**, *8*, 1053–1065. [[CrossRef](#)]
19. Fathizad, H.; Ali Hakimzadeh Ardakani, M.; Sodaiezhadeh, H.; Kerry, R.; Taghizadeh-Mehrjardi, R. Investigation of the spatial and temporal variation of soil salinity using random forests in the central desert of Iran. *Geoderma* **2020**, *365*, 114233. [[CrossRef](#)]
20. Dias, L.E.; Jucksch, I.; RICCI, M.S.F.; Alvarez, V.H. Comparação de diferentes métodos de determinação de carbono orgânico em amostras de solos. *Rev. Bras. Ciência Solo* **1991**, *15*, 157–162.
21. Poggio, L.; De Sousa, L.M.; Batjes, N.H.; Heuvelink, G.B.; Kempen, B.; Ribeiro, E.; Rossiter, D. SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *Soil* **2021**, *7*, 217–240. [[CrossRef](#)]
22. Sreenivas, K.; Sujatha, G.; Sudhir, K.; Kiran, D.V.; Fyzee, M.A.; Ravisankar, T.; Dadhwal, V.K. Spatial assessment of soil organic carbon density through random forests based imputation. *J. Indian Soc. Remote Sens.* **2014**, *42*, 577–587. [[CrossRef](#)]
23. Nabiollahi, K.; Shahlaee, S.; Zahedi, S.; Taghizadeh-Mehrjardi, R.; Scholten, T. Land use and soil organic carbon stocks-change detection over time using digital soil assessment, a case study from Kamyaran Region, Iran (1988–2018). *Agronomy* **2021**, *11*, 597. [[CrossRef](#)]

24. Were, K.; Bui, D.T.; Dick, Ø.B.; Singh, B.R. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecol. Ind.* **2015**, *52*, 394–403. [[CrossRef](#)]
25. Ottoy, S.; De Vos, B.; Sindayihebura, A.; Hermy, M.; Van Orshoven, J. Assessing soil organic carbon stocks under current and potential forest cover using digital soil mapping and spatial generalisation. *Ecol. Ind.* **2017**, *77*, 139–150. [[CrossRef](#)]
26. Jin, H.; Peng, J.; Bi, R.; Tian, H.; Zhu, H.; Ding, H. Comparing Laboratory and Satellite Hyperspectral Predictions of Soil Organic Carbon in Farmland. *Agronomy* **2024**, *14*, 175. [[CrossRef](#)]
27. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
28. Bureau, A.; Dupuis, J.; Hayward, B.; Falls, K.; Van Eerdewegh, P. Mapping complex traits using Random Forests. *BMC Genet.* **2003**, *4*, 1–5.
29. Zhang, L.; Yang, L.; Cai, Y.; Huang, H.; Shi, J.; Zhou, C. A multiple soil properties oriented representative sampling strategy for digital soil mapping. *Geoderma* **2022**, *406*, 115531. [[CrossRef](#)]
30. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2016.
31. He, X.; Yang, L.; Li, A.; Zhang, L.; Shen, F.; Cai, Y.; Zhou, C. Soil organic carbon prediction using phenological parameters and remote sensing variables generated from Sentinel-2 images. *Catena* **2021**, *205*, 105442. [[CrossRef](#)]
32. Zhao, W.; Zhang, R.; Cao, H.; Tan, W. Factor contribution to soil organic and inorganic carbon accumulation in the Loess Plateau, Structural equation modeling. *Geoderma* **2019**, *352*, 116–125. [[CrossRef](#)]
33. Lefcheck, J.S. piecewiseSEM: Piecewise Structural Equation Modeling in R for Ecology, Evolution, and Systematics. *Methods Ecol. Evol.* **2016**, *7*, 573–579. [[CrossRef](#)]
34. Wang, S.; Adhikari, K.; Wang, Q.; Jin, X.; Li, H. Role of environmental variables in the spatial distribution of soil carbon (C), nitrogen (N), and C, N ratio from the northeastern coastal agroecosystems in China. *Ecol. Indic.* **2018**, *84*, 263–272. [[CrossRef](#)]
35. Yan, X.; Cai, Z.; Wang, S.; Smith, P. Direct measurement of soil organic carbon content change in the croplands of China. *Global Change Biol.* **2011**, *17*, 487–496. [[CrossRef](#)]
36. Meng, X.; Bao, Y.; Luo, C.; Zhang, X.; Liu, H. SOC content of global Mollisols at a 30m spatial resolution from 1984 to 2021 generated by the novel ML-CNN prediction model. *Remote Sens. Environ.* **2024**, *300*, 113911. [[CrossRef](#)]
37. Chen, F.; Feng, P.; Harrison, M.T.; Wang, B.; Liu, K.; Zhang, C.; Hu, K. Cropland carbon stocks driven by soil characteristics, rainfall and elevation. *Sci. Total Environ.* **2023**, *862*, 160602. [[CrossRef](#)] [[PubMed](#)]
38. Homann, P.S.; Kapchinske, J.S.; Boyce, A. Relations of mineral-soil C and N to climate and texture, regional differences within the conterminous USA. *Biogeochemistry* **2007**, *85*, 303–316. [[CrossRef](#)]
39. Zhang, L.; Zeng, G.; Tong, C. A review on the effects of biogenic elements and biological factors on wetland soil carbon mineralization. *Acta Ecol. Sin.* **2011**, *31*, 5387–5395.
40. Craine, J.M.; Fierer, N.; McLaughlan, K.M. Widespread coupling between the rate and temperature sensitivity of organic matter decay. *Nat. Geosci.* **2010**, *3*, 854–857. [[CrossRef](#)]
41. Wolka, K.; Biazin, B.; Martinsen, V.; Mulder, J. Soil organic carbon and associated soil properties in Enset (*Ensete ventricosum* Welw. Cheesman)-based homegardens in Ethiopia. *Soil Tillage Res.* **2021**, *205*, 104791. [[CrossRef](#)]
42. Yang, L.; He, X.; Shen, F.; Zhou, C.; Zhu, A.X.; Gao, B.; Chen, Z.; Li, M. Improving prediction of soil organic carbon content in croplands using phenological parameters extracted from NDVI time series data. *Soil Tillage Res.* **2020**, *196*, 104465. [[CrossRef](#)]
43. Núñez, A.; Schipanski, M. Changes in soil organic matter after conversion from irrigated to dryland cropping systems. *Agric. Ecosyst. Environ.* **2023**, *347*, 108392. [[CrossRef](#)]
44. Emde, D.; Hannam, K.D.; Most, I.; Nelson, L.M.; Jones, M.D. Soil organic carbon in irrigated agricultural systems, A meta-analysis. *Global Change Biol.* **2021**, *27*, 3898–3910. [[CrossRef](#)]
45. Hu, Q.; Liu, T.; Ding, H.; Li, C.; Tan, W.; Yu, M.; Liu, J.; Cao, C. Effects of nitrogen fertilizer on soil microbial residues and their contribution to soil organic carbon and total nitrogen in a rice-wheat system. *Appl. Soil Ecol.* **2023**, *181*, 104648. [[CrossRef](#)]
46. Wu, X.; Zhao, L.; Hu, G.; Liu, G.; Li, W.; Ding, Y. Permafrost and land cover as controlling factors for light fraction organic matter on the southern Qinghai-Tibetan Plateau. *Sci. Total Environ.* **2018**, *613*, 1165–1174. [[CrossRef](#)] [[PubMed](#)]
47. Liu, X.; Li, S.; Wang, S.; Bian, Z.; Zhou, W.; Wang, C. Effects of farmland landscape pattern on spatial distribution of soil organic carbon in Lower Liaohe Plain of northeastern China. *Ecol. Indic.* **2022**, *145*, 109652. [[CrossRef](#)]
48. Martin, M.P.; Orton, T.G.; Lacarce, E.; Meersmans, J.; Saby, N.P.A.; Paroissien, J.B.; Jolivet, C.; Boulonne, L.; Arrouays, D. Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale. *Geoderma* **2014**, *223–225*, 97–107. [[CrossRef](#)]
49. Huang, J.; Wu, C.; Minasny, B.; Roudier, P.; McBratney, A.B. Unravelling scale-and location-specific variations in soil properties using the 2-dimensional empirical mode decomposition. *Geoderma* **2017**, *307*, 139–149. [[CrossRef](#)]
50. Ramifehiarivo, N.; Brossard, M.; Grinand, C.; Andriamananjara, A.; Razafimbelo, T.; Rasolohery, A.; Razafimahatratra, H.; Seyler, F.; Ranaivoson, N.; Rabenarivo, M.; et al. Mapping soil organic carbon on a national scale, towards an improved and updated map of Madagascar. *Geoderma Reg.* **2017**, *9*, 29–38. [[CrossRef](#)]

51. Hemingway, J.D.; Rothman, D.H.; Grant, K.E.; Rosengard, S.Z.; Elinton, T.I.; Derry, L.A.; Galy, V.V. Mineral protection regulates long-term global preservation of natural organic carbon. *Nature* **2019**, *570*, 228–231. [[CrossRef](#)]
52. Schaeffer, S.M.; Ziegler, S.E.; Belnap, J.; Evans, R.D. Effects of *Bromus tectorum* invasion on microbial carbon and nitrogen cycling in two adjacent undisturbed arid grassland communities. *Biogeochemistry* **2012**, *111*, 427–441. [[CrossRef](#)]
53. Topa, D.; Cara, I.G.; Jităreanu, G. Long term impact of different tillage systems on carbon pools and stocks, soil bulk density, aggregation and nutrients, A field meta-analysis. *Catena* **2021**, *199*, 105102. [[CrossRef](#)]
54. Hengl, T.; Mendes de Jesus, J.; Heuvelink, G.B.M.; Ruiperez Gonzalez, M.; Kilibarda, M.; Blagotić, A.; Shangguan, W.; Wright, M.N.; Geng, X.; Bauer-Marschallinger, B.; et al. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE* **2017**, *12*, e016974. [[CrossRef](#)] [[PubMed](#)]

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.