Ecological Indicators 84 (2018) 263–272

Contents lists available at ScienceDirect

Ecological Indicators

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

Research paper

Role of environmental variables in the spatial distribution of soil carbon (C), nitrogen (N), and C:N ratio from the northeastern coastal agroecosystems in China

Shuai Wang[a], Kabindra Adhikari[b], Qiubing Wang[a,*], Xinxin Jin[a], Hongdan Li[a]

[a] Shenyang Agricultural University, College of Land and Environment, No. 120 Dongling Road, Shenhe District, Shenyang, Liaoning 110866, China
[b] University of Arkansas, Department of Crop, Soil, and Environmental Sciences, Fayetteville, AR 72701, United States

ARTICLE INFO

Keywords:
Digital soil mapping
Boosted regression trees
Soil organic carbon
Total nitrogen
C:N ratio
Coastal agroecosystems

ABSTRACT

Soil organic carbon (SOC) and total nitrogen (TN) influence physical, chemical and biological properties and process in soils that determine soil fertility and enhance or maintain agricultural productivity and food security. Knowledge on the spatial distribution of SOC and TN is necessary for sustainable soil resource management. In this study, we applied boosted regression trees (BRT) model to predict and map the spatial distribution of SOC and TN in the northeastern coastal areas of China, and C:N ratio map was generated from the predicted maps of SOC and TN. A total of 149 topsoil (0–20 cm) samples and 12 environmental variables that included topography, landuse and remote sensing indices, and climatic variables were selected. The performance of the models was evaluated based on cross-validation using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2). The BRT model was run for 100 iterations where the average of 100 SOC and TN maps was considered as final predicted maps and its standard deviation as prediction uncertainty. The BRT model showed a good predictive performance due to its higher R^2 and lower MAE, and RMSE indices. The model explained approximately 64% and 56% of the total SOC and TN variability. Topographic variables showed a maximum influence in the prediction of SOC and TN, followed by vegetation and climate. The SOC and TN contents were reduced from the northwest towards the southeast of the study area; average predicted SOC and TN contents were 15.4 g kg^{-1} and 1.01 g kg^{-1}, respectively. The spatial distribution of C:N ratio was closely related to landuse types in a decreasing order: woodland > orchard > cultivated land > grassland. We found topographic variables as main environmental indicators of SOC and TN distribution, and recommend to include them in future SOC and TN mapping studies in the coastal agroecosystems in China.

1. Introduction

Soil organic carbon (SOC) and total nitrogen (TN) influence physical, chemical and biological properties and process in soils that determine soil fertility and enhance and maintain agricultural productivity and food security (Batjes, 1996; Reeves, 1997; Wang et al., 2011; Lal, 2004). SOC and TN constitute a major terrestrial carbon (C) and nitrogen (N) pool (Bronson et al., 2004; Wang et al., 2011; Winowiecki et al., 2016). Its decomposition can increase atmospheric CO_2 and N_2O levels, while it‘s sequestration can help mitigate global climate change (Powers and Schlesinger, 2002; Bronson et al., 2004; Lal, 2004; Noponen et al., 2013). Ratio of SOC and TN (C:N ratio hereafter) can be a sensitive index of soil quality and also an index to measure the balance of SOC and TN in soils. In addition, the spatial variation of SOC, TN, and C:N ratio in an agroecosystem influences regional SOC and TN balances and may contribute to global C and N cycles (Huang et al., 2006; Elbasiouny et al., 2014; Deng et al., 2016). Therefore, there has been an interest to understand spatial variations and controlling factors of SOC, TN, and C:N ratio for soil quality, SOC accounting, and environmental applications (Jobbágy and Jackson, 2000; Wiesmeier et al., 2014; Yang et al., 2016a; Wang et al., 2016).

The spatial distribution of SOC, TN and C:N ratio in an agroecosystem is not only affected by natural ecological processes but is also impacted by several factors, including climate, soil type, topography and land use etc., thus posing great challenges for accurate predictive mapping at regional scales (Wang et al., 2016; Yang et al., 2016). Digital soil mapping (DSM) provides a rapid and inexpensive way to estimate soil properties from sparse sampling points into continuous surfaces (McBratney et al., 2003). Most DSM methods are based on soil landscape models, which build quantitative relationships between soil

* Corresponding author.
E-mail address: wangqbsy@yahoo.com (Q. Wang).

http://dx.doi.org/10.1016/j.ecolind.2017.08.046
Received 17 May 2017; Received in revised form 18 August 2017; Accepted 21 August 2017
Available online 25 September 2017

attributes and easily obtained environmental covariates such as elevation, precipitation, temperature, landuse, parent material etc. Numerous DSM techniques have been used to estimate SOC and TN, including geographically weighted regression (Mishra et al., 2010), mixed linear regression (Zhao et al., 2014), random forest (RF) model (Wiesmeier et al., 2011; Yang et al., 2016a), artificial neural networks (Were et al., 2015), kriging (Elbasiouny et al., 2014), boosted regression tree (BRT) model (Martin et al., 2014; Yang et al., 2016a,b; Wang et al., 2016), and regression kriging based on regression rules (Adhikari et al., 2014). However, selecting a best prediction model in a given agroecosystem has always been a great challenge in DSM.

Of the DSM techniques mentioned above, tree models have been widely used to map the spatial distribution of SOC and TN (Yang et al., 2016a; Wang et al., 2016; Martin et al., 2011). However, due to improper parameter settings, simplicity rules and tree instability, appropriating single trees are difficult to construct. These problems have been effectively avoided since the development of bagging, boosting and random methods in tree-based models (Skurichina and Duin, 2002). The BRT model is one of the tree-based models that combine a large number of simple trees into an influential tree model (Martin et al., 2011). Compared to the single-tree model, BRT is more powerful and efficient (Wang et al., 2016). As a data mining method, the BRT model has several advantages that are flexible to deal with, can handle linear, polynomial, exponential, logistic, periodic, or general nonlinear problem, thus averted over-fitting of the model and improved prediction effectively (Carslaw and Taylor, 2009; Yang et al., 2016; Heung et al., 2016). Owing to these advantages, BRT has been used in various research areas such as environmental science (Carslaw and Taylor, 2009), ecology (Müller et al., 2013), remote sensing (Vasques et al., 2014; Dhingra et al., 2014), microbiology (Yang et al., 2015), epidemiology (Cheong et al., 2014), and soil science (Martin et al., 2014; Wang et al., 2016). Despite the use of BRT in various fields, its application in SOC and TN predictions are still limited and rarely reported in the literature.

This study aimed to characterize the spatial variations of SOC, TN, and C:N radio from the northeastern coastal agroecosystems of China. The specific objectives included:

(1.) To develop a BRT model to quantify spatial relationship between SOC, TN contents and environmental variables;
(2.) To determine key environmental variables in the prediction of SOC and TN content;
(3.) To map the spatial distribution of SOC, TN and C:N ratio in Dalian City.

2. Materials and methods

2.1. Description of the study area

The study area is located in the southern tip of Liaodong Peninsula (latitude: 39.02° − 39.07° N, longitude: 121.73° − 121.82° E) in northeastern China. It spreads between Bohai and Huanghai Seas covering a total area of 13,237 km^2 (Fig. 1). The study area is dominated by a warm temperate continental monsoon climate with four distinct seasons (namely spring, summer, autumn and winter). Mean annual precipitation ranges from 550 mm in the northeast to 950 mm in the southwest, 60–70% of which falls between June and August, mostly as a high intensity rainstorm. Mean annual temperature is 10.5 °C, with mean maximum and minimum temperatures 37.8 °C and −19.1 °C, respectively. Annual total solar radiation is 143.3 kcal cm^{-2} with an average annual sunlight between 2600–2900 h and frost free period is 165-185 days. Surrounded by a range of mountains, the landform types of the study area include mountainous region, hills and coast landforms, and hills and gullies. The elevation ranges between 0–1127 m above sea level.

Eutric Cambisols and Cutanic Luvisols (IUSS Working Group WRB, 2014) are the main soil types found in the study area and are mostly distributed along the Low Mountain and hilly areas. Major landuse types include woodland, orchard, cultivated land and grasslands. Cultivated land is more common in the valleys and coastal plains and it accounts for 22.8% of the study area. The main economic crops grown include rice, sweet potatoes, cherries and apples.

2.2. Soil sampling and laboratory analysis

Soil sampling was performed between August to September 2011 considering variations in elevation, climate, and parent material during the identification of sampling locations. A total of 149 soil samples were collected from the topsoil depth (0–20 cm) excluding litter layer, if present. Geographical coordinates of each sampling site were recorded with a hand-hold GPS (global positioning system). The samples were transported to the Key Laboratory of Agricultural Resources and Environment of Liaoning Province, Shenyang Agricultural University, Shenyang, China and were air dried at room temperature. Samples were then crushed and passed through a 2-mm sieve removing non-soil materials like gravel, plant roots etc. The fine earth fraction was then sieved through a 0.147-mm sieve to prepare samples for soil organic carbon (SOC) and total nitrogen (TN) content determination. SOC and TN content (g kg^{-1}) were determined by dry combustion using a Vario EL III elemental analyzer (Elementar Analysen systeme GmbH, Hanau, Germany).

2.3. Environmental variables

Environmental variables used in this study included topographic and climatic variables derived from a digital elevation model (DEM), vegetation indices derived from remote sensing (RS) data. These variables were considered as environmental indicators that influence the spatial distribution of SOC and TN content in the study area. They were collected from different sources and were converted to a raster grid of 90 m resolution used in this study. Measured data on SOC and TN, and all the predictors were brought into a geographic information system (GIS) in a common projection system (Krasovsky_1940_Albers) for further geospatial processing and analysis.

2.3.1. Topographic variables

Topographic variables were derived from a 90 m digital Elevation Model (DEM) which was acquired from Shuttle Radar Topography Mission (SRTM). The DEM was downloaded from Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn). From the DEM three primary terrain attributes ¨Delevation, slope aspect and slope gradient`¨Dand two secondary terrain attributes ¨Dtopographic wetness index (TWI), and catchment area (CA) ¨Dwere derived. All the geospatial analysis including DEM processing and extraction of terrain attributes were done in ArcGIS 10.1 (Environmental Systems Research Institute, Redlands, CA, 2012) and SAGA GIS (Conrad et al., 2015) environment.

2.3.2. Climatic variables

Climatic variables used in this study came from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/) and it included mean annual precipitation (mm), and maximum and minimum annual average temperatures (°C). These data were derived as 1-km grid layers interpolated from point measurements collected at 14 meteorological stations around and within the study area and were downscaled to 90 m grid to use in the study. Interpolation was based on ordinary kriging.

2.3.3. Vegetation index

Normalized Difference Vegetation Index (NDVI) was the main source of vegetation index used in this study. NDVI was derived from Landsat 5 TM imagery provided by the International Scientific & Technical Data Mirror Site, Computer Network Information