Received: 16 December 2016 | Revised: 3 May 2017 | Accepted: 30 September 2017
DOI: 10.1002/ldr.2833

WILEY

RESEARCH ARTICLE

Large-scale soil organic carbon mapping based on multivariate modelling: The case of grasslands on the Loess Plateau

Yinyin Wang[1,2] | Lei Deng[3] | Gaolin Wu[3] | Kaibo Wang[4] | Zhouping Shangguan[1,2]

[1] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Yangling, Shaanxi 712100, PR China

[2] University of Chinese Academy of Sciences, Beijing 100049, PR China

[3] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, Shaanxi 712100, PR China

[4] State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an, Shaanxi 710061, PR China

Correspondence
Zhouping Shangguan, State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, Xinong Road 26, Yangling, Shaanxi 712100, PR China.
Email: shangguan@ms.iswc.ac.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 41390463; National Key Technology R&D Program of China, Grant/Award Number: 2016YFC0501605

Abstract

The Loess Plateau is considered one of the world's regions with severe soil erosion. Grasslands are widely distributed on the Loess Plateau, accounting for approximately 40% of the total area. Soil organic carbon (SOC) plays an important role in the terrestrial carbon cycle in this region. We compiled more than 1,000 measurements of plant biomass and SOC content derived from 223 field studies of grasslands on the Loess Plateau. Combined with meteorological factors (precipitation and air temperature) and the photosynthetically active radiation factor, the topsoil SOC contents of grasslands were predicted using the random forest (RF) regression algorithm. Predicted grassland SOC content (1.70–40.34 g kg^{-1}) decreased from the southeast to the northwest of the Loess Plateau, with approximately 1/5 of the grassland exhibiting values lower than 4 g kg^{-1}. Observed SOC content was positively correlated with observed plant biomass, and for predicted values, this correlation was strong in the desert steppe and the steppe desert of rocky mountains. Air temperature was the most important factor affecting SOC contents in the RF model. Moreover, the residual error of observations and predictions increased as the grazing intensity varied from none to very severe in the temperate desert steppe, and this RF model may not perform well in plains. The use of the RF model for SOC prediction in Loess Plateau grasslands provides a reference for C storage studies in arid and semi-arid regions, and aboveground biomass and temperature should receive more attention due to increasing C sequestration.

KEYWORDS

aboveground biomass, belowground biomass, grassland, Loess Plateau, soil organic carbon

1 | INTRODUCTION

Terrestrial soil stores large amounts of carbon (C) and constitutes one of the primary atmospheric C pools (Eswaran et al., 2000; Q. Li, Yu, Li, Zhou, & Chen, 2014). Soil organic carbon (SOC) plays non-negligible roles in C cycling and mitigating global warming (Davidson & Janssens, 2006; Lal, 2004; Post & Kwon, 2000). Soil erosion is recognized as a key driver of SOC dynamics (Lal, 2005; Starr et al., 2000), and redistribution of SOC is one cause of decreased soil quality (Dominati, Mackay, & Patterson, 2010). Furthermore, soil erosion or SOC redistribution is closely related to land degradation and could be used as an indicator of land degradation (Vågen, Winowiecki, Abegaz, & Hadgu, 2013). Over the last 10 years, reliable measurement of SOC and related C pools at different spatial scales has remained a challenge (Lal, 2009). This challenge is not only due to the limitations of traditional observation methods but also due to the immaturity and uncertainty of new methods, such as the introduction of data mining techniques for estimation.

Applying data mining techniques in simulations of SOC content has become a popular methodology. Ancillary factors are indispensable in these simulations. Geomorphology affects SOC content, and Kuhn, Hoffmann, Schwanghart, and Dotterweich (2009) proposed an ecogeomorphologic approach to study the movement of organic C through landscapes. Specifically, elevation can be regarded as an ancillary factor in SOC content simulation (Cécillon et al., 2009). In addition to elevation, vegetation types and coverage also control SOC content (Ding, Yang, Song, & Zhang, 2012; Neufeldt, Resck, & Ayarza, 2002; S. W. Zhang et al., 2013). Other natural factors, such as rainfall (Beguería, Angulo-Martínez, Gaspar, & Navas, 2015; Chamizo, Rodríguez-Caballero, Román, & Cantón, 2017) and temperature (Ghosh et al., 2016; Qi et al., 2016), have varying effects on SOC content. Concerning human factors, land use has the greatest effect on the SOC pool (Gelaw, Singh, & Lal, 2014). In cropping systems, the cover cropping frequency (Brennan & Acosta-Martinez, 2017), long-term fertilization (J. Zhang, Hu, Li, Zheng, & Li, 2017), and winter tillage (G. S. Zhang & Ni, 2017) are the main drivers of SOC content dynamics. In grasslands, grazing is the main driver of the SOC pool (W. Li, Cao, et al., 2017; Stavi, Argaman, & Zaady, 2016). Beginning in 1999, the Chinese government implemented the "Grain for Green" project, aimed at improving the ecological environment in arid and semi-arid regions (L. Wang, Wei, Horton, & Shao, 2011). Under this programme, croplands have been converted to grasslands and forests as a systematic management measure, and grazing has been inhibited in closed areas. As of 2013, these vegetation restoration measures had primarily been successful, and the vegetation coverage on the Loess Plateau had doubled compared with the 1900s (Chen et al., 2015). As grassland afforestation is a main aim of the "Grain for Green" programme on the Loess Plateau (Q. Li et al., 2016; K. Zhang, Dang, Tan, Cheng, & Zhang, 2010; B. Zhang, He, Burnham, & Zhang, 2016), it is necessary to understand the present situation regarding soil quality, especially concerning the SOC contents of grasslands. However, it remains challenging to effectively estimate the SOC content of grasslands at this large scale. To solve this problem, attempts have been made to employ remote sensing (RS) images for such analyses.

RS has greatly contributed to SOC mapping (Grinand et al., 2017; Maynard & Levi, 2017). To integrate multivariate factors acquired using RS data, random forest (RF) models can be employed. RF models are a type of multidecision tree model whose advantages include the ability to model numerical and categorical factors without considering the multicollinearity among factors (Breiman, 2001; Crookston & Finley, 2008). Sreenivas et al. (2014) used an RF model to map SOC density in the states of Andhra Pradesh and Karnataka (India) at a spatial resolution of 1 km, and the model performed well. Therefore, it is feasible to employ this model to predict the SOC contents of grasslands on the Loess Plateau, combined with the use of RS images and an RF model.

The soil C sequestration capability varies in different types of grassland (Fornara & Tilman, 2008; J. H. Li, Yang, et al., 2014). Ma, He, Yu, Wen, and Peng (2016) estimated that meadows stocked more SOC than the steppes in China. Steppes and deserts stock less SOC but are widely distributed, especially on the Loess Plateau. Researchers have studied SOC storage (W. Li, Cao, et al., 2017; Zhang & Shangguan, 2016) and distribution (W. Sun, Zhu, & Guo, 2015; J. Wang, Yang, & Bai, 2015; Xin, Qin, & Yu, 2016; Zhao et al., 2016) on the Loess Plateau, but the scale of this research was limited to nature reserves or watersheds.

To integrate auxiliary factors and solve the problem mentioned above in an efficient manner, we utilized an RF model to predict the SOC content of grasslands on the Loess Plateau. This method not only exploits the potential of the RS technique but also expands the scope of the application of data mining and machine learning. In addition to SOC contents, plant biomass was also predicted on the Loess Plateau, as a contrast. We addressed the following questions: (a) How can the SOC content and plant biomass of grasslands across the Loess Plateau be predicted; (b) what is the spatial distribution pattern of grassland SOC contents; (c) which factors are correlated with grassland SOC content; and (d) what is the relationship between grassland SOC content and plant biomass.

2 | MATERIALS AND METHODS

2.1 | Study area and sampling sites

The Loess Plateau (100°52′–114°33′E, 33°41′–41°16′N) covers an area of 646,200 km^2 and is located in the middle of China. There are six types of topography on the Loess Plateau (Figure S1): plains, blown sandy hills, rocky mountains, rocky hills, plateaus and gullies, and hills and gullies. Rainfall on the Loess Plateau is affected by the Asian monsoon and decreases from southeast to northwest (X. P. Sun & Wang, 2005). Due to the loose texture of loess, extreme rainfall aggravates soil erosion in sparsely vegetated areas. Grasslands cover an area of approximately 250,000 km^2 on the Loess Plateau and play an important role in vegetation restoration. These grasslands can be partitioned into six types (Figure S2): typical temperate steppe, temperate desert steppe, temperate forest steppe, northern warm temperate tussock and shrub tussock, southern warm temperate tussock and shrub tussock, and temperate steppe desert.

We compiled more than 1,000 measurements of SOC content and plant biomass, acquired from 223 field studies (Figure 1) conducted on grasslands. Each field study was set in a plot with a width and length of 100 m, and along the diagonal of each plot, several quadrats (1 × 1 m) with intervals of 10 m were identified as representatives along a 100-m linear transect. The measured data in this study were the means for the quadrats in the corresponding plot. As topsoil is generally applied to establish the soil spectral library in RS techniques (Castaldi et al., 2016), we studied the SOC content at a soil depth of 10 cm. Both the aboveground biomass (AGB) and belowground biomass (BGB) were dried and weighed in the laboratory, and BGB was measured in the soil profiles to a depth of 1 m, as most grass roots are distributed no deeper than 1 m on the Loess Plateau.

2.2 | Factor selection and modelling approach

In this study, an RF regression model was applied to predict the SOC content, AGB, BGB, and total biomass (TB) of grasslands. Grassland TB was calculated as the sum of AGB, BGB, and litters. The RF model provides an evaluation of important variables that can be used to select essential variables among various factors.

Topography, geographical location, vegetation index, photosynthetically active radiation (FPAR), temperature, and rainfall were chosen as auxiliary factors in this study (Table 1). We downloaded RS images from the United States Geological Survey (https://www.usgs.gov/), acquired climate data from the China Meteorological Administration (http://data.cma.cn/), and interpolated the climate data from various stations into continuous raster images using the AUSPLINE method (Hutchinson, 2001). As shown in Table 1, images with three spatial resolutions (30, 100, and 500 m) were used. Considering the time

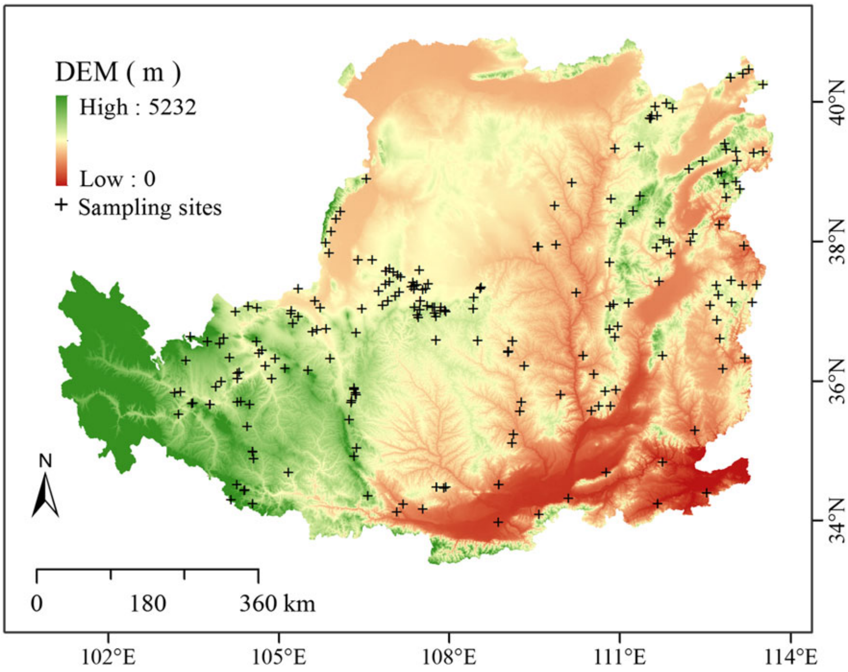

FIGURE 1 Study area and sampling sites. Digital elevation models (DEMs) represent the elevation on the Loess Plateau. DEM data were downloaded from the United States Geological Survey and are free for use by the public. Figure 1 was generated with ArcGIS 10.0 (http://www.esri.com/) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Factor descriptions and data sources used in this study

Factors definitions	Abbreviation(s)	Source data set	Resolution (m)
Topography			
Elevation Slope	Elev Slope	Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)	30
Geographic location			
Longitude Latitude	x y	Sample Data	
Vegetation indexes			
Normalized difference vegetation index (NDVI) based TM data	TNDVI	Landsat 5 Thematic	30
Summer average Leaf Area Index (LAI) of 2011/2012/2013 Average of L2011, L2012 and L2013	L2011/L2012/L2013 LM	Moderate Resolution Imaging Spectroradiometer (MODIS):	500
FPAR			
Summer average FPAR of 2011/2012/2013 Average of F2011, F2012 and F2013	F2011/F2012/F2013 FM	MODIS: MOD15A2H	500
Temperature			
Daily maximum temperature of 2011/2012/2013 Average of HT2011, HT2012 and HT2013 Daily minimum temperature of 2011/2012/2013 Average of LT2011, LT2012 and LT2013 Summer Monthly average temperature of 2011/2012/2013 Average of ST2011, ST2012 and ST2013 Monthly average temperature of 2011/2012/2013 Average of T2011, T2012 and T2013	HT2011/HT2012/HT2013 HTM LT2011/LT2012/LT2013 LTM ST2011/ST2012/ST2013 ST T2011/T2012/T2013 TM	ANUSPIN interpolation based on 64 recording stations (http://www.nmic.gov.cn)	100
Precipitation			
Monthly average precipitation of 2011/2012/2013 Average of R2011, R2012 and R2013 Summer monthly average precipitation of 2011/2012/2013 Average of SR2011, SR2012 and SR2013	R2011/R2012/R2013 RM SR2011/SR2012/SR2013 SRM	ANUSPIN interpolation based on 64 recording stations (http://www.nmic.gov.cn)	100

required for the computer processing of the RF model, we set the spatial resolution of the predicted data to 500 m.

The variables utilized to build the RF model can be ordered by their importance value. The degree of variance explained refers to the percentage of explained variance for SOC content or plant biomass in RF model processing. First, the most important variable was added to the RF model, followed by the second; the next-most important variables were then sequentially added until the 20th variable. In the stepwise addition of variables, the most recently added variable should be abandoned if the degree of variance explained decreases upon the variable's addition; otherwise, it should be retained, and the process should be continued. Using this goal-directed method, the combination of factors with the highest degree of explanation for grassland SOC content and plant biomass prediction was selected (Figure 2).

2.3 | Predicted accuracy assessment

To validate the accuracy of the predicted grassland SOC content and plant biomass, we calculated the root mean square error (RMSE) and the normalized mean square error (NMSE) for 10-fold cross-validation. The formulas for RMSE and NMSE are expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2}, \tag{1}$$

$$\mathrm{NMSE} = \frac{\sum_{n=0}^{N-1}\sum_{i=0}^{i-1}|O_i - P_i|^2}{\sum_{n=0}^{N-1}\sum_{i=0}^{i-1}|O_i|^2}, \tag{2}$$

where O_i represents the observed value and P_i represents the predicted value. In addition, Pearson correlation coefficients (R) were calculated to assess the consistency of the observed and predicted data. In this calculation, we randomly sampled 50% of observed data as the validation dataset. To ensure precision, we repeated the sampling process 7 times and calculated the mean. The formula used to calculate R is expressed as follows:

$$R = \frac{\sum_{i=1}^{N}(O_i - \overline{O})(P_i - \overline{P})}{\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}\sqrt{\sum_{i=1}^{N}(P_i - \overline{P})^2}}. \tag{3}$$

In Formula 3, O_i and P_i have the same definitions as above. R was also calculated to determine the correlation between grassland SOC content and plant biomass under different geography and vegetation types.

3 | RESULTS

3.1 | Spatial distribution of SOC content and plant biomass in grasslands

The SOC contents predicted by the RF model in grasslands on the Loess Plateau are shown in Figure 3. Areas with SOC contents lower than 4.00 g kg^{-1} accounted for 19.65% of the total grasslands (Appendix S1) and were mainly distributed in blown sandy hills (Figure S1). In general, except in the southwest, grassland SOC content decreased from the southeast to the northwest and was not evenly distributed in the east and west of the Loess Plateau.

The plant biomasses predicted by the RF model for grasslands on the Loess Plateau are shown in Figure 4. The output maps of plant biomass were expressed as grassland AGB (Figure 4a), BGB (Figure 4b), TB (Figure 4c), and root/shoot ratio maps (Figure 4d). The root/shoot ratio is negatively correlated with mean annual precipitation (Waring & Powers, 2017). Generally, grasses distributed in arid or semi-arid areas exhibit higher root/shoot ratios than those in humid areas. The spatial distribution of AGB decreased steadily from east to west (Figure 4a), and the root/shoot ratio increased from south to north (Figure 4d), which differed from what was observed for SOC content in the grasslands.

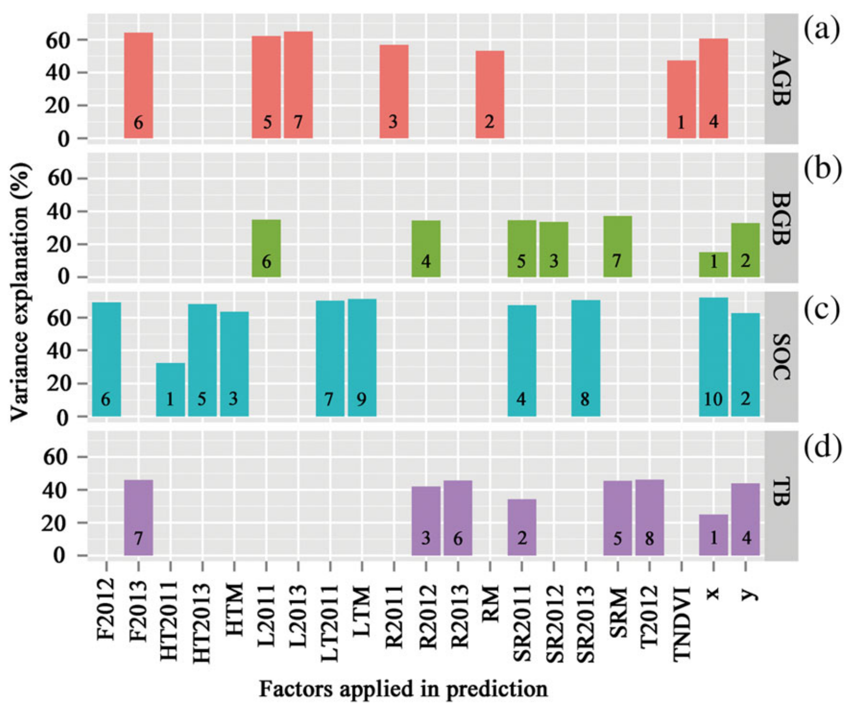

FIGURE 2 Factors applied in the prediction of grassland soil organic carbon (SOC) and biomass. For factor definitions, please refer to Table 1. The numbers on the bars refer to the sequence in which the variables were added to the random forest model. Smaller numbers correspond to higher variable importance values. The variable explanation refers to the combined degree of explanation of the added variables and not that of single variables. AGB = aboveground biomass; BGB = belowground biomass; TB = total biomass [Colour figure can be viewed at wileyonlinelibrary.com]

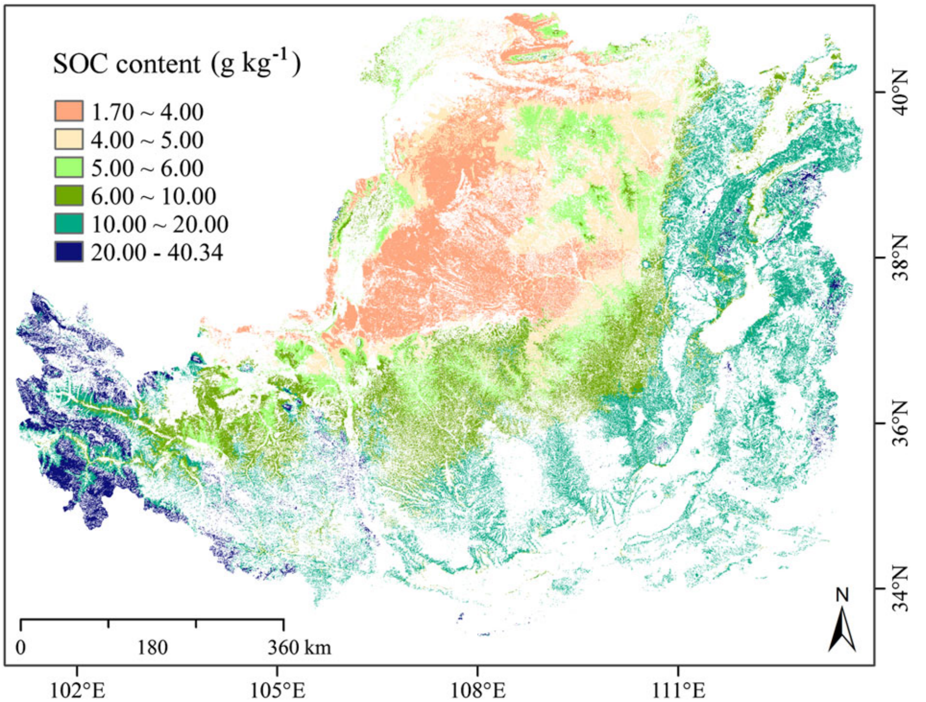

FIGURE 3 Spatial distribution of predicted grassland soil organic carbon (SOC) on the Loess Plateau [Colour figure can be viewed at wileyonlinelibrary.com]

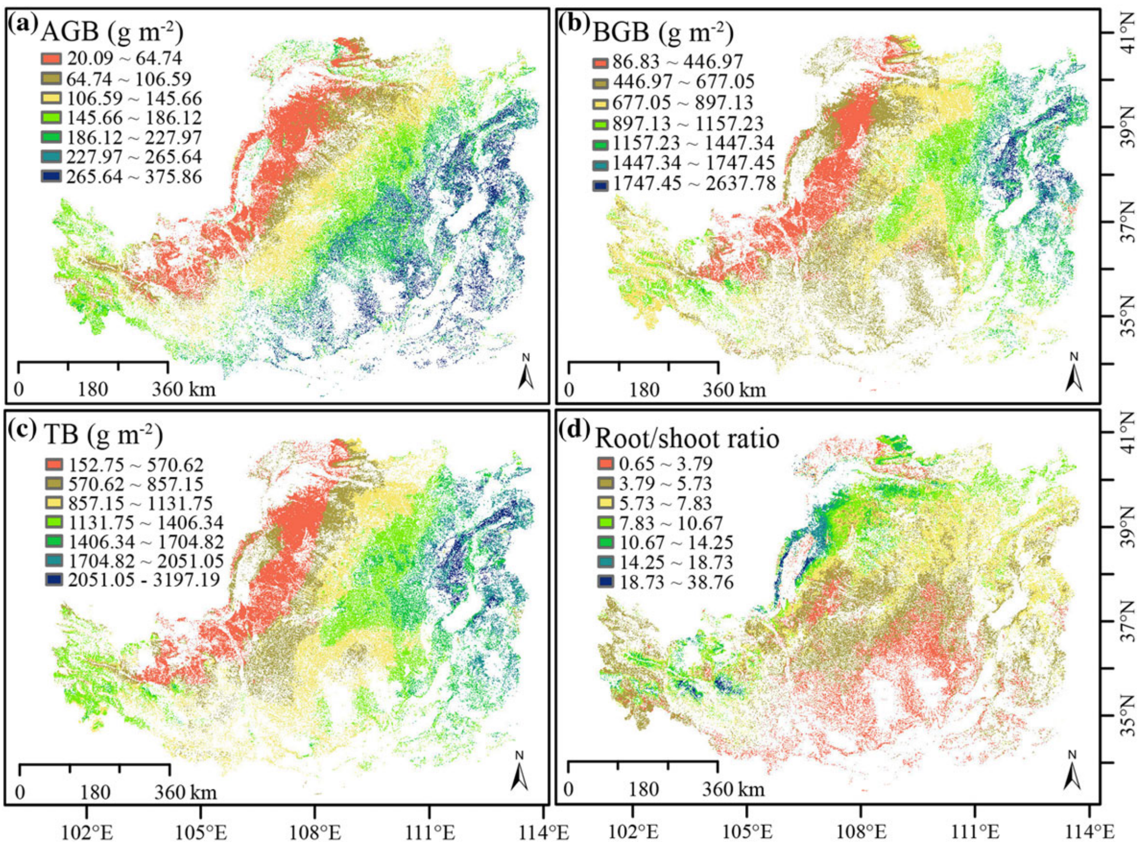

FIGURE 4 Spatial distribution of plant biomass and root/shoot ratios in grasslands on the Loess Plateau. AGB = aboveground biomass; BGB = belowground biomass; TB = total biomass [Colour figure can be viewed at wileyonlinelibrary.com]

3.2 | Regional SOC content under different topographies and grassland types

Predicted SOC contents for the grasslands (24,000 km^2) were calculated for the six topography and vegetation types (Figure 5). As shown in Figure 5a,c–f, the median SOC content differed substantially among the different topographical types. The grasslands on the plains exhibited the lowest SOC in typical temperate steppe (Figure 5d), temperate desert steppe (Figure 5e), and temperate steppe desert areas (Figure 5c). For the temperate forest steppe ecosystem, the lowest SOC contents were observed on blown sandy hills (Figure 5f). In tussock and shrub tussock ecosystems (Figure 5a,b), the highest SOC content was found on rocky mountains, and the closer to the south, the lower the sensitivity of the SOC content to the topography (Figure 5b).

3.3 | SOC contents under different grazing intensities of sampling sites

Figure 6 showed the effect of grazing intensities (none, light, moderate, severe, and very severe) on grassland SOC contents. We divided the measured SOC contents of sampling sites into two equal parts, one for training and the other for validation in RF model. In general, the residual error of observations and predictions increased with the increasing of grazing intensities in the training set, but this trend was not obviously in the validation set (Figure 6c).

4 | DISCUSSION

4.1 | Natural factors affecting the prediction of SOC contents

As shown in Figure 3, the daily maximum temperature in 2011 (HT2011) was the most important factor influencing the SOC content. According to Figure 7a, temperature (including the daily maximum and minimum temperature) was negatively correlated with the observed SOC content, with the daily minimum temperature having a weaker effect than the maximum temperature. Increased temperatures promote soil warming, and warming significantly decreases SOC contents (Fang, Smith, Moncrieff, & Smith, 2005; Qi et al., 2016). Increased levels of soil respiration and mineralized C associated with increasing temperature may be responsible for SOC consumption (Allison, Wallenstein, & Bradford, 2010; Hou, Ouyang, Maxim, Wilson, & Kuzyakov, 2016; Lefèvre et al., 2014; Wu, Dijkstra, Koch, Peñuelas, & Hungate, 2011).

Rainfall was another factor that affected grassland SOC contents. In humid tropical climate zones, rainfall facilitates SOC sequestration (Velayutham, Pal, & Bhattacharyya, 2000). Qiu, Hao, and Wu (2017) reported that in a rain-fed agro-ecosystem, SOC accumulation is enhanced by precipitation but impaired by increased temperature. However, C. Wang et al. (2017) argued that on the Loess Plateau, the trade-off of SOC and soil moisture is determined by the precipitation gradient. Moreover, the effect of rainfall may differ among different topography units (Fissore et al., 2017).

Latitude (*y*) was the second important factor affecting SOC content, and longitude (*x*) was the least important (Figure 3). These findings indicate that the SOC content varied to a greater extent across latitude than across longitude. This difference may be due to differences in topography. Furthermore, the effect of topography on SOC may be attributed to the vegetation and soil conditions associated with land use (W. Sun et al., 2015). As shown in Figure 7a, FPAR was positively correlated with the observed SOC content, because high FPAR values reflect the high photosynthetic transformation efficiency of grass.

4.2 | Relationship between predicted SOC content and plant biomass

Vegetation biomass plays a non-negligible role in determining SOC stocks (L. F. Chen et al., 2016). Vegetation recovery on the Loess Plateau favours SOC sequestration (Deng, Liu, & Shangguan, 2014;

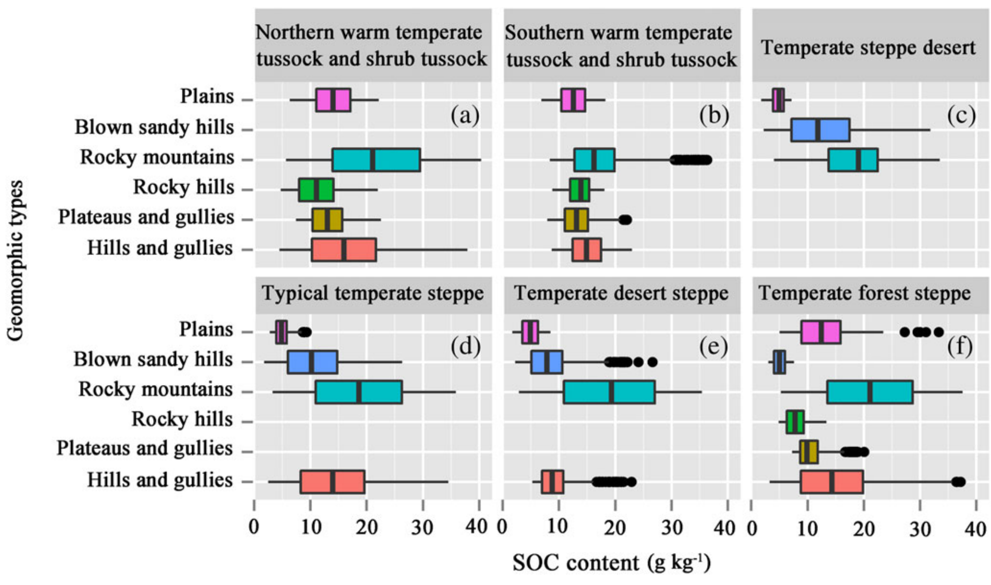

FIGURE 5 Variations in the predicted grassland soil organic carbon (SOC) for different topography and grassland types on the Loess Plateau [Colour figure can be viewed at wileyonlinelibrary.com]

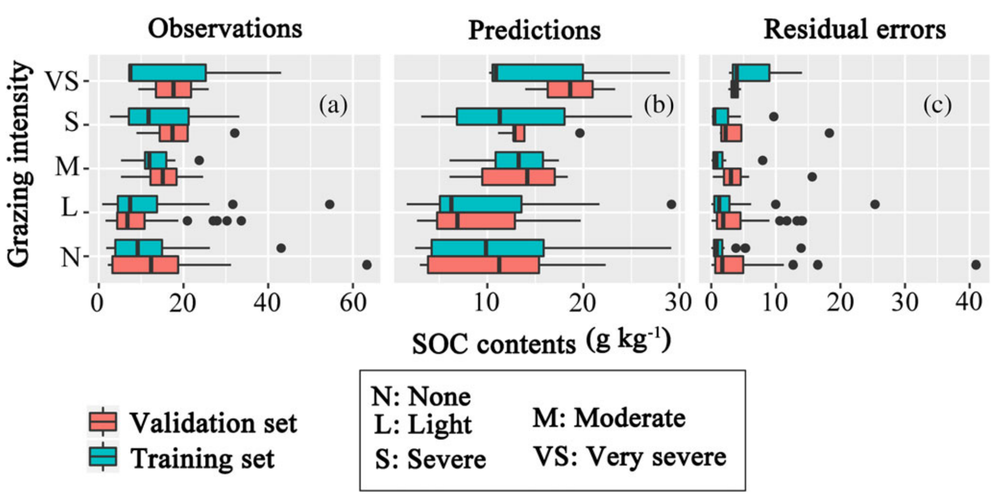

FIGURE 6 Grassland soil organic carbon (SOC) contents and residual errors of predictions and observations under five grazing intensities on the Loess Plateau. Data of 50% of the sampling sites were randomly sampled as training set; accordingly, validation set was the other half. The residual error was the absolute value of the result of subtracting the prediction from the observation [Colour figure can be viewed at wileyonlinelibrary.com]

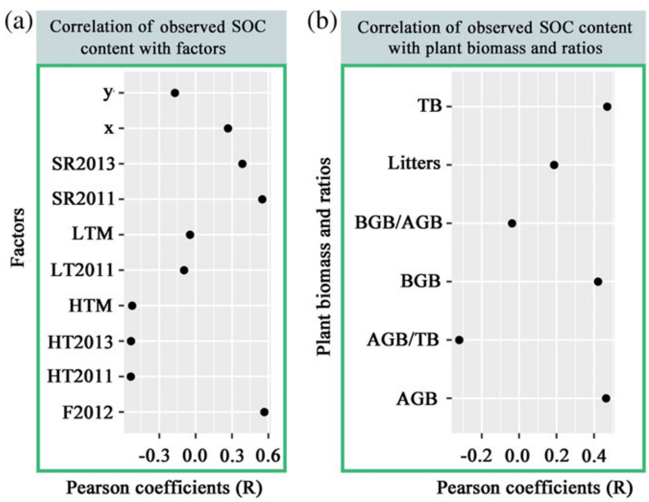

FIGURE 7 Correlations between the observed grassland soil organic carbon (SOC) and factors/biomass on the Loess Plateau. For factor definitions, please refer to Table 1. AGB = aboveground biomass; BGB = belowground biomass; TB = total biomass [Colour figure can be viewed at wileyonlinelibrary.com]

Deng & Shangguan, 2017; Deng, Shangguan, Wu, & Chang, 2017; Liu, Dang, Tian, Wang, & Wu, 2016). Pearson correlation coefficients and associated significance levels were utilized to identify the relationship between grassland SOC and biomass. These coefficients were calculated at two levels: the point level (observations) and the regional level (predictions). The correlation between the observed grassland SOC content and plant biomass was first calculated for the sampling points (Figure 7b) and then for different combinations of topography and grassland types (Figure 8).

The transfer of organic matter from plants to soil can increase SOC stocks (Cardinael et al., 2017). As shown in Figure 7b, the observed grassland SOC content was positively correlated with plant biomass ($R < .500$). J. Sun and Wang (2016) showed that the trade-off between AGB and BGB increased with an increasing SOC gradient across an alpine grassland; however, in this study, the observed grassland SOC was not significantly correlated with the root/shoot ratio ($R = -.038$) on the Loess Plateau. As we studied SOC contents across six grassland ecosystems, this trade-off may manifest in a single ecosystem but may be blurred among multi-ecosystems.

When considering the correlation between the predicted SOC content and plant biomass in grasslands (24,000 km^2), the situation differed among the different topography and grassland types (Figure 8). These changes were attributable to the significant differences between the sites and regions. The observed data reflected the true values at the sampling sites. In contrast, the predicted data were approximately fitted to the observed data and reflected the regional averages of SOC content and plant biomass.

In rocky mountains, grassland AGB exhibited a strong positive correlation with SOC content in both temperate desert steppe (.912**) and temperate steppe desert (.839**) ecosystems (Figure 8a). Xin et al. (2016) investigated 180 topsoil sites (0–20 cm) in the Luoyugou watershed of the Loess Plateau and found that the SOC content of the mountain ridge area (a high-elevation area) was higher than on the sides of the valley. This observation suggests that the high grassland SOC in rocky mountains is mainly transformed from AGB. Zhang and Shangguan (2016) studied the soil water content (SWC) in the Ziwuling forest region and found that grasslands exhibited a higher SWC than shrublands and forestlands. In addition, Liu et al. (2016) studied the Qiaozigou watershed and found that the SOC content in sediments was positively correlated with SWC ($p < .05$). A high grassland SWC might favour SOC accumulation in mountain areas. In the steppe of rocky mountains, grassland BGB (Figure 8c) and root/shoot ratios (Figure 8d) were negatively correlated with the SOC content. Lai et al. (2017) studied the contribution of fine root dynamics to SOC on a degraded steppe in China. They found that herbaceous fine roots in shrublands exhibited the fastest turnover and decomposition rates, followed by herbaceous plants in steppe areas. However, fine roots were scarce in the surface layer (Imada, Matsuo, Acharya, & Yamanaka, 2015), which may be the reason that we did not identify a contribution of fine roots to SOC accumulation, although this has been observed in other studies.

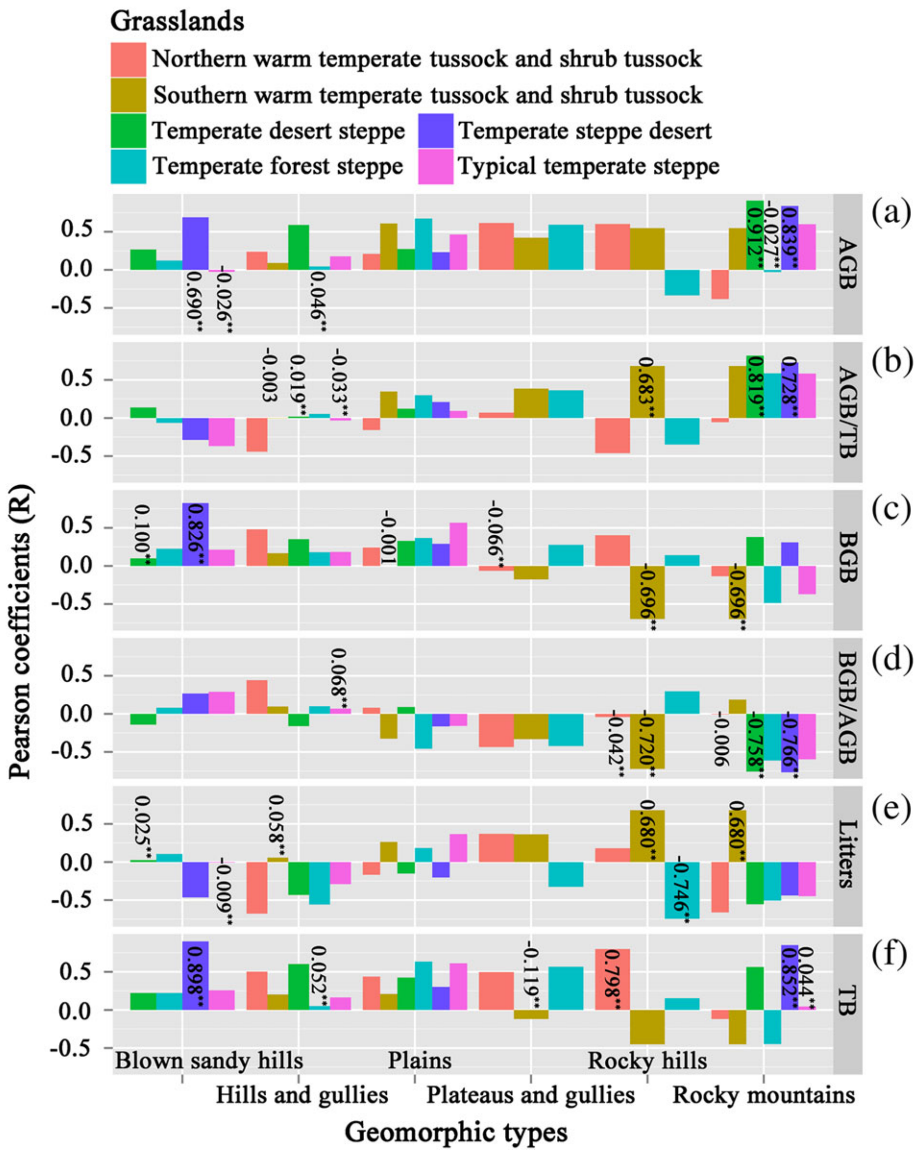

FIGURE 8 Correlation of predicted grassland soil organic carbon and biomass on the Loess Plateau (**significance is defined at .01). AGB = aboveground biomass; BGB = belowground biomass; TB = total biomass [Colour figure can be viewed at wileyonlinelibrary.com]

On blown sandy hills, the effect of biomass on SOC was significant (AGB: .690**; BGB: .826**) in temperate steppe desert. This result suggests that revegetation could stop the process of desertification (Y. L. Chen et al., 2017). In Chen's study (2017), herbaceous cover mainly influenced fine root length in the 0- to 40-cm soil layer, whereas shrub cover affected fine root length at depths of 0–300 cm. These differences might explain why BGB was positively correlated with the SOC content in this area but negatively correlated with the SOC content in shrub tussocks on rocky mountains (Figure 8c). In other types of steppe, plant biomass had little effect on SOC content (Figure 8); because the soil was more sensitive to desertification than to plant biomass (Tang, An, & Shangguan, 2015), the variation of SOC was not in accordance with plant succession.

On rocky hills, litter was significantly negatively correlated with the grassland SOC content in temperate forest steppe areas (−.746**; Figure 8e). In the southern warm temperate tussock and shrub tussock ecosystems, litter was positively correlated (.680**; Figure 8e) with the grassland SOC content, reflecting the effect of ecosystem deviation on grassland SOC contents. An increase in litter promoted SOC accumulation in the shrub tussock ecosystem but had the opposite effect in the steppe ecosystem. This finding might have occurred because the rate of C transfer is largely controlled by the quality of litter in different ecosystems (Walela et al., 2014).

4.3 | Assessment of accuracy

RMSE, NMSE, and R were utilized to assess the accuracy of the predicted grassland SOC and biomass (Table 2). The RMSE and NMSE values shown in Table 2 are the mean values obtained from 10-fold cross-validation. We randomly sampled 50% of the observed data 7 times to calculate the R values of the observed and predicted data. The R value for the predicted and observed grassland SOC content was highest (.72**), and the RMSE of the predicted and observed SOC content was 3.99 g kg^{-1}. The predicted TB exhibited a relatively

TABLE 2 Predicted accuracy of grassland SOC and biomass on the Loess Plateau

	R (training set = 50%, prediction set = 50%, 7-fold cross-validation)	RMSE (mean RMSE of 10-fold cross-validation)	NMSE (mean NMSE of 10-fold cross-validation)
AGB (g/m^2)	0.62**	89.37	0.64
BGB (g/m^2)	0.53**	279.72	0.23
TB (g/m^2)	0.54**	271.31	0.17
SOC (g kg^{-1})	0.72**	3.99	0.21

Note. RMSE referred to as root mean square error of observed and predicted SOC/biomass. NMSE referred to as normalized mean square error of observed and predicted SOC/biomass. AGB = aboveground biomass; BGB = belowground biomass; TB = total biomass; SOC = soil organic carbon.

**Significance is defined at .01.

weak correlation with the observed TB (.54**) but presented the lowest NMSE value. Gomez, Viscarra Rossel, and McBratney (2008) predicted pasture SOC contents using RS data in Australia and obtained R^2 values of .42–.43. Castaldi et al. (2016) estimated SOC contents based on multispectral and hyperspectral images, and the obtained R^2 values for the predicted and observed SOC content ranged from .49 to .67. Vaudour, Gilliot, Bel, Lefevre, and Chehdi (2016) predicted cropland SOC contents using visible near-infrared airborne hyperspectral images of the Versailles Plain and Alluets Plateau areas of France. These authors obtained R^2 values of 0.29–0.44 and RMSE values of 4.04–4.05 g kg^{-1}. Relative to other studies, the scale of the present study, and taking into account the complex geographies of the Loess Plateau, the accuracy of our predicted results was acceptable.

4.4 | Grazing intensities affecting prediction accuracy of grassland SOC contents

The residual error of the observed and predicted SOC contents varied under different grazing intensities (Figure 9). Residual error was high in the temperate forest steppe if there was no grazing, but when the grazing intensity was light, the residual error was higher in the south warm temperate tussock and shrub tussock. Moreover, the highest residual error was located in the temperate desert steppe with moderate grazing, and typically, the residual error increased with the increasing of grazing intensity (Figure 9a). In Figure 9b, the residual error was high in plains with severe and very severe grazing, followed by the error located in blown sandy hills with severe grazing. It implies that this model may not perform well in plains or in desert steppe, and the bigger the grazing intensity, the higher the uncertainty of the prediction of grassland SOC contents. It also proved that grazing is the main driver of the SOC pool (Z. Li, Liu, et al., 2017), and the prediction accuracy may be improved if this factor is put into the RF model.

4.5 | Challenges in assessing land degradation at a large scale

At present, there is a lack of large-scale predictions (Maillard, McConkey, & Angers, 2017) Although the application of RS and multivariate methods for large-scale predictions has improved data utilization efficiency, large-scale predictions display a lower predictive

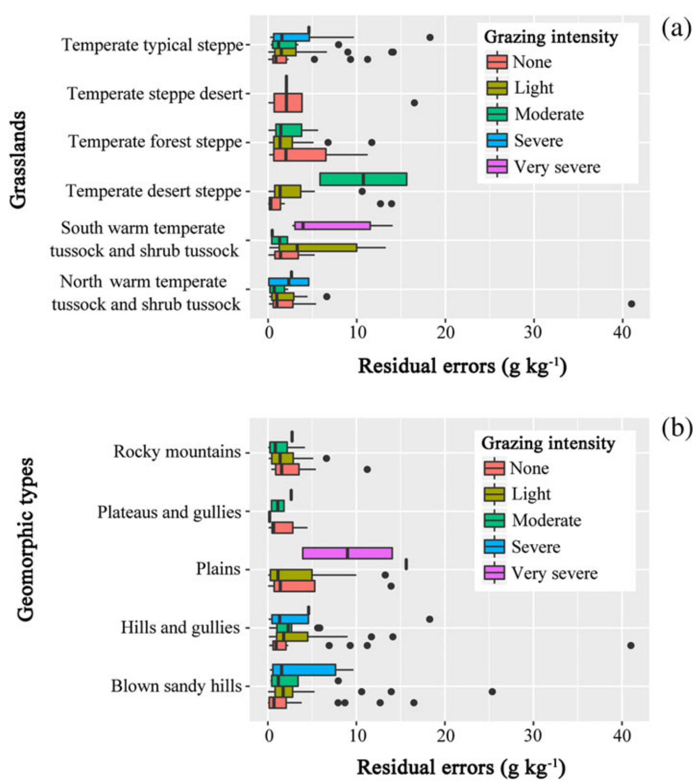

FIGURE 9 The distribution of residual errors in different geomorphic and grassland types under five grazing intensities. The residual error was the absolute value of the result of subtracting the prediction from the observation (referred to Figure 6) [Colour figure can be viewed at wileyonlinelibrary.com]

accuracy than small-scale predictions (Croft, Kuhn, & Anderson, 2012; de Araujo Barbosa, Atkinson, & Dearing, 2015; Grinand et al., 2017). It is not easy to map SOC contents at a very high resolution (<1 m) at a large scale for either data acquisition or model robustness. As SOC content can be regarded as an indicator of land degradation (Vågen et al., 2013), in combination with plant biomass and other soil physical or chemical properties, it could be feasible to assess the risk of land degradation at a large scale. Nevertheless, a great deal of work needs to be done to improve the prediction accuracy, especially in regions with severe land degradation.

5 | CONCLUSIONS

In this study, using the RF model and considering topography, geographic location, vegetation indexes (normalized difference vegetation index and leaf area index), precipitation, temperature, and FPAR, we predicted grassland SOC content and plant biomass on the Loess Plateau, providing a preliminary exploration of SOC content prediction at a large scale.

The trends in the spatial variation of grassland SOC content and plant biomass were not consistent across the Loess Plateau. The root/shoot ratio generally decreased from the south to the north. Overall, the SOC content was highest on rocky mountains. In most cases, increased grassland AGB promoted SOC sequestration. BGB slightly or negatively affected topsoil SOC content in shrub tussock ecosystems but positively affected SOC content in steppe desert areas on blown sandy hills. Litter increased the SOC content in shrub tussocks but reduced the SOC content in the steppe due to different C transfer rates. On rocky mountains, the relationship between grassland SOC content and plant biomass was stronger than it was in other topography types. Grazing increased the uncertainty of the prediction of SOC contents in grasslands.

In the future, we intend to decrease the error propagation in such predictions. In addition, we may explore C density and C stocks in grassland ecosystems. Based on the predicted SOC content and plant biomass maps, by standardizing the level of soil quality, it would be possible to identify the potential area and extent of SOC loss on the Loess Plateau and optimize grassland management strategies. Combined with multiple factors related to land degradation, the results of this study could be further used to assess land degradation risk at large scales.

CONFLICT OF INTEREST

All the authors declare no conflicts of interest.

ORCID

Yinyin Wang http://orcid.org/0000-0002-2804-7252
Gaolin Wu http://orcid.org/0000-0002-5449-7134

REFERENCES

Allison, S. D., Wallenstein, M. D., & Bradford, M. A. (2010). Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, *3*, 336–340. https://doi.org/10.1038/ngeo846

Beguería, S., Angulo-Martínez, M., Gaspar, L., & Navas, A. (2015). Detachment of soil organic carbon by rainfall splash: Experimental assessment on three agricultural soils of Spain. *Geoderma*, *245–246*, 21–30. https://doi.org/10.1016/j.geoderma.2015.01.010

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Brennan, E. B., & Acosta-Martinez, V. (2017). Cover cropping frequency is the main driver of soil microbial changes during six years of organic vegetable production. *Soil Biology and Biochemistry*, *109*, 188–204. https://doi.org/10.1016/j.soilbio.2017.01.014

Cardinael, R., Chevallier, T., Cambou, A., Béral, C., Barthès, B. G., Dupraz, C., … Chenu, C. (2017). Increased soil organic carbon stocks under agroforestry: A survey of six different sites in France. *Agriculture, Ecosystems & Environment*, *236*, 243–255. https://doi.org/10.1016/j.agee.2016.12.011

Castaldi, F., Palombo, A., Santini, F., Pascucci, S., Pignatti, S., & Casa, R. (2016). Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sensing of Environment*, *179*, 54–65. https://doi.org/10.1016/j.rse.2016.03.025

Cécillon, L., Barthès, B. G., Gomez, C., Ertlen, D., Genot, V., Hedde, M., … Brun, J. J. (2009). Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). *European Journal of Soil Science*, *60*, 770–784. https://doi.org/10.1111/j.1365-2389.2009.01178.x

Chamizo, S., Rodríguez-Caballero, E., Román, J. R., & Cantón, Y. (2017). Effects of biocrust on soil erosion and organic carbon losses under natural rainfall. *Catena*, *148*, 117–125. https://doi.org/10.1016/j.catena.2016.06.017

Chen, L. F., He, Z. B., Zhu, X., Du, J., Yang, J., & Li, J. (2016). Impacts of afforestation on plant diversity, soil properties, and soil organic carbon storage in a semi-arid grassland of northwestern China. *Catena*, *147*, 300–307. https://doi.org/10.1016/j.catena.2016.07.009

Chen, Y., Wang, K., Lin, Y., Shi, W., Song, Y., & He, X. (2015). Balancing green and grain trade. *Nature Geoscience*, *8*, 739–741. https://doi.org/10.1038/ngeo2544

Chen, Y. L., Zhang, Z. S., Huang, L., Zhao, Y., Hu, Y., Zhang, P., … Zhang, H. (2017). Co-variation of fine-root distribution with vegetation and soil properties along a revegetation chronosequence in a desert area in northwestern China. *Catena*, *151*, 16–25. https://doi.org/10.1016/j.catena.2016.12.004

Croft, H., Kuhn, N. J., & Anderson, K. (2012). On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. *Catena*, *94*, 64–74. https://doi.org/10.1016/j.catena.2012.01.001

Crookston, N. L., & Finley, A. O. (2008). YaImpute: An R package for kNN imputation. *Journal of Statistical Software*, *23*, 1–16. https://doi.org/10.18637/jss.v023.i10

Davidson, E. A., & Janssens, I. A. (2006). Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature*, *440*, 165–173. https://doi.org/10.1038/nature04514

de Araujo Barbosa, C. C., Atkinson, P. M., & Dearing, J. A. (2015). Remote sensing of ecosystem services: A systematic review. *Ecological Indicators*, *52*, 430–443. https://doi.org/10.1016/j.ecolind.2015.01.007

Deng, L., Liu, G. B., & Shangguan, Z. P. (2014). Land-use conversion and changing soil carbon stocks in China's 'Grain-for-Green' program: A synthesis. *Global Change Biology*, *20*, 3544–3556. https://doi.org/10.1111/gcb.12508

Deng, L., & Shangguan, Z. P. (2017). Afforestation drives soil carbon and nitrogen changes in China. *Land Degradation & Development*, *28*, 151–165. https://doi.org/10.1002/ldr.2537

Deng, L., Shangguan, Z. P., Wu, G. L., & Chang, X. F. (2017). Effects of grazing exclusion on carbon sequestration in China's grassland. *Earth-Science Reviews*, *173*, 84–95. https://doi.org/10.1016/j.earscirev.2017.08.008

Ding, Y. K., Yang, J., Song, B. Y., & Zhang, L. (2012). Effect of different vegetation types on soil organic carbon in Mu Us Desert. *Acta Prataculturae Sinica*, *21*, 18–25. http://cyxb.lzu.edu.cn/EN/article/searchArticle.do

Dominati, E., Mackay, A., & Patterson, M. (2010). Modelling the provision of ecosystem services from soil natural capital. In *Proceedings of the 19th World Congress of Soil Science: Soil solutions for a changing world* (pp. 32–35). Brisbane, Australia, 1–6 August 2010. Congress Symposium 2: Soil Ecosystem Services. http://www.iuss.org/19th%20WCSS/Symposium/pdf/1841.pdf

Eswaran, H., Reich, P. F., Kimble, J. M., Beinroth, F. H., Padmanabhan, E., & Moncharoen, P. (2000). Global carbon stocks. In R. Lal, J. M. Kimble, H. Eswaran, & B. A. Stewart (Eds.), *Global change and pedogenic carbonate* (pp. 15–25). Boca Raton, FL: CRC Press.

Fang, C., Smith, P., Moncrieff, J. B., & Smith, J. U. (2005). Similar response of labile and resistant soil organic matter pools to changes in temperature. *Nature*, *433*, 57–59. https://doi.org/10.1038/nature03138

Fissore, C., Dalzell, B. J., Berhe, A. A., Voegtle, M., Evans, M., & Wu, A. (2017). Influence of topography on soil organic carbon dynamics in a Southern California grassland. *Catena*, *149*, 140–149. https://doi.org/10.1016/j.catena.2016.09.016

Fornara, D. A., & Tilman, D. (2008). Plant functional composition influences rates of soil carbon and nitrogen accumulation. *Journal of Ecology*, *96*, 314–322. https://doi.org/10.1111/j.1365-2745.2007.01345.x

Gelaw, A. M., Singh, B. R., & Lal, R. (2014). Soil organic carbon and total nitrogen stocks under different land uses in a semi-arid watershed in Tigray, Northern Ethiopia. *Agriculture, Ecosystems & Environment*, *188*, 256–263. https://doi.org/10.1016/j.agee.2014.02.035

Ghosh, A., Bhattacharyya, R., Dwivedi, B. S., Meena, M. C., Agarwal, B. K., Mahapatra, P., … Agnihorti, R. (2016). Temperature sensitivity of soil organic carbon decomposition as affected by long-term fertilization under a soybean based cropping system in a sub-tropical alfisol. *Agriculture, Ecosystems & Environment*, *233*, 202–213. https://doi.org/10.1016/j.agee.2016.09.010

Gomez, C., Viscarra Rossel, R. A., & McBratney, A. B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, *146*, 403–411. https://doi.org/10.1016/j.geoderma.2008.06.011

Grinand, C., Le Maire, G. L., Vieilledent, G., Razakamanarivo, H., Razafimbelo, T., & Bernoux, M. (2017). Estimating temporal changes in soil carbon stocks at ecoregional scale in Madagascar using remote-sensing. *International Journal of Applied Earth Observation and Geoinformation*, *54*, 1–14. https://doi.org/10.1016/j.jag.2016.09.002

Hou, R. X., Ouyang, Z., Maxim, D., Wilson, G., & Kuzyakov, Y. (2016). Lasting effect of soil warming on organic matter decomposition depends on tillage practices. *Soil Biology and Biochemistry*, *95*, 243–249. https://doi.org/10.1016/j.soilbio.2015.12.008

Hutchinson, M. F. (2001). *Anusplin version 4.2 userguide*. Cauberra: Australian National University.

Imada, S., Matsuo, N., Acharya, K., & Yamanaka, N. (2015). Effects of salinity on fine root distribution and whole plant biomass of *Tamarix ramosissima* cuttings. *Journal of Arid Environments.*, *114*, 84–90. https://doi.org/10.1016/j.jaridenv.2014.11.011

Kuhn, N. J., Hoffmann, T., Schwanghart, W., & Dotterweich, M. (2009). Agricultural soil erosion and global carbon cycle: Controversy over? *Earth Surface Processes and Landforms*, *34*, 1033–1038. https://doi.org/10.1002/esp.1796

Lai, Z., Liu, J., Zhang, Y., Wu, B., Qin, S., Sun, Y., … Bai, Y. (2017). Introducing a shrub species in a degraded steppe shifts fine root dynamics and soil organic carbon accumulations, in northwest China. *Ecological Engineering*, *100*, 277–285. https://doi.org/10.1016/j.ecoleng.2017.01.001

Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma*, *123*, 1–22. https://doi.org/10.1016/j.geoderma.2004.01.032

Lal, R. (2005). Soil erosion and carbon dynamics. *Soil & Tillage Research*, *81*, 137–142. https://doi.org/10.1016/j.still.2004.09.002

Lal, R. (2009). Challenges and opportunities in soil organic matter research. *European Journal of Soil Science*, *60*, 158–169. https://doi.org/10.1111/j.1365-2389.2008.01114.x

Lefèvre, R., Barré, P., Moyano, F. E., Christensen, B. T., Bardoux, G., Eglin, T., … Chenu, C. (2014). Higher temperature sensitivity for stable than for labile soil organic carbon—Evidence from incubations of long-term bare fallow soils. *Global Change Biology*, *20*, 633–640. https://doi.org/10.1111/gcb.12402

Li, J. H., Yang, Y. J., Li, B. W., Li, W. J., Wang, G., & Knops, J. M. H. (2014). Effects of nitrogen and phosphorus fertilization on soil carbon fractions in alpine meadows on the Qinghai-Tibetan Plateau. *PLoS One*, *9*, e103266. https://doi.org/10.1371/journal.pone.0103266

Li, Q., Chen, D., Zhao, L., Yang, X., Xu, S., & Zhao, X. (2016). More than a century of Grain for Green program is expected to restore soil carbon stock on alpine grassland revealed by field 13C pulse labeling. *Science of the Total Environment*, *550*, 17–26. https://doi.org/10.1016/j.scitotenv.2016.01.060

Li, Q., Yu, P., Li, G., Zhou, D., & Chen, X. (2014). Overlooking soil erosion induces underestimation of the soil closs in degraded land. *Quaternary International*, *349*, 287–290. https://doi.org/10.1016/j.quaint.2014.05.034

Li, W., Cao, W., Wang, J., Li, X., Xu, C., & Shi, S. (2017). Effects of grazing regime on vegetation structure, productivity, soil quality, carbon and nitrogen storage of alpine meadow on the Qinghai-Tibetan Plateau. *Ecological Engineering.*, *98*, 123–133. https://doi.org/10.1016/j.ecoleng.2016.10.026

Li, Z., Liu, C., Dong, Y., Chang, X., Nie, X., Liu, L., … Zeng, G. (2017). Response of soil organic carbon and nitrogen stocks to soil erosion and land use types in the Loess hilly–gully region of China. *Soil and Tillage Research*, *166*, 1–9. https://doi.org/10.1016/j.still.2016.10.004

Liu, Y., Dang, Z. Q., Tian, F. P., Wang, D., & Wu, G. L. (2016). Soil organic carbon and inorganic carbon accumulation along a 30-year grassland restoration chronosequence in semi-arid regions (China). *Land Degradation & Development*, *28*, 189–198. https://doi.org/10.1002/ldr.2632

Ma, A., He, N., Yu, G., Wen, D., & Peng, S. (2016). Carbon storage in Chinese grassland ecosystems: Influence of different integrative methods. *Scientific Reports*, *6*, 21378. https://doi.org/10.1038/srep21378

Maillard, É., McConkey, B. G., & Angers, D. A. (2017). Increased uncertainty in soil carbon stock measurement with spatial scale and sampling profile depth in world grasslands: A systematic analysis. *Agriculture, Ecosystems & Environment*, *236*, 268–276. https://doi.org/10.1016/j.agee.2016.11.024

Maynard, J. J., & Levi, M. R. (2017). Hyper-temporal remote sensing for digital soil mapping: Characterizing soil-vegetation response to climatic variability. *Geoderma*, *285*, 94–109. https://doi.org/10.1016/j.geoderma.2016.09.024

Neufeldt, H., Resck, D. V. S., & Ayarza, M. A. (2002). Texture and land-use effects on soil organic matter in Cerrado Oxisols, Central Brazil. *Geoderma*, *107*, 151–164. https://doi.org/10.1016/S0016-7061(01)00145-8

Post, W. M., & Kwon, K. C. (2000). Soil carbon sequestration and land-use change: Processes and potential. *Global Change Biology*, *6*, 317–327. https://doi.org/10.1046/j.1365-2486.2000.00308.x

Qi, R., Li, J., Lin, Z., Li, Z., Li, Y., Yang, X., … Zhao, B. (2016). Temperature effects on soil organic carbon, soil labile organic carbon fractions, and soil enzyme activities under long-term fertilization regimes. *Applied Soil Ecology*, *102*, 36–45. https://doi.org/10.1016/j.apsoil.2016.02.004

Qiu, L., Hao, M., & Wu, Y. (2017). Potential impacts of climate change on carbon dynamics in a rain-fed agro-ecosystem on the Loess Plateau of China. *Science of the Total Environment*, *577*, 267–278. https://doi.org/10.1016/j.scitotenv.2016.10.178

Sreenivas, K., Sujatha, G., Sudhir, K., Kiran, D. V., Fyzee, M. A., Ravisankar, T., & Dadhwal, V. K. (2014). Spatial assessment of soil organic carbon density through random forests based imputation. *Journal of the Indian*

Society of Remote Sensing, *42*, 577–587. https://doi.org/10.1007/s12524-013-0332-x

Starr, G. C., Lal, R., Malone, R., Hothem, D., Owens, L., & Kimble, J. (2000). Modeling soil carbon transported by water erosion processes. *Land Degradation & Development*, *11*, 83–91. https://doi.org/10.1002/(SICI)1099-145X(200001/02)11:1%3C83::AID-LDR370%3E3.0.CO;2-W

Stavi, I., Argaman, E., & Zaady, E. (2016). Positive impact of moderate stubble grazing on soil quality and organic carbon pool in dryland wheat agro-pastoral systems. *Catena*, *146*, 94–99. https://doi.org/10.1016/j.catena.2016.02.004

Sun, J., & Wang, H. (2016). Soil nitrogen and carbon determine the trade-off of the above- and below-ground biomass across alpine grasslands, Tibetan Plateau. *Ecological Indicators*, *60*, 1070–1076. https://doi.org/10.1016/j.ecolind.2015.08.038

Sun, W., Zhu, H., & Guo, S. (2015). Soil organic carbon as a function of land use and topography on the Loess Plateau of China. *Ecological Engineering*, *83*, 249–257. https://doi.org/10.1016/j.ecoleng.2015.06.030

Sun, X. P., & Wang, P. (2005). How old is the Asian monsoon system?—Palaeobotanical records from China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *222*, 181–222. https://doi.org/10.1016/j.palaeo.2005.03.005

Tang, Z. S., An, H., & Shangguan, Z. P. (2015). The impact of desertification on carbon and nitrogen storage in the desert steppe ecosystem. *Ecological Engineering*, *84*, 92–99. https://doi.org/10.1016/j.ecoleng.2015.07.023

Vågen, T. G., Winowiecki, L. A., Abegaz, A., & Hadgu, K. M. (2013). Landsat-based approaches for mapping of land degradation prevalence and soil functional properties in Ethiopia. *Remote Sensing of Environment.*, *134*, 266–275. https://doi.org/10.1016/j.rse.2013.03.006

Vaudour, E., Gilliot, J. M., Bel, L., Lefevre, J., & Chehdi, K. (2016). Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra. *International Journal of Applied Earth Observation and Geoinformation*, *49*, 24–38. https://doi.org/10.1016/j.jag.2016.01.005

Velayutham, M., Pal, D. K., & Bhattacharyya, T. (2000). Organic carbon stock in soils of India. In R. Lal, J. M. Kimble, & B. A. Stewart (Eds.), *Global climate change and tropical ecosystem* (pp. 71–95). Boca Raton, FL: CRC Press.

Walela, C., Daniel, H., Wilson, B., Lockwood, P., Cowie, A., & Harden, S. (2014). The initial lignin: Nitrogen ratio of litter from above and below ground sources strongly and negatively influenced decay rates of slowly decomposing litter carbon pools. *Soil Biology and Biochemistry*, *77*, 268–275. https://doi.org/10.1016/j.soilbio.2014.06.013

Wang, C., Wang, S., Fu, B., Li, Z., Wu, X., & Tang, Q. (2017). Precipitation gradient determines the tradeoff between soil moisture and soil organic carbon, total nitrogen, and species richness in the Loess Plateau, China. *Science of the Total Environment*, *575*, 1538–1545. https://doi.org/10.1016/j.scitotenv.2016.10.047

Wang, J., Yang, R., & Bai, Z. (2015). Spatial variability and sampling optimization of soil organic carbon and total nitrogen for minesoils of the Loess Plateau using geostatistics. *Ecological Engineering*, *82*, 159–164. https://doi.org/10.1016/j.ecoleng.2015.04.103

Wang, L., Wei, S., Horton, R., & Shao, M. (2011). Effects of vegetation and slope aspect on water budget in the hill and gully region of the Loess Plateau of China. *Catena*, *87*, 90–100. https://doi.org/10.1016/j.catena.2011.05.010

Waring, B. G., & Powers, J. S. (2017). Overlooking what is underground: Root : shoot ratios and coarse root allometric equations for tropical forests. *Forest Ecology and Management*, *385*, 10–15. https://doi.org/10.1016/j.foreco.2016.11.007

Wu, Z., Dijkstra, P., Koch, G. W., Peñuelas, J., & Hungate, B. A. (2011). Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. *Global Change Biology*, *17*, 927–942. https://doi.org/10.1111/j.1365-2486.2010.02302.x

Xin, Z., Qin, Y., & Yu, X. (2016). Spatial variability in soil organic carbon and its influencing factors in a hilly watershed of the Loess Plateau, China. *Catena*, *137*, 660–669. https://doi.org/10.1016/j.catena.2015.01.028

Zhang, B., He, C., Burnham, M., & Zhang, L. (2016). Evaluating the coupling effects of climate aridity and vegetation restoration on soil erosion over the Loess Plateau in China. *Science of the Total Environment*, *539*, 436–449. https://doi.org/10.1016/j.scitotenv.2015.08.132

Zhang, G. S., & Ni, Z. W. (2017). Winter tillage impacts on soil organic carbon, aggregation and CO_2 emission in a rainfed vegetable cropping system of the mid-Yunnan Plateau, China. *Soil and Tillage Research*, *165*, 294–301. https://doi.org/10.1016/j.still.2016.09.008

Zhang, J., Hu, K., Li, K., Zheng, C., & Li, B. (2017). Simulating the effects of long-term discontinuous and continuous fertilization with straw return on crop yields and soil organic carbon dynamics using the DNDC mode. *Soil and Tillage Research*, *165*, 302–314. https://doi.org/10.1016/j.still.2016.09.004

Zhang, K., Dang, H., Tan, S., Cheng, X., & Zhang, Q. (2010). Change in soil organic carbon following the 'Grain-for-Green' programme in China. *Land Degradation and Development*, *21*, 13–23. https://doi.org/10.1002/ldr.954

Zhang, S. W., Shen, C. Y., Chen, X. Y., Ye, H. C., Huang, Y. F., & Lai, S. (2013). Spatial interpolation of soil texture using compositional kriging and regression kriging with consideration of the characteristics of compositional data and environment variables. *Journal of Integrative Agriculture*, *12*, 1673–1683. https://doi.org/10.1016/S2095-3119(13)60395-0

Zhang, Y.-W., & Shangguan, Z. P. (2016). The coupling interaction of soil water and organic carbon storage in the long vegetation restoration on the Loess Plateau. *Ecological Engineering*, *91*, 574–581. https://doi.org/10.1016/j.ecoleng.2016.03.033

Zhao, W., Zhang, R., Huang, C., Wang, B., Cao, H., Koopal, L. K., & Tan, W. (2016). Effect of different vegetation cover on the vertical distribution of soil organic and inorganic carbon in the Zhifanggou watershed on the Loess Plateau. *Catena*, *139*, 191–198. https://doi.org/10.1016/j.catena.2016.01.003

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Wang Y, Deng L, Wu G, Wang K, Shangguan Z. Large-scale soil organic carbon mapping based on multivariate modelling: The case of grasslands on the Loess Plateau. *Land Degrad Dev*. 2018;29:26–37. https://doi.org/10.1002/ldr.2833