SOILS, SEC 1 • SOIL ORGANIC MATTER DYNAMICS AND NUTRIENT CYCLING • RESEARCH ARTICLE

Soil organic carbon in semiarid alpine regions: the spatial distribution, stock estimation, and environmental controls

Meng Zhu [1,2] • Qi Feng [1] • Mengxu Zhang [2] • Wei Liu [1] • Ravinesh C. Deo [3] • Chengqi Zhang [1] • Linshan Yang [1]

Abstract

Purpose Soil organic carbon (SOC) in alpine regions is characterized by a strong local heterogeneity, which may contribute to relatively large uncertainties in regional SOC stock estimation. However, the patterns, stock, and environmental controls of SOC in semiarid alpine regions are still less understood. Therefore, the purpose of this study is to comprehensively quantify the stock and controls of SOC in semiarid alpine regions.

Materials and methods Soils from 138 study sites across a typical semiarid alpine basin (1755–5051 m, $\sim 1 \times 10^4$ km^2) are sampled at 0–10, 10–20, 20–40, and 40–60 cm. SOC content, bulk density, soil texture, and soil pH are determined. Both a classical statistical model (i.e., a multiple linear regression, MLR) and a machine learning technique (i.e., a random forest, RF) are applied to estimate the SOC stock at a basin scale. The study further quantifies the environmental controls of SOC based on a general linear model (GLM) coupled with the structural equation modeling (SEM).

Results and discussion SOC density varies significantly with topographic factors, with the highest values occurring at an elevation zone of ~3400 m. The results show that the SOC is more accurately estimated by the RF compared to the MLR model, with a total stock of 219.33 Tg C and an average density of 21.25 kg C m^{-2} at 0–60 cm across the study basin. The GLM approach reveals that the topography is seen to explain about 58.11% of the total variation in SOC density at 0–10 cm, of which the largest two proportions are attributable to the elevation (44.32%) and the aspect factor (11.25%). The SEM approach further indicates that, of the climatic, vegetative, and edaphic factors examined, the mean annual temperature, which is mainly shaped by topography, exerts the most significant control on SOC, mainly through its direct effect, and also, through indirect effect as delivered by vegetation type.

Conclusions The results of this study highlight the presence of high stocks of organic carbon in soils of semiarid alpine regions, indicating a fundamental role played by topography in affecting the overall SOC, which is mainly attained through its effects on the mean annual temperature.

Keywords Random forest · Semiarid alpine regions · Soil organic carbon · Structural equation modeling · Topography

✉ Qi Feng
  qifeng@lzb.ac.cn

1 Key Laboratory of Ecohydrology of Inland River Basin, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

2 University of Chinese Academy of Sciences, Beijing 100049, China

3 School of Agricultural Computational and Environmental Sciences, Centre for Sustainable Agricultural Systems and Centre for Applied Climate Sciences, University of Southern Queensland, Springfield, QLD 4300, Australia

1 Introduction

The soils store about 1500 Pg (1 Pg = 10^{15} g) of organic carbon for the upper 1 m depth, and this amount is about two times larger than that present in the atmosphere (Batjes 1996). Thus, even a small fraction of the organic carbon pool being converted into greenhouse gases and then released into the atmosphere could trigger significant positive feedback towards global warming (Melillo et al. 2002; Stockmann et al. 2013; Schuur et al. 2015). A comprehensive evaluation to the spatial variation, stocks, and environmental controls of the soil organic carbon (SOC) may therefore enable one to better parameterize the carbon cycle models which have been used to

predict feedbacks of soil carbon to global environmental change. Subsequently, this can help support new measures for improved carbon management and climate change mitigation policy (Yang et al. 2008; Liu et al. 2011; Scharlemann et al. 2014; Li et al. 2018).

Soils in alpine regions are of great importance in the global carbon cycle, mainly because of the high organic carbon density and the potential carbon losses as influenced by a warming climate (Prietzel et al. 2016; Chen et al. 2016a; Zhu et al. 2018). It has been suggested that the warming of the climate system in a mountainous region is likely to be elevation-dependent (Pepin et al. 2015); thus, organic carbon loss rate at a higher elevation zone may be subsequently enhanced. Additionally, as the carbon losses are highly dependent on the size of the standing organic carbon stock (Yang et al. 2010), an accurate estimation to the regional organic carbon stocks in alpine soils should therefore be regarded as a major research priority in the soil science area. However, due to insufficient field measurements resulting from a limited accessibility as well as a strong topography-induced spatial heterogeneity, the estimation to SOC stocks in alpine regions is still characterized by relatively large uncertainties (Chen et al. 2016a; Zhu et al. 2017; Qin et al. 2018). To reduce these uncertainties, a comprehensive soil sampling procedure is necessary, and appropriate upscaling and soil modeling methods should also be introduced to facilitate the estimation of SOC.

The most commonly used methods for upscaling the site-level measurements of SOC to the regional scale include the kriging interpolations approach based on geostatistical models, the classical statistical model that assume linear or non-linear relationships, and the more recently adopted machine learning techniques (Meersmans et al. 2008; Prasad et al. 2018a, b; Ghorbani et al. 2019). The ordinary kriging interpolation in particular has been used to predict the spatial patterns of SOC in regions that have a reasonably less strong local heterogeneity (Yang et al. 2010; Li et al. 2018). However, in regions characterized by complex landscape, where the SOC profile is likely to vary sharply at a very short distance, the interpolations-based approach (e.g., kriging) must be further calibrated by using other environmental covariates that influence the SOC, to improve the prediction accuracy (Song et al. 2016). The classical statistical models, on the other hand, such as those based on general linear model (Rial et al. 2016) and its modified counterpart known as the generalized linear model (Yang et al. 2008), including linear mixed models (Doetterl et al. 2013), and general addictive models have also been employed to predict the SOC content (de Brogniez et al. 2015). In recent years, with a rapid development of artificial intelligence approaches, machine learning techniques have evolved. Among them, artificial neural network (Yang et al. 2014), support vector machines (Ding et al. 2016), boosted regression tree (Yang et al. 2016), and the

random forest models (Akpa et al. 2016; Kouadio et al. 2018) have been the most used technique in mapping and modeling SOC and related soil parameters. These machine learning techniques were demonstrated to have a better performance compared to the other methods, since these approaches have a good ability to fit the non-linear relationships between SOC and any of its related environmental covariates (Grimm et al. 2008; Ding et al. 2016).

As for the environmental controls of SOC in a mountainous region, the topography, which substantially shapes the climatic, edaphic, and ecological processes, may also be regarded as a fundamental control parameter (Yimer et al. 2006; Lozano-García et al. 2016; Zhu et al. 2017). Generally, topographic factors do not directly affect the SOC accumulation, whereas they may shape the spatial patterns of SOC through directly affecting climatic factors, which can further alter the plant production and decomposition rates of the organic matter. Although many studies have highlighted the role of topography in influencing SOC at hillslopes and catchment scales (e.g., Lozano-García and Parras-Alcántara 2014; Chen et al. 2016a; Bangroo et al. 2017; Zhu et al. 2018), only a few have quantified how topography can affect the SOC through exerting its direct or indirect influence on other climatic, vegetative, and edaphic factors, especially in semiarid alpine regions. This knowledge, if well researched, is likely to enable researchers to better understand the topography-climate-vegetation-carbon relationships under complex topography. Furthermore, it will help reduce uncertainties in modeling the relative direction of change and the size of the carbon-climate feedback in a semiarid alpine region.

The Qilian Mountains, which is the focus of this study, has elevation ranging from 1600 to 5500 m and is a typical semiarid alpine region in northwestern China. These mountains have a long history of land use activities focused on grazing pastures. However, due to their importance in providing the main source of water flow for the oases in the middle and lower reaches (Yang et al. 2017), human activities in the study area are relatively limited since establishment of the National Nature Reserve of the Qilian Mountains in 1988. Across the range of this mountain, the topography-induced vegetation patterns are characterized both the vertical zonality and the horizontal patchiness. Within each vegetation and elevation zone, the aspect- and slope-induced microclimates are able to shape the community occurrence and plant productions at a smaller scale (Zhu et al. 2017; Zhu et al. 2018). This complex topography-induced vegetation patterns could further promote a much stronger spatial heterogeneity in the SOC profile within a relatively short distance (Qin et al. 2016; Zhu et al. 2017). Taken together, the unique vegetation patterns and strong spatial heterogeneity, together with the limited anthropogenic disturbances, make the Qilian Mountains an ideal semiarid alpine region for evaluating the spatial variability, stocks, and environmental controls of SOC.

Considering this importance and a paucity of this information over any semiarid alpine region, this study has conducted a soil sampling procedure by the topography across a typical basin in the middle Qilian Mountains. The main objectives of this research study are threefold: (1) to investigate the distribution of SOC density in respect to its topography factors, (2) accurately estimate the regional SOC stocks based on a more appropriate upscaling method, and (3) determine how the topography is able to shape the spatial patterns of SOC density by directly exerting influences on the other environmental controls.

2 Materials and methods

2.1 Study area

This research study was carried out in the Yingluoxia basin (98°–101°E, 38°–42°N), located in the middle Qilian Mountains. This region lies at the northern edge of the Qinghai-Tibetan Plateau (Fig. 1a). The Yingluoxia basin covers an area of 1.05×10^4 km^2 with elevation ranging from 1755 to 5051 m ASL (Fig. 1a). The basin is characterized by a typical semiarid alpine climate with regional mean annual temperature and precipitation around -4.6 °C and 396 mm, respectively. The spatial patterns of climatic factors mainly depend on the elevation, with a lapse rate of 0.58 °C for temperature changes, and about 15.5–22.1 mm per 100 m increasing rate of precipitation, respectively (Chen et al. 1992). Vegetation zone is strongly stratified by the elevation, and it varies from being that of a desert steppe (< 2500 m), a forest steppe (2500–3300 m), a subalpine shrub-meadow (3300–3600 m), an alpine meadow (3600–4000 m) to the alpine desert zone (> 4000 m) (Fig. 1b, c). Dominant species in the study area include *Stipa breviflora*, *Agropyron cristatum*, *Polygonum viviparum*, and *Kobresia* spp. for grasslands; *Potentilla fruticosa*, *Berberis diaphana*, *Caragana jubata*, and *Salix gilashanica* for shrublands; and *Sabina przewalskii* and *Picea crassifolia* for forests.

The soil parent materials are mainly eolian deposits and residual, slope, diluvial, alluvial, and fluvioglacial deposits made of limestone, sand-shale stone, or conglomerate. According to the Chinese soil genetic classification, the soils along the elevation gradient are defined as sierozems, kastanozems, chernozems, gray cinnamon, subalpine meadow soils, subalpine shrubby meadow soils, alpine meadow soils, and alpine frigid desert soils (Chen et al. 1992). These are roughly referred as luvic calcisols, kastanozems, chernozems, haplic greyzems, mollic leptosols, eutric leptosols, gelic leptosols, and glacic cryosols, respectively, according to the World Reference Base for Soil Resources 2014 (IUSS Working Group WRB 2015). In the alpine desert zone, the soil depths are generally less than 20 cm, and there appear to be rock compositions, mainly of the limestone type that lies below 60 cm in the sloping areas of shrublands and forests zones. As for the steppe and desert steppe zones, the soil profile comprises of lime concretions below the 60-cm depth.

2.2 Soil sampling and analysis

During the summer period of 2012–2015, researchers conducted soil sampling campaigns across the middle of the Qilian Mountains, and a total of 138 sampling sites were selected according to the topography and its vegetation type (Fig. 1). At each sampling site, the geographic and topographic information (i.e., the longitude, latitude, elevation, aspect, and slope) was recorded by a hand-held GPS and a compass. Following this, five soil profiles in a 10-m × 10-m plot were excavated, and the bulk density samples of 5.00 cm in diameter and 5.05 cm in height were collected by using the core ring method (100 cm^3 core volume) at a depth of 5, 15, 30, and 50 cm. Soil samples with three replicates were collected for an analysis of the soil chemical properties next to each soil profile using a soil auger (3.5 cm in diameter) at 0–10, 10–20, 20–40, and 40–60 cm. Finally, the soil samples from each plot were pooled to generate a composite sample from each of the four depth intervals.

In this study, the soil samples collected in the field were first sealed in aluminum specimen boxes and then transported to the research laboratory. Soil samples were then air-dried and then passed through a 2-mm sieve before any chemical analysis. The SOC content was determined by means of the wet oxidation with dichromate according to the Walkley-Black method (Nelson and Sommers 1982). Soil texture (i.e., sand, silt, and clay content) was determined by a laser diffraction approach using a Mastersizer 2000 (Malvern Instruments, Malvern, England). Soil bulk density was calculated as the ratio of the oven-dry soil mass (after 24 h desiccation at 105 °C) to the core volume (100 cm^3). Soil pH was measured in a 1:2.5 soil-to-deionized water mixture by a pH electrode (PB-10, Sartorius, Germany). SOC density (SOC amount per unit area) for a given depth at each site was calculated using Eq. (1) as (Yang et al. 2008):

$$\text{SOCD} = \sum_{i=1}^{n} C_i D_i B_i (1 - G_i)/100 \quad (1)$$

where SOCD is the SOC density (kg C m^{-2}), n is the number of layers divided in each soil profile, and C_i, D_i, B_i, and G_i represent the SOC content (g kg^{-1}), thickness (cm), bulk density (g cm^{-3}), and fractional percentage (%) of gravel larger than 2 mm in diameter of layer i, respectively.

2.3 Environmental variables

The elevation, aspect, slope, longitude, and latitude data in their raster formats at a 90-m resolution were derived
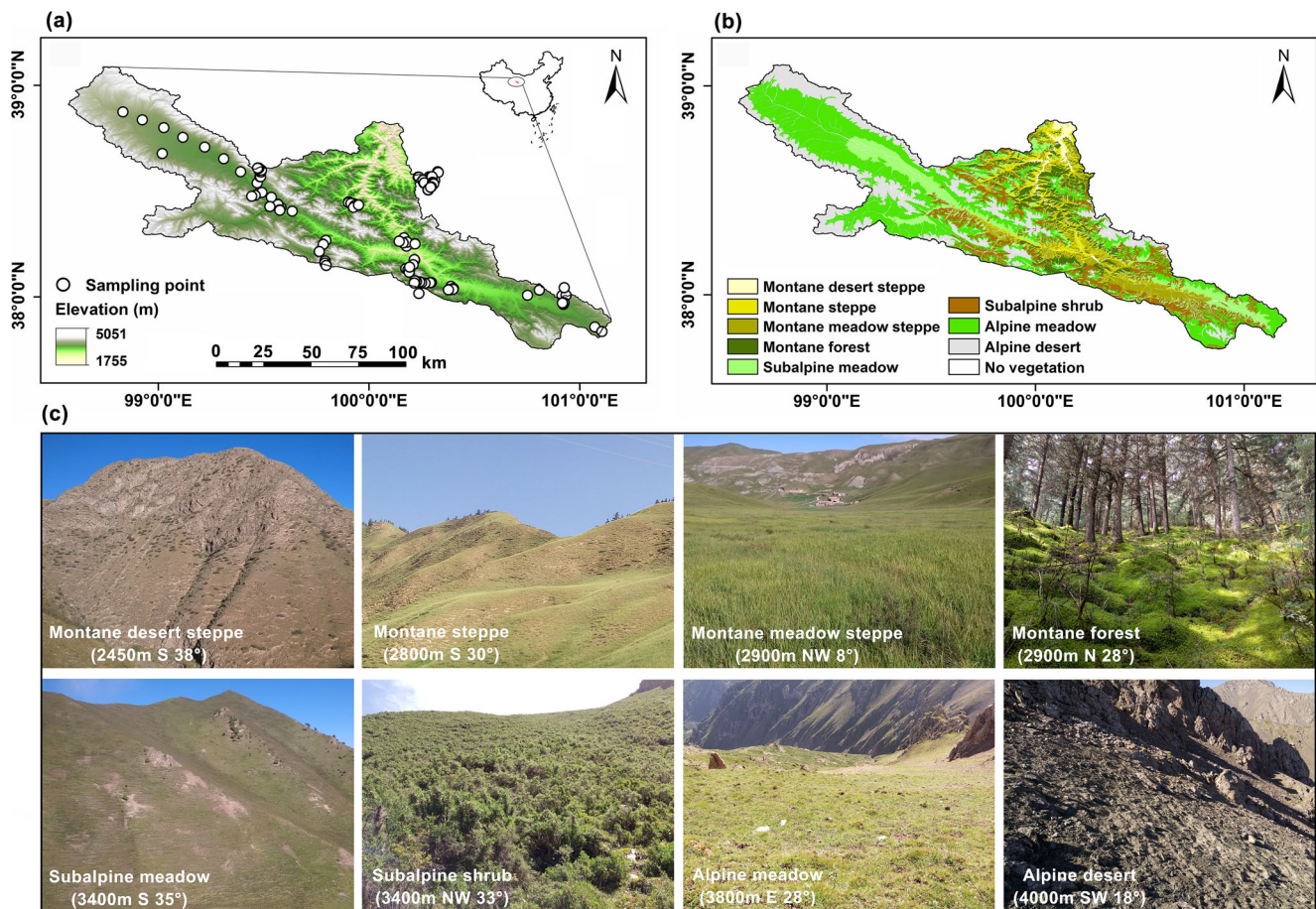
Fig. 1 Distribution of sampling sites across the Yingluoxia basin in the Qilian Mountains, northwestern China (**a**). Distribution of vegetation type in the basin (**b**). Photographs of the eight vegetation types with elevation, aspect, and slope added in the brackets (**c**). The S, NW, N, E, and SW in the brackets in **c** represent the south-, northwest-, north-, east-, and southwest-facing slopes, respectively

from the SRTM 90-m digital elevation model (DEM) (USGS 2006) using ArcGIS 10.2 (ESRI Inc., USA). The monthly compositing normalized difference vegetation index (NDVI) products at a 30-m resolution (2011 to 2014) were derived from the HJ/CCD (Li et al. 2017). The vegetation type data at the 1:100000 scale was also obtained from the Cold and Arid Regions Science Data Center at Lanzhou (http://westdc.westgis.ac.cn). The monthly NDVI products to a 90-m resolution were resampled, and subsequently, they were averaged over the growing season (from May to September) to generate their relative seasonal values (Ding et al. 2016).

In the present study area, the mean values of annual temperature and precipitation mainly depended on the topographic factors. These were obtained for each sampling site according to the following empirical equations (Zhao et al. 2005, 2006):

$$MAT = 20.957 - 0.00549H - 0.166Y + 0.0089X, R^2 = 0.98 \quad (2)$$

$$MAP = 1680.6235 + 0.119H - 75.264Y + 12.405X, R^2 = 0.92 \quad (3)$$

where MAT, MAP, H, Y, and X is the mean annual temperature (°C), mean annual precipitation (mm), elevation (m), latitude (°), and longitude (°) at the sampling sites, respectively. Note that the latitude, longitude, and elevation for each sampling site were derived by using a hand-held GPS.

The potential solar insolation for each sampling site was calculated based on the following equation (McCune and Keon 2002):

$$Rad = 0.339 + 0.808\cos(Y)\cos(S) - 0.196\sin(L)\sin(S) \quad (4)$$

$$-0.482\cos(A)\sin(S), R^2 = 0.98$$

where Rad represents the potential solar insolation (MJ cm^{-2} year^{-1}), and Y, S, and A is the latitude (°), slope (°), and aspect (°) at each of the sampling site, respectively.

2.4 Model inputs

In accordance with previous studies, the vegetation index and topographic factors are reasonably good indicators of the SOC under complex topographic systems, especially in an alpine

region (Yang et al. 2008; Song et al. 2016; Yang et al. 2016). In addition, it is important to note that an ideal model must have a series of covariates that are reasonably well correlated with the dependent variable (e.g., SOC), while the covariates themselves must have negligible correlation with each other. In this study, a pre-analysis showed that the climatic factors were highly depended on topographic factors as per Eqs. (2)–(4). Consequently, to minimize the uncertainties encountered in the estimation of SOC, and also to ensure computational efficiency of the models is not compromised due to a redundancy of the model's potential inputs, this study has only used satellite-derived vegetation index (i.e., the NDVI), elevation, aspect, slope, and latitude as model inputs to predict the spatial patterns of SOC density at 0–10, 10–20, 20–40, and 40–60 cm.

2.5 Model development process

In this research study, both a multiple linear regression (MLR) and a random forest (RF) regression model were developed using site-level measurements to predict the spatial patterns of SOC density at a depth of 0–10, 10–20, 20–40, and 40–60 cm across the study basin. The study has utilized the MLR model, as a classical regression method belonging to the general linear model, in which the SOC density can be modeled as a linear combination of the respective environmental covariates. By contrast, the RF regression model is a relatively newly developed machine learning technique especially in soil science area (Breiman 2001). Despite this, it has been increasingly used, for example, in digital soil mapping (Vågen et al. 2013; Heung et al. 2016), forecasting soil moisture (Prasad et al. 2018a, 2018b), modeling coffee yield in respect to soil fertility parameters (Kouadio et al. 2018), and the prediction of soil electrical conductivity (Ghorbani et al. 2019). According to these studies, RF regression model was seen to exhibit a relatively good performance in predicting SOC, mainly because it is able to well accommodate the nonlinear relationships between SOC and environmental covariates (Grimm et al. 2008; Yang et al. 2016; Akpa et al. 2016). In addition, the RF model is characterized by two randomization procedures of bootstrapping and random input selection and carries out a bagging of the predictions, which can subsequently improve its prediction ability (Suuster et al. 2012; Heung et al. 2016).

The best-fit equation of the MLR model used in this study can be stated as follows:

$$\ln(\text{SOCD}) = a_0 + a_1 NDVI + a_2 Y + a_3 H + a_4 A + a_5 AS \quad (5)$$

where SOCD is the SOC density (kg C m^{-2}), a_0, a_1, a_2, a_3, a_4, and a_5 are the intercept, regression coefficient for the NDVI, latitude (Y), elevation (H), aspect (A), and the interaction between aspect and slope (AS), respectively.

The regression coefficients, which are normally used to explain the response variable against a set of exploratory variables, were estimated by using the ordinary least squares method. In addition, we log-transformed the SOC density to meet the normal distribution requirements of the MLR approach, which helped avoid any negative (i.e., false) predictions of the tested SOC dataset.

It should be noted that the elevation term in Eq. (3) had to be first transformed, mainly to ensure a linear relationship is established between elevation and the ln(SOCD), so as to further improve the performance of the prescribed MLR model. The transformation was also conducted according to Eqs. (3)–(6) for a depth of 0–10, 10–20, 20–40, and 40–60 cm, respectively, detailed as follows:

For 0−10 cm, H

$$= 2.0248 \exp\left(-0.5 |(h-3385.9358)/797.3748|^{2.1113}\right) \quad (6)$$

For 10−20 cm, H

$$= 1.8412 \exp\left(-0.5 |(h-3402.6992)/775.0846|^{1.7827}\right) \quad (7)$$

For 20−40 cm, H

$$= 2.4320 \exp\left(-0.5 |(h-3426.2270)/910.6737|^{1.0000}\right) \quad (8)$$

For 40−60 cm, H

$$= 2.2589 \exp\left(-0.5 |(h-3472.1136)/830.1236|^{1.0000}\right) \quad (9)$$

In Eqs. (6)–(9), the term H and h refers to the transformed and real (or raw) elevation (m), respectively. These equations above were obtained by regressing the relationships between ln(SOCD) at the four selected depths with the real (or raw) elevation dataset (Fig. 2a–d). In addition, as the raw values of aspect can be relatively less useful variable in building models (Zhu et al. 2017), transformed transformation of the aspect with the cosine function was implemented, as a data pre-processing technique before conducting the MLR modeling procedure.

To fit the RF regression model in this study, researchers employed the same model inputs used in the MLR model. As for the key parameters incorporated into the regression model, one third of the total number of covariates were used to grow each of the random forest's tree, with the minimum number of terminal nodes set to 5, and the number of trees in the forest set at 1000. The "lm" function and "randomForest" package in the R software was used for modeling (R version 3.5.1, R Development Core Team 2018).

2.6 Model evaluation criteria

For any predictive model, the evaluation of its performance with both the visual display measures (e.g., plots of modeled
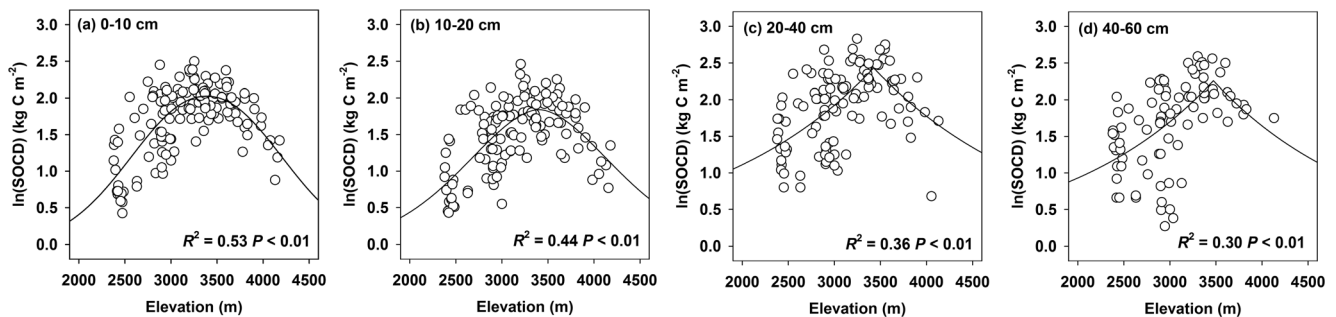
Fig. 2 Relationships between SOC density (SOCD) and elevation at 0–10 (**a**), 10–20 (**b**), 20–40 (**c**), and 40–60 cm (**d**). Note that the SOC density is log-transformed to meet the normal distribution premise of the fit models, and also to avoid negative predictions in SOC density at very lower or higher elevations

and real values) and overall statistical metrics is important. To validate the accuracy of the prescribed MLR and the RF regression models used in predicting SOC at the four different depths, this study has applied a 10-fold cross-validation procedure (Ali et al. 2018; Yin et al. 2018). In the 10-fold cross-validation approach, the original dataset was randomly divided into ten equal-sized subsets. Out of these ten subsets, a single subset was retained as the validation data for testing the developed model, and the remaining nine subsets were used as the model training dataset. This cross-validation procedure was then repeated ten times, with each of the ten subsets used exactly once as the validation data. The ten sets of results were then averaged to generate a single estimation model truly representative of a well-optimized model for SOC. The advantage of this procedure, which follows from a 50-fold model developed earlier (Yin et al. 2018) entailed from its repeated random sub-sampling approach where all of the observations were used in both the training and the validation process, and each of these observations were also used for the validation process exactly once.

After each round of a 10-fold cross-validation procedure, the root mean square error (*RMSE*), coefficient of determination (R^2), and Lin's concordance correlation coefficient (*LCCC*) (Lin 1989) were calculated for the two prescribed models, according to the following equations (Yang et al. 2016):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_i - O_i)^2} \qquad (10)$$

$$R^2 = \frac{\sum_{i=1}^{n} \left(P_i - \overline{O}\right)^2}{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2} \qquad (11)$$

$$LCCC = \frac{2r\sigma_O\sigma_P}{\sigma_O^2 + \sigma_P^2 + \left(\overline{O} - \overline{P}\right)^2} \qquad (12)$$

where P_i and O_i are the prediction and observation at sampling site i, respectively; n, σ_O, σ_P, and r are the sample size,

variance of the observations, variance of predictions, and Pearson correlation coefficient between predictions and observations, respectively; $\overline{O}$ and $\overline{P}$ are the means of observations and predictions, respectively. To ensure the precision of the above statistics for the two models, we totally conducted 100 times of the 10-fold cross-validation procedure and then calculated the means. The predictions of SOC became increasingly optimal as the *RMSE* and R^2 getting closer to 0 and 1, respectively. *LCCC* represented the degree to which the predicted and observed values follow the 45° line.

To further quantify the relative uncertainties in the bootstrapped predictions generated by trees or rules from different models in RF regression, we calculated the interquartile value (i.e., the difference between the 75th and 25th percentile) of the predicted SOC by performing 100 simulations of the RF regression for each pixel (Ding et al. 2016). As for the uncertainties in the MLR model, the 50% confidence interval of the regression results was first calculated, and then the upper limits with the lower limits was subtracted to obtain the interquartile for each pixel. The summed up quartiles were then used to assess the uncertainty of the SOC for each vegetation type. In addition, the relative uncertainty was obtained by dividing the interquartile by mean value for each pixel.

2.7 Statistical analyses

Analysis of variance (ANOVA) was used to examine the variation in SOC density in respect to the topographic factors. For this, the Duncan's new multiple range tests were performed for multiple comparisons when significant differences were detected by the ANOVA approach. In addition, the general linear models (GLMs) were used to quantify the contribution of aspect, elevation, and slope to the overall variation in SOC density. Furthermore, to explore the relationships between SOC and the topography-induced climatic factors, as well as the vegetative and edaphic factors, the Pearson correlation analysis was first conducted. Then, the relative importance of each control on SOC density based on the structural equation modeling (SEM) was quantified according to the known relationships between SOC and its key controls. The SEM

was a multivariate statistical method that can test the plausibility of a hypothetical model, which was based on a priori information regarding the relationships among particular variables. Additionally, the SEM method was also able to partition the direct and indirect effects that one environmental factor may have on another, and also to estimate the size of the multiple effects (Chen et al. 2016b).

Considering the non-linear relationships between SOC density and mean annual temperature and precipitation in this study, the two climatic factors with functions derived from the fit equations between SOC density and them were transformed, before conducting the Pearson correlation and SEM analyses. Also, the real (raw) elevation and aspect with Eqs. (6)–(9) and cosine function, respectively, were transformed and then standardized, before performing the GLM analyses. Both the ANOVA and the GLM analyses were performed employing the GLM procedure in SAS 9.2 (SAS Institute, Inc., Cary, NC, USA). The results of the Pearson correlation analysis were also obtained and visualized by the "cor" function and "corrplot" package, respectively, in R software. The SEM was performed using AMOS 20.0 software (IBM Corp., Armonk, New York, USA).

3 Results

3.1 Spatial distribution of SOC density

In general, the spatial distribution of SOC density was significantly shaped by the topographic factors (Table 1 and Fig. 2).

Specifically, the SOC density of 3000–3500 m was close to 3500–4000 m, and it was also significantly larger than that at higher or lower elevation zones ($P < 0.01$). The fit lines in Fig. 2 further indicated that SOC density had reached its peak at an elevation zone of ~3400 m. Along the aspect gradient, the SOC density actually decreased by 28%, 36%, 42%, and 46% as the aspect was transformed from north (0–45°) to south (135–180°) at a depth of 0–10, 10–20, 20–40, and 40–60 cm, respectively. SOC density was less varied with slope while still presented a slightly decreasing trend with increasing slope (Table 1).

3.2 SOC stock estimation

The R^2 values of the RF regression were higher than that of the MLR except at a depth of 0–10 cm, where the R^2 value of the MLR was higher than that of RF regression by about 0.01 (Fig. 3a–h). The average uncertainties in the MLR model reached about 23.91%, 28.52%, 18.99%, and 21.47% for the 0–10, 10–20, 20–40, and 40–60 cm, respectively, with large uncertainties mainly existent in regions with lower SOC values (Fig. 4a–d). By contrast, the RF regression provided a more robust prediction at most parts of the basin, with average uncertainties of about 6.48%, 6.40%, 7.94%, and 8.07% at the four depths, respectively (Fig. 4e–h).

According to SOC density map estimated by the RF regression (Fig. 5a–h), the total SOC stock at 0–60 cm was estimated to be about 219.33 Tg C (1 Tg = 10^{12} g) with an average density of 21.25 kg C m^{-2} (Table 2). The SOC stock at a depth

Table 1 Spatial distribution of SOC density

Type	Subgroup	Sample size	SOC density (kg C m^{-2})			
			0–10 cm	10–20 cm	20–40 cm	40–60 cm
Elevation (m)	< 2500	16	2.64 ± 0.26 d	2.34 ± 0.22 c	4.33 ± 0.39 c	4.00 ± 0.38 b
	2500–3000	36	5.61 ± 0.37 b	4.52 ± 0.28 b	6.88 ± 0.53 b	5.36 ± 0.54 b
	3000–3500	53	7.38 ± 0.26 a	6.01 ± 0.28 a	9.37 ± 0.50 a	7.74 ± 0.63 a
	3500–4000	26	6.93 ± 0.32 a	6.01 ± 0.32 a	9.09 ± 0.76 a	7.69 ± 0.68 a
	> 4000	7	4.08 ± 0.41 c	2.98 ± 0.25 c	–	–
Aspect (°)	0–45	55	7.03 ± 0.26 a	5.86 ± 0.24 a	8.89 ± 0.48 a	7.25 ± 0.54 a
	45–90	24	6.68 ± 0.53 a	5.38 ± 0.48 ab	8.07 ± 0.90 a	6.40 ± 0.81 a
	90–135	22	5.06 ± 0.51 b	4.43 ± 0.47 bc	5.50 ± 0.82 b	4.10 ± 0.82 b
	135–180	37	5.04 ± 0.37 b	3.76 ± 0.30 c	5.14 ± 0.60 b	3.91 ± 0.60 b
Slope (°)	0–10	40	6.64 ± 0.33 a	5.37 ± 0.28 a	7.57 ± 0.78 a	5.93 ± 0.83 a
	10–20	31	6.72 ± 0.43 a	5.68 ± 0.38 a	7.39 ± 0.69 a	5.70 ± 0.82 a
	20–30	33	5.85 ± 0.37 ab	4.81 ± 0.37 ab	7.27 ± 0.75 a	5.60 ± 0.73 a
	>30	34	5.22 ± 0.46 b	4.18 ± 0.38 b	7.16 ± 0.66 a	5.79 ± 0.63 a

The soil layer above 4000 m is less than 20 cm in depth; thus, there are no values for SOC density below 20 cm. All data are presented as mean ± standard error, and statistically different mean values within each type are indicated by different lowercase letters ($P < 0.05$)
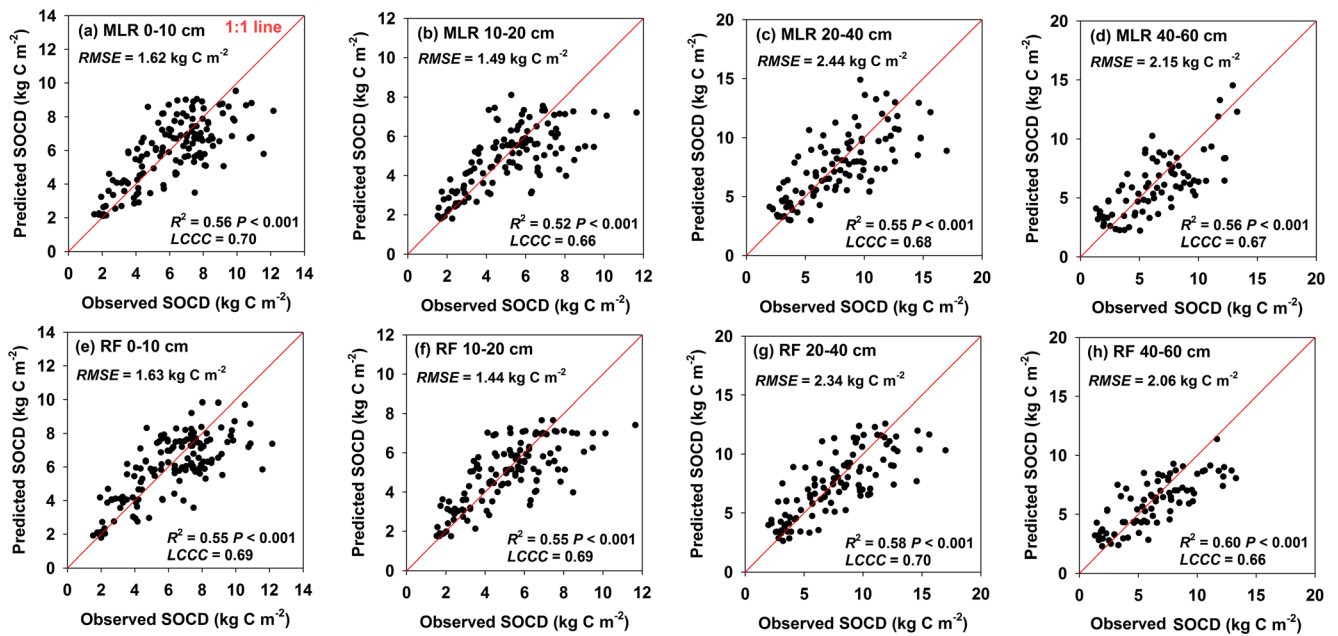
Fig. 3 Model evaluation for both the MLR (**a**–**d**) and the RF regression (**e**–**h**) in predicting SOC density (SOCD) at 0–10, 10–20, 20–40, and 40–60 cm. Note that the predicted SOC density in the scatter plots are averages of the predictions from 100 runs of the 10-fold cross-validation

of 0–60 cm in the alpine meadow, subalpine shrub, and subalpine meadow was estimated to be about 77.78, 39.59, and

38.94 Tg C, respectively, which together accounted for about 71.27% of the total stock over the basin (Table 2).

Fig. 4 Relative uncertainties of SOC density at a resolution of 90 m × 90 m across the basin for different depths, estimated by the MLR (**a**–**d**) and the RF regression (**e**–**h**). The relative uncertainties are calculated as the ratio between the interquartile range (difference between the 75th and 25th percentiles) and the mean
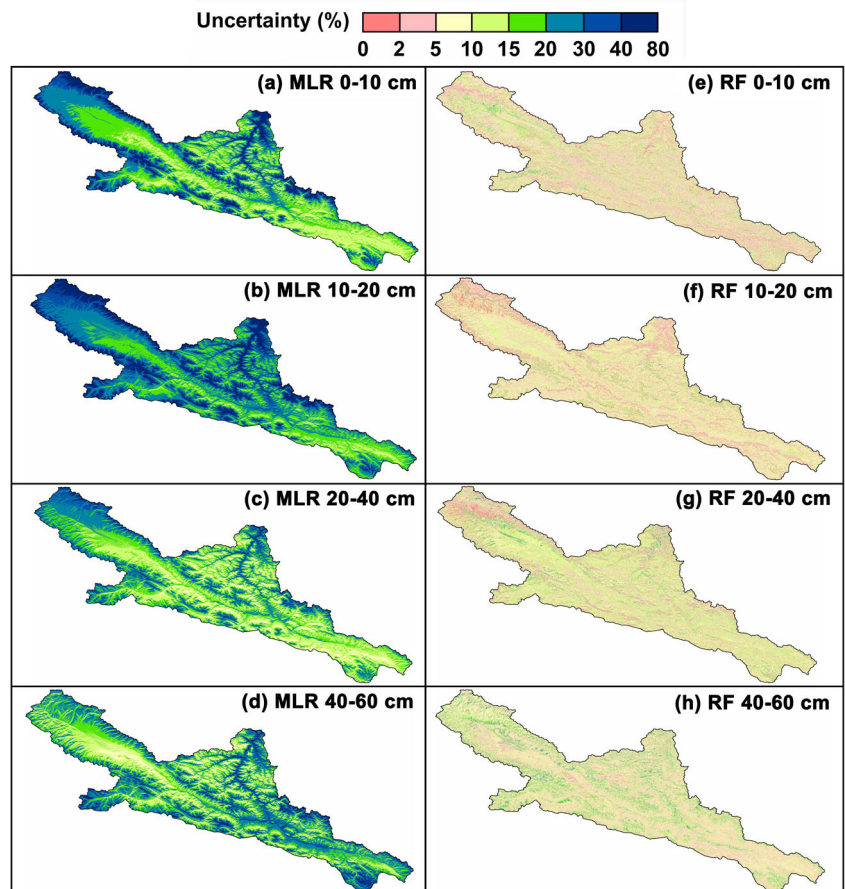
Fig. 5 SOC density (SOCD) at different depths estimated by the MLR (**a**–**d**) and the RF regression (**e**–**h**) at a resolution of 90 m × 90 m across the basin
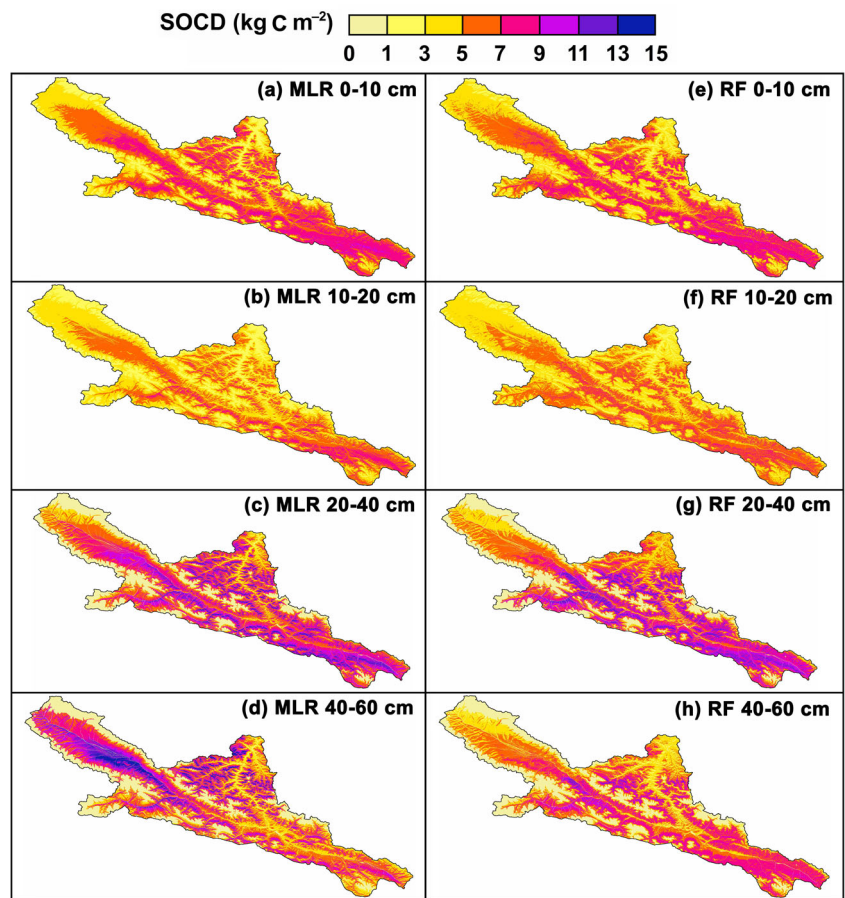
3.3 Environmental controls of SOC density

The general linear model (GLM) indicated that the topography could explain about 58.11%, 54.66%, 56.98%, and 55.32% of the total variation in SOC density at the 0–10, 10–20, 20–40, and 40–60 cm, respectively (Table 3). Out of the topographic factors examined in this study, the elevation and aspect factors were able to explain the largest two proportions (i.e., 44.39% and 11.25%, respectively) of the variation in SOC density at 0–10 cm, while the slope and the interactions between aspect and slope was able to only explain about 1.07% and 1.40%, respectively. Similar proportions

contributed by each topographic factor were also obtained at a depth of 10–20, 20–40, and 40–60 cm (Table 3).

The SOC density showed somewhat different correlations with the climatic, vegetative, and edaphic factors at the four depths (Fig. 6a–d). Specifically, the SOC density at 0–10 cm was significantly positively correlated with mean annual precipitation (correlation coefficient = 0.59), clay (0.42) and silt (0.21) content, and vegetation type (0.65), while negatively correlated with mean annual temperature (− 0.66), solar insolation (− 0.29), soil pH (− 0.39), and sand content (− 0.41) (Fig. 6a). Similar correlations were also found for deeper layers (Fig. 6b–d).

Table 2 Summary of estimated density and stock of SOC (with interquartile range) to a depth of 60 cm by the RF regression for different vegetation types across the basin

Vegetation type	Area (km^2)	SOC density (kg C m^{-2})	SOC stock (Tg C)
Montane steppe	658.8	17.95 (17.47–18.26)	11.83 (11.51–12.03)
Montane meadow steppe	726.2	24.93 (24.19–25.50)	18.10 (17.57–18.52)
Montane forest	691.2	26.77 (26.11–27.36)	18.50 (18.05–18.91)
Subalpine meadow	1367.3	28.48 (27.76–29.11)	38.94 (37.96–39.80)
Subalpine shrub	1303.7	30.37 (29.67–31.02)	39.59 (38.68–40.44)
Alpine meadow	3558.3	21.86 (21.29–22.36)	77.78 (75.76–79.56)
Alpine desert	1870.8	6.64 (6.25–6.76)	12.42 (11.69–12.65)
Total	10,321.6	21.25 (20.66–21.71)	219.33 (213.24–224.08)

Table 3 Contribution of topographic factors to the total variation of SOC density at different depths

| Source	0–10 cm			10–20 cm			20–40 cm			40–60 cm		
	Df	MS	SS (%)	df	MS	SS (%)	df	MS	SS (%)	df	MS	SS (%)
Elevation	1	60.81	44.39**	1	23.80	37.09**	1	67.72	37.97**	1	41.11	37.79**
Aspect	1	15.41	11.25**	1	10.02	15.61**	1	31.30	17.55**	1	17.96	16.51**
Slope	1	1.47	1.07	1	0.83	1.29	1	1.31	0.73	1	0.61	0.56
Aspect×slope	1	1.92	1.40*	1	0.43	0.67	1	1.30	0.73	1	0.50	0.46
Error	133	0.43	41.89	128	0.23	45.34	112	0.68	43.02	86	0.57	44.68

df, MS, and SS are degrees of freedom, mean squares, and proportion of the variance explained by variable, respectively. The results are obtained from the general linear model

*$P < 0.05$; **$P < 0.01$

The structural equation modeling (SEM) further showed that the climatic, vegetative, and edaphic factors were able to explain about 54%, 50%, 56%, and 62% of the total variation in SOC density at a depth of 0–10, 10–20, 20–40, and 40–60 cm. Specifically, the solar insolation appeared to have a direct effect on SOC density for a depth of 20–40 cm, and it also had an indirect effect on SOC density by negatively influencing vegetation type for all considered depths. Mean

annual precipitation indirectly affected SOC density through its positive correlation with vegetation type (Fig. 7a–d). Mean annual temperature exerted the strongest direct effect on SOC density, with a standardized coefficient (β) of about −0.43, −0.33, and −0.35 at a depth of 0–10, 10–20, and 20–40 cm, respectively (Fig. 7a–c). As for the vegetation type, which was strongly affected by mean annual temperature ($\beta = -0.65$), the results showed the strongest direct effect on SOC density

Fig. 6 Pearson correlation coefficients (r) for SOC density (SOCD) at 0–10 (**a**), 10–20 (**b**), 20–40 (**c**), and 40–60 cm (**d**) and other environmental controls. Correlation coefficients larger than $r_{(0.05)}$ are statistically significant at the 0.05 level. The Veg, MAP, MAT, and Rad represent vegetation type, mean annual precipitation, mean annual temperature, and solar insolation, respectively. As the vegetation type is not a numerical variable, we first rank it from 1 to 8 before correlation analysis, which sequentially represent the montane desert steppe, alpine desert, montane steppe, alpine meadow, subalpine meadow, montane forest, subalpine shrub, and montane meadow steppe. The rank is conducted according to the average NDVI value under each vegetation type, and higher code value corresponds to higher vegetation cover and potentially higher plant production
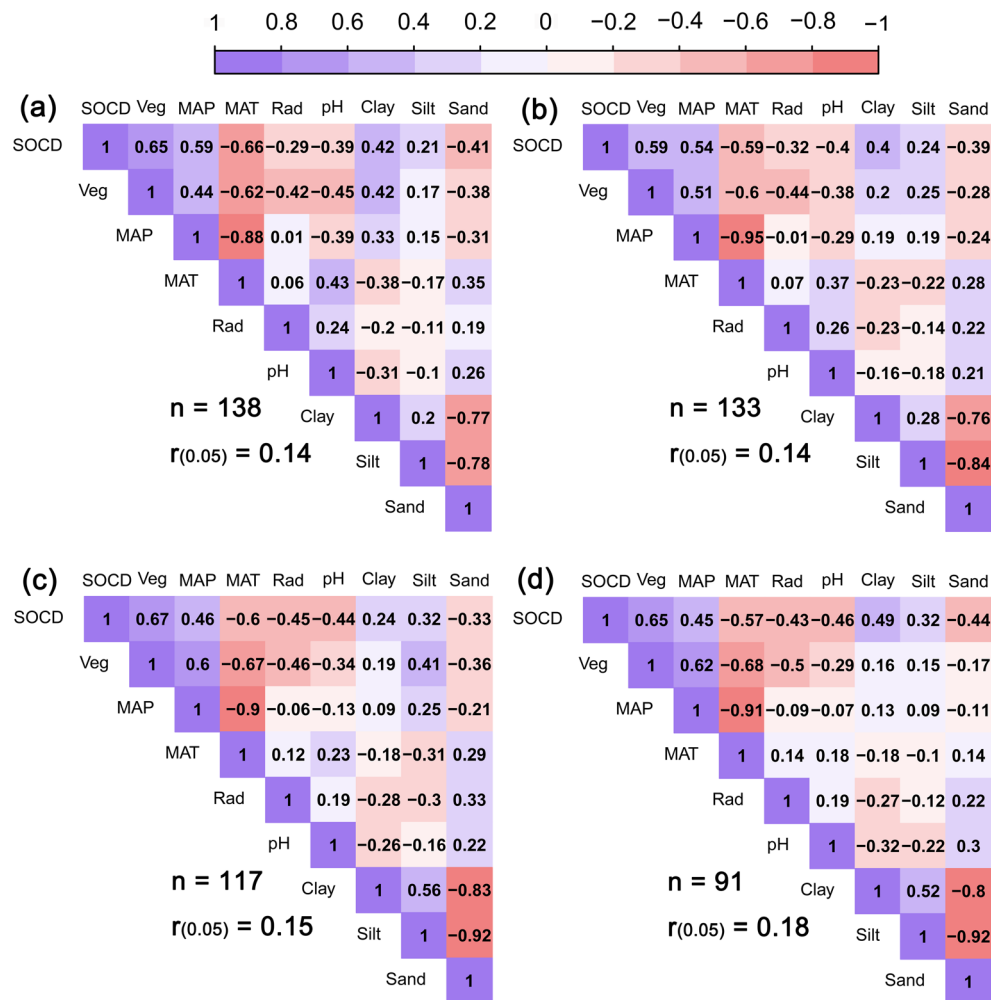
Fig. 7 Structure equation modeling (SEM) examining the climatic, vegetative, and edaphic effects on SOC density (SOCD) at 0–10 (**a**), 10–20 (**b**), 20–40 (**c**), and 40–60 cm (**d**). Single-headed arrows represent the hypothesized direction of causation. Numbers next to single-headed arrows are standardized path coefficients, which indicate the effect size of the relationship. Double-headed arrows indicate covariance between related variables. Red and purple arrows indicate negative and positive relationships, respectively, and arrow width is proportional to the strength of the relationship. The proportion of variance (r^2) explained appears below each response variables in the model. Only statistically significant ($P < 0.05$) relationships were shown. The Veg, MAP, MAT, and Rad represent vegetation type, mean annual precipitation, mean annual temperature, and solar insolation, respectively. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$
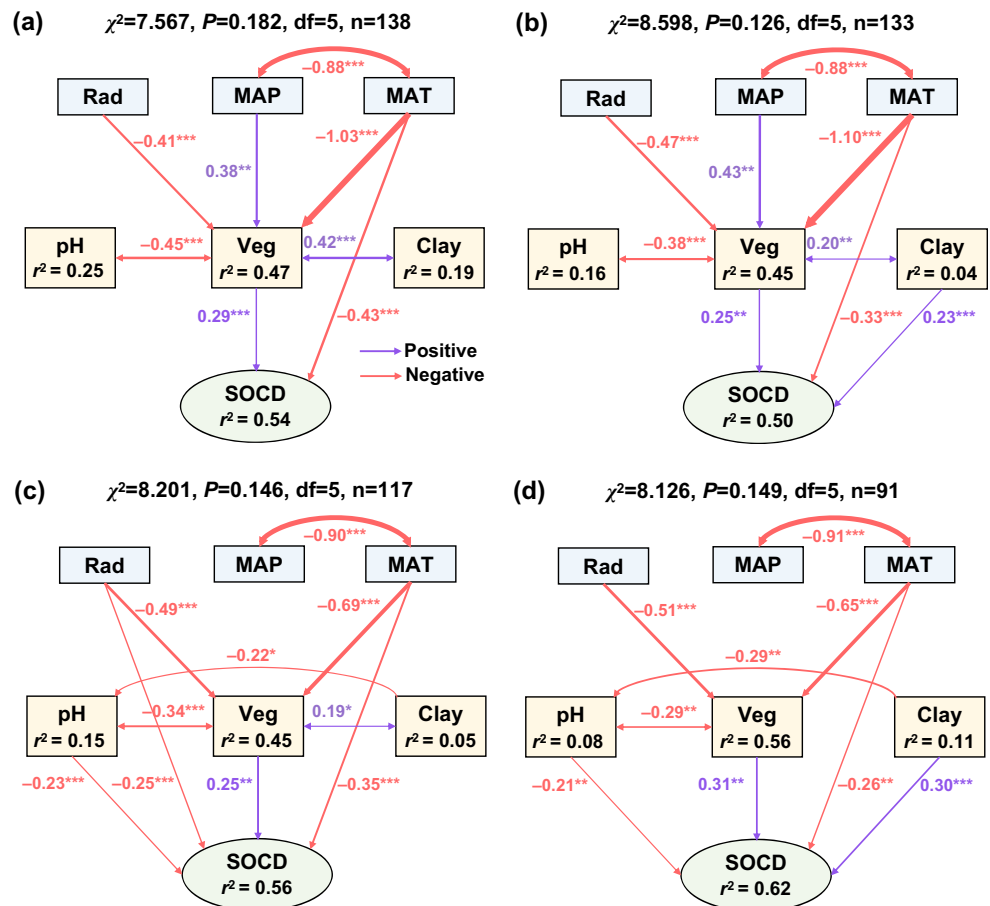
4 Discussion

4.1 Spatial distribution of SOC density

In any mountainous region, SOC tends to exhibit an increasing trend with an increase in its elevation mainly due to the lower decomposition rate at higher elevations (Parras-Alcantára et al. 2015; Zhao et al. 2017). Evidence can be drawn from the study of Prietzel and Christophel (2014) which suggested that the soils at generally high-elevation sites are likely to have a particularly large SOC density, for example, the forest soils of German Alps. However, in this study, the researchers have found that the SOC density increased with an increase in elevation below ∼ 3400 m, and then it showed a decreasing trend with elevation above that height (Fig. 2). A similar result was obtained by Chen et al. (2016a), who had found that the SOC density increased significantly with elevation from 2650 to 3400 m on the north-facing slopes. However, it showed a significantly decreasing trend

from 3400 to 3700 m. Hence, as a good indicator of SOC in alpine regions, the elevation should first be calibrated to assure its linear relationship with SOC, as suggested by Eqs. (6)–(9) in this study, so as to improve the accuracy of the linear models used to predict SOC.

The SOC density on the north-facing slopes decreased by 44% in comparison with the south-facing slopes across the basin (Table 1). Similar results were obtained by several recent studies, which showed that SOC was higher on the north-facing slopes (Lozano-García et al. 2016; Qin et al. 2016). In addition, in regions where the natural pattern of vegetation type is strongly shaped by aspect, the spatial variability of SOC may be further enhanced. Zhu et al. (2017) has highlighted the importance of aspect in markedly shaping the spatial patterns of SOC at a hillslope scale in the forest steppe zone of the Xishui Forest Reserve, northwestern China, where SOC on the north-facing slopes was larger than the south-facing slopes by a factor of 3.2.

Generally, the slope is likely to affect the variability of SOC at the hillslope scale (Zhu et al. 2018). Slopes may thus alter the distribution of SOC along the hillslope through soil erosion and decomposition processes (Zhang et al. 2013; Ma et al. 2016). In this study, we found that SOC showed a slightly decreasing trend with increasing slope. This trend was consistent with that of many previous studies (Karchegani et al.

at a depth of 40–60 cm (Fig. 7d). In addition, the clay content also had a significant direct effect on SOC density at 10–20 and 40–60 cm.

2012; Ajami et al. 2016), whereas the variation of SOC by slope in this study was more trivial (Table 1). This was probably associated with the limited soil erosion processes along hillslopes in the study area, mainly due to higher vegetation cover, fewer extreme precipitation events, and less disturbances to soils as a result of limited human activities. In addition, the effect of vegetation type may mask the slope effect on SOC, as a previous study conducted only in grasslands suggested that, among the examined topographic factors, slope positions could contribute to almost 10% of the total variation in SOC (Zhu et al. 2018).

4.2 SOC stock estimation

The 10-fold cross-validation and the uncertainty analysis suggested that the RF regression exhibited better performance (higher R^2 values and lower uncertainties) in predicting SOC density than the MLR (Figs. 3 and 4), although we have preprocessed the raw elevation and aspect data for the MLR. This could be ascribed to the two advantages of the RF regression. First, compared with linear regression models, the RF could better deal with the non-linear relationships between SOC and environmental covariates (Grimm et al. 2008; Akpa et al. 2016). Second, as a machine learning technique, the RF is not constrained by any statistical premise such as normality and independence, and also overcomes the limitations of parametric and nonparametric statistical methods, such as spatial autocorrelation (Breiman 2001; Were et al. 2015). In this study, although we used topographic factors and NDVI as predictors in the RF regression, the R^2 value obtained by the 10-fold cross-validation still reached to 0.55–0.60, which was close to that of previous studies conducted at a similar spatial scale (Wiesmeier et al. 2011; Yang et al. 2016). Taken together, our results highlighted that the RF regression combined with topographic and vegetative factors as model inputs is an effective approach for spatially explicit SOC prediction in semiarid alpine regions.

Based on the spatially gridded SOC predicted by the RF regression, we calculated the SOC stock across the basin. Results showed that the total SOC stock was estimated at 219.33 Tg C for the upper 60 cm, with an average density of 21.25 kg C m^{-2} (Table 2). Compared with regions under similar semiarid climates, we found that SOC density in the middle Qilian Mountains was much higher. For example, the mean SOC density in the semiarid Loess Plateau regions of China was estimated at 7.70 kg C m^{-2} at 0–100 cm (Liu et al. 2011). As for the semiarid Horqin grasslands, the SOC density at 0–100 cm was about 6.84 kg C m^{-2} (Li et al. 2018). Albaladejo et al. (2013) found that SOC density reached to 7.18 kg C m^{-2} in the semiarid region in Spain. Fernández-Romero et al. (2014) suggested that the maximum SOC density in the forest soils of Mediterranean natural areas reached to 15.87 kg C m^{-2}. In the mountainous regions of northern

Iran, the SOC density was estimated at about 18.58 kg C m^{-2} (Ajami et al. 2016).

The higher SOC density in this study area was largely associated with the cold climate under alpine environments and also with the poorly drained conditions in valleys, which together enhanced the accumulation of organic matter in soils. Our result confirmed the importance of alpine soils in regional carbon cycle due to the much higher carbon density in comparison with the surrounding zonal regions. In addition, we found that the amount of organic carbon stored in soils of the alpine meadow, subalpine shrub, and subalpine meadow totally accounted for 71.27% of the total stock across the basin. The large contribution of these alpine vegetation types to regional carbon pool further suggested that protection to vegetation at high-elevation zones should be given priority in terms of carbon management, due to the potentially large amount of carbon losses under global environmental change.

4.3 The topography-climate-vegetation-carbon relationships in semiarid alpine regions

In this study, the researchers have found that the climatic, vegetative, and edaphic factors were significantly correlated with SOC density at different depths (Fig. 6). These results were similar with many previous studies (Feng et al. 2002; Wynn et al. 2006; Prietzel and Christophel 2014; Song et al. 2016; Zhu et al. 2019). However, in semiarid alpine regions where human activities are limited, topography is supposed to be a fundamental factor in affecting the spatial patterns of SOC density (Chen et al. 2016a; Zhu et al. 2019). In this study, the GLM analyses showed that topography could explain about 58.11% of the total variation in SOC density for the upper 10 cm (Table 3). Additionally, this proportion was close to that obtained from the SEM analyses, which showed that the variation in climatic, vegetative, and edaphic factors could totally explain 54% of the total variation in SOC density at 0–10 cm (Fig. 7a). Generally, the elevation, aspect, and slope do not directly affect the SOC density, whereas they could directly shape the patterns of mean annual precipitation and temperature as well as solar insolation in semiarid alpine regions (Zhao et al. 2005, 2006; McCune and Keon 2002), which further alter the plant production and decomposition rate of soil organic matter, and thus contribute to the variability in SOC density (Qin et al. 2016; Zhu et al. 2018). In this study area, zonal climatic factors like mean annual precipitation and temperature largely depend on elevation (Eqs. (2) and (3)), while solar insolation on the hillslope mainly depend on aspect and slope (Eq. (4)). That is to say, topography substantially affects both the zonal climatic conditions and local microclimate in the study area.

The SEM analyses in this study suggested that the topography-induced variation in mean annual temperature exerted the largest control on SOC density, mainly by its direct

effect, and also indirect effect as delivered by the vegetation type (Fig. 7). This was mainly because that the temperature not only directly affected the decomposition rate of organic matter in soils but also lowered the plant production at very high elevations because of reduced length of the growing season and accumulative temperature. The SEM analyses also showed that although the topography-induced variation in precipitation did not directly affect SOC density, it could exert indirect effect by its positive correlation with vegetation type. This could be ascribed to the semiarid climates in our sampling basin, where water availability was supposed to be a dominant environmental control of plant production (Yang et al. 2008; Qin et al. 2016). Higher precipitation may stimulate the growth of forest and shrub species characterized by much higher plant production than the steppe and desert steppe, and subsequently enhance the litter input into soils.

Actually, soil temperature and soil moisture as affected by microclimates along the hillslope may contribute to the spatial variability of SOC density at a smaller scale (Qin et al. 2016; Zhu et al. 2019). Importantly, this variability may not be explained by the zonal climatic factors, i.e., the SOC density could still vary significantly under the same mean annual temperature and precipitation. In this study, we used the solar insolation, as calculated by aspect and slope, to reflect this variability at the local scale. The SEM analyses suggested that solar insolation exerted direct negative effect as well as indirect effect through its negative correlation with vegetation type, on SOC density at 20–40 cm. The direct effect of solar radiation was mainly associated with its control on the heat patterns characterized by higher soil temperature on the south-facing slopes in comparison with the north-facing slopes (Qin et al. 2016), which shaped the decomposition rate of organic matter in soils. The indirect effect was mainly associated with the response of vegetation type to aspect. In the middle Qilian Mountains, the aspect-induced variability in solar insolation under such semiarid climates significantly affected soil water conditions at the hillslope scale, which further sharply shape the vegetation type along the aspect gradient (Zhu et al. 2017, 2019). As a result, steppe and desert steppe mainly occupied the south-facing slopes, while shrublands and forests were mainly distributed on the north-facing slopes. These unique patterns of vegetation type, primarily resulting from the aspect-induced variability in solar insolation, further enhance the variability of SOC density within the same elevation/climatic zone.

5 Conclusions

In this study, the higher SOC density appears to highlight the importance of soils in alpine regions in terms of the balance of regional carbon cycle. The better performance of the RF regression confirmed its effectiveness in estimating SOC stock in semiarid alpine regions. The GLM and the SEM analyses suggested the fundamental roles of elevation and aspect in shaping the spatial patterns of SOC density. In addition, the elevation-induced variation in mean annual temperature was supposed to be the most important control of SOC, highlighting the potentially large amount of carbon losses as affected by the elevation-dependent warming under future climate change.

References

Ajami M, Heidari A, Khormali F, Gorji M, Ayoubi S (2016) Environmental factors controlling soil organic carbon storage in loess soils of a subhumid region, northern Iran. Geoderma 281:1–10

Akpa SIC, Odeh IOA, Bishop TFA, Hartemink AE, Amapu IY (2016) Total soil organic carbon and carbon sequestration potential in Nigeria. Geoderma 271:202–215

Albaladejo J, Ortiz R, Garcia-Franco N, Ruiz Navarro A, Almagro M, Garcia Pintado J, Martinez-Mena M (2013) Land use and climate change impacts on soil organic carbon stocks in semi-arid Spain. J Soils Sediments 13:265–277

Ali M, Deo RC, Downs NJ, Maraseni T (2018) An ensemble-ANFIS based uncertainty assessment model for forecasting multi-scalar standardized precipitation index. Atmos Res 207:155–180

Bangroo SA, Najar GR, Rasool A (2017) Effect of altitude and aspect on soil organic carbon and nitrogen stocks in the Himalayan Mawer Forest range. Catena 158:63–68

Batjes NH (1996) Total carbon and nitrogen in the soils of the world. Eur J Soil Sci 47:151–163

Breiman L (2001) Random forests. Mach Learn 45:5–32

Chen LH, Qu YG, Chen HS, Li FX (1992) Water and land resources and their rational development and utilization in the Hexi region. Science Press, Beijing (in Chinese)

Chen LF, He ZB, Du J, Yang JJ, Zhu X (2016a) Patterns and environmental controls of soil organic carbon and total nitrogen in alpine ecosystems of northwestern China. Catena 137:37–43

Chen LY, Liang JY, Qin SQ, Liu L, Fang K, Xu YP, Ding JZ, Li F, Luo YQ, Yang YH (2016b) Determinants of carbon release from the active layer and permafrost deposits on the Tibetan plateau. Nat Commun 7:13046

de Brogniez D, Ballabio C, Stevens A, Jones RJA, Montanarella L, van Wesemael B (2015) A map of the topsoil organic carbon content of Europe generated by a generalized additive model. Eur J Soil Sci 66: 121–134

Development Core Team R (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Ding JZ, Li F, Yang GB, Chen LY, Zhang BB, Liu L, Fang K, Qin SQ, Chen YL, Peng YF, Ji CJ, He HL, Smith P, Yang YH (2016) The permafrost carbon inventory on the Tibetan plateau: a new evaluation using deep sediment cores. Glob Chang Biol 22:2688–2701

Doetterl S, Stevens A, van Oost K, Quine TA, van Wesemael B (2013) Spatially-explicit regional-scale prediction of soil organic carbon stocks in cropland using environmental variables and mixed model approaches. Geoderma 204:31–42

Feng Q, Endo KN, Cheng GD (2002) Soil carbon in desertified land in relation to site characteristics. Geoderma 106:21–43

Fernández-Romero ML, Lozano-García B, Parras-Alcántara L (2014) Topography and land use change effects on the soil organic carbon stock of forest soils in Mediterranean natural areas. Agric Ecosyst Environ 195:1–9

Ghorbani MA, Deo RC, Kashani MH, Shahabi M, Ghorbani S (2019) Artificial intelligence-based fast and efficient hybrid approach for spatial modelling of soil electrical conductivity. Soil Till Res 186:152–164

Grimm R, Behrens T, Maerker M, Elsenbeer H (2008) Soil organic carbon concentrations and stocks on Barro Colorado Island-digital soil mapping using random forests analysis. Geoderma 146:102–113

Heung B, Ho HC, Zhang J, Knudby A, Bulmer CE, Schmidt MG (2016) An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. Geoderma 265:62–77

IUSS Working group WRB (2015) World Reference Base for Soil Resources 2014, update 2015 International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO. Rome. http://www.fao.org/3/i3794en/I3794en.pdf. Accessed 19 Feb 2019

Karchegani PM, Ayoubi S, Mosaddeghi MR, Honarjoo N (2012) Soil organic carbon pools in particle-size fractions as affected by slope gradient and land use change in hilly regions, western Iran. J Mt Sci 9:87–95

Kouadio L, Deo RC, Byrareddy V, Adamowski JF, Mushtaq S, Nguyen VP (2018) Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties. Comput Electron Agr 155:324–338

Li X, Liu SM, Xiao Q, Ma MG, Jin R, Che T, Wang WZ, Hu XL, Xu ZW, Wen JG, Wang LX (2017) A multiscale dataset for understanding complex eco-hydrological processes in a heterogeneous oasis system. Sci Data 4:170083

Li YQ, Wang XY, Niu YY, Lian J, Luo YQ, Chen YP, Gong XW, Yang H, Yu PD (2018) Spatial distribution of soil organic carbon in the ecologically fragile Horqin grassland of northeastern China. Geoderma 325:102–109

Lin LI (1989) A concordance correlation coefficient to evaluate reproducibility. Biometrics 45:255–268

Liu ZP, Shao MA, Wang YQ (2011) Effect of environmental factors on regional soil organic carbon stocks across the Loess Plateau region, China. Agric Ecosyst Environ 142:184–194

Lozano-García B, Parras-Alcántara L (2014) Variation in soil organic carbon and nitrogen stocks along a toposequence in a traditional Mediterranean olive grove. Land Degrad Dev 25:297–304

Lozano-García B, Parras-Alcántara L, Brevik EC (2016) Impact of topographic aspect and vegetation (native and reforested areas) on soil organic carbon and nitrogen budgets in Mediterranean natural areas. Sci Total Environ 544:963–970

Ma WM, Li ZW, Ding KY, Huang B, Nie XD, Lu YM, Xiao HB, Zeng GM (2016) Stability of soil organic carbon and potential carbon sequestration at eroding and deposition sites. J Soils Sediments 16:1705–1717

McCune B, Keon D (2002) Equations for potential annual direct incident radiation and heat load. J Veg Sci 13:603–606

Meersmans J, De Ridder F, Canters F, De Baets S, Van Molle M (2008) A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143:1–13

Melillo JM, Steudler PA, Aber JD, Newkirk K, Lux H, Bowles FP, Catricala C, Magill A, Ahrens T, Morrisseau S (2002) Soil warming and carbon-cycle feedbacks to the climate system. Science 298:2173–2176

Nelson DW, Sommers LE (1982) Total carbon, organic carbon and organic matter. In: Page AL, Miller RH, Keeney D (eds) Methods of soil analysis, part2. Chemical and microbiological properties agronomy monograph, vol 9. ASA and SSSA, Madison, pp 539–579

Parras-Alcantára L, Lozano-García B, Galán-Espejo A (2015) Soil organic carbon along an altitudinal gradient in the Despenaperros Natural Park, southern Spain. Solid Earth 6:125–134

Pepin N et al (2015) Elevation-dependent warming in mountain regions of the world. Nat Clim Chang 5:424–430

Prasad R, Deo RC, Li Y, Maraseni T (2018a) Ensemble committee-based data intelligent approach for generating soil moisture forecasts with multivariate hydro-meteorological predictors. Soil Till Res 181:63–81

Prasad R, Deo RC, Li Y, Maraseni T (2018b) Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. Geoderma 330:136–161

Prietzel J, Christophel D (2014) Organic carbon stocks in forest soils of the German Alps. Geoderma 221:28–39

Prietzel J, Zimmermann L, Schubert A, Christophel D (2016) Organic matter losses in German Alps forest soils since the 1970s most likely caused by warming. Nat Geosci 9:543–550

Qin YY, Feng Q, Holden NM, Cao JJ (2016) Variation in soil organic carbon by slope aspect in the middle of the Qilian Mountains in the upper Heihe River basin, China. Catena 147:308–314

Qin Y, Yi SH, Ding YJ, Xu GW, Chen JJ, Wang ZW (2018) Effects of small-scale patchiness of alpine grassland on ecosystem carbon and nitrogen accumulation and estimation in northeastern Qinghai-Tibetan plateau. Geoderma 318:52–63

Rial M, Martinez Cortizas A, Rodriguez-Lado L (2016) Mapping soil organic carbon content using spectroscopic and environmental data: a case study in acidic soils from NW Spain. Sci Total Environ 539:26–35

Scharlemann JPW, Tanner EVJ, Hiederer R, Kapos V (2014) Global soil carbon: understanding and managing the largest terrestrial carbon pool. Carbon Manag 5:81–91

Schuur EAG, McGuire AD, Schaedel C, Grosse G, Harden JW, Hayes DJ, Hugelius G, Koven CD, Kuhry P, Lawrence DM, Natali SM, Olefeldt D, Romanovsky VE, Schaefer K, Turetsky MR, Treat CC, Vonk JE (2015) Climate change and the permafrost carbon feedback. Nature 520:171–179

Song XD, Brus DJ, Liu F, Li DC, Zhao YG, Yang JL, Zhang GL (2016) Mapping soil organic carbon content by geographically weighted regression: a case study in the Heihe River basin, China. Geoderma 261:11–22

Stockmann U, Adams MA, Crawford JW, Field DJ, Henakaarchchi N, Jenkins M, Minasny B, McBratney AB, Courcelles VR, Singh K, Wheeler I, Abbott L, Angers DA, Baldock J, Bird M, Brookes PC, Chenu C, Jastrow JD, Lal R, Lehmann J, O'Donnell AG, Parton WJ, Whitehead D, Zimmermann M (2013) The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric Ecosyst Environ 164:80–99

Suuster E, Ritz C, Roostalu H, Kolli R, Astover A (2012) Modelling soil organic carbon concentration of mineral soils in arable land using legacy soil data. Eur J Soil Sci 63:351–359

USGS (2006) Shuttle Radar Topography Mission, 1 Arc Second scenes. In: Global Land Cover facility-University of Maryland (es), Maryland, US

Vågen TG, Winowiecki LA, Abegaz A, Hadgu KM (2013) Landsat-based approaches for mapping of land degradation prevalence and soil functional properties in Ethiopia. Remote Sens Environ 134:266–275

Were K, Bui DT, Dick OB, Singh BR (2015) 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 Indic 52:394–403

Wiesmeier M, Barthold F, Blank B, Koegel-Knabner I (2011) Digital mapping of soil organic matter stocks using random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 340:7–24

Wynn JG, Bird MI, Vellen L, Grand-Clement E, Carter J, Berry SL (2006) Continental-scale measurement of the soil organic carbon pool with climatic, edaphic, and biotic controls. Glob Biogeochem Cycles 20: 1–12

Yang YH, Fang JY, Tang YH, Ji CJ, Zheng CY, He JS, Zhu B (2008) Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. Glob Chang Biol 14:1592–1599

Yang YH, Fang JY, Ma WH, Smith P, Mohammat A, Wang SP, Wang W (2010) Soil carbon stock and its changes in northern China's grasslands from 1980s to 2000s. Glob Chang Biol 16:3036–3047

Yang YH, Li P, Ding JZ, Zhao X, Ma WH, Ji CJ, Fang JY (2014) Increased topsoil carbon stock across China's forests. Glob Chang Biol 20:2687–2696

Yang RM, Zhang GL, Liu F, Lu YY, Yang F, Yang F, Yang M, Zhao YG, Li DC (2016) Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. Ecol Indic 60:870–878

Yang LS, Feng Q, Yin ZL, Wen XH, Si JH, Li CB, Deo CR (2017) Identifying separate impacts of climate and land use/cover change on hydrological processes in upper stream of Heihe River, Northwest China. Hydrol Process 31:1100–1112

Yimer F, Ledin S, Abdelkadir A (2006) Soil organic carbon and total nitrogen stocks as affected by topographic aspect and vegetation in the Bale Mountains, Ethiopia. Geoderma 135:335–344

Yin ZL, Feng Q, Wen XH, Deo RC, Yang LS, Si JH, He ZB (2018) Design and evaluation of SVR, MARS and M5Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. Stoch Environ Res Risk Assess 32:2457–2476

Zhang X, Li ZW, Tang ZH, Zeng GM, Huang JQ, Guo W, Chen XL, Hirsh A (2013) Effects of water erosion on the redistribution of soil organic carbon in the hilly red soil region of southern China. Geomorphology 197:137–144

Zhao CY, Nan ZR, Cheng GD (2005) Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian mountains, China. Ecol Model 189:209–220

Zhao CY, Nan ZR, Cheng GD, Zhang JH, Feng ZD (2006) GIS-assisted modelling of the spatial distribution of Qinghai spruce (Picea crassifolia) in the Qilian Mountains, northwestern China based on biophysical parameters. Ecol Model 191:487–500

Zhao BH, Li ZB, Li P, Xu GC, Gao HD, Cheng YT, Chang EH, Yuan SL, Zhang Y, Feng ZH (2017) Spatial distribution of soil organic carbon and its influencing factors under the condition of ecological construction in a hilly-gully watershed of the Loess Plateau, China. Geoderma 296:10–17

Zhu M, Feng Q, Qin YY, Cao JJ, Li HY, Zhao Y (2017) Soil organic carbon as functions of slope aspects and soil depths in a semiarid alpine region of Northwest China. Catena 152:94–102

Zhu M, Feng Q, Zhang MX, Liu W, Qin YY, Deo RC, Zhang CQ (2018) Effects of topography on soil organic carbon stocks in grasslands of a semiarid alpine region, northwestern China. J Soils Sediments. https://doi.org/10.1007/s11368-018-2203-0

Zhu M, Feng Q, Qin YY, Cao JJ, Zhang MX, Liu W, Deo RC, Zhang CQ, Li RL, Li BF (2019) The role of topography in shaping the spatial patterns of soil organic carbon. Catena 176:296–305

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.