### Estimation of soil organic carbon under different vegetation types on a hillslope of China’s northern Loess Plateau using state-space approach

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Soil Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID:</td>
<td>CJSS-2017-0042.R2</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>15-Jun-2017</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Zhang, Qingyin; Yao, Yufei; Institute of Soil and Water Conservation, Chinese Academy of Sciences and Ministry of Water Resources, State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau; University of Chinese Academy of Sciences, Jia, Xiaoxu; Shao, Mingan; Institute of Geographic Sciences and Natural Resources Research,</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Soil organic carbon, state-space model, land-use type, Loess Plateau</td>
</tr>
</tbody>
</table>
Estimation of soil organic carbon under different vegetation types on a hillslope of China’s northern Loess Plateau using state-space approach

Qingyin Zhang\textsuperscript{1,2}, Yufei Yao\textsuperscript{1,3}, Xiaoxu Jia\textsuperscript{1,3,4*}, Ming’an Shao\textsuperscript{1,2,3,4*}

\textsuperscript{1}State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau, Northwest A&F University, Yangling 712100, China

\textsuperscript{2}College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China

\textsuperscript{3}College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, 100190, China

\textsuperscript{4}Key Laboratory of Ecosystems Network Observation and Modeling, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

\textbf{Corresponding authors:} Xiaoxu Jia (jiaxx@igsnrr.ac.cn), Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

Ming’an Shao (shaoma@igsnrr.ac.cn), College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China
Abstract

Soil organic carbon (SOC) plays a critical role in revegetation of semi-arid areas. The accurate estimation of SOC under various land-use types is fundamental to sustaining ecosystem productivity. Thus the dominant factors of the spatial distribution of SOC in shallow soil layers were determined at hillslope scale. The state-space modeling approach was used to quantify the relationship between SOC stock and land use type, soil properties, topographic features and fine root biomass at 0–20 and 20–40 cm soil layers of a hillslope on the Loess Plateau. The best state-space models explained more than 96% of the variations in SOC stocks on the hillslope. The best multivariate state-space models including land-use type, fine root biomass, soil pH and total nitrogen were optimal for 0–20 and 20–40 cm soil layers. Land-use type was the dominant factor for identification of localized variation in SOC in the 0–40 cm soil layer. The results underscored the importance of land-use type in SOC variation on the hillslopes of the Loess Plateau. It also provided a useful insight into the accurate estimation of SOC using state-space modeling approach driven by other easily obtainable variables.

Key words: Soil organic carbon; state-space model; land-use type; Loess Plateau

Introduction

As a key control on soil fertility and ecosystem productivity (Tiessen et al. 1994), soil is the largest pool of terrestrial organic carbon (OC) and contains at least three times as much carbon (C) as terrestrial plants or the atmosphere (Schlesinger 1997). The dynamic equilibrium of gains and losses of soil organic carbon (SOC) is sensitive to climate change and human disturbance (Knorr
et al. 2005; Wu et al. 2010; Wang et al. 2015a). The feedback of these dynamics, such as carbon
dioxide (CO$_2$) concentrations in the atmosphere, greatly affects the rate of climate change
(Trumbore et al. 1996; Muñoz-Rojas et al. 2013). Land-use change is another factor of the global
C budget which influences SOC (Foley et al. 2005). The accurate estimation of SOC distribution
provides essential data for both SOC monitoring and process-based carbon cycle modeling (Lal
2004). Therefore, a clear knowledge on SOC stocks is essential for accurate application of
modeling and measurement data in prediction and evaluation of the effects of land-use change on

Heterogeneity in soil and biotic factors can create a high spatial variability of SOC (Liu et al.
2012; Jia et al. 2017). To better understand this, traditional methods such as linear regression and
analysis of variance have been used to estimate SOC at local (Zhang et al. 2013; Liu et al. 2016),
regional (Wang et al. 2015b; Willaarts et al. 2016) and global (Jobbágy and Jackson 2000; Lal
2004; Van Minnen et al. 2009) scales. However, most of the estimations have apparently ignored
the spatial dependence of the variables (Goovaerts 1999). Therefore, autoregressive state-space
models have been used to investigate spatial variation of soil variables, taking into account their
autocorrelation in space and the correlations of point-to-point spatial processes between variables
(Heuvelink and Webster 2001). The state-space modeling approach has proven more effective than
classical statistical methods in identifying localized variations in spatial series of soil properties,
crop production and nutrient status (Liu et al. 2012; Wendroth et al. 2003; Zhao et al. 2016). For
example, yields of wheat and barley can be estimated by soil properties, topographic properties,
and aerial infrared photographs (Wendroth et al. 2003).

Using the state-space approach and artificial neural network technique, Zhao et al. (2016)
estimated soil water content driven by relevant soil properties and climatic conditions. For SOC estimation, the state-space approach is much better at large scales than the equivalent linear regression (Liu et al. 2012; Jia et al. 2017). Although She et al. (2014) focused on the relationship between topographic properties and SOC, the relative contribution of biotic and abiotic factors to SOC variation was not clarified. As a spatial variable, SOC can be influenced by various factors and its spatial processes analyzed from more easily measured variables, including climatic factors (e.g., precipitation and temperature), edaphic factors (e.g., soil texture or soil nutrient), topographic factors (e.g., relative elevation and aspect), land-use types (cropland, grassland and fallow land) and land management practices (Jobbágy and Jackson 2000; Booker et al. 2013; Zhang et al. 2013; Wang et al. 2015a, 2016; Liu et al. 2016). Specifically SOC is highly influenced by land-use type (Qiu et al. 2012; Wei et al. 2013). However, the factors that contribute to spatial variability of SOC storage in surface soil layers of a hillslope have not been well addressed.

The Loess Plateau, best known for intensive soil erosion and heavy sediment loads, is the most fragile ecosystem in China (Tang 2000). To reduce the fragility of the natural ecosystem on the Loess Plateau, vegetation restoration and environmental protection programs (“Grain-for-Green” project) were established in 1999 (Fu et al. 2006). Since restoration of vegetation must be implemented on a relatively small scale (e.g. hillslope or small watershed) due to the high variability of the landform, land use pattern and soil type, research on spatial variability of SOC in surface soil layer of hillslopes and the related driving factors are critical for understanding and enhancing carbon sequestration on the Loess Plateau (Chen et al. 2007). Data for characterization of the relationship between SOC and environmental factors are also useful for
the management of land-use and the recovery of soil fertility. Large-scale studies exist on spatial variability of SOC and the driving environmental factors using state-space modeling (Liu et al. 2012; Jia et al. 2017). However, there is little small-scale (e.g., hillslope scale) research addressing spatial variation in SOC and its driving factors using the same state-space modeling.

In this study, a total of 48 sampling points were collected across four land-use types to estimate the spatial variation in SOC on a hillslope. The specific objectives of the study were to: i) collect environmental and SOC data under different vegetation types on a hillslope of China’s northern Loess Plateau; ii) estimate SOC stocks at hillslope scale using a state-space modeling approach driven by land-use type, vegetation biomass, and soil and topographic properties; and iii) determine the relative importance of biotic and abiotic factors in influencing spatial distribution of SOC in surface soil layers. Since land-use type has significant effect on biotic (i.e., fine root biomass, microbial community, etc.) and abiotic (i.e., soil physical and chemical factors) factors in semiarid ecosystems, it was hypothesized that land-use type is the dominant factor for spatial variation of SOC on hillslopes as soil type and topography (i.e., slope aspect and gradient) are relatively homogeneous. It was next hypothesized that artificial shrubland and grassland can increase SOC stocks, of which magnitude varies with vegetation type, since vegetation restoration in cultivated sloping lands enhances above- and below-ground biomass and decreases soil erosion.

Materials and methods

Study area

The study was conducted in Liudaogou watershed (38°46′–38°51′N, 110°21′–110°23′E and 1094–1274 m above sea level), which is 14 km west of Shenmu County, Shaanxi Province, China
The watershed is located in the water-wind erosion crisscross of the Loess Plateau and covers an area of 6.89 km$^2$. It is a semiarid region with a continental monsoon climate. Based on climatic data for 1950–2010, mean annual precipitation in the study area is 437.4 mm, approximately 70% of which occurs in June through September. The mean annual temperature is 8.4 °C, which can be as low as -9.7 °C in January and as high as 23.7 °C in July. Based on the FAO/UNESCO system, the soil is classified as Calcaric Regosol. The soil developed from low fertility loess and has a weak cohesion, high infiltration, low water retention and easy erodibility. Throughout the area, perennial vegetation, including purple alfalfa (*Medicago sativa*) and korshinsk peashrub (*Caragana korshinskii Kom*) has been planted in the past decade to control soil degradation. The natural biomes in the area include bunge needlegrass (*Stipa bungeana Trin*) and *Poa sphondylodes* (*Poa sphondylodes Trin*). Millet (*Setaria italica* L.) and potato (*Solanum tuberosum* L.) are the main grain crops.

**Experimental design**

Four parallel experimental plots, each $61 \times 5$ m with plot spacing of 0.8 m, were constructed on a hillslope of the Loess Plateau in 2004 (*Fu et al. 2010*). The hillslope was cultivated with millet for decades, but abandoned for two years before the experiment. The study included four land-use types: korshinsk peashrub (shrubland), purple alfalfa (grassland), natural fallow land (fallow land), and cropland under millet (cropland). Both korshinsk peashrub and alfalfa grow extensively in the region. Natural fallow occurs on abandoned fields left to recover with natural vegetation, which is a traditional land use type. The uniform conditions of slope gradient (12°), aspect (northwest), soil type (aeolian loess) and soil texture (loamy) ensured that the differences in SOC at the same depth
among the four plots were largely influenced by land use type. The selected physical and chemical properties of the 0–10-cm soil layer on the hillslope were 15.6% of clay, 44% of sand, 0.27 g/kg of soil organic carbon and 0.19 g/kg of total nitrogen (Fu et al. 2013). Further details on the background information are documented by Zeng et al. (2011).

Field sampling and measurements

As the state-space model is designed for one-dimensional analysis, the experiment was arrayed as shown in Fig. 1. The hillslope upper region was planted with korshinsk peashrub and the bottom region planted with millet. In each plot, 12 sampling points were taken along the plot axle at an interval of 5 m with the first and 12th points taken at 3 m from the upper and lower boundaries. To preclude boundary effect, a distance of 0.5 m was eliminated from the border of the plot. Thus, the distance between the last point in a plot and the first point in the next plot was 5 m. A total of 48 points were selected for soil sampling.

From 26 July to 4 August 2016, a total of 13 growing seasons after the start of the experiment, soils samples were collected at 0–20 and 20–40 cm soil depths using a 5 cm diameter auger. While the depth was not sufficient for a complete picture of SOC stock, the response of SOC to land use change was detectable since SOC mainly occurs in the upper soil horizon (Wei et al. 2013; Qiu et al. 2012). All large pieces of undecomposed organic material were removed from the soil samples. The roots of korshinsk peashrub, purple alfalfa, fallow land, and millet were extracted manually at the 0–20 and 20–40 cm soil depths in late August 2016. Three sampling points were randomly selected in each plot using a 50 cm × 50 cm quadrat and the collected roots washed in flowing water. The fine roots (< 2 mm in diameter) were separated and oven-dried at 70 °C to a constant
mass and weighed to determine fine root biomass (FRB). The remaining soil was air-dried and ground to pass through 1.00 mm and 0.25 mm nylon screens prior to laboratory analysis. Samples passed through the 0.25 mm mesh were used to determine SOC and total nitrogen (TN) with the Walkey-Black (Nelson and Sommers 1982) and Kjeldahl (Bremner and Mulvaney 1982) methods, respectively. Samples passed through the 1.00 mm mesh were used to determine soil pH and soil mechanical composition (%). Soil pH was measured at a soil to water ratio by mass of 1:1.25 using a pH meter equipped with a calibrated combined glass electrode. Soil mechanical composition was determined by laser diffraction using Mastersizer 2000 (Malvern Instruments, Malvern, England).

Undisturbed soil samples were taken using a stainless steel cutting ring (5 cm in diameter by 5 cm in height) at mid-height for 0–20 cm and 20–40 cm soil layers. The soil core samples were slowly saturated from the base with deionized water before measuring saturated hydraulic conductivity (K_s) by the constant head method (Kanwar et al. 1990), and then were oven-dried at 105 °C before weighing to determine soil bulk density.

**State-space model theory**

The state-space approach is a multivariate autoregressive technique that can characterize the state of a system (a set of unobservable variables) at location i to its state at location i-h, where h = 1, 2, 3, …, n (Shumway 1988; Wendroth et al. 2003). The state-space model consists of two vector equations (a state equation and an observation equation) which represent the relationship between the input and output of a dynamic system (Timm et al. 2003). For h = 1, the state equation with the structure of the first order autoregressive model is given as:
\[ Z_p(x_i) = \Phi_{pp}Z_p(x_{i-1}) + \omega_{Zp}(x_i) \]  

(1)

where \( Z_p(x_i) \) is the state vector, a set of \( p \) variables at location \( i \); \( \Phi_{pp} \) is a \( p \times p \) state coefficient matrix (transition matrix), which indicates the measure of the regression; and \( \omega_{Zp}(x_i) \) is the uncorrelated zero mean model error.

In the observation equation, the observation vector \( Y_p(x_i) \) is related to the state vector \( Z_p(x_i) \) by an observation matrix \( M_{pp}(x_i) \) (usually known as an identity matrix) and the uncorrelated zero mean observation error vector \( v_{yp}(x_i) \) as:

\[ Y_p(x_i) = M_{pp}(x_i)Z_p(x_i) + v_{yp}(x_i) \]  

(2)

It is assumed that \( \omega_{Zp}(x_i) \) and \( v_{yp}(x_i) \) are normally distributed and independent of each other. The state vector is the true state of the system and the observation vector needs not to be true, but can be seen as indirect measurements reflecting the true state of the variable added to an error (Wendroth et al. 1999). The state coefficients of the matrix \( \Phi_{pp} \) and model error \( \omega_{Zp}(x_i) \) are estimated through EM algorithm with an iterative procedure given by Shumway and Stoffer (1982), and are optimized using the Kalman filter (Kalman 1960).

Prior to the state-space analysis, data \( Z_p(x_i) \) are normalized with respect to the mean \( m \) and standard deviation \( s \) as follows:

\[ z_p(x_i) = \left[ Z_p(x_i) - (m - 2s) \right]/4s \]  

(3)

where \( z_p(x_i) \) are the normalized values, dimensionless and with mean of 0.5 and standard deviation of 0.25 (Wendroth et al. 2003). The scaling avoids numerical problems when two or more variables differ by orders of magnitude. This transformation allows state coefficients of \( \Phi_{pp} \) having magnitudes directly proportional to their contributions to each state variable to be included in the state-space analysis (Timm et al. 2003).
Data analysis

Analysis of variance (ANOVA) and post hoc least significant difference test were conducted to evaluate differences in SOC among the land use types for 0–20 and 20–40 cm soil depths. A paired t-test was used to determine any significant difference between the two soil depths for the same land use type. Pearson correlation coefficient was used to determine the strength of the correlations between SOC stocks with soil physical and chemical properties, plant growth, topography and land use variables in SPSS 18.0 (SPSS for windows, Chicago, IL, USA).

The spatial characteristics of SOC in the 0–20 and 20–40 cm soil profile from 48 points, including four land use types, were modeled using the state-space method. The estimated values of SOC and transition coefficients of the state equation were computed using the Applied Statistical Time Series Analysis (ASTSA) software developed by Shumway (1988). In order to evaluate the prediction quality of the different state-space model approaches, the coefficients of determination ($R^2$) was calculated from the linear regression between observations and estimations, and the regression residuals ($RSS_{avg}$) and Akaike Information Criterion ($AIC$) (Akaike 1973) determined using Eq. (3) and (4), respectively.

$$RSS_{avg} = \frac{1}{n} \sum_{i=1}^{n} (y_i - y_i^*)^2$$

(4)

$$AIC = \ln(RSS_{avg}) + \frac{2k}{n}$$

(5)

where $n$ is the number of observations; $y_i$ and $y_i^*$ are the observation and estimation at location $i$, respectively; and $k$ is the number of regression variables. The lower $RSS_{avg}$ and $AIC$, the better the estimation quality of the model.
Results and discussions

SOC variation with land use

The descriptive statistical parameters for SOC content and the nine environmental factors from the 48 sampling points are shown in Table 1. The stock of SOC in both the 0–20 and 20–40 cm soil depths is smaller than the average for China (Wu et al. 2003; Yu et al. 2007). The low level of SOC stock was attributed to the severe soil erosion and low biomass productivity due to the low and erratic rainfall in the study area (Wang et al. 2009). The stocks of SOC were 91.48%, 113.33%, 99.25% and 79.32% higher in the 0–20 cm soil layer than the 20–40 cm soil layer under korshinsk peashrub, purple alfalfa, fallow land, and millet, respectively ($p < 0.05$, Fig. 2). The higher SOC stocks in the 0–20 cm soil layer can be related to plant and soil characteristics at the local scale as both factors affect the balance of C input from plant growth and output from plant decomposition (Wu et al. 2010; Willaarts et al. 2016).

Although SOC for all four land-use types were relatively small, restoration of vegetation (korshinsk peashrub, purple alfalfa and fallow land) still affected SOC stock to varying degrees. SOC stock was highest under korshinsk peashrub, followed by purple alfalfa, fallow land, and millet, much in agreement with the findings of Chen et al. (2007). The significantly lower SOC under purple alfalfa and fallow land than that under korshinsk peashrub can be explained by the differences in aboveground and belowground biomass (Yimer et al. 2007). For example, studies show that the maximum aboveground biomass of korshinsk peashrub is significantly larger than that of purple alfalfa (Fu et al. 2010); and that fine root biomass of korshinsk peashrub is greater than that of purple alfalfa (Fig. 3a). Purple alfalfa was cut near the ground surface, meaning that the input of SOC in shed litter was relatively small. Compared with purple alfalfa, fallow land,
and millet hardly had any effect on the increase in SOC, especially in the 20–40 cm layer at the lower slope position (Fig. 2). Studies at the same experimental plots show that runoff coefficient and erosion rate under millet are greater than those under korshinsk peashrub, purple alfalfa and fallow land (Zeng 2006). The greater soil erosion rates and the relatively high concentration of SOC in the 0–20 cm soil layer suggested that more SOC could be lost to surface runoff in the future, leading to a reduction in soil fertility under millet (Fu et al. 2010). Overall, even if purple alfalfa and fallow land had no direct effect on the increase in SOC stock as korshinsk peashrub, there was still the potential to enhance SOC sequestration in the long-term because of their ability to hold soil on sloping lands, subsequently reducing SOC loss through runoff and sediment transport.

**SOC correlation analyses**

The Pearson correlation analysis showed that SOC stock in the 0–20 and 20–40 cm soil layers was closely correlated with land use, fine root biomass, soil pH and TN (Table 2). It is interesting to note that soil pH was negatively correlated with SOC stock. Soil pH influences SOC by influencing plant nutrient uptake and biomass production, and by influencing microbial population in decomposing organic matter (Liu et al. 2012). When SOC decomposes, it releases organic acids that in turn affect soil pH (Evrendilek et al. 2004). There was no significant correlation between SOC and soil bulk density for the 0–20 cm layer, but there was a significant negative correlation for the 20–40 cm soil layer (p < 0.01). No significant correlation existed among soil saturated hydraulic conductivity, field capacity, soil texture and SOC for any of the investigated soil layers.

Although not significant, topographic factors such as relative elevation were negatively
correlated with SOC stock. Thus, the effect of topography on the spatial distribution of SOC in the 0–40 cm soil profile was generally weak. Given the scale of sampling, the limited altitudinal gradient selected in this study could have been small to detect any relationship between SOC and relative elevation (She et al. 2014).

All observations across the 48 sampling points were arrayed in one dimension for state-space analysis, and the SOC data plotted against the sampling point numbers in Fig. 4. Based on CV analysis (Table 1), SOC in the two soil layers obvious fluctuated for most of the points, with moderate variability (Nielsen and Bouma 1985). The spatial pattern of SOC was characterized by remarkable local variation, suggesting that it was better represented by state-space models than by space-independent models (Wendroth et al. 1999; Timm et al. 2003). From the correlation analyses, four variables (land use type, fine root biomass, soil pH and TN) were selected for further analyses in relation to the 0–20 cm soil layer using the state-space models. The four variables and soil bulk density were again used in analysis for SOC in the 20–40 cm soil layer.

**State-space models for SOC**

Autoregressive state-space models with boundary conditions (land use type, fine root biomass, soil pH, soil TN and/or soil bulk density) were used to determine the spatial characteristics of SOC and the main driving factors for each soil layer on the hillslopes of China’s Loess Plateau. Based on $R_{ss_{avg}}$ and $AIC$ analysis, optimal state-space equations were established for one to four regression variables (Table 3). By using all the data points, the spatial characteristics of SOC at the two soil depths were explained by the optimal state equations (Fig. 5). This was used to determine how closely the boundary conditions explained the spatial characteristics of SOC.
The accuracy of the state-space models in explaining SOC stock varied with the number of regression variables used. The state-space model with four regression variables (land use, fine root biomass, soil pH and TN) had the most negative AIC values: -11.121 and -9.634, respectively, for the 0–20 and 20–40 cm soil layers (Table 3). By adding soil bulk density into the state-space model, AIC increased by 1.784 for the 20-40 cm soil layer, indicating that soil bulk density was not suitable for the prediction of SOC stock. The standard error at the 95% confidence interval for the value of each estimated variable at position i was ±2 (Fig. 5). The closer the value to the 95% confidence interval limit of SOC was lower for the 0–20 cm than the 20–40 cm soil layer, indicating that the prediction accuracy was better for surface soils than the sub-surface soils (Jia et al. 2017). R² of the optimal state-space models for the 0–40 cm soil profile was 0.967–1.000 (p < 0.01; Table 3). This suggested that the state-space models explained 96.7–100.0% of the spatial variations in SOC stock along the hillslope of the Loess Plateau. It was therefore concluded that the model functions well described the variations in SOC in different soil layers and with a good predictive ability as indicated by AIC and R² values.

For instance, the spatial pattern of SOC in the 0–20 cm soil layer was described by the state equation as: 
\[
SOC_i = -0.215SOC_{i-1} + 1.078LU_{i-1} + 0.302TN_{i-1} - 0.064pH_{i-1} - 0.070FRB_{i-1} + w_i.
\]
This equation comprises of the preceding values of soil organic carbon (SOC_{i-1}) up to 0.215, land use type (LU_{i-1}) up to 1.078, soil total nitrogen (TN_{i-1}) up to 0.302, soil pH (pH_{i-1}) down to -0.064 and fine root biomass (FRB_{i-1}) up to 0.070 (Table 3, Fig. 5). The coefficient before each variables was obtained from the transition matrix \( \phi_{pp} \). As all the variables included in the analysis were normalized, the magnitudes of the transition coefficients for the autoregressive equations directly
reflected their relative contributions to the estimated SOC stock (Timm et al. 2003; Yang and Wendroth 2014). Therefore, to compare the effect of each boundary condition on the spatial characteristics of SOC in different soil layers, the relative contributions were calculated (Fig. 6).

The impact of land use type, soil total nitrogen, soil pH and fine root biomass on SOC stock was included in the description for the 0-40 cm soil layer. The sum of the relative contributions of the four factors was 0.876 and 1.000 for the 0–20 cm and 20–40 cm soil layers, respectively. The effect of land use on SOC in the 0–40 cm soil layer was greater than that of the other factors. This implied that the spatial pattern of SOC in shallow soil layers (0–40 cm) along the slopes of the Loess Plateau was mainly driven by land use type. The corresponding weights for the 0–20 and 20–40 cm soil layers were 0.623 and 0.880, respectively. The increasing effect of land use on SOC with increasing depth of the soil profile was probably related to plant root distribution in shrubland, grassland and cropland areas (Jia et al. 2017). This observation was consistent with that of Wang et al. (2015b), where land use (probably by differences in root distribution patterns) significantly influenced SOC distribution in deep soil layers. Generally, crops have the shallowest root profile on the Loess Plateau, followed by grass and shrubs. For example, in the study, CV (117.2 %) for fine root biomass in the 20–40 cm soil layer was higher than that (66.9%) in the 0–20 cm soil layer (Table 1). Another possible reason for the increasing effect of land use on SOC with increasing soil depth was differences in soil water content. Jia et al. (2016) noted that spatial variations in soil water content increase with increasing soil depth along the transect of the Loess Plateau. In this study therefore, land use significantly influenced the spatial pattern of SOC in deep soil layers; suggesting that it was a critical factor in SOC analysis in deep soil layers.

Soil TN, pH and fine root biomass contributed little to the state-space model estimated SOC.
distribution in the 0‒40 cm soil layer. The corresponding weights of the variables on SOC stock dropped from 0.175 to 0.029. More importantly, soil pH can determine the communities and dynamics of soil microorganisms, which control the transformation process of organic carbon to nitrogen in soils (Baath and Anderson 2003; Aciego Pietri and Brookes 2008). It was concluded that land-use type was the dominant factor driving the spatial patterns of SOC in shallow (0‒40 cm) soil layers. This is critical for estimating C flux and C stock in shallow soil layers.

Conclusions

Generally, the rank order of SOC stock in the 0‒40 cm soil layer under the investigated land-use types was: shrubland > grassland > fallowland > cropland. This suggested that shrubland was more favorable than grassland and fallowland for SOC sequestration in the semi-arid study area. The stock of SOC was higher in the 0‒20 cm than in the 20‒40 cm soil layer under the four land use types. An optimal state-space model that with no more than four environmental factors (land-use type, soil TN, soil pH and fine root biomass) was established for accurate determination of SOC in each soil layer. All the state-space models well described the spatial characteristics of SOC, explaining over 96% of its variations in the hillslopes of China’s Loess Plateau. Based on the weight and relative contribution of each variable, land use type was the dominant factor driving spatial variability of SOC stock in the 0‒40 cm soil layer, which effect increased with increasing soil depth. The state-space approach comprehensively explained the variations in SOC along the hillslope of China’s Loess Plateau region. It provided a data-guided insight into the requirements for the accurate estimation of SOC using easy-to-obtain variables for state-space model simulation.
Acknowledgements

This study was supported by the National Key Project for Research and Development (2016YFC0501605), the National Natural Science Foundation of China (41571130081 and 41390461), the Open Research Fund of the State Key Laboratory of Soil Erosion and Dryland Farming on the Loess Plateau (A314021402-1503), the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2017076), and the Program for Bingwei Excellent Talents from the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences (2015RC204). We thank the editors and reviewers for the constructive comments on the manuscript.

References


Table captions

Table 1 Descriptive statistics of soil organic carbon (SOC) and the related variables for the 48 sampling points. BD: soil bulk density; Ks: soil saturated hydraulic conductivity; FC: field capacity; TN: Total nitrogen; FRB: fine root biomass; SD: standard deviation; and CV: coefficient of variation.

Table 2 Pearson correlation analyses for soil organic carbon (SOC) stock in two soil layers as driven by land use (LU) and environmental variables.

Table 3 Optimal state-space equation for SOC stock in each soil layer and the respective RSS$_{avg}$, $AIC$ and $R^2$ based on land use (LU), soil bulk density (BD), soil total nitrogen (TN), soil pH and fine root biomass (FRB). SOC$_{i-1}$ indicates that SOC in the objective layer at location $i$-1; $k$ is the number of regression variables, and ** indicates significance at $p < 0.01$. 


Table 1 Descriptive statistics of soil organic carbon (SOC) and the related variables for the 48 sampling points. BD: soil bulk density; Ks: soil saturated hydraulic conductivity; FC: field capacity; TN: Total nitrogen; FRB: fine root biomass; SD: standard deviation; and CV: coefficient of variation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Soil depth (cm)</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOC stocks</td>
<td>0-20</td>
<td>3.9</td>
<td>15.5</td>
<td>7.8</td>
<td>7.3</td>
<td>2.4</td>
<td>30.6</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>2.6</td>
<td>7.5</td>
<td>3.5</td>
<td>4.0</td>
<td>1.1</td>
<td>29.4</td>
</tr>
<tr>
<td>Clay (%)</td>
<td>0-20</td>
<td>5.5</td>
<td>13.0</td>
<td>8.9</td>
<td>9.0</td>
<td>1.6</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>6.1</td>
<td>13.5</td>
<td>9.3</td>
<td>9.4</td>
<td>1.8</td>
<td>20.0</td>
</tr>
<tr>
<td>Silt (%)</td>
<td>0-20</td>
<td>27.3</td>
<td>52.0</td>
<td>39.6</td>
<td>40.0</td>
<td>5.5</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>29.0</td>
<td>51.2</td>
<td>40.5</td>
<td>40.7</td>
<td>5.4</td>
<td>13.5</td>
</tr>
<tr>
<td>Sand (%)</td>
<td>0-20</td>
<td>36.1</td>
<td>67.1</td>
<td>51.3</td>
<td>50.9</td>
<td>6.9</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>35.3</td>
<td>64.4</td>
<td>50.0</td>
<td>49.8</td>
<td>7.1</td>
<td>14.3</td>
</tr>
<tr>
<td>BD (g cm⁻³)</td>
<td>0-20</td>
<td>1.2</td>
<td>1.5</td>
<td>1.4</td>
<td>1.4</td>
<td>0.1</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>1.3</td>
<td>1.5</td>
<td>1.4</td>
<td>1.4</td>
<td>0.1</td>
<td>3.4</td>
</tr>
<tr>
<td>Ks (cm h⁻¹)</td>
<td>0-20</td>
<td>0.4</td>
<td>2.7</td>
<td>1.2</td>
<td>1.2</td>
<td>0.4</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>0.3</td>
<td>1.8</td>
<td>0.9</td>
<td>0.9</td>
<td>0.4</td>
<td>40.6</td>
</tr>
<tr>
<td>FC (%)</td>
<td>0-20</td>
<td>19.8</td>
<td>32.1</td>
<td>24.5</td>
<td>24.3</td>
<td>2.3</td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>18.1</td>
<td>36.8</td>
<td>23.0</td>
<td>23.1</td>
<td>2.8</td>
<td>12.2</td>
</tr>
<tr>
<td>TN (g kg⁻¹)</td>
<td>0-20</td>
<td>0.1</td>
<td>0.5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>28.3</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>0.11</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>33.0</td>
</tr>
<tr>
<td>pH</td>
<td>0-20</td>
<td>8.3</td>
<td>8.8</td>
<td>8.6</td>
<td>8.5</td>
<td>0.1</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>8.4</td>
<td>8.9</td>
<td>8.6</td>
<td>8.6</td>
<td>0.1</td>
<td>1.8</td>
</tr>
<tr>
<td>FRB (g m⁻²)</td>
<td>0-20</td>
<td>1.7</td>
<td>226.8</td>
<td>81.7</td>
<td>71.0</td>
<td>54.7</td>
<td>66.9</td>
</tr>
<tr>
<td></td>
<td>20-40</td>
<td>1.1</td>
<td>95.1</td>
<td>19.1</td>
<td>11.8</td>
<td>22.4</td>
<td>117.2</td>
</tr>
</tbody>
</table>
Table 2: Pearson correlation analyses for soil organic carbon (SOC) stock in two soil layers as driven by land use (LU) and environmental variables.

<table>
<thead>
<tr>
<th>Soil depth</th>
<th>LU</th>
<th>RE</th>
<th>clay</th>
<th>silt</th>
<th>sand</th>
<th>BD</th>
<th>K_s</th>
<th>FC</th>
<th>TN</th>
<th>pH</th>
<th>FRB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-20 cm</td>
<td>-0.523**</td>
<td>-0.209</td>
<td>-0.182</td>
<td>-0.106</td>
<td>0.120</td>
<td>-0.080</td>
<td>-0.091</td>
<td>-0.167</td>
<td>0.984**</td>
<td>-0.398**</td>
<td>0.433**</td>
</tr>
<tr>
<td>20-40 cm</td>
<td>-0.470**</td>
<td>-0.155</td>
<td>0.105</td>
<td>0.150</td>
<td>-0.143</td>
<td>0.534**</td>
<td>-0.08</td>
<td>-0.172</td>
<td>0.925**</td>
<td>-0.385**</td>
<td>0.355*</td>
</tr>
</tbody>
</table>

* Significant correlation at the 0.05 probability level (2-tailed); ** Significant correlation at the 0.01 probability level (2-tailed); RE is relative elevation; FRB is fine root biomass; BD is soil bulk density; K_s is soil hydraulic conductivity; FC is soil field capacity; and TN is total nitrogen.
Table 3 Optimal state-space equation for SOC stock in each soil layer and the respective RSS_{avg}, AIC and $R^2$ based on land use (LU), soil bulk density (BD), soil total nitrogen (TN), soil pH and fine root biomass (FRB). SOC$_{i-1}$ indicates that SOC in the objective layer at location $i$-1; k is the number of regression variables, and ** indicates significance at $p < 0.01$.

<table>
<thead>
<tr>
<th>k</th>
<th>state equation</th>
<th>RSS$_{avg}$</th>
<th>AIC</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20-40 cm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$SOC_i = -0.777SOC_{i-1} + 1.671TN_{i-1} + w_i$</td>
<td>0.0002</td>
<td>-8.475</td>
<td>0.997**</td>
</tr>
<tr>
<td>2</td>
<td>$SOC_i = -0.267SOC_{i-1} + 0.150TN_{i-1} + 1.081FRB_{i-1} + w_i$</td>
<td>0.0001</td>
<td>-8.980</td>
<td>0.998**</td>
</tr>
<tr>
<td>3</td>
<td>$SOC_i = -0.201SOC_{i-1} + 1.025LU_{i-1} + 0.284TN_{i-1} + 0.077FRB_{i-1} + w_i$</td>
<td>0.0001</td>
<td>-9.681</td>
<td>0.999**</td>
</tr>
<tr>
<td>4</td>
<td>$SOC_i = -0.215SOC_{i-1} + 1.078LU_{i-1} + 0.302TN_{i-1} - 0.064pH_{i-1} - 0.070FRB_{i-1} + w_i$</td>
<td>0.0000</td>
<td>-11.121</td>
<td>1.000**</td>
</tr>
<tr>
<td></td>
<td>20-40 cm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$SOC_i = 0.546SOC_{i-1} + 0.456FRB_{i-1} + w_i$</td>
<td>0.0201</td>
<td>-3.863</td>
<td>0.967**</td>
</tr>
<tr>
<td>2</td>
<td>$SOC_i = 0.049SOC_{i-1} + 0.992LU_{i-1} + 0.008FRB_{i-1} + w_i$</td>
<td>0.0003</td>
<td>-8.038</td>
<td>0.997**</td>
</tr>
<tr>
<td>3</td>
<td>$SOC_i = 0.066SOC_{i-1} + 1.014LU_{i-1} + 0.029BD_{i-1} - 0.020TN_{i-1} + w_i$</td>
<td>0.0001</td>
<td>-9.114</td>
<td>0.999**</td>
</tr>
<tr>
<td>4</td>
<td>$SOC_i = -0.001SOC_{i-1} + 1.066LU_{i-1} + 0.035TN_{i-1} - 0.090pH_{i-1} + 0.021FRB_{i-1} + w_i$</td>
<td>0.0001</td>
<td>-9.634</td>
<td>0.999**</td>
</tr>
<tr>
<td>5</td>
<td>$SOC_i = 0.009SOC_{i-1} + 1.071LU_{i-1} - 0.038BD_{i-1} + 0.008TN_{i-1} - 0.062pH_{i-1} + 0.044FRB_{i-1} + w_i$</td>
<td>0.0003</td>
<td>-7.850</td>
<td>0.995**</td>
</tr>
</tbody>
</table>

** indicates significant at $p < 0.01$. 

https://mc.manuscriptcentral.com/cjss-pubs
Figure captions

**Fig. 1** Sampling sites of different soil layers arrayed in one dimension for state-space approach study. From left to right, KOP denotes korshinsk peashrub, ALF is purple alfalfa, NAF is natural fallow and MIL is millet.

**Fig. 2** Soil organic carbon (SOC) stock under KOP (korshinsk peashrub), ALF (purple alfalfa), NAF (natural fallow) and MIL (millet) in different soil layers. Significant differences among land-use types for the same soil layer are denoted by different lowercase letters. The least significant difference test was done at $p < 0.05$.

**Fig. 3** Fine root biomass (FRB) and soil pH in different soil layers under KOP (korshinsk peashrub), ALF (purple alfalfa), NAF (natural fallow land) and MIL (millet) land-use types. Significant differences among land-use types within the same soil layer are denoted by different lowercase letters. The least significant difference test was done at $p < 0.05$.

**Fig. 4** Spatial distribution of SOC among 48 sampling points in one dimensional array.

**Fig. 5** Spatial characteristics of SOC stock in the 0–40 cm soil layer as determined by the optimal state-space model driven by land use (LU), fine root biomass (FRB), soil pH and total nitrogen (TN). The respective state equations and AIC values are also given.

**Fig. 6** Relative contribution of each variable in the optimal state-space model for SOC stock in each soil layer. LU is land use, TN is soil total nitrogen, and FRB is fine root biomass.
Fig. 1
Fig. 2

![Image of a bar chart showing soil organic carbon (SOC) at different depths and treatments.

- Legend:
  - KOP
  - ALF
  - NAF
  - MIL

- X-axis: Depth (cm)
  - 0-20
  - 20-40

- Y-axis: SOC (Mg ha⁻¹)

- Data points labeled with letters to indicate significant differences.
Fig. 3

[Graph showing data for FFB (g m$^{-2}$) and pH at different depths (0-20 and 20-40 cm). The graph compares KOP, ALF, NAF, and MIL treatments. Significant differences are indicated by letters (a, b, c, d) above the bars.]

https://mc.manuscriptcentral.com/cjss-pubs
Fig. 5

\[
\text{SOC}_i = -0.215 \text{SOC}_{i-1} + 1.078 \text{LU}_{i-1} - 0.070 \text{FRB}_{i-1} - 0.064 pH_{i-1} + 0.302 TN_{i-1} w_i
\]

- Observed
- Estimated
- Estimated ± 2SE

0-20 cm

\[
\text{SOC}_i = -0.001 \text{SOC}_{i-1} + 1.066 \text{LU}_{i-1} - 0.021 \text{FRB}_{i-1} - 0.090 pH_{i-1} + 0.035 TN_{i-1} w_i
\]

20-40 cm
Fig. 6