Geoderma 314 (2018) 8–19

Contents lists available at ScienceDirect

Geoderma

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

Spatial variations of soil organic carbon stocks in a coastal hilly area of China

Shuai Wang[a,b], Qianlai Zhuang[b,*], Shuhai Jia[a], Xinxin Jin[a,b], Qiubing Wang[a,**]

[a] College of Land and Environment, Shenyang Agricultural University, Shenyang, 110866, Liaoning Province, China
[b] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, USA

A R T I C L E I N F O

Handling Editor: A.B. McBratney

Keywords:
Soil organic carbon
Geographically weighted regression
Spatial variability

A B S T R A C T

Quantification of soil organic carbon (SOC) stocks and their spatial variations at regional scales is a foundation to adequately assess plant productivity and soil carbon sequestration potentials so as to establish better practices for land use and land management. This study evaluated the spatial variation of SOC stocks from 1982 to 2012 in Wafangdian, Liaoning Province, China. To map SOC stock, we used geographically weighted regression (GWR) and regression kriging (RK) methods and a large set of soil samples, in which nine topographic and remote sensing variables were observed. The GWR approach performed better than the RK approach as the former has smaller absolute mean errors (AME), mean errors (ME), root mean square errors (RMSE) in comparison with observational data. Our results indicated that SOC stocks have an increasing trend in northeast and southwest mountainous areas in our study periods. Land-use changes caused by returning cultivation land to forest promoted SOC accumulation. The total SOC stocks of cultivation land, grasslands and forests within 0–0.2 m of soils were estimated to be 5.25 and 5.40 Tg in 1982 and 2012, respectively. This study provided important information of spatial variations in SOC stocks to agencies and communities in this region to evaluate soil quality and assess carbon sequestration potentials and carbon credits.

1. Introduction

Soil has been recognized as a large sink of atmospheric CO_2 (Scholes and Andreae, 2000; Wang et al., 2004). Carbon storage within 1 m of soil depth is about twice more carbon than stored in the atmosphere (Watson et al., 2000; Kumar et al., 2012). SOC is a vital constituent in carbon capture and storage to alleviate rising atmospheric CO_2 concentrations. Globally, soils stored about 1500 Pg C (1 Pg = 10^{15} g) within 1 m depth (Lal, 2004). In addition, estimation of SOC stock is also important to assessing soil quality and plant productivity under a changing climate so as to develop effective land management policies (Jobbagy and Jackson, 2000; Mondini and Sequi, 2008; Don et al., 2011; Li et al., 2012). Cost-efficient techniques for mapping SOC stock are therefore indispensable (Mishra et al., 2010; Wang et al., 2016; Minasny and McBratney, 2016).

Geographic or purely spatial approaches have been used to predict soil properties at un-sampled locations since the late 1960s (McBratney et al., 2003). SOC is affected by both natural vegetation and human activities (Elbasiouny et al., 2014). However, due to spatial heterogeneity and lack of extensive sampling data, some approaches are often not capable of accurately mapping C stocks (Batjes, 1996; Wang et al., 2016; Wang et al., 2017). Since the advent of geographic information systems (GIS) and high-precision remote sensing data, climate data, terrain data and those derived variables have been widely used to estimate SOC stock (Kumar et al., 2012). Multiple linear regression (MLR), regression kriging (RK), and ordinary cokriging (OCK) are often combined with these auxiliary environmental variables to map soil properties (Robinson and Metternicht, 2006; Grimm et al., 2008). Consequently, the selection of prediction variables is one of the necessary steps to accurately map SOC stocks (Mishra et al., 2010).

Spatial variability of SOC stocks can be estimated by using various techniques, which can be merged into two categories: (1) the measure and multiply model (MMA), and (2) the soil landscape modeling (SLM) model. In the MMA model, an average SOC stock is allocated to each map unit of soil type or land-use type in an area (Batjes, 1996; Bernoux et al., 2002; Guo et al., 2006). However, this approach results in constant values within each map unit that cannot show its large spatial heterogeneity of SOC stock and the error of estimated SOC are due to using a few SOC stock data points. In contrast, the SLM model can produce more detailed spatial variations of SOC stocks with assistance

* Correspondence to: Q. Zhuang, Purdue University, West Lafayette, IN 47907, USA.
** Correspondence to: Q. Wang, College of Land and Environment, Shenyang Agricultural University, No. 120 Dongling Road, Shenhe District, Shenyang, Liaoning Province 110866, China.
E-mail addresses: qzhuang@purdue.edu (Q. Zhuang), wangqbsy@yahoo.com (Q. Wang).

http://dx.doi.org/10.1016/j.geoderma.2017.10.052
Received 31 May 2017; Received in revised form 13 October 2017; Accepted 28 October 2017

of auxiliary environmental variables including topography, climate, vegetation, and remote sensing imagery. Compared with the MMA model that does not consider the effects of environment variables in the study area, the SLM model has lower prediction errors (Tompson and Kolka, 2005; Mishra et al., 2010).

Since the late 1990s, a simple approach known as geographically weighted regression (GWR) has attracted much attention and was introduced for the study of digital soil mapping (DSM) (Brunsdon et al., 1996; Fotheringham et al., 2002; Song et al., 2016). GWR can be seen as an extension of a spatial non-stationarity regression approach at different locations (Kumar et al., 2012). Compared to a traditional regression model, GWR is more powerful and efficient (Song et al., 2016). Specifically, GWR is an extension of the traditional multiple linear regression toward a local regression, in which regression coefficients are specific to a location rather than being globally estimated. This model provides a flexible parameter estimation method for the spatial non-stationarity of regression coefficients between the target variable and explanatory variables by measuring those coefficients locally using local data. Owing to these merits, GWR has been applied to explore the spatial relationships among the environmental variables (Kumar et al., 2012), estimate complex spatial variation in parameters (Kumar et al., 2012), model spatially heterogeneous processes (Lloyd, 2010; Mishra et al., 2010; Song et al., 2016), and forecast the SOC stock (Mishra et al., 2010; Kumar et al., 2012; Wang et al., 2012; Song et al., 2016).

Using GWR for SOC stock mapping has been applied in various studies at different scales. Mishra et al. (2010) compared three models of GWR, MLR and RK in the Midwest of the United States. In those studies GWR outperformed MLR and RK. GWR caused a reduction in root mean square errors (RMSE) of 22% and 2% over MLR and RK. In China, Wang et al. (2013) compared the prediction performance of GWR and MLR and showed that the RMSE was reduced by 11%. Song et al. (2016) compared GWR to MLR, geographically weighted ridge regression (GWRR), kriging with an external drift (KED), and GWR plus ordinary kriging of model residuals (GWRSK) for predicting the spatial distribution of SOC in the Heihe basin, China. Eventually, they found that GWR better captured the spatial variability of SOC and improving its prediction accuracy.

Geostatistical models based on global regression coefficients are not absolutely inferior to GWR model (Lloyd, 2010; Harris and Juggins, 2011; Song et al., 2016). It has not been shown if GWR model outperforms the RK model (Song et al., 2016). The RK model parameters are determined using the restricted maximum likelihood (REML) method with two separate steps: (1) using the least square method to determine the regression coefficient; (2) using method-of-moments from the regression model residuals to determine the variogram parameters. These two steps are iterated to achieve the best fitting. This process produces suboptimal parameters so as to produce suboptimal prediction results (Song et al., 2016). Therefore, comparing GWR with RK is essential to evaluating the benefits of local regression coefficients in mapping SOC stock.

This study used a GWR approach to evaluate the spatial variability of the SOC stocks in topsoil (0–20 cm) at a regional scale. The specific objectives were to: (1) map SOC stocks in 1982 and 2012; (2) compare the performance of GWR and RK models; and (3) investigate temporal dynamics of SOC stocks from 1982 to 2012.

2. Materials and methods

2.1. Site description

This study was conducted in Wafangdian, Liaoning province, China (121°13′-122°16′ E, 39°20′- 40°07′) (Fig. 1), covering a total area of 3827 km^2. Seventy-one percent of the study area was under agriculture and the rest mainly for garden plots and urban land. The chief crops of study area are corn, rice, and sorghum in the mid-west plain region, and fruit orchards in the upland areas. The elevation of this area increased from southwest to northeast, with a range from 0 m to 772 m above sea level. The study region has warm temperate continental monsoon climate, and it is the warmest area in the Northeast of China. The annual mean temperature (MAT) is 9.3 °C, with the highest temperature of 37.8 °C in summer and the lowest temperature of 19.3 °C in winter. The annual mean precipitation (MAP) ranges from 580 to 750 mm and 60%–70% of the MAP is in the rainy season (June–August), accompanied by heavy rainfall. Garden and forest lands are the main types that are suitable for re-development. However, soil fertility is poor or medium (Wang et al., 2016). The main geomorphic units are characterized by complex and undulating hills systems intersected by river valleys. According to the classification of World Reference Base for Soil Resources (WRB) (IUSS Working Group, 2014), the dominant soil types are Cambisols (58%) and Fluvisols (13%) in the study area.

2.2. Soil sampling

2.2.1. Soil survey data in 1982

Typical soil profiles were obtained from the Second National Soil Survey of Liaoning Province conducted between 1979 and 1990 (OSSLP, 1990). Soil profile data include information on parent material, cropping system, land use and soil physical and chemical properties. However, our research only focused on the topsoil (0–20 cm) SOC and bulk density (BD). A total of 978 topsoil data was obtained to represent all soil types and land use types in the study area, and we randomly selected 80% of these as the training data (782), and the remaining were the testing data (196). The create-subset function in the geostatistical module of ArcGIS 10.2 (ESRI Inc., USA) software was used for training and testing the model. The unavailable measurements of soil bulk density (BD) were calculated from SOC content using a pedotransfer function (PTF):

$$BD = 1.46 - 0.09^{*}\sqrt{SOC} \quad (R^2 = 0.78, P < 0.001) \tag{1}$$

2.2.2. Soil sampling in 2012

A total of 1195 (956 for training, 239 for testing) topsoil (0–20 cm) samples were collected in a new survey on a 1.6 × 1.6 km grid across the study area in 2012 (Fig. 1, right). The coordinates of sampling sites were determined by a hand-held Global Positioning System (GPS). Each sample site was a mixed sample based on the four corners and center points of the 1 × 1 m square. A subsample of 1 kg per mixed sample was isolated for laboratory analysis. SOC content of the samples was determined by a wet oxidation method (Walkley–Black method) (Nelson and Sommers, 1982) in Key Laboratory of Agricultural Resources and Environment of Liaoning Province, Shenyang Agricultural University. To estimate dry bulk density, 100 cm^3 of undisturbed soil cores were collected from topsoil layers and then were dried for 48 h at 105 °C for bulk density measurement.

2.3. Environmental variables

A suite of 9 environmental covariates representing topographic and remote sensing variables were used as predictors in this study. Environmental variables were collected and converted to raster data through ArcGIS 10.2 (ESRI Inc., USA). Considering the widespread extent of the data, we believed that covariates at a 30 × 30 m resolution were sufficient to meet our needs.

2.3.1. Topographic variables

Digital elevation model (DEM) data covering 30 × 30 m resolution of the entire study area were obtained from the United States Geological Survey (USGS, Reston, VA, USA). The elevation gradient varies from 0 m to 722 m. The low-elevation area is mainly in the west and southwest coastal areas (0 m), and the corresponding high-elevation areas are mainly in the northeast mountain area (722 m). Three