Catena 164 (2018) 23–31

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

Estimates of carbon storage in grassland ecosystems on the Loess Plateau

Yinyin Wang[a,b], Lei Deng[c], Gaolin Wu[c], Kaibo Wang[d], Zhouping Shangguan[a,b,*]

[a] 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, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling, Shaanxi 712100, China
[d] State Key Lab. of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an, Shaanxi 710061, China

ARTICLE INFO

Keywords:
Decision tree
Linear model
Biomass carbon, grassland ecosystems

ABSTRACT

Grassland ecosystems play an important role in the carbon (C) balance of arid and semi–arid regions. These ecosystems provide C for grass growth and soil microbial activities and represent one of the main sources of atmospheric C. In this study, we estimated the C density and storage of 223 sampling sites in grassland ecosystems on the Loess Plateau using elevation, vegetation indexes, precipitation, air temperature, day and night land surface temperature (LSTd and LSTn, respectively), evapotranspiration (ET), percent tree cover and the non–vegetated area to build decision regression tree and generalized linear regression models (GLMs). The results showed that the C density decreased from south to north and ranged from 0.22 to 29.29 kg C/m^2. The average amount of C stored in the ecosystems was 1.46 Pg. The typical steppe and forest steppe stored the most C, and the steppe desert stored the least. The soil (0–1 m) stored most of the organic C, accounting for > 90%, and the belowground biomass (BGB) contained > 3 times the amount of C as the aboveground biomass (AGB). This study provides reference information for the loss of C and associated mitigation strategies on the Loess Plateau.

1. Introduction

Grasslands store > 10% of the total carbon (C) in the biosphere (Nosberger et al., 2000) and represent a non–negligible component of terrestrial ecosystems (Soussana et al., 2004). Ecosystem responses are a stronger driver of the inter–annual variability in C fluxes in grasslands than climate variability (Zhang et al., 2016). Moreover, drought (Gang et al., 2016), afforestation (Chen et al., 2016), land degradation (Wang et al., 2017a), species diversity (Rutledge et al., 2017a), land use (Petrie et al., 2015; Ward et al., 2017), irrigation (Moinet et al., 2017) and pasture (Tanentzap and Coomes, 2012; Rutledge et al., 2017b) affect the C sequestration capacity of grasslands. These factors increase the uncertainty of grassland C stock assessments at different spatial scales and in different sampling sites (Zhang et al., 2011; Maillard et al., 2017). In ecologically fragile regions, such as the Loess Plateau (Yang et al., 2015), the assessment of grassland C storage is especially important for land management and further vegetation restoration.

The Chinese government implemented the "Grain for Green" project in 1999 (Deng et al., 2014a) to restore vegetation by converting croplands to grasslands or forestlands. Researchers have found that this project drives land use change and C sequestration on the Loess Plateau (Deng et al., 2014b). Chang et al. (2011) noted that the "Grain for Green" project can significantly increase the storage of soil organic C (SOC) on the Loess Plateau; thus, it is important to determine the size of the C reserves in this region.

Grasslands currently cover approximately 1/3 of the Loess Plateau, but few studies have attempted to assess their C storage. Vegetation types, topography (Wang et al., 2017b) and climate–driven changes (Chen et al., 2017a) affect SOC, of which the differences among grasses could be represented by vegetation indexes such as the leaf area index (LAI), the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI). In addition to vegetation indexes, percent tree cover and non–vegetated area information may be useful in somewhere mixed grasses with trees or bare lands. Precipitation determines surface soil moisture on the Loess Plateau, and soil moisture is the main disturbance controlling grassland aboveground biomass (AGB) (Zhang et al., 2016; Hu et al., 2016) and its stored C (AGBC). To quantify the climate–driven factors, we considered precipitation and air temperature as primary factors for C estimation. Because it is difficult to acquire regional soil moisture data for the Loess Plateau, we instead utilized the day land surface temperature (LSTd), the night land surface temperature (LSTn) and evapotranspiration (ET).

* Corresponding author at: Xinong Rd. 26, Institute of Soil and Water Conservation, Yangling, Shaanxi 712100, China.
E-mail addresses: wangkb@ieecas.cn (K. Wang), shangguan@ms.iswc.ac.cn (Z. Shangguan).

https://doi.org/10.1016/j.catena.2018.01.007
Received 30 May 2017; Received in revised form 31 December 2017; Accepted 3 January 2018
Available online 20 February 2018
0341-8162/

These factors were used to build decision regression tree models and generalized linear models (GLMs) for C storage assessment. A decision tree is a stable machine leaning algorithm for data mining (Czajkowski and Kretowski, 2016; Rahmatian et al., 2017; Tayefi et al., 2017) and handling multivalued numerical response variables (D'Ambrosio et al., 2017), and they have various applications (Hong et al., 2015; Fei et al., 2017; Chen et al., 2017b). Generalized linear regression models are commonly used and intuitive (Höskuldsson, 2015; Ross, 2017). We selected decision regression tree and GLM models as they were relatively robust, widely used and easy to transfer.

Furthermore, we collected the results of other research for comparison (Ma et al., 2016). In order to make our results comparable, we extrapolated previous grassland C observation data to the entire Loess Plateau by decision tree and GLM regression models. The objectives of this study were to (1) predict the spatial distribution of grassland AGBC, belowground biomass C (BGBC) and SOC density on the Loess Plateau; (2) estimate the C storage in the grassland ecosystem; and (3) assess the accuracy of our predictions and determine the deviation between our results and previous results.

2. Materials and methods

2.1. Study area

The study area was the Loess Plateau (Fig. 1), which is a landscape with uneven loess in the middle of China that covers an area of 646,200 km^2. The Yellow River cuts across the north of the plateau. The loess stratum can reach > 300 m, and because of the long–term runoff caused by extreme climatic events (Gao et al., 2015), the surface of the loess deposit area has become fragmented (Wu et al., 2017). In addition to the soil erosion under these conditions, the loss of C has become increasingly problematic (He et al., 2016; Li et al., 2017). Grasslands are widely distributed on the hills and plains of the Loess Plateau, and because of its poor soil water–holding capacity, the grass in this arid region is always deeply rooted in the soil to reach water. Grassland map used in this study was supported from "Loess Plateau Data Center, National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China. (http://loess.geodata.cn)".

2.2. Experimental design

In the summers from 2011 to 2013, we established 223 sampling plots (100 × 100 m) in grasslands on the Loess Plateau that were representative of the local area. In each sampling plot, we established a transect (100 m) along the diagonal, and several subplots (typical steppe: 1 × 1 m; other grasslands: 1 × 1 m – 2 × 2 m) were set up at 10–m intervals. Biomass C and SOC density were quantified from > 1000 grassland profiles and were summarized for the 223 sampling plots (average subplot $_{\text{C density}}$ → plot $_{\text{C density}}$).

Grass samples were dried in a drying oven at 65 °C, ground by a cup crusher, and filtered through a 10–mesh sieve until all the dried grass samples were sized. The sized samples were reground in a ball mill after freezing, and filtered again through an 80–100–mesh sieve. Soils in plots were sampled at depths of 0–10 cm, 10–20 cm, 20–30 cm, 30–50 cm, and 50–100 cm with two samples taken from each layer. The samples were filtered through a 2–mm sieve to reject plant residues and dried to a constant weight in a drying oven at 105 °C. Dried soil samples were frozen and ground in a ball mill and filtered again through an 80–100–mesh sieve. All the samples were weighed and placed into tin capsules, and their C contents were assayed by dichromate oxidation.

SOC densities were calculated (Penman et al., 2003) according to Formula 1, and C storage values were calculated by Formula 2. The SOC density for a soil profile with k layers was calculated as follows:

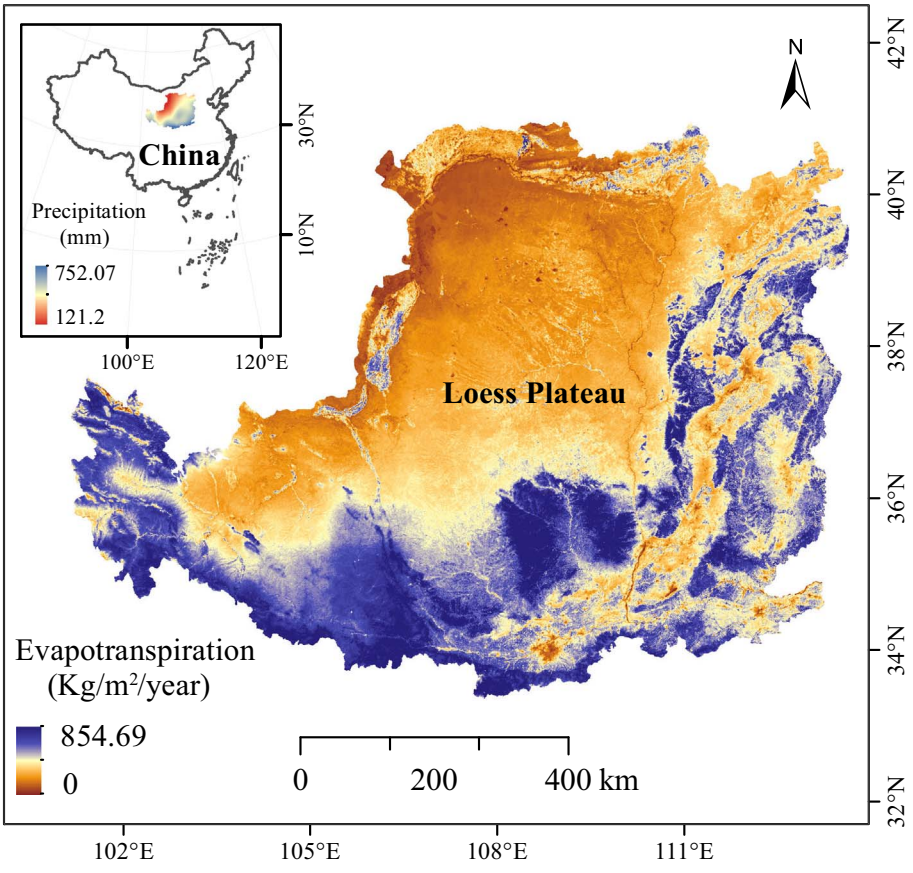

Fig. 1. Study area: the Loess Plateau. The background images are average precipitation and evapotranspiration from 2004 to 2014.

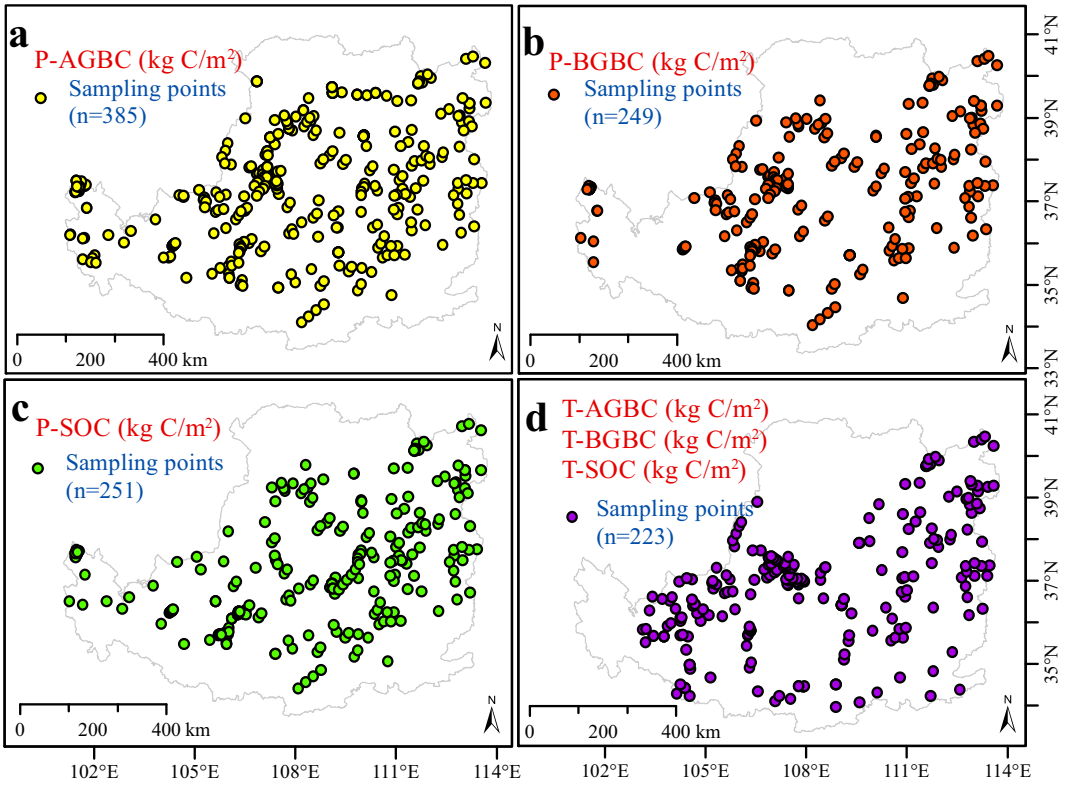

Fig. 2. Spatial locations of the two observation data sets.

$$SOC_{density} = \sum_{i=1}^{k} (\mathrm{SOC}_i \times BD_i \times \mathrm{D}_i \times (1 - \mathrm{CF}_i)) \quad (1)$$

where SOC density is the total amount of SOC per unit area (kg C/m^2); SOC_i is the SOC content in layer i (g/kg); BD_i is the bulk density of layer i (Mg m^{-3}); D_i is the thickness of layer i (m); and CF_i is the fraction of coarse fragment volumes > 2 mm in layer i ($0 \le \mathrm{CF}_i < 1$).

$$\text{Grassland } \mathrm{SOC}_{storage} = \mathrm{SOC}_{density} \times \text{grassland area} \quad (2)$$

2.3. Decision regression tree model and GLM

Observation data sets (Fig. 2). Two C density observation data sets were utilized to build the regression models: one from the 223 sampling sites (T: 2011–2013) and the other from published papers (P: 2004–2014) (Ma et al., 2016).

2.3.1. Models

The decision regression tree model and GLM were run using Marine Geospatial Ecology Tools 0.8a64 (http://mgel2011–kvm.env.duke.edu/mget/), which was developed as a plug–in component of the ARCGIS 10.0 Toolbox (http://www.esri.com/arcgis/about–arcgis). ① The decision regression tree model were based on the rpart (recursive partitioning and regression trees: https://cloud.r–project.org/) framework, which implements the CART (classification and regression trees) methodology by Breiman (Breiman et al., 1983). In a CART tree, factor with minimum sum of error square was split firstly, and then, the next, until the last. Tree was constantly generated by this rule. Resulting models could be shown as binary trees. ② The GLM was generated by multiplying each regression coefficient with its related predictor variable (Guisan and Zimmermann, 2000) and was developed and implemented using the glm2 package (Marschner, 2011) in R. glm2 package provides greater stability than glm package in model convergence. As C storage obeyed a Gaussian distribution, we predicted the C storage of the Loess Plateau according to the formulas generated in GLM.

2.3.2. Variable selection

To optimize regression models, some predictor variables must be remove. In this study, we used two groups of predictor variables corresponding to two C density observation data sets (response variable). The predictor variables were data sets (images) downloaded from websites (Table 1) and contained vegetation indexes (EVI, NDVI: http://www.nasa.gov/; LAI: http://earthexplorer.usgs.gov/), elevation (extracted from digital elevation models (DEMs): http://earthexplorer.usgs.gov/), climate (precipitation and air temperature: https://gis.ncdc.noaa.gov/), percent tree cover, non-vegetated area, ET, LSTd and LSTn (http://www.nasa.gov/). We stepwise selected predictor variables according to the mean squared error (MSE) (Formula 3) and the normalized MSE (NMSE) (Formula 4) generated by 10—fold cross validations (CV). If errors increased when adding a variable, the variable was abandoned and the next was added until all the variables had been considered. The selected predictor variables are shown in Table 2; before modelling, all the variables were resampled to 500 m.

$$\mathrm{MSE} = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2 \quad (3)$$

$$\mathrm{NMSE} = \frac{\sum_{n=0}^{N-1} \sum_{i=0}^{i-1} |O_i - P_i|^2}{\sum_{n=0}^{N-1} \sum_{i=0}^{i-1} |O_i|^2} \quad (4)$$

2.4. Comparison and assessment

We used gridded soil data (250 m) from the World Soil Information Service (WoSIS) as a contrast to the predictions (http://www.isric.org/data/wosis). The global WoSIS dataset was introduced by Batjes et al. (2016), but it needs to be validated and updated based on current practices. The downloaded soil grids of the Loess Plateau were extracted from the global soil map and were resampled to 500 m. The grids included SOC density profile data at depths of 0–5 cm, 5–15 cm,

Table 1
Factor definitions and data sources.

Factors	Explanations	Definitions	Data sources	Resolutions	Unit
EVI	Enhanced Vegetation Index	Average annual EVI (2011 − 2013)	MOD13A3	1000 m	
NDVI	Normalized difference vegetation index	Average annual NDVI (2011–2013)			
LAI	Leaf area index	Average annual LAI (2011–2013)	MOD15A2H	500 m	
Treecover	Percent Tree Cover	Percent tree cover per pixel (2011–2013; 2004–2014)	MOD44B	250 m	%
Nonveg	Percent Non Vegetated	Percent non vegetated area per pixel (2011–2013; 2004–2014)			
Elevation	Elevation	Elevation acquired from Digital Elevation Models (DEMs) of the Loess Plateau	ASTER GDEM	30 m	m
LSTd	Day Land surface temperature	Average annual LSTd (2011–2013)	MOD11B3	5600 m	˚C
LSTn	Night Land surface temperature	average annual LSTn (2011–2013)			
Prep	Precipitation	Average annual precipitation (2011–2013; 2004–2014)	NOAA (National centers for environmental information)	500 m	mm
Temp	Air temperature	Average annual air temperature (2011–2013; 2004–2014)			˚C
ET	Evapotranspiration	Average annual ET (2011–2013; 2004–2014)	MOD16A3		kg/m^2/ year

15–30 cm, 30–60 cm, and 60–100 cm. We summarized the profile data across 0–100 cm for comparison.

The correlation coefficient r (Formula 5) was used to measure the degree of fit of the observations and predictions, MSE (Formula 3) and NMSE (Formula 4), which were used to quantify the deviation errors between the observations and predictions. To assess the prediction accuracy of the decision tree regression model and GLM, all the r, MSE and NMSE values were generated by 10–fold model CV.

$$r = \sqrt{\frac{\sum_{i=1}^{n}(P_i - \overline{O}_i)^2}{\sum_{i=1}^{n}(P_i - \overline{O}_i)^2 + \sum_{i=1}^{n}(O_i - P_i)^2}} \quad (5)$$

In Formulas 3, 4 and 5, O_i represents observed C density, and P_i represents predicted C density.

3. Results

3.1. Carbon density predicted by the decision tree model

The grassland C densities predicted by the decision regression tree model are shown in Fig. 3 and Fig. 4. The spatial patterns of AGBC density showed an obvious decrease from east to west across the study area (Fig. 3a). Most of the western parts had an AGBC density lower than 0.09 kg C/m^2, and most of the eastern parts had values higher than 0.09 kg C/m^2. The predicted BGBC density (Fig. 3b; Fig. 4b) exhibited a pattern similar to the AGBC density, and both the AGBC and BGBC densities were lower than the SOC density (Fig. 3 and Fig. 4). As shown in Fig. 3d and 4d, the spatial variation in ecosystem C density was mostly determined by the SOC density on the Loess Plateau.

Table 2
Predictor variables in regression models.

Response variables: sampling C density (kg/m^2)					
AGBC (n = 223)		BGBC (n = 223)		SOC (0–1 m) (n = 223)	
Decision tree	GLM	Decision tree	GLM	Decision tree	GLM
Nonveg	Nonveg	LSTn	LSTn	LSTn	LSTd
LSTn	LSTn	Prep	Elevation	Treecover	Prep
Elevation	Elevation	ET	Temp	Temp	LSTn
Temp	Temp	EVI	Prep	ET	Treecover
Treecover	LSTd	NDVI	Treecover	NDVI	Nonveg
LSTd	LAI	LAI	ET	EVI	Temp
ET	EVI	LSTd	EVI	Elevation	ET
EVI	NDVI		NDVI		LAI
NDVI			LAI		NDVI
			LSTd		EVI
					Elevation
Response variables: C density collected from published papers (Ma et al., 2016)					
AGBC (n = 385)		BGBC (n = 249)		SOC (0–1 m) (n = 251)	
Decision tree	GLM	Decision tree	GLM	Decision tree	GLM
Elevation	Elevation	Temp	Nonveg	Temp	Treecover
Treecover	Treecover	ET	Temp	Elevation	Nonveg
Temp	ET	Prep	Treecover		Prep
Prep	Temp		Elevation		ET
	Prep		ET		Temp
			Prep		Elevation

Note: The meanings of the abbreviations are the same as in Table 1.

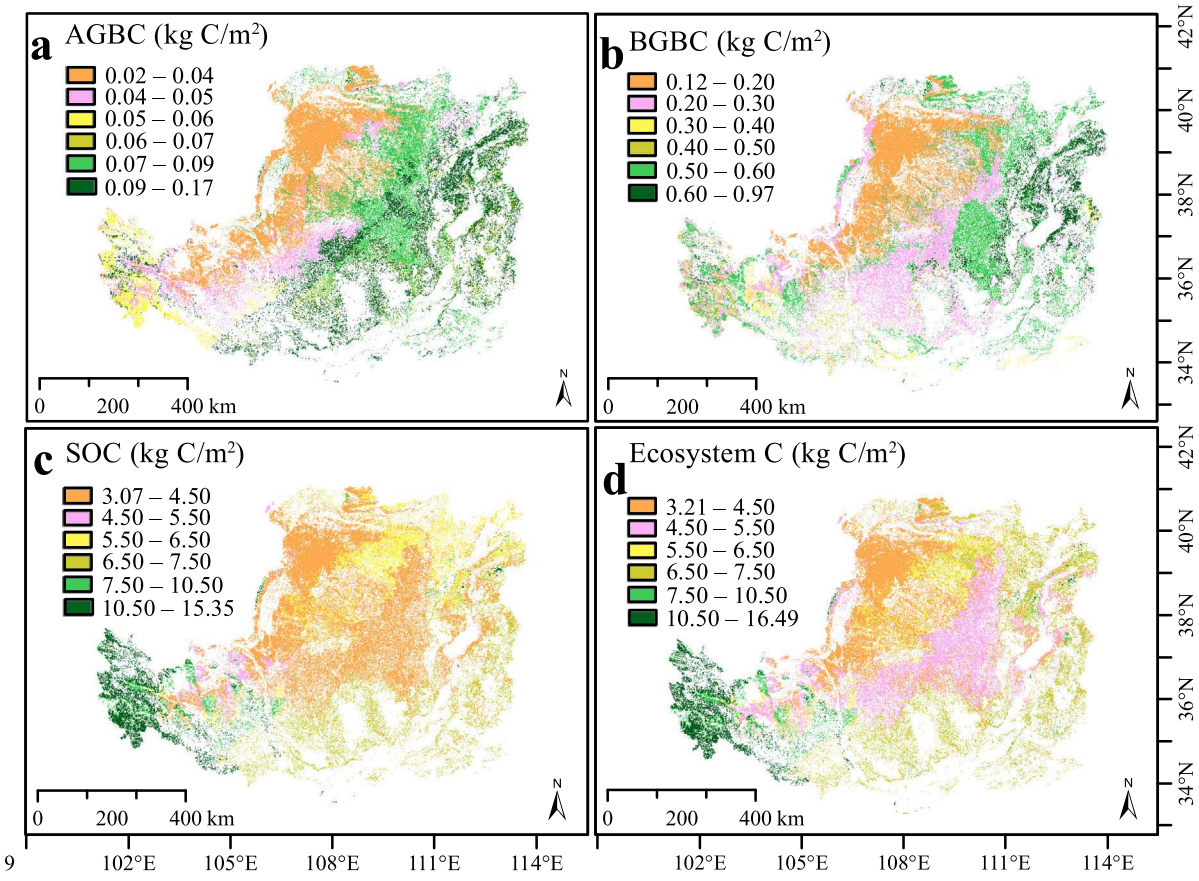

Fig. 3. Spatial distribution of grassland C density on the Loess Plateau as predicted by the decision regression tree model (response variable = T data set: 2011–2013).

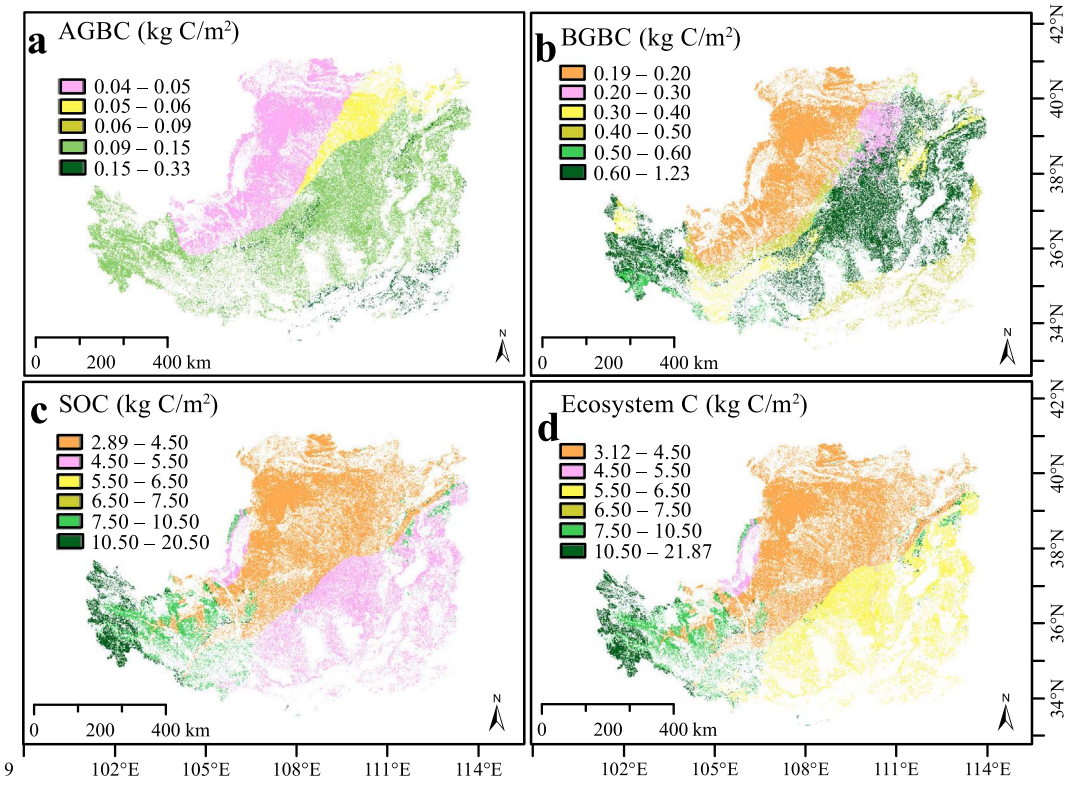

Fig. 4. Spatial distribution of grassland C density on the Loess Plateau as predicted by the decision regression tree model (response variable = P data set: 2004–2014).

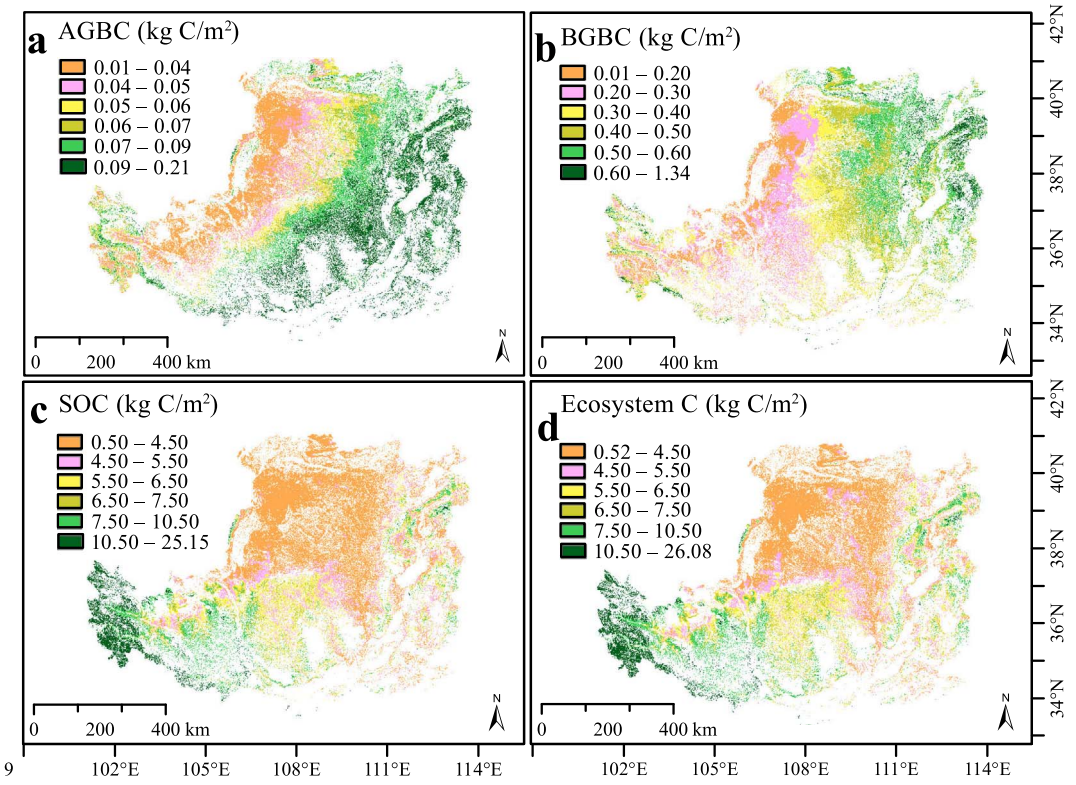

Fig. 5. Spatial distribution of grassland C density on the Loess Plateau as predicted by GLM (response variable = T data set).

3.2. Carbon density predicted by the GLM

The grassland C densities predicted by the GLM are shown in Fig. 5 and Fig. 6. The BGBC density (Fig. 5b and Fig. 6b) had a wider range (0.01–1.34 kg C/m^2 and 0.01–1.47 kg C/m^2, respectively) than that shown in Fig. 3b (0.12–0.97 kg C/m^2) and Fig. 4b (0.19–1.23 kg C/m^2). Similarly, the map of SOC density (Fig. 5c) had many features in common with the C density map of the Loess Plateau grasslands

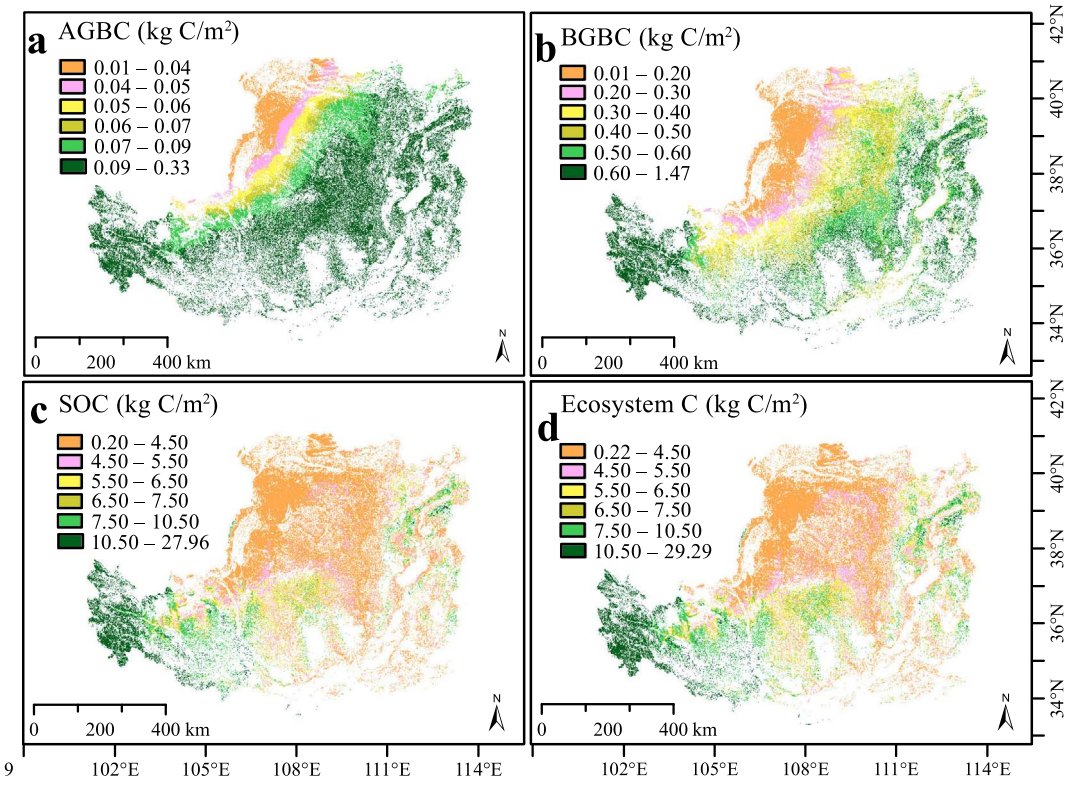

Fig. 6. Spatial distribution of grassland C density on the Loess Plateau as predicted by GLM (response variable = P data set).

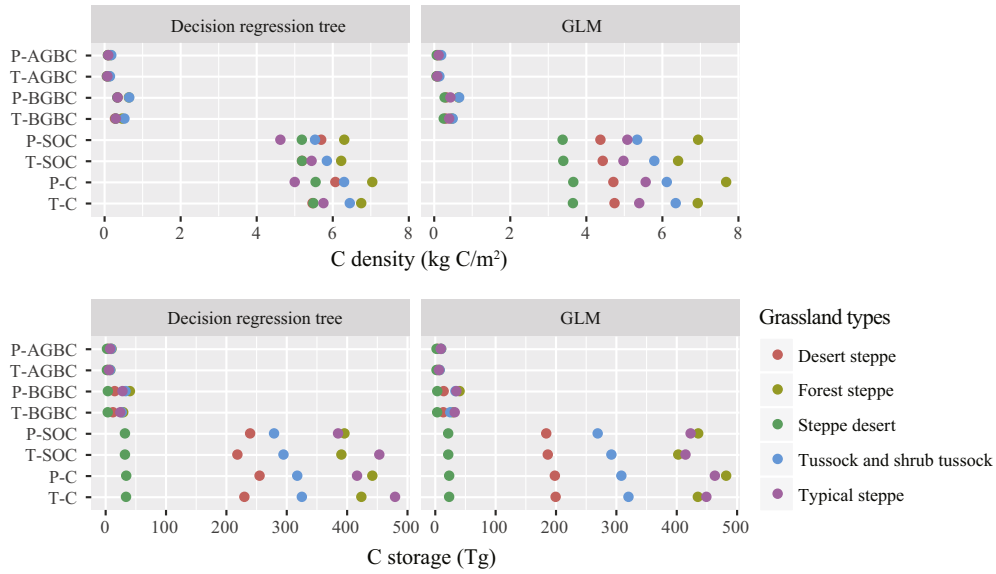

Fig. 7. Carbon density and storage in different grassland ecosystems.

(Fig. 5d). The map AGBC density (Fig. 5a) showed similar spatial patterns as in Fig. 3a, but the decreasing trend from east to west was clearer.

It could be concluded that the predicted ecosystem C density was lower than 30 kg C/m^2 (Figs. 3–6) and was high in the southwest corner of the Loess Plateau (adjacent to the Tibetan Plateau). AGBC and BGBC density obviously decreased from east to west, but the spatial patterns of SOC density determined that of ecosystem C density.

3.3. Carbon density and storage of different grasslands

Grassland C density on the Loess Plateau was predicted by the decision regression tree model and the GLM, and C storage was calculated from C density (Formula 2), and Fig. 7 shows the predicted C density and storage in different grasslands. Predicted AGBC and BGBC density were lower than 2 kg C/m^2, and predicted SOC density (< 7 kg C/m^2) and ecosystem C (< 8 kg C/m^2) were highest in forest steppe. In the GLM, the predicted ecosystem C density was lowest in steppe desert, but which grassland had the lowest predicted C density could not be determined by decision tree modelling.

Biomass C storage (AGBC + BGBC) in each grassland was lower than 200 Tg, and ecosystem C storage (AGBC + BGBC + SOC) in each grassland was lower than 500 Tg. The predicted ecosystem C storage (T – C) in grasslands could be ordered as follows: typical steppe > forest steppe > tussock and shrub tussock > desert steppe > steppe desert. However, the order differed if the observation data set was changed (P – C): forest steppe > typical steppe > tussock and shrub tussock > desert steppe > steppe desert. In both cases, steppe desert had the lowest C storage, and the C storage in forest steppe and typical steppe were highest.

As shown in Table 3, the SOC storage on the grasslands of the Loess Plateau reached 2537.76 Tg according to the WoSIS data, which was nearly twice the model–predicted SOC storage. The calculated grassland C storage in each section was as follows: AGBC, 16.53–25.26 Tg; BGBC, 86.12–116.44 Tg; SOC, 1307.81–1379.64 Tg; and ecosystem C, 1417.59–1482.44 Tg.

4. Discussion

4.1. The WoSIS soil grid database overestimated grassland SOC storage on the Loess Plateau

The WoSIS SOC density data (0–100 cm) were compared with the results of this study (Fig. 8). Fig. 8 shows that the highest SOC density in the WoSIS data set was 158.00 kg C/m^2, which was much higher than the highest predicted values (Figs. 3–6: 0.2–27.96 kg C/m^2). In Fig. 8, the highest predicted C density was also in the southwest of the Loess Plateau, which may be affected by the high SOC density of the Tibetan Plateau (Chen et al., 2017a).

As the WoSIS SOC data were contributed by numerous researchers from around the world, its accuracy still needed to be validated by field surveys. We used two sources of survey data (our observations and published studies) to illustrate this question (Fig. 8); in general, the WOSIS SOC density of grasslands on the Loess Plateau was much higher than those of the other sources. Most of the predicted SOC densities were lower than 5.5 kg C/m^2 (Figs. 3–6), but most SOC densities ranged from 6.50–27.5 kg C/m^2 on the WoSIS SOC density map.

Matsuura et al. (2012) found that the SOC density of four types of grassland (natural, semi–natural, meadows/pastures and artificial grasslands) in Japan was 11.4 kg C/m^2. Considering that the soil fertility on the Loess Plateau is poor (Xu et al., 2014; Liu et al., 2013; Jiao et al., 2011), the SOC density could be lower than that observed in

Table 3
Grassland ecosystem C storage on the Loess Plateau.

Method	C storage (Tg): 1Tg = 10^{12} g							
	Observation set: T				Observation set: P			
	AGBC	BGBC	SOC	Ecosystem C	AGBC	BGBC	SOC	Ecosystem C
WoSIS			2537.76				2537.76	
Decision tree	16.67	86.12	1379.64	1482.44	23.65	110.24	1321.9	1455.79
GLM	16.53	93.25	1307.81	1417.59	25.26	116.44	1323.96	1465.67

Note: T = C density from sampling 223 study sites; P = C density collected from published papers (Ma et al., 2016).

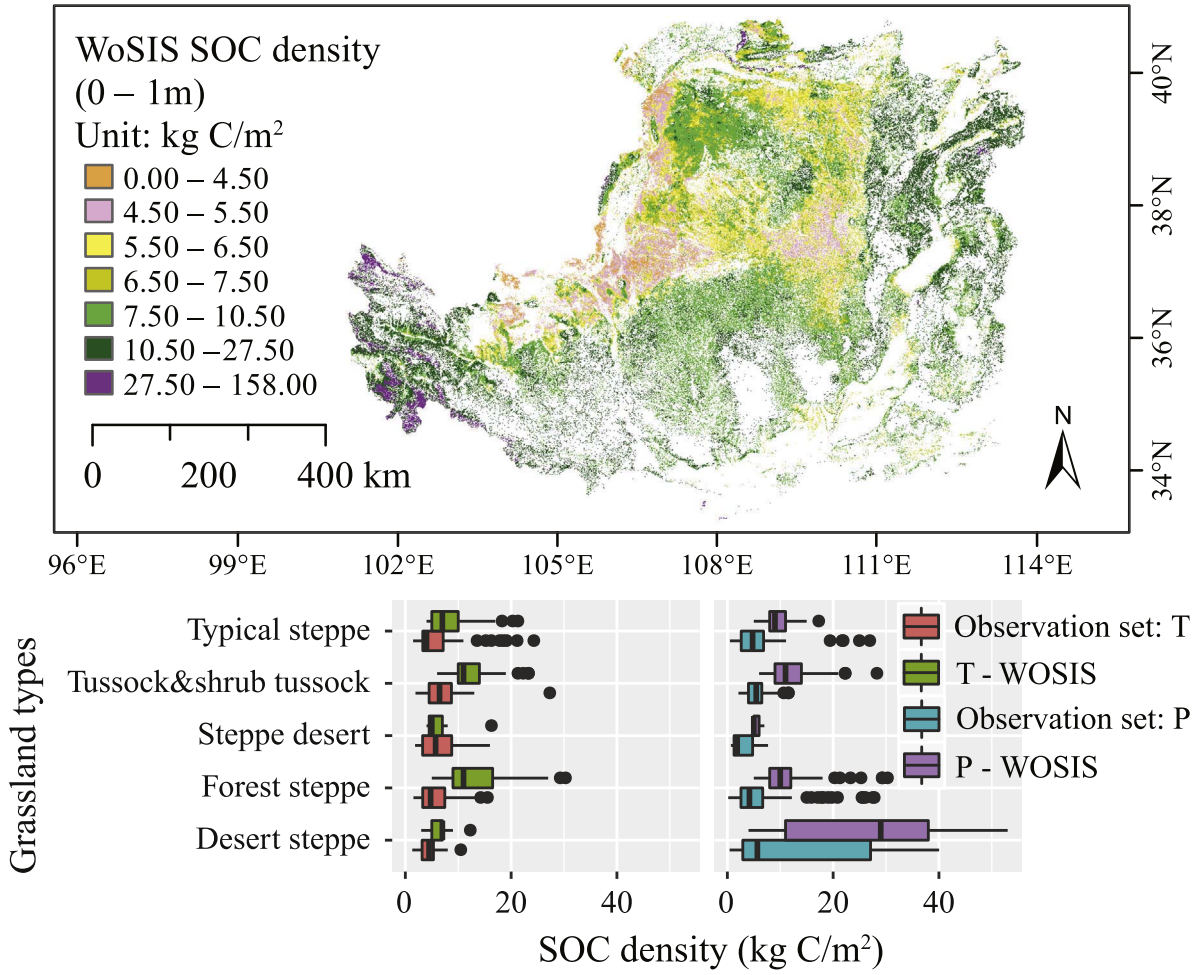

Fig. 8. Comparison of the observed SOC density and WOSIS SOC density (soil depth: 0–1 m) in grasslands of the Loess Plateau.

Japan. Overall, the spatial patterns of the predicted C density were consistent with the WoSIS map, although the actual C densities require further validation.

4.2. Accuracy assessment of the predicted C density

Table 4 shows the prediction accuracy of grassland C densities across the Loess Plateau. We used 10–fold CV to assess the prediction accuracy, and the MSE, NMSE and r were the means of the results. The AGBC density (observation set T) had the lowest prediction error (MSE: 0.001 kg C/m^2), and the SOC density (observation set P) had the highest prediction error (MSE: 33.126 kg C/m^2). Overall, the prediction accuracy of the decision regression tree model and the GLM were similar, and the r based on the decision regression tree model could be a little better than that based on the GLM.

Table 4
Prediction accuracy for C density (MSE, NMSE and r were the means of the results after 10–fold CV).

Observation set	C	Decision tree			GLM		
		MSE (kg C/m^2)	NMSE	r	MSE (kg C/m^2)	NMSE	r
T	AGBC	0.002	0.593	0.73	0.001	0.501	0.733
	BGBC	0.073	0.871	0.634	0.06	0.689	0.606
	SOC	11.263	0.908	0.647	10.047	0.769	0.63
P	AGBC	0.008	0.815	0.606	0.008	0.85	0.427
	BGBC	0.158	0.908	0.608	0.143	0.828	0.508
	SOC	33.126	0.834	0.721	30.908	0.799	0.62

4.3. Uncertainty of Loess Plateau grassland ecosystem C estimations

There was uncertainty in the C storage predictions (Yuen et al., 2016). Zhan et al. (2013) indicated that SOC decreased from the southeast to the northwest of the Loess Plateau and estimated the SOC storage of the 440,000 km^2 area to be 3.629 Pg (1 Pg = 1000 Tg), but the spatial location of their study was different than ours. The highest estimated SOC storage in this study was 1.38 Pg, but the WoSIS data indicated that the SOC storage was 2.538 Pg (Table 3). Both estimates were lower than that of Zhan's study. Thus, additional studies must be performed to confirm the source of the uncertainty and to improve the accuracy of C storage predictions at the landscape scale.

Ma et al. (2017) concluded that SOC storage was larger than that of biomass C regardless of grassland type in north China, and this was supported by our study. We also found that soils on the Loess Plateau stored 10 times more C than grass biomass. Ma et al. (2017) showed that desert grassland contributed the least to total C storage, and we found that steppe desert contributed the least, followed by desert steppe.

5. Conclusions

In this study, we utilized a decision regression tree model and GLM to predict the C density and storage of grasslands on the Loess Plateau. The AGBC density decreased from east to west, but the highest

ecosystem C density was observed in the southwest. Furthermore, the predicted AGBC density was the most accurate, followed by the predicted SOC.

The steppe desert had the lowest C storage, and typical steppe and forest steppe were the main contributors to the grassland C pool. For all grasslands, the BGB contained more C than the AGB, and the soil stored the greatest amount of C.

Estimating the grassland ecosystem C storage on the Loess Plateau could provide a reference for the dynamic monitoring of C losses in further studies and could help elucidate whether C storage conditions would favour regional C transfer. Taken together, these results could provide insights into efficient environmental management strategies for the Loess Plateau.

Acknowledgements

The study was financially supported by the National Key Technology R&D Program of China, Grant/Award Number: 2016YFC0501605; National Natural Science Foundation of China, Grant/Award Number: 41390463; National Sci-Tech Basic Program of China, Grant/Award Number: 2014FY210100.

References

Batjes, N., Ribeiro, E., van Oostrum, Ad, Leenaars, J., Hengl, T., Mendes de Jesus, J., 2016. WoSIS: serving standardised soil profile data for the world. Earth. Syst. Sci. Data. http://dx.doi.org/10.5194/essd–2016–34. (Published on line, 2016).

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1983. Classification and Regression Trees. Wadsworth, Belmont, CA.

Chang, R.Y., Fu, B.J., Liu, G.H., Liu, S.G., 2011. Soil carbon sequestration potential for "Grain for Green" Project in Loess Plateau, China. Environ. Manag. 48, 1158–1172.

Chen, L.F., He, Z.B., Zhu, X., Du, J., Yang, J.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.

Chen, L., Jing, X., Flynn, D.F.B., Shi, Y., Kühn, P., Scholten, T., He, J.S., 2017a. Changes of carbon stocks in alpine grassland soils from 2002 to 2011 on the Tibetan Plateau and their climatic causes. Geoderma 288, 166–174.

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D.T., Duan, Z., Ma, J., 2017b. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena 151, 147–160.

Czajkowski, M., Kretowski, M., 2016. The role of decision tree representation in regression problems – an evolutionary perspective. Appl. Soft. Comput. 48, 458–475.

D'Ambrosio, A., Aria, M., Iorio, C., Siciliano, R., 2017. Regression trees for multivalued numerical response variables. Expert Syst. Appl. 69, 21–28.

Deng, L., Liu, G.B., Shangguan, Z.P., 2014a. Land use conversion and changing soil carbon stocks in China's "Grain–for–Green" Program: a synthesis. Glob. Chang. Biol. 20, 3544–3556.

Deng, L., Shangguan, Z.P., Sweeney, S., 2014b. "Grain for Green" driven land use change and carbon sequestration on the Loess Plateau, China. Sci. Rep. 4, 7039. http://dx.doi.org/10.1038/srep07039.

Fei, Y., Gao, K., Hu, J., Tu, J., Li, W., Wang, W., Zong, G., 2017. Predicting the incidence of portosplenomesenteric vein thrombosis in patients with acute pancreatitis using Classification and Regression Tree (CART) algorithm. J. Crit. Care 39, 124–130. Available online. https://doi.org/10.1016/j.jcrc.2017.02.019.

Gang, C., Wang, Z., Chen, Y., Yang, Y., Li, J., Cheng, J., Qi, J., Odeh, I., 2016. Drought–induced dynamics of carbon and water use efficiency of global grasslands from 2000 to 2011. Ecol. Indic. 67, 788–797.

Gao, Y., Feng, Q., Liu, W., Lu, A., Wang, Y., Yang, J., Cheng, A., Wang, Y., Su, Y., Liu, L., Ma, Q., 2015. Changes of daily climate extremes in Loess Plateau during 1960–2013. Quat. Int. 371, 5–21.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

He, S., Liang, Z., Han, R., Wang, Y., Liu, G., 2016. Soil carbon dynamics during grass restoration on abandoned sloping cropland in the hilly area of the Loess Plateau, China. Catena 137, 679–685.

Hong, H., Pradhan, B., Xu, C., Bui, D.T., 2015. Spatial prediction of landslide hazard at the Yihuang area (China) using two–class kernel logistic regression, alternating decision tree and support vector machines. Catena 133, 266–281.

Höskuldsson, A., 2015. Common framework for linear regression. Chemom. Intell. Lab. Syst. 146, 250–262.

Hu, Q., Jiang, D., Lang, X., Xu, B., 2016. Moisture sources of the Chinese Loess Plateau during 1979–2009. Palaeogeogr. Palaeoclimatol. Palaeoecol Available online. https://doi.org/10.1016/j.palaeo.2016.12.030.

Jiao, F., Wen, Z.M., An, S.S., 2011. Changes in soil properties across a chronosequence of vegetation restoration on the Loess Plateau of China. Catena 86, 110–116.

Li, Z., Liu, C., Dong, Y., Chang, X., Nie, X., Liu, L., Xiao, H., Lu, Y., 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 Tillage Res. 166, 1–9.

Liu, Z.P., Shao, M.A., Wang, Y.Q., 2013. Spatial patterns of soil total nitrogen and soil total phosphorus across the entire Loess Plateau region of China. Geoderma 197–198, 67–78.

Ma, A., He, N., Yu, G., Wen, D., Peng, S., 2016. Carbon storage in Chinese grassland ecosystems: influence of different integrative methods. Sci. Rep. 6, 21378. http://dx.doi.org/10.1038/srep21378.

Ma, Q., Kuang, W., Liu, Z., Hu, F., Qian, J., Liu, B., Zhu, J., Cao, C., Wu, J., Li, X., Zhou, Q., Kou, Z., Shou, W., 2017. Spatial pattern of different component carbon in varied grasslands of northern China. Geoderma 303, 27–36.

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. Agric. Ecosyst. Environ. 236, 268–276.

Marschner, I.C., 2011. glm2: fitting generalized linear models with convergence problems. R J. 3/2, 12–15.

Matsuura, S., Sasaki, H., Kohyama, K., 2012. Organic carbon stocks in grassland soils and their spatial distribution in Japan. Grassl. Sci. 58, 79–93.

Moinet, G.Y.K., Cieraad, E., Turnbull, M.H., Whitehead, D., 2017. Effects of irrigation and addition of nitrogen fertiliser on net ecosystem carbon balance for a grassland. Sci. Total Environ. 579, 1715–1725.

Nosberger, J., Blum, H., Fuhrer, J., 2000. Crop ecosystem responses to climatic change: productive grasslands. In: Hodges, H.F. (Ed.), Climate Change and Global Crop Productivity. CAB International, Wallingford, UK, pp. 271–291.

Penman, J., Gytarsky, M., Hiraishi, T. (Eds.), 2003. Good Practice Guidance for Land Use, Land–Use Change and Forestry. Institute for Global Environmental Strategies (IGES) for the IPCC, Kanagawa (1.1–LR.5).

Petrie, M.D., Collins, S.L., Swann, A.M., Ford, P.L., Litvak, M.E., 2015. Grassland to shrubland state transitions enhance carbon sequestration in the northern Chihuahuan Desert. Glob. Chang. Biol. 21, 1226–1235.

Rahmatian, M., Chen, Y.C., Palizban, A., Moshref, A., Dunford, W.G., 2017. Transient stability assessment via decision trees and multivariate adaptive regression splines. Electr. Power Syst. Res. 142, 320–328.

Ross, S.M., 2017. Chapter 12: linear regression. In: Introductory Statistics, Fourth edition. pp. 519–584.

Rutledge, S., Wall, A.M., Mudge, P.L., Troughton, B., Campbell, D.I., Pronger, J., Joshi, C., Schipper, L.A., 2017a. The carbon balance of temperate grasslands part I: the impact of increased species diversity. Agric. Ecosyst. Environ. 239, 310–323.

Rutledge, S., Wall, A.M., Mudge, P.L., Troughton, B., Campbell, D.I., Pronger, J., Joshi, C., Schipper, L.A., 2017b. The carbon balance of temperate grasslands part II: the impact of pasture renewal via direct drilling. Agric. Ecosyst. Environ. 239, 132–142.

Soussana, J.–.F., Loiseau, P., Vuichard, N., Ceschia, E., Balesdent, J., Chevallier, T., Arrouays, D., 2004. Carbon cycling and sequestration opportunities in temperate grasslands. Soil Use Manag. 20, 219–230.

Tanentzap, A.J., Coomes, D.A., 2012. Carbon storage in terrestrial ecosystems: do browsing and grazing herbivores matter? Biol. Rev. 87, 72–94.

Tayefi, M., Tajfard, M., Saffar, S., Hanachi, P., Amirabadizadeh, A.R., Esmaeily, H., Taghipour, A., Ferns, G.A., Moohebati, M., Ghayour-Mobarhan, M., 2017. hs–CRP is strongly associated with coronary heart disease (CHD): a data mining approach using decision tree algorithm. Comput. Methods Prog. Biomed. 141, 105–109.

Wang, Z., Deng, X.Z., Song, W., Li, Z., Chen, J., 2017a. What is the main cause of grassland degradation? A case study of grassland ecosystem service in the middle–south Inner Mongolia. Catena 150, 100–107.

Wang, Z., Hu, Y., Wang, R., Guo, S., Yao, Z., 2017b. Soil organic carbon on the fragmented Chinese Loess Plateau: combining effects of vegetation types and topographic positions. Soil Tillage Res. 174, 1–5.

Ward, A., Yin, K., Dargusch, P., Fulton, E.A., Aziz, A.A., 2017. The impact of land use change on carbon stored in mountain grasslands and shrublands. Ecol. Econ. 135, 114–124.

Wu, J., Miao, C., Wang, Y., Duan, Q., Zhang, X., 2017. Contribution analysis of the long–term changes in seasonal runoff on the Loess Plateau, China, using eight Budyko–based methods. J. Hydrol. 545, 263–275.

Xu, M., Zhang, J., Liu, G.B., Yamanaka, N., 2014. Soil properties in natural grassland, *Caragana korshinskii* planted shrubland, and *Robinia pseudoacacia* planted forest in gullies on the hilly Loess Plateau, China. Catena 119, 116–124.

Yang, X., Jiang, W., Yang, S., Kong, Z., Luo, Y., 2015. Vegetation and climate changes in the western Chinese Loess Plateau since the Last Glacial Maximum. Quat. Int. 372, 58–65.

Yuen, J.Q., Fung, T., Ziegler, A.D., 2016. Review of allometric equations for major land covers in SE Asia: uncertainty and implications for above– and below–ground carbon estimates. For. Ecol. Manag. 360, 323–340.

Zhan, C., Cao, J., Han, Y., Huang, S., Tu, X., Wang, P., An, Z., 2013. Spatial distributions and sequestrations of organic carbon and black carbon in soils from the Chinese loess plateau. Sci. Total Environ. 465, 255–266.

Zhang, L., Wylie, B.K., Ji, L., Gilmanov, T.G., Tieszen, L.L., Howard, D.M., 2011. Upscaling carbon fluxes over the Great Plains grasslands: sinks and sources. J. Geophys. Res. Biogeosci. 116 (G3) (2005–2012).

Zhang, X., Zhao, W., Liu, Y., Fang, X., Feng, Q., 2016. The relationships between grasslands and soil moisture on the Loess Plateau of China: a review. Catena 145, 56–67.