remote sensing

MDPI

Article

Soil Texture and Its Relationship with Environmental Factors on the Qinghai–Tibet Plateau

Yadong Liu [1,2], Xiaodong Wu [1,2,*], Tonghua Wu [1,3], Lin Zhao [1,4], Ren Li [1,2], Wangping Li [5], Guojie Hu [1], Defu Zou [1], Jie Ni [6], Yizhen Du [6], Mengjuan Wang [7], Zhihong Li [5], Xianhua Wei [1,2] and Xuchun Yan [1,2]

1 Cryosphere Research Station on the Qinghai-Tibet Plateau, State Key Laboratory of Cryospheric Science, Northwest Institute of Eco-Environment and Resource, Chinese Academy of Sciences, Lanzhou 730000, China
2 College of Resources and Environment, University of Chinese Academy Sciences, Beijing 100049, China
3 Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou 511458, China
4 School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210000, China
5 School of Civil Engineering, Lanzhou University of Technology, Lanzhou 730050, China
6 College of Tourism, Resources and Environment, Zaozhuang University, Zaozhuang 277160, China
7 School of Civil Engineering, Jiangxi University of Engineering, Xinyu 338000, China
* Correspondence: wuxd@lzb.ac.cn

Abstract: Soil texture data are the basic input parameters for many Earth System Models. As the largest middle–low altitude permafrost regions on the planet, the land surface processes on the Qinghai–Tibet Plateau can affect regional and even global water and energy cycles. However, the spatial distribution of soil texture data on the plateau is largely unavailable due to the difficulty of obtaining field data. Based on collection data from field surveys and environmental factors, we predicted the spatial distribution of clay, silt, and sand contents at a 1 km resolution, from 0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm soil depth layers. The random forest models were constructed to predict the soil texture according to the relationships between environmental factors and soil texture data. The results showed that the soil particles of the QTP are dominated by sand, which accounts for more than 70% of the total particles. As for the spatial distribution, silt and clay contents are high in the southeast plateau, and low values of silt and clay mainly appeared in the northwest plateau. Climate and NDVI values are the most important factors that affect the spatial distribution of soil texture on the QTP. The results of this study provide the soil texture data at different depths for the whole plateau at a spatial resolution of 1 km, and the dataset can be used as an input parameter for many Earth System Models.

Keywords: digital soil mapping; machine learning; soil particle size; random forest; soil pits

Citation: Liu, Y.; Wu, X.; Wu, T.; Zhao, L.; Li, R.; Li, W.; Hu, G.; Zou, D.; Ni, J.; Du, Y.; et al. Soil Texture and Its Relationship with Environmental Factors on the Qinghai–Tibet Plateau. *Remote Sens.* **2022**, *14*, 3797. https://doi.org/10.3390/rs14153797

Academic Editor: Ulrich Kamp

Received: 30 May 2022
Accepted: 4 August 2022
Published: 6 August 2022

1. Introduction

Soil texture is one of the most basic physical properties of soil [1,2]. The composition and content of coarse and fine particles can affect the hydraulic characteristics of water retention and conductivity in soil, as well as thermal parameters such as the heat transfer coefficient and heat capacity [3–5]. Soil texture can also affect evaporation and infiltration rates [6,7]. Soil texture has an important impact on the soil's moisture. Studies have shown that the spatial heterogeneity of a soil's texture is responsible for the spatial variation of moisture, which further increases the differences in the spatial distribution of moisture over time [8,9]. As a relatively stable natural attribute, soil texture also has a close relationship with soil carbon stocks [10–13]. Therefore, soil texture is an important input parameter for land surface process models, hydrological models, and atmospheric models [14,15]. In the warming climate, there is an increasing need for soil texture data in studies that deal with global and regional issues such as climate change, soil degradation, water resource shortage, environmental pollution, agriculture, and sustainable development.

The Qinghai-Tibet Plateau (QTP) has an average elevation of more than 4000 m and has been called the roof of the world [16]. The total area of the plateau is about 2.5 million km^2 [17]. As a typically cold region, the QTP hosts the largest permafrost region in the low and middle latitudes, with a permafrost area of 1.06×10^6 km^2 [18,19]. The QTP is sensitive to global climate change, and increases in air temperature since the 1980s is at a rate of twice the global average [20,21]. As an important influencing factor and process parameter of the global water and energy cycle, soil moisture affects the distribution of surface energy, directly affecting the formation and development of soil [22]. In the past few decades, the permafrost on the QTP has been experiencing a rapid degradation, such as an increasing active layer, increasing ground temperature, and decreasing permafrost area [23–25]. Permafrost degradation not only presents threats to engineering construction [23,26], but also stimulates the decomposition of the soil organic matter and further affects ecological succession and the carbon cycle [21]. To quantify the environmental changes in the QTP and their feedback to climate, many land surface models have been used to simulate the changing processes, and have attempted to predict the future changes in the water, carbon, and energy cycles on the QTP. Most of these models require multiple input parameters, and soil texture data are one of the most important of these parameters.

Due to the importance of knowledge concerning soil texture data, many efforts have been conducted to create the spatial distribution of soil texture on the QTP. Previous studies have used spatial interpolation methods to upscale data from a site scale to a regional scale. The Food and Agriculture Organization of the United Nations (FAO) developed a global soil texture dataset with a data ratio of 1:100,000 [27], which has been widely used in many land surface models. This dataset contains very rare soil sample data in China, but there are almost no data in the QTP. According to the field survey data from the Second National Soil Survey of China, the conventional linkage method was used to obtain the spatial distribution of sand, silt, and clay in China, with a spatial resolution of 1 km [28]. The dataset compiled many observation data, and the accuracy was improved. However, the sampling sites on the QTP during the Second National Soil Survey were mainly distributed in seasonally-frozen ground areas, and the data in rural permafrost regions were largely lacking. Since the permafrost regions occupy nearly half of the QTP, there is much room for improvement in obtaining the spatial distribution of soil textures on the QTP.

Recently, several machine learning methods have been widely used for soil texture mapping. For example, the spatial distribution of soil texture fractions in Denmark, predicted at six depth layers at a resolution of 30 m, was performed using the Cubist decision tree algorithm [1]. A 100 m resolution model of the soil particle size distribution in Chile has been created using the classification and regression tree algorithm [29]. A soil texture map of different depths was also created for China, with a spatial resolution of 90 m, and used the random forest model [6]. However, there are still few field sites in the permafrost regions on the QTP.

In order to obtain the spatial distribution of soil texture data on the QTP, we collected a dataset across vast areas of both permafrost and nonpermafrost regions on the QTP. By combining high-resolution, remotely-sensed data and interpolated meteorological datasets, we simulated the spatial distribution of soil texture on the QTP (At depths of 0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm) using the random forest method. We also examined the relationship between the soil texture and environmental factors. The results can serve as a basic input parameter for Earth System Models to study the processes of land surfaces on the QTP.

2. Materials and Methods

2.1. Study Area

The QTP (Figure 1) spans more than 3000 km from east to west and more than 1500 km from north to south [17]. The plateau is located in the semi-arid and arid alpine climate zone, with a mean annual precipitation of about 250 mm [30]. The mean annual average

temperature on the QTP is below 2.8 °C [31]. The air temperature on the QTP shows great heterogeneities, with the coldest values in the north of the plateau and the warmest values in the southeast of the plateau [30].

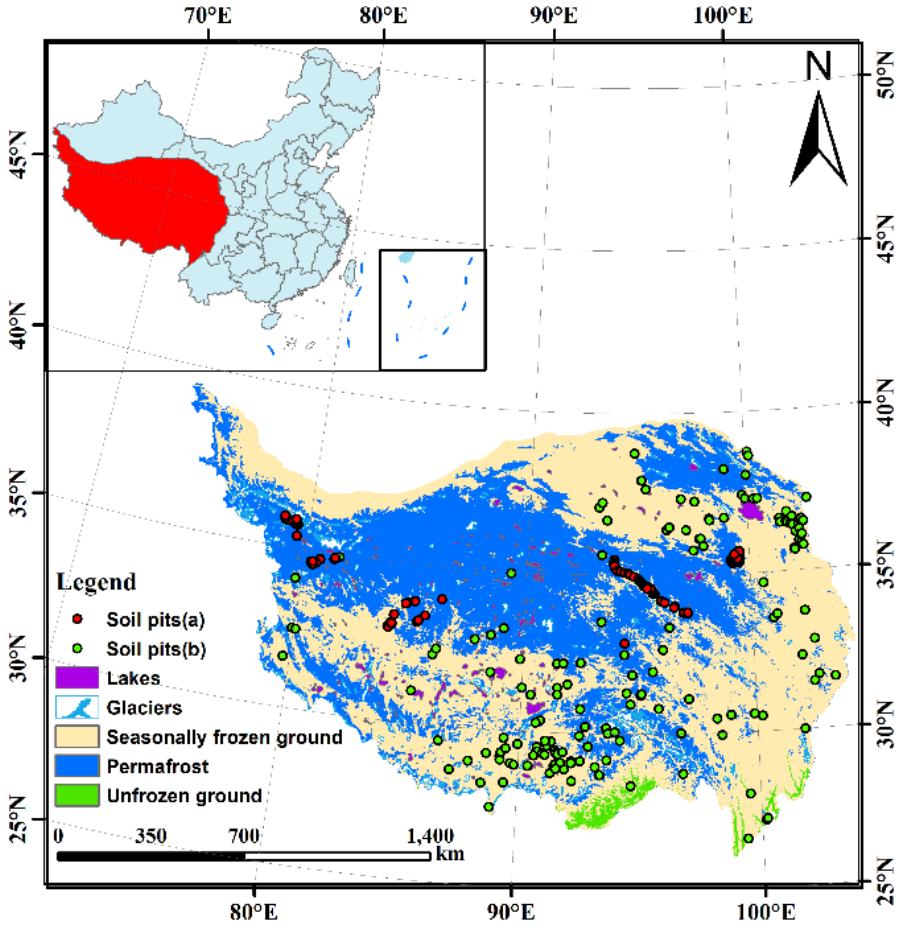

Figure 1. Study area and the distribution of sample sites (the frozen ground map was derived from [18]).

2.2. The Observation Data

There were two sources for the data used in this study: (1) The field observation data of typical permafrost regions on the QTP from 2009 to 2012. The sampling sites were selected based on environmental gradients such as altitude, slope, aspect, and geophysical prospecting. During the field work, soil pits (a) were excavated and soil samples were collected according to the soil genetic horizons. (2) The data collected from the National Soil Series Survey and Compilation of Soil Series of China, conducted from 2009 to 2019 (Soil pits (b)). Most of these samples were located in the seasonally-frozen soil area of the QTP (Figure 1). The two sources of soil sample data mentioned above included valuable samples collected both in the permafrost area of the QTP and in seasonally frozen areas, which provide supporting data for the study of the spatial distribution of soil texture on the QTP. All of the soil profiles were recorded by the genetic horizons, and soil samples were collected at each horizon. The soil variables for different depths were resampled using the mathematical mean method. The soil properties were measured, including the soil texture, soil organic carbon, bulk density, pH, cation exchange capacity, total nitrogen, total phosphorus, total potassium, available phosphorus, and available potassium. Among them, we determined the soil texture fractions, namely the percentages of sand (2–0.05 mm), silt (0.05–0.002 mm), and clay (<0.002 mm), and these percentages were determined using a combination of the wet sieving technique and a particle size analyzer (Sedimat 4–12, UGT, Müncheberg, Germany) [32].

2.3. Data Processing and Analysis

As the measured particle composition is based on the occurrence layer, the number of layers divided by different profiles and the depth of each layer are both different, thus the depth of the acquired profile layers needs to be standardized (0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm). The method is as follows:

$$PC_{t-b} = \sum_{i=1}^{n} PC_i \times H_i \quad (1)$$

PC_{t-b} is the content of a certain fraction of the normalized level t–b, where t and b represent the top and bottom depths of the level, respectively; n represents the profile and the number of occurrence layers. PC_i represents the content of a specific fraction of the occurrence layer i, and H_i represents the proportion of the depth of layer i in the normalized layer t–b.

The soil particle size data were resampled at different depth intervals, i.e., at 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, and 100–200 cm. Among the data, there were 316 samples for 0–5 cm and 5–15 cm, 314 samples for 15–30 cm, 302 samples for 30–60 cm, 266 samples for 60–100 cm, and 184 samples for 100–200 cm.

2.4. Environmental Covariates

The topography of the QTP is complex, meaning that the effects of environmental factors on soil texture are complicated. The soil is a product of climate, biology, topography, parent material, and time [33,34]. This theory provides a theoretical basis to select environmental factors to predict the soil texture, i.e., the climate, vegetation, topography, and parent materials. Although it largely remains unknown which climate variables are most important in affecting soil textures, the annual mean temperature and annual mean precipitation data are closely associated with soil textures. Therefore, we selected these variables as climate variable input parameters (Table 1).

Table 1. Description of the selected variables.

Variables	Description	Resolution	Soil Forming Factors
Elevation	Elevation above sea level (km)	1 km	r
Slope	Slope gradient	1 km	r
Aspect	Aspect gradient	1 km	r
Plan	Plan curvature	1 km	r
Profile	Profile curvature	1 km	r
TWI	Topographic wetness index	1 km	r
TCA	Total catchment area	1 km	r
RSP	Relative slope position	1 km	r
CI	Convergence index	1 km	r
VD	Valley depth	1 km	r
CND	Channel network distance	1 km	r
CNB	Channel network base level	1 km	r
LS	Slope length and steepness factor	1 km	r
MAT	Annual mean temperature (°C)	1 km	c
MAP	Annual precipitation (mm)	1 km	c
NDVI	Mean NDVI during the growing season	1 km	o, c
QGT	Quaternary geological type	1 km	p
ST	Soil type	1 km	o

The DEM data were downloaded from the International Scientific Data Service Platform (http://datamirror.csdb.cn, accessed on 5 August 2021). Using the SAGA GIS software (http://www.saga-gis.org, accessed on 29 May 2022), we calculated the derivative variables of the DEM data (Table 1). Normalized difference vegetation index (NDVI) data were obtained from the United States Geological Survey (http://modis.gsfc.nasa.gov/, accessed on 10 August 2021). To avoid the possible errors that could be caused by rapid climate

change since 2000, we used the multi-year (2000–2015) average data as the input parameters. The mean annual air temperature (MAT) and mean annual precipitation (MAP) data were downloaded from WorldClim version 2.1 (https://www.worldclim.org, accessed on 15 August 2021). These datasets were observed from global meteorological stations in the years 1970–2000. The parent material data were retrieved from the Quaternary geological map, which were download from the National Glacier, Permafrost and Desert Science Data Center (http://www.ncdc.ac.cn, accessed on 20 August 2021). Soil type data were retrieved from the Resource and Environment Science and Data Center (https://www.resdc.cn/, accessed on 22 August 2021).

2.5. Random Forest Model

During the past decades, various machine learning methods have been applied to digital soil mapping research. The classic machine learning method used is the neural network, despite being a computationally expensive method. The classification tree algorithm was established in the 1980s, and this method greatly reduced the amount of calculation needed. The random forest (RF) algorithm is a widely used machine learning algorithm based on the decision tree model [35–37]. The RF algorithm mainly consists of classification and regression models [38,39]. The RF algorithm has a high prediction accuracy, good tolerance to outliers and noise, and is not prone to overfitting [40,41]. Therefore, the algorithm is a natural nonlinear modeling tool, and it is currently one of the most useful tools in data mining and geographic information systems [42,43]. The random forest (RF) algorithm has been implemented into the R software (https://www.r-project.org, accessed on 22 February 2019). On the R platform, the random forest regression model uses the so-called 'randomForest' and 'caret' packages to study spatial distribution. The random forest regression model uses the %IncMSE index to describe the importance of the variables [44]. The greater the value of the %IncMSE, the greater the importance of the environmental variables in the model.

2.6. Accuracy Validation

In order to verify the random forest model, a ten-fold cross-validation was performed to evaluate its accuracy. The coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) were used to assess the prediction's accuracy. Higher R^2 and correspondingly lower MAE and RMSE values indicate a higher prediction accuracy. The calculation methods are as follows [6,45]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (P_i - O_i)^2}{\sum_{i=1}^{n} (P_i - \overline{O})^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |(O_i - P_i)| \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2} \quad (4)$$

The P_i and O_i are the predicted and observed values of a soil texture fraction at the sample site i; $\overline{O}$ is the average observation value.

3. Results

3.1. Statistical Summary of Soil Texture Samples

For all of the soil particle data, the mean values of sand were significantly higher than those of silt and clay, and the values of silt and clay decreased while sand increased with depth (Table 2). The standard error (SD) values of clay were lower than those of sand and silt, ranging from 7.11 to 8.13. The SD values of silt and sand were 11.96–12.72 and 16.80–18.55, respectively. The coefficient of variation (CV) for sand ranged 0.22–0.25, which was a lower range than those of silt and clay. The skewness and kurtosis of silt, clay, and

sand were all near zero, indicating that the soil particle distribution largely satisfies the normal distribution.

Table 2. Statistical analysis of the clay, silt and sand percentages at different depths.

Depth (cm)	Mean (%)	SD (%)	CV	Skewness	Kurtosis
Clay					
0–5	11.56	7.25	0.63	0.66	0.01
5–15	11.55	7.36	0.64	0.76	0.26
15–30	11.33	7.69	0.68	0.93	0.84
30–60	10.81	7.99	0.74	1.03	0.81
60–100	9.90	8.13	0.82	1.25	1.19
100–200	8.00	7.11	0.89	1.70	2.59
Silt					
0–5	16.94	12.72	0.75	0.87	−0.08
5–15	16.71	12.16	0.73	0.71	−0.73
15–30	16.51	12.58	0.76	0.85	−0.33
30–60	15.56	12.19	0.78	0.89	−0.28
60–100	14.37	12.05	0.84	1.19	0.55
100–200	12.41	11.96	0.96	1.88	3.18
Sand					
0–5	71.48	17.14	0.24	−0.60	−0.41
5–15	71.73	16.80	0.23	−0.49	−0.83
15–30	72.14	18.02	0.25	−0.63	−0.63
30–60	73.61	18.23	0.25	−0.77	−0.43
60–100	75.64	18.55	0.25	−0.99	0.29
100–200	79.58	17.15	0.22	−1.66	0.36

3.2. Model Accuracy Evaluation

The ten-fold cross-validation showed that the R^2 values of sand and silt were significantly higher than that of clay (Table 3). The R^2 range of values for silt and sand were 0.44–0.57 and 0.45–0.62, respectively. This indicates that the random forest model can explain 44–57% of the total variances in silt and 45–62% of the total variance in sand. The simulation of sand and silt had higher accuracies than those of clay, which had R^2 values in the range of 0.28–0.50.

Table 3. Accuracy assessment of the soil particle size distribution predictions.

Depth (cm)	R^2	RMSE	MAE
Clay			
0–5	0.36	5.85	4.30
5–15	0.33	6.10	4.45
15–30	0.30	6.49	4.83
30–60	0.28	6.91	5.06
60–100	0.32	6.84	4.86
100–200	0.50	5.02	3.49
Silt			
0–5	0.49	9.11	6.46
5–15	0.50	8.64	6.49
15–30	0.47	9.29	6.87
30–60	0.44	9.30	7.02
60–100	0.49	8.77	6.19
100–200	0.57	8.05	5.43
Sand			
0–5	0.52	12.01	8.68
5–15	0.51	11.95	8.91
15–30	0.45	13.54	10.15
30–60	0.48	13.29	10.12
60–100	0.49	13.28	9.47
100–200	0.62	10.67	7.21

3.3. Spatial Patterns of Predictions

The contents of silt and clay in soil are largely high in the southeast and low in the northwest, while sand showed the opposite spatial patterns, with high values in the northwest and low values in the southeast (Figures 2–4). Similar with the statistical results from the field observation data, the sand content increases with depth (Table 2).

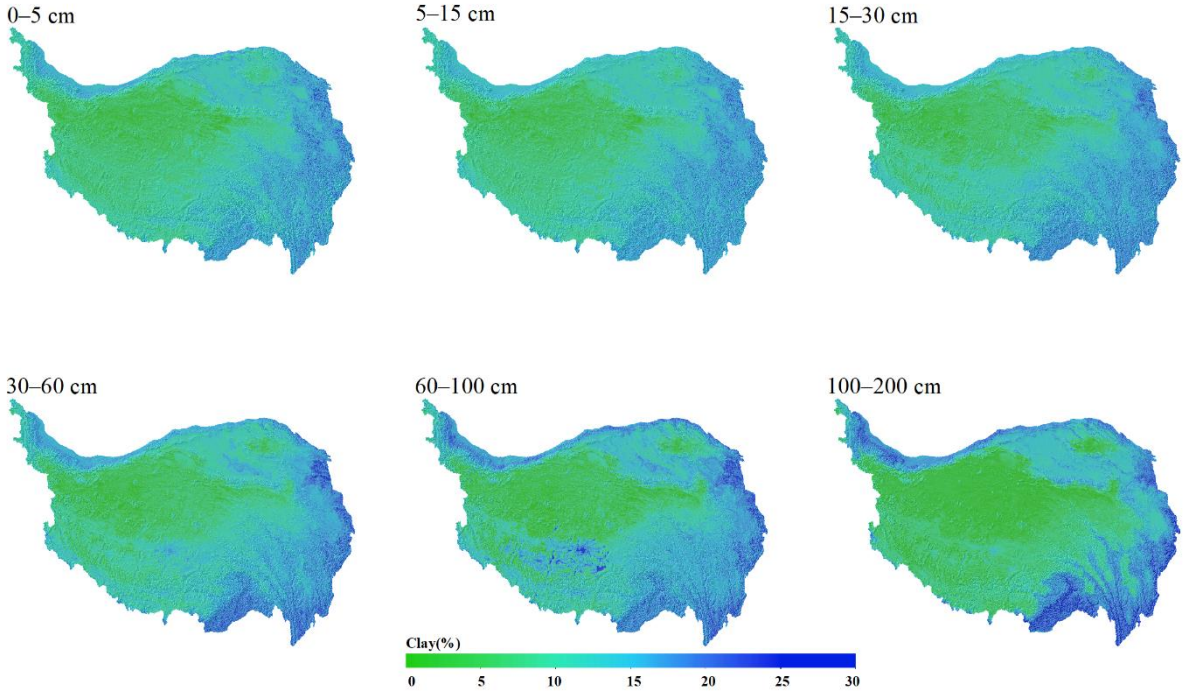

Figure 2. Spatial distribution of clay content at the six depth intervals.

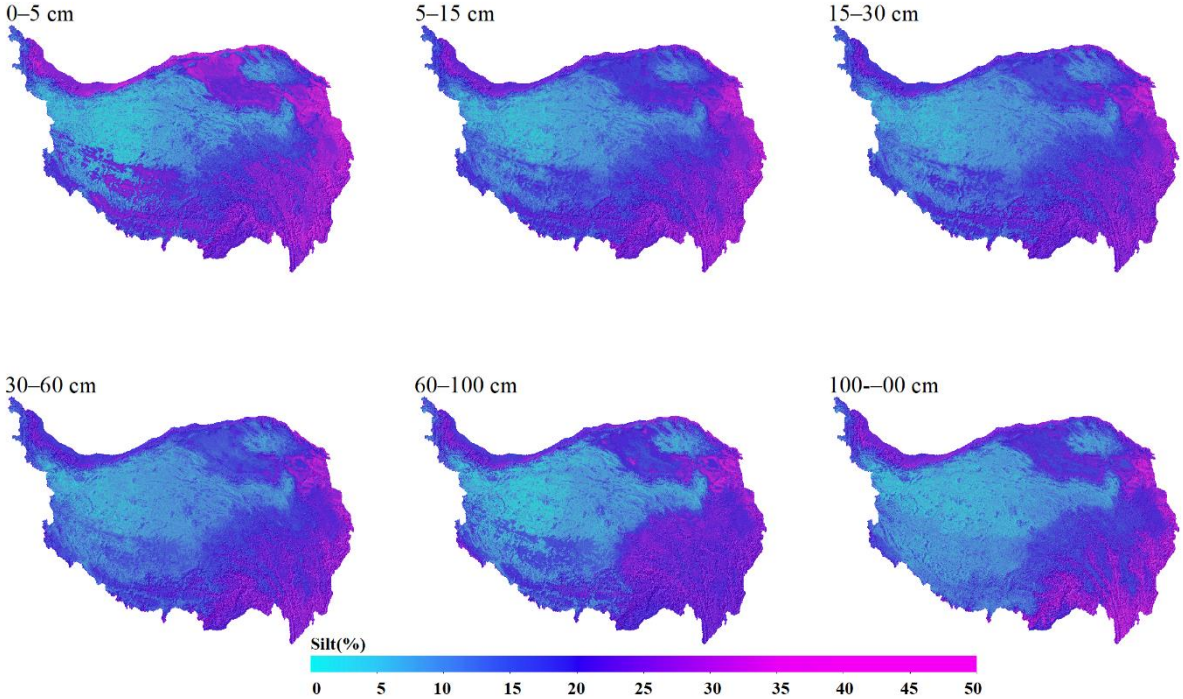

Figure 3. Spatial distribution of silt content at the six depth intervals.

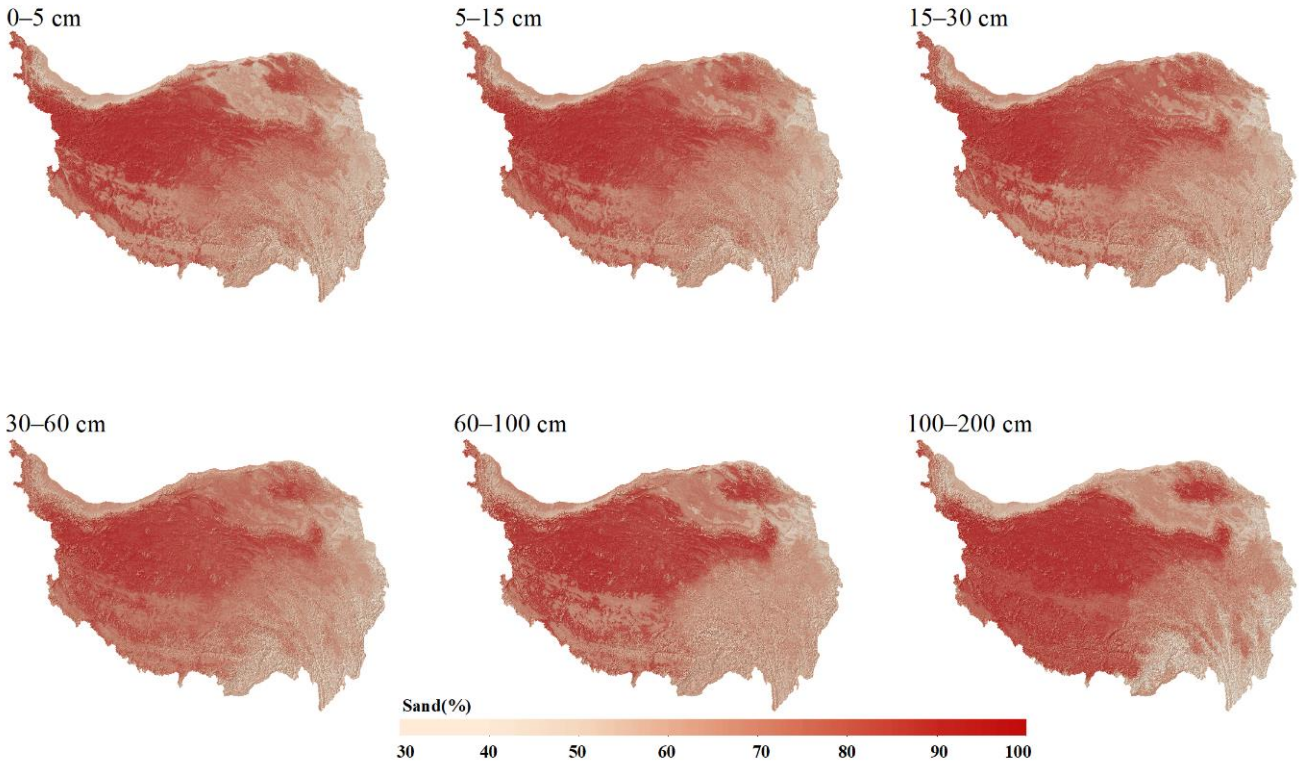

Figure 4. Spatial distribution of sand content at the six depth intervals.

3.4. Controlling Factors of QTP Soil Texture Patterns

The importance of environmental factors for soil texture in each layer were also calculated by the RF algorithm. We listed the relative importance of the first 12 important variables at multiple depths (Figures 5–7). For the clay content in the soils, the most important environmental factors affecting the spatial distribution of clay for the depth intervals within 0–60 cm were the annual mean temperature, channel network base level, and NDVI. For the 60–100 cm, the most important environmental factors were the channel network base level, annual mean temperature, NDVI, and elevation. For the 100–200 cm depth, the channel network base level, elevation, and annual mean temperature were the most important factors. For the silt content in the soil, the annual precipitation and annual mean temperature were the two most important factors in each layer. In the depth layer intervals within 0–60 cm, soil type, channel network base level and NDVI were also relatively important in addition to annual mean temperature and annual precipitation. In the depth layer intervals within 60–200 cm, elevation also played an important role. For the sand content in the soil, the annual mean temperature, annual precipitation, and NDVI were the main environmental factors that affect the spatial distribution of sand. For the 100–200 cm layer, soil type and elevation showed the greatest impact in sand contents.

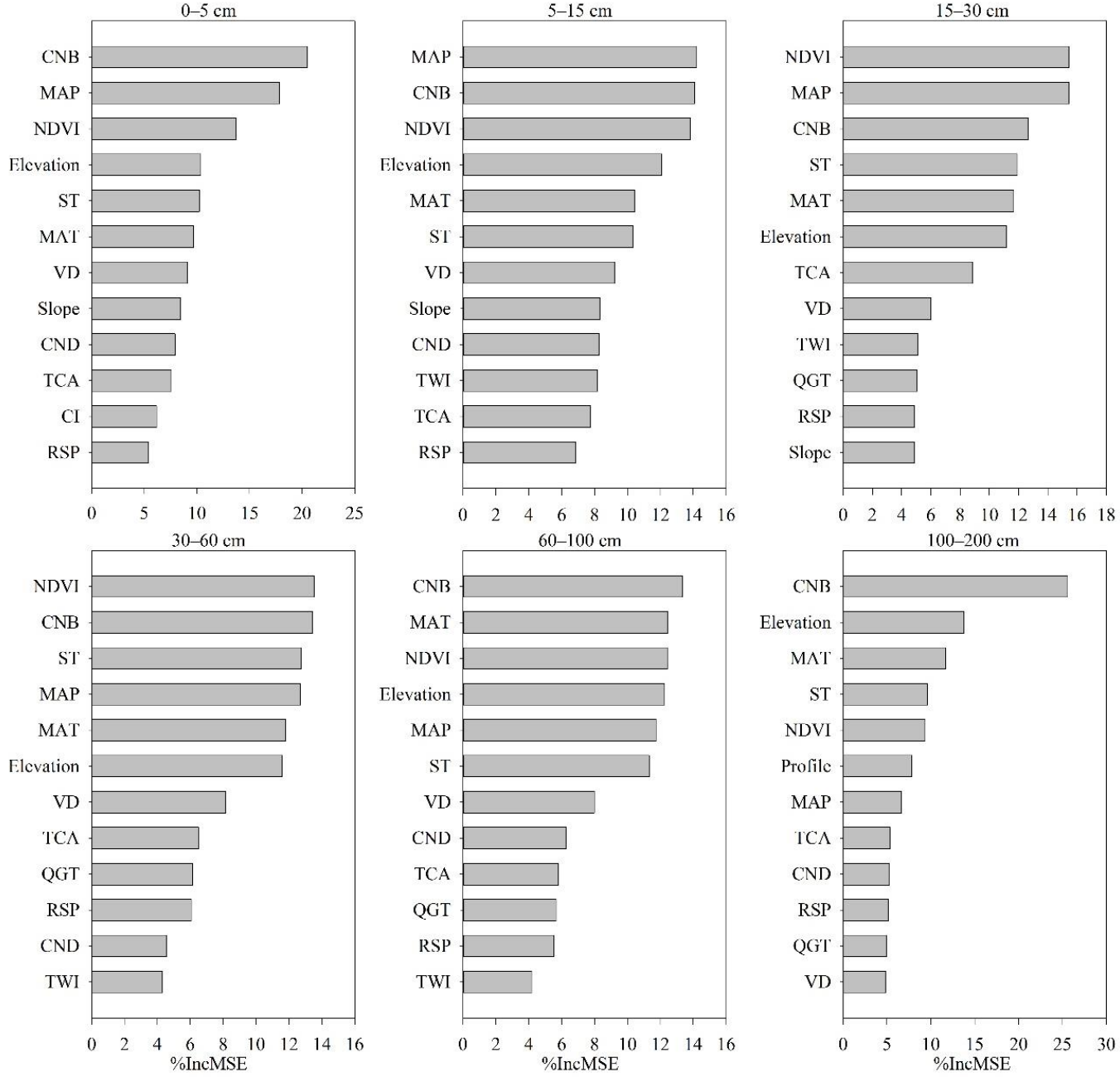

Figure 5. Relative importance (%IncMSE) of the covariates for clay predictions.

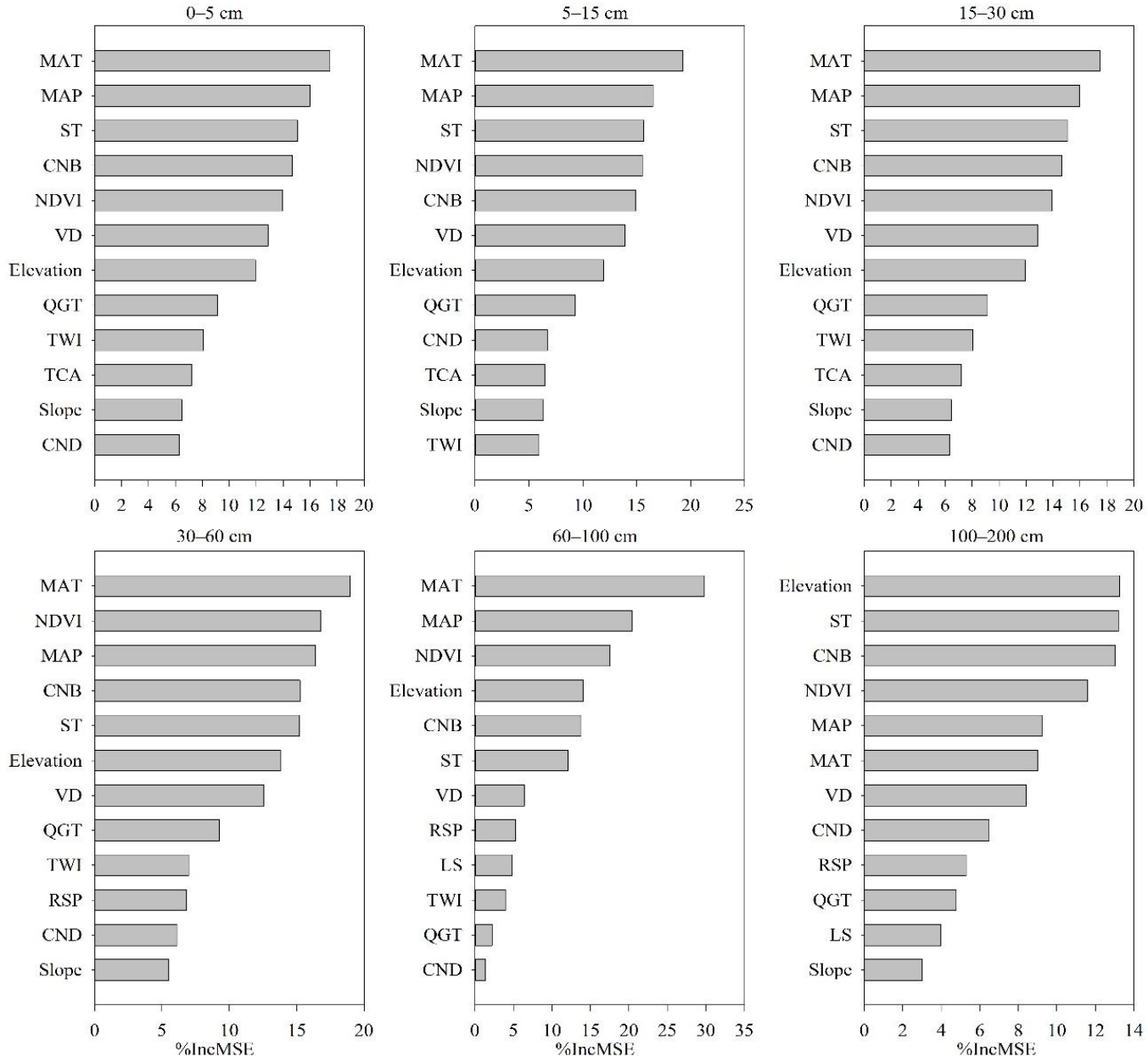

Figure 6. Relative importance (%IncMSE) of the covariates for silt predictions.

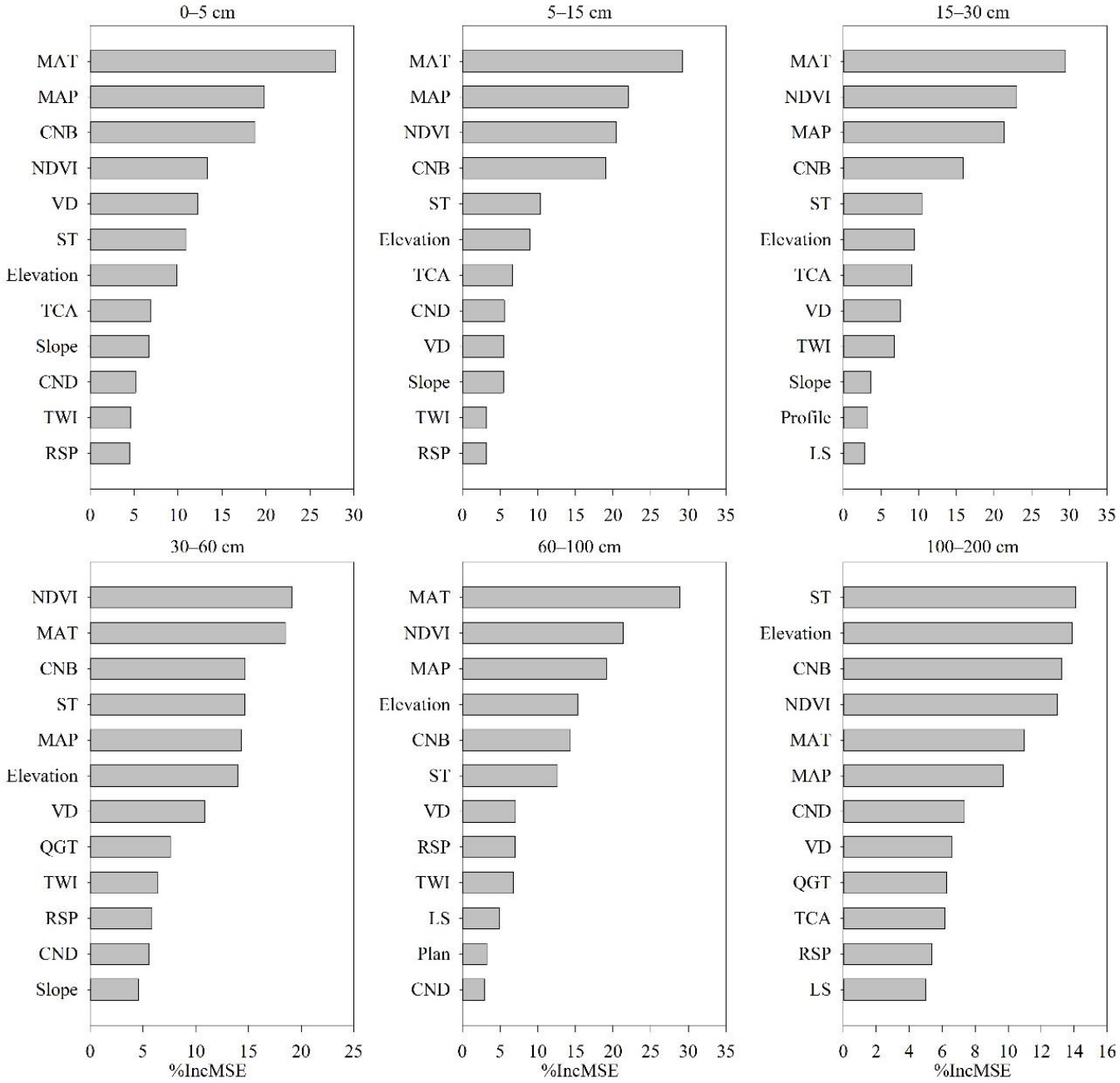

Figure 7. Relative importance (%IncMSE) of the covariates for sand predictions. (MAT: Annual mean temperature; MAP: Annual precipitation; CNB: channel network base level; NDVI: Mean NDVI during the growing season; VD: valley depth; ST: Soil type; Elevation: Elevation above sea level; TCA: total catchment area; Slope: Slope gradient; CND: channel network distance; TWI: topographic wetness index; LS: Slope length and steepness factor; RSP: relative slope position; CI: convergence index; QGT: quaternary geological type; Aspect: Aspect gradient; Plan: Plan curvature; Profile: Profile curvature).

4. Discussion

Our study showed that the soil particles on the QTP were mainly composed of coarse particles of sand, which accounted for more than 70% of the total content, along with low contents of silt and clay. As for the spatial distribution, silt and clay were higher in the east of the plateau and lower in the west of the plateau. Sand showed an opposite spatial pattern. This finding is in agreement with previous reports that the soil on the QTP largely consist of coarse particles [32,46,47]. The QTP is the youngest plateau of the world,

and the harsh natural conditions have hampered vegetation growth [48]. As a result, the weak pedogenesis leads to the coarse soil particle sizes distribution. Comparing with the permafrost map of the QTP [18], it can be found that higher sand contents mainly appear in the permafrost regions, while silt and clay are mainly distributed in seasonally-frozen ground regions. This can be explained as the fact that the permafrost regions on the QTP have a low air temperature and precipitation, but a high evapotranspiration [49,50]. Under these conditions, pedogenesis is usually weak [51], resulting in the coarse soil particle distribution that was observed.

In some studies that are based on regression models, collecting sample data in the study area is relatively easy, and can be obtained in large quantities. Because of this, a part of the sample data can be reserved for the accuracy evaluation of the mapping results and can be used in the comparative analysis of previous research results [6,52–54]. However, the QTP has a large altitude gradient and a harsh environment, especially in the permafrost regions, making data collection difficult. Therefore, our sample points were few and unevenly distributed relative to the huge QTP. We used all of the sample points of each standard layer for the establishment of the random forest model, and used a ten-fold cross-validation to evaluate the model's accuracy. The ten-fold cross-validation showed that the random forest model has a good accuracy in the prediction of sand and silt. In many studies, the simulation for sand and silt usually had higher accuracies than that of clay [53,55,56]. This is reasonable, as the clay contents usually were low and had large heterogeneities in horizontal and vertical distribution. For the predictions of silt and sand, our results had R^2 values in the range of 0.44–0.57 and 0.45–0.62 for silt and sand, respectively, and these values were high. For the soil texture mapping of China, the ten-fold cross-validated R^2 values of sand and silt are 0.45–0.50 and 0.44–0.49, respectively [6]. Using the Cubist decision tree algorithm, the accuracy of soil texture predictions in Denmark had the R^2 values of 0.27–0.44, 0.29–0.55, and 0.26–0.54 for sand, silt and clay, respectively [1]. In France, the R^2 values of silt, clay, and sand predictions were are 0.21–0.42, 0.19–0.33, and 0.25–0.44 for sand, silt, and clay, respectively [53]. Through the horizontal comparison with the regression model R^2 of the above studies in different regions worldwide, we found that we achieved good results in the prediction and mapping of the soil texture stratification in the QTP using the random forest model.

For clay, within the depth layers of 0–30 cm, CNB, MAP, and NDVI are the main control environmental factors, but in the layers of 30–60 cm, 60–100 cm, and 100–200 cm, the main control environmental factors have no obvious regularity. For silt, within the depth layers of 0–30 cm, MAT, MAP, and ST are the main control environmental factors, and at 30–60 and 60–100 cm, MAT, NDVI, and MAP play the most important role. For sand, within the 0–30 cm depths, MAT, MAP, and NDVI play a decisive role, while at 30–60 and 60–100 cm, NDVI and MAT are the main control factors. Therefore, MAT, MAP, and NDVI can generally be regarded as the most important environmental factors affecting the spatial distribution of soil texture on the QTP. Studies have shown that the precipitation decreases from east to west, and that the temperature decreases when moving from southeast to northwest on the QTP [57,58], which has great synergy with the spatial distribution of soil texture that we predicted. Climate can control the formation and development of soil by affecting vegetation type [59]. Under arid and semi-arid climates, the QTP has a sparse soil vegetation coverage, and the surface soil is vulnerable to wind and water erosion, meaning that the organic matter accumulation and profile development are weak. Therefore, vegetation develops better in areas on the QTP with higher temperatures and more precipitation, which promotes the accumulation of organic matter and the development of soil profiles, and results in finer soil particles; however, in areas with low temperature and low precipitation, the vegetation development is very poor, the surface is mostly bare, the soil development degree is very low, and the soil particles are mostly composed of large particles. Studies have shown that the grassland, which occupies nearly 75% of the plateau, is the dominant land cover type of the QTP, and the vegetation coverage on the plateau decreases when

moving from the southeast to the northwest [30,60]. Consequently, the NDVI was found to be an important factor affecting the distribution of soil particles. We stress that our study shows the importance of mean annual precipitation and mean annual air temperature in the determination of the soil texture; it is possible that other climate variables, such as monthly mean air temperature and precipitation, may be better predictors of soil texture. Future studies should be conducted to find more useful climate indices to increase the accuracy of soil texture predictions.

The QTP has long been recognized as an important geomorphic unit on Earth [61]. The energy, water, and carbon cycles in this region has global implications [21]. Affected by factors such as altitude, temperature, and precipitation, the composition of soil particles contains obvious regional characteristics. However, the existing soil texture data on this plateau prior to this study were only part of data on a national or global scale [6]. Previous reports only included a few sample sites from the QTP, and especially in permafrost regions where collection is limited by terrain, traffic, etc., sample sites were few. In addition, the models to predict soil particle size from a national or global scale likely result in a lower prediction accuracy on the QTP, which can be attributed to the fact that the QTP has very different climatic and geological conditions. Thus, the results obtained by machine learning have a great level of uncertainty in the QTP region. In this study, we collected a large amount of field soil profile data on the QTP to overcome the shortcomings of previous studies.

5. Conclusions

We collected field observation data on the QTP and examined the spatial distribution of the soil texture on the QTP using the random forest regression model. The ten-fold cross-validation results show that the simulation of silt and sand have high accuracy, with the R^2 values being 0.44–0.57 and 0.45–0.62 for silt and sand, respectively. On the QTP, the soil particles were dominated by sand, which accounted for more than 70% of the total composition. The sand contents were mostly high in the northwest and low in the southeast portion of the plateau. Conversely, clay and silt contents were high in the southeast and low in the northwest. The high sand contents were mainly found in permafrost areas with high altitudes and harsh climatic conditions. In contrast, the high values of silt and clay mainly appeared in the seasonally-frozen ground areas. The climate and NDVI values were determined to be the most important environmental factors affecting the spatial distribution of soil texture on the QTP. It is worth noting that it is difficult to obtain soil samples in the 'no man's land' hinterland of the QTP due to the limited traffic access. Therefore, more field observation data studies are needed to improve the simulation accuracy at the plateau scale.

Author Contributions: Conceptualization: X.W. (Xiaodong Wu) and W.L.; formal analysis and investigation: L.Z., T.W., R.L., G.H., D.Z., J.N. and Y.D.; writing—original draft: Y.L.; funding acquisition: X.W. (Xiaodong Wu) and W.L.; writing—review and editing: M.W., Z.L., X.W. (Xianhua Wei) and X.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant No. 41871060, 32061143032, and 41601066), the State Key Laboratory of Cryospheric Sciences (Under grant SKLCS-ZZ-2022), and the West Light Foundation of the Chinese Academy of Sciences.

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

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

References

1. Adhikari, K.; Kheir, R.B.; Greve, M.B.; Bøcher, P.K.; Malone, B.P.; Minasny, B.; McBratney, A.B.; Greve, M.H. High-Resolution 3-D Mapping of Soil Texture in Denmark. *Soil Sci. Soc. Am. J.* **2013**, *77*, 860–876. [CrossRef]
2. Fernandez-Illescas, C.P.; Porporato, A.; Laio, F.; Rodriguez-Iturbe, I. The ecohydrological role of soil texture in a water-limited ecosystem. *Water Resour. Res.* **2001**, *37*, 2863–2872. [CrossRef]
3. Safari, Y.; Boroujeni, I.E.; Kamali, A.; Salehi, M.H.; Bodaghabadi, M.B. Mapping of the soil texture using geostatistical method (a case study of the Shahrekord plain, central Iran). *Arab. J. Geosci.* **2012**, *6*, 3331–3339. [CrossRef]
4. Shwetha, P.; Varija, K. Soil Water Retention Curve from Saturated Hydraulic Conductivity for Sandy Loam and Loamy Sand Textured Soils. *Aquat. Procedia* **2015**, *4*, 1142–1149. [CrossRef]
5. Abu-Hamdeh, N.H. Thermal Properties of Soils as affected by Density and Water Content. *Biosyst. Eng.* **2003**, *86*, 97–102. [CrossRef]
6. Liu, F.; Zhang, G.-L.; Song, X.; Li, D.; Zhao, Y.; Yang, J.; Wu, H.; Yang, F. High-resolution and three-dimensional mapping of soil texture of China. *Geoderma* **2020**, *361*, 114061. [CrossRef]
7. Lehmann, P.; Merlin, O.; Gentine, P.; Or, D. Soil Texture Effects on Surface Resistance to Bare-Soil Evaporation. *Geophys. Res. Lett.* **2018**, *45*, 10398–10405. [CrossRef]
8. Srivastava, A.; Saco, P.M.; Rodriguez, J.F.; Kumari, N.; Chun, K.P.; Yetemen, O. The role of landscape morphology on soil moisture variability in semi-arid ecosystems. *Hydrol. Processes* **2020**, *35*, e13990. [CrossRef]
9. Famiglietti, J.S.; Ryu, D.; Berg, A.A.; Rodell, M.; Jackson, T.J. Field observations of soil moisture variability across scales. *Water Resour. Res.* **2008**, *44*, 5804. [CrossRef]
10. Plante, A.F.; Conant, R.T.; Stewart, C.E.; Paustian, K.; Six, J. Impact of Soil Texture on the Distribution of Soil Organic Matter in Physical and Chemical Fractions. *Soil Sci. Soc. Am. J.* **2006**, *70*, 287–296. [CrossRef]
11. Nh, A.; Sm, B.; Rv, C. Exploring the driving forces and digital mapping of soil organic carbon using remote sensing and soil texture. *Catena* **2019**, *182*, 104141.
12. Liao, K.; Xu, S.; Wu, J.; Zhu, Q. Spatial estimation of surface soil texture using remote sensing data (Soil physics). *Soil Sci. Plant Nutr.* **2013**, *59*, 488–500. [CrossRef]
13. De Carvalho Conceição Telles, E.; de Camargo, P.B.; Martinelli, L.A.; Trumbore, S.E.; da Costa, E.S.; Santos, J.; Higuchi, N.; Oliveira, R.C. Influence of soil texture on carbon dynamics and storage potential in tropical forest soils of Amazonia. *Glob. Biogeochem. Cycles* **2003**, *17*, 1953.
14. Wang, D.-C.; Zhang, G.-L.; Pan, X.-Z.; Zhao, Y.-G.; Zhao, M.-S.; Wang, G.-F. Mapping Soil Texture of a Plain Area Using Fuzzy-c-Means Clustering Method Based on Land Surface Diurnal Temperature Difference. *Pedosphere* **2012**, *22*, 394–403. [CrossRef]
15. Braud, I.; Roux, H.; Anquetin, S.; Maubourguet, M.-M.; Manus, C.; Viallet, P.; Dartus, D. The use of distributed hydrological models for the Gard 2002 flash flood event: Analysis of associated hydrological processes. *J. Hydrol.* **2010**, *394*, 162–181. [CrossRef]
16. Luo, D.L.; Jin, H.J.; He, R.X.; Wang, X.F.; Muskett, R.R.; Marchenko, S.S.; Romanovsky, V.E. Characteristics of Water-Heat Exchanges and Inconsistent Surface Temperature Changes at an Elevational Permafrost Site on the Qinghai-Tibet Plateau. *J. Geophys. Res. Atmos.* **2018**, *123*, 10057–10075. [CrossRef]
17. Li, X.L.; Gao, J.; Brierley, G.; Qiao, Y.M.; Zhang, J.; Yang, Y.W. Rangeland Degradation on the Qinghai-Tibet Plateau: Implications for Rehabilitation. *Land Degrad. Dev.* **2013**, *24*, 72–80. [CrossRef]
18. Zou, D.; Zhao, L.; Sheng, Y.; Chen, J.; Hu, G.; Wu, T.; Wu, J.; Xie, C.; Wu, X.; Pang, Q.; et al. A new map of permafrost distribution on the Tibetan Plateau. *Cryosphere* **2017**, *11*, 2527–2542. [CrossRef]
19. Zhao, L.; Zou, D.; Hu, G.; Du, E.; Pang, Q.; Xiao, Y.; Li, R.; Sheng, Y.; Wu, X.; Sun, Z.; et al. Changing climate and the permafrost environment on the Qinghai–Tibet (Xizang) plateau. *Permafr. Periglac. Process.* **2020**, *31*, 396–405. [CrossRef]
20. Poloczanska, E. *The IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*; Intergovernmental Panel on Climate Change: Geneva, Switzerland, 2019.
21. Mu, C.; Abbott, B.W.; Norris, A.J.; Mu, M.; Fan, C.; Chen, X.; Jia, L.; Yang, R.; Zhang, T.; Wang, K.; et al. The status and stability of permafrost carbon on the Tibetan Plateau. *Earth-Sci. Rev.* **2020**, *211*, 103433. [CrossRef]
22. Dickinson, R.E. Land surface processes and energy balance climate-surface albedos and Energy Balance. *Adv. Geophys.* **1983**, *25*, 305–353.
23. Wu, Q.; Zhang, T.; Liu, Y.J.G.; Change, P. Permafrost temperatures and thickness on the Qinghai-Tibet Plateau. *Glob. Planet. Chang.* **2010**, *72*, 32–38. [CrossRef]
24. Li, R.; Zhao, L.; Ding, Y.J.; Wu, T.H.; Xiao, Y.; Du, E.J.; Liu, G.Y.; Qiao, Y. Temporal and spatial variations of the active layer along the Qinghai-Tibet Highway in a permafrost region. *Chin. Sci. Bull.* **2012**, *57*, 4609–4616. [CrossRef]
25. Cheng, G.; Wu, T. Responses of permafrost to climate change and their environmental significance, Qinghai-Tibet Plateau. *J. Geophys. Res.* **2007**, *112*, 631. [CrossRef]
26. Niu, F.; Cheng, G.; Ni, W.; Jin, D. Engineering-related slope failure in permafrost regions of the Qinghai-Tibet Plateau. *Cold Reg. Sci. Technol.* **2005**, *42*, 215–225. [CrossRef]

27. Nachtergaele, F.; Velthuizen, H.V.; Verelst, L.; Batjes, N.H.; Dijkshoorn, K.; Engelen, V.; Fischer, G.; Jones, A.; Montanarela, L. The Harmonized World Soil Database. 2010. Available online: https://library.wur.nl/WebQuery/wurpubs/fulltext/154132 (accessed on 29 May 2022).
28. Wei, S.; Dai, Y.; Liu, B.; Ye, A.; Hua, Y.J.G. A soil particle-size distribution dataset for regional land and climate modelling in China. *Geoderma* **2012**, *171–172*, 85–91.
29. Padarian, J.; Minasny, B.; McBratney, A.B. Chile and the Chilean soil grid: A contribution to GlobalSoilMap. *J Geoderma Regional.* **2016**, *9*, 17–28. [CrossRef]
30. Ke, H.; Yangjian, Z.; Juntao, Z.; Yaojie, L.; Jiaxing, Z.; Jing, Z. The Influences of Climate Change and Human Activities on Vegetation Dynamics in the Qinghai-Tibet Plateau. *Remote Sens.* **2016**, *8*, 876.
31. Lin, Z.; Ping, C.L.; Yang, D.; Cheng, G.; Ding, Y.; Liu, S.J.G.; Change, P. Changes of climate and seasonally frozen ground over the past 30 years in Qinghai–Xizang (Tibetan) Plateau, China. *Glob. Planet. Chang.* **2004**, *43*, 19–31.
32. Wu, X.; Fang, H.; Zhao, Y.; Smoak, J.M.; Li, W.; Shi, W.; Sheng, Y.; Zhao, L.; Ding, Y. A conceptual model of the controlling factors of soil organic carbon and nitrogen densities in a permafrost-affected region on the eastern Qinghai-Tibetan Plateau. *J. Geophys. Res. Biogeosci.* **2017**, *122*, 1705–1717. [CrossRef]
33. Gilluly, J.; Waters, A.C.; Woodford, A.O. *Principles of Geology*; Penguin Classics: London, UK, 1951.
34. McBratney, A.B.; Santos, M.M.; Minasny, B.J.G. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [CrossRef]
35. Wang, F.; Yang, S.; Wei, Y.; Shi, Q.; Ding, J. Characterizing soil salinity at multiple depth using electromagnetic induction and remote sensing data with random forests: A case study in Tarim River Basin of southern Xinjiang, China. *Sci. Total Environ.* **2021**, *754*, 142030. [CrossRef]
36. Van Beijma, S.; Comber, A.; Lamb, A. Random forest classification of salt marsh vegetation habitats using quad-polarimetric airborne SAR, elevation and optical RS data. *Remote Sens. Environ.* **2014**, *149*, 118–129. [CrossRef]
37. LI, Z.-W.; Xin, X.-P.; Huan, T.; Fan, Y.; Chen, B.-R.; Zhang, B.-H. Estimating grassland LAI using the Random Forests approach and Landsat imagery in the meadow steppe of Hulunber, China. *J. Integr. Agric.* **2017**, *16*, 286–297. [CrossRef]
38. Breiman, L.J.M.l. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
39. Ham, J.; Chen, Y.; Crawford, M.M.; Ghosh, J.; Sensing, R. Investigation of the random forest framework for classification of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 492–501. [CrossRef]
40. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]
41. Strobl, C.; Boulesteix, A.L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [CrossRef]
42. Touw, W.G.; Bayjanov, J.R.; Overmars, L.; Backus, L.; Boekhorst, J.; Wels, M.; van Hijum, S.A.F.T. Data mining in the Life Sciences with Random Forest: A walk in the park or lost in the jungle? *Brief. Bioinform.* **2013**, *14*, 315–326. [CrossRef]
43. Zabihi, M.; Pourghasemi, H.R.; Pourtaghi, Z.S.; Behzadfar, M. GIS-based multivariate adaptive regression spline and random forest models for groundwater potential mapping in Iran. *Environ. Earth Sci.* **2016**, *75*, 665. [CrossRef]
44. Ye, T.; Zhao, N.; Yang, X.; Ouyang, Z.; Liu, X.; Chen, Q.; Hu, K.; Yue, W.; Qi, J.; Li, Z. Improved population mapping for China using remotely sensed and points-of-interest data within a random forests model. *Sci. Total Environ.* **2019**, *658*, 936–946. [CrossRef]
45. Avdeef, A. Do you know your r2? *ADMET DMPK* **2021**, *9*, 69–74. [CrossRef]
46. Baumann, F.; HE, J.S.; Schmidt, K.; Kuehn, P.; Scholten, T. Pedogenesis, permafrost, and soil moisture as controlling factors for soil nitrogen and carbon contents across the Tibetan Plateau. *Glob. Chang. Biol.* **2009**, *15*, 3001–3017. [CrossRef]
47. Wu, X.; Zhao, L.; Fang, H.; Zhao, Y.; Smoak, J.M.; Pang, Q.; Ding, Y. Environmental controls on soil organic carbon and nitrogen stocks in the high-altitude arid western Qinghai-Tibetan Plateau permafrost region. *J. Geophys. Res. Biogeosci.* **2016**, *121*, 176–187. [CrossRef]
48. Deng, T.; Ding, L. Paleoaltimetry reconstructions of the Tibetan Plateau: Progress and contradictions. *Natl. Sci. Rev.* **2015**, *2*, 417–437. [CrossRef]
49. Feng, W.; Lu, H.; Yao, T.; Yu, Q. Drought characteristics and its elevation dependence in the Qinghai–Tibet plateau during the last half-century. *Sci. Rep.* **2020**, *10*, 14323. [CrossRef]
50. Qu, Y.; Zhu, Z.; Chai, L.; Liu, S.; Montzka, C.; Liu, J.; Yang, X.; Lu, Z.; Jin, R.; Li, X. Rebuilding a microwave soil moisture product using random forest adopting AMSR-E/AMSR2 brightness temperature and SMAP over the Qinghai–Tibet Plateau, China. *Remote Sens.* **2019**, *11*, 683. [CrossRef]
51. Mu, C.; Zhang, T.; Zhang, X.; Cao, B.; Peng, X.; Cao, L.; Su, H. Pedogenesis and physicochemical parameters influencing soil carbon and nitrogen of alpine meadows in permafrost regions in the northeastern Qinghai-Tibetan Plateau. *Catena* **2016**, *141*, 85–91. [CrossRef]
52. Poggio, L.; Gimona, A. 3D mapping of soil texture in Scotland. *Sci. Rep.* **2017**, *9*, 5–16. [CrossRef]
53. Mulder, V.L.; Lacoste, M.; Richer-de-Forges, A.C.; Arrouays, D. GlobalSoilMap France: High-resolution spatial modelling the soils of France up to two meter depth. *Sci. Total Environ.* **2016**, *573*, 1352–1369. [CrossRef]
54. Dobarco, M.R.; Orton, T.G.; Arrouays, D.; Lemercier, B.; Paroissien, J.-B.; Walter, C.; Saby, N.P. Prediction of soil texture using descriptive statistics and area-to-point kriging in Region Centre (France). *Geoderma Reg.* **2016**, *7*, 279–292. [CrossRef]
55. Ramcharan, A.; Hengl, T.; Nauman, T.; Brungard, C.; Waltman, S.; Wills, S.; Thompson, J. Soil property and class maps of the conterminous United States at 100-meter spatial resolution. *Soil Sci. Soc. Am. J.* **2018**, *82*, 186–201. [CrossRef]

56. Rossel, R.V.; Chen, C.; Grundy, M.; Searle, R.; Clifford, D.; Campbell, P.J. The Australian three-dimensional soil grid: Australia's contribution to the GlobalSoilMap project. *Soil Res.* **2015**, *53*, 845–864. [CrossRef]
57. Chen, X.; An, S.; Inouye, D.W.; Schwartz, M.D. Temperature and snowfall trigger alpine vegetation green-up on the world's roof. *Glob. Chang. Biol.* **2015**, *21*, 3635–3646. [CrossRef]
58. An, S.; Chen, X.; Zhang, X.; Lang, W.; Ren, S.; Xu, L. Precipitation and minimum temperature are primary climatic controls of alpine grassland autumn phenology on the Qinghai-Tibet Plateau. *Remote Sens.* **2020**, *12*, 431. [CrossRef]
59. Braskerud, B. The influence of vegetation on sedimentation and resuspension of soil particles in small constructed wetlands. *J. Environ. Qual.* **2001**, *30*, 1447–1457. [CrossRef]
60. Wang, Z.-W.; Wang, Q.; Zhao, L.; Wu, X.-D.; Yue, G.-Y.; Zou, D.-F.; Nan, Z.-T.; Liu, G.-Y.; Pang, Q.-Q.; Fang, H.-B. Mapping the vegetation distribution of the permafrost zone on the Qinghai-Tibet Plateau. *J. Mt. Sci.* **2016**, *13*, 1035–1046. [CrossRef]
61. Li, J.; Zhou, S.; Zhao, Z.; Zhang, J. The Qingzang movement: The major uplift of the Qinghai-Tibetan Plateau. *Sci. China Earth Sci.* **2015**, *58*, 2113–2122. [CrossRef]