Geoderma 303 (2017) 133–142

Contents lists available at ScienceDirect

Geoderma

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

Combining reflectance spectroscopy and the digital elevation model for soil oxidizable carbon estimation

Radim Vašát*, Radka Kodešová, Luboš Borůvka, Ondřej Jakšík, Aleš Klement, Lukáš Brodský

Czech University of Life Sciences Prague, Faculty of Agrobiology, Food and Natural Resources, Department of Soil Science and Soil Protection, Kamýcká 129, 165 00 Prague-Suchdol, Czech Republic

ARTICLE INFO

Keywords:
Continuum removal
Predictive modeling
Soil organic carbon
SOC
vis–NIR

ABSTRACT

As soil oxidizable carbon (C_{ox}) has several different absorptions in the visible– and near infrared (vis–NIR) region due to its complex composition, multivariate calibration techniques such as multiple linear regression (MLR), partial least squares regression (PLSR), support vector machines (SVM) or random forest (RF) can be advantageously used to obtain a good prediction. Besides that, the content of C_{ox} is often affected by the character of the terrain, mainly due to prevailing water regime with associated transport and sedimentation processes. Therefore, the question arises; if predictive models calibrated by combining vis–NIR diffuse reflectance spectroscopy (vis–NIR DRS, 350–2500 nm) and the digital elevation model (DEM) derivatives will provide a more accurate estimate of C_{ox}. Focused on a sloping arable land (100 ha) affected by distinct water erosion, we tested for this purpose two conceptually different predictive approaches that differ in the nature of the spectroscopic predictor variables. In Approach A that relied on absorption feature (AF) parameters, the inclusion of DEM derivatives resulted in improved C_{ox} prediction using all the tested calibration techniques, i.e. MLR, RF, PLSR and SVM. The MLR prediction that was the most accurate among all others improved from R^2_{cv} = 0.81 (vis–NIR DRS dataset) or 0.50 (DEM derivatives dataset) to 0.84 (combination of both). For that prediction especially AF centered at 500, 700, 900, 1800, 1900, 2200 and 2400 nm, as well as elevation, LS factor and plan curvature were important. In contrast, in the Approach B that relied on reflectance (RF and PLSR) or normalized reflectance (SVM) values at each wavelength, no positive effect of inclusion of DEM derivatives was observed.

1. Introduction

Over the last three decades, remote (satellite or airborne) and proximal sensing technologies have become extremely popular in soil assessment applications, as they provide rapid and cost-effective measurements of the Earth's surface. Among such methods, diffuse reflectance spectroscopy in the visible– and near infrared region (vis–NIR DRS, 350–2500 nm), in particular, has proved as a powerful and promising analysis tool for highly accurate estimation of many different chemical (e.g. soil pH, organic matter content, content of nutrients, cation exchange capacity, etc.) and physical soil properties (e.g. soil texture, soil moisture, bulk density, etc.), as well as other soil characteristics (e.g. biological) and constituents (Stenberg et al., 2010). From all the possibly predictable soil properties, especially soil organic carbon (SOC) is important as a major indicator of soil quality (e.g. Zádorová et al., 2015). The prediction accuracy of SOC models is usually ranked highly as it is often rated as reliable or even excellent (e.g. Chang et al., 2001; Shepherd and Walsh, 2002; Brown et al., 2006; Gholizadeh et al., 2013; Vašát et al., 2015a; Vašát et al., 2015b).

As SOC responses multiply along the spectrum scan (it has several different absorptions in vis–NIR range due to its complex composition), mainly multivariate statistically based approaches such as partial least squares regression (PLSR), support vector machines (SVM), random forest (RF), multivariate adaptive regression splines or artificial neural networks, usually provide accurate prediction (Viscarra Rossel and Behrens, 2010; Gholizadeh et al., 2013). However, relatively simpler physically based approaches, i.e. simple or multiple linear regression (MLR), that are focusing on the relationship with absorption features (AF) (as derived from spectra normalized via continuum removal (CR) (Clark and Roush, 1984)) may occasionally provide comparative results (Bayer et al., 2012). Moreover, if the best subset of AF parameters (e.g. area, width and depth) for the MLR calibration is selected based on statistical indices, such approach may outperform the statistically based PLSR (Vašát et al., 2014). On top of that, the addition of two more AF parameters (left and right hand side area) (Vašát et al., 2015b) to those three already taken into considerations (area, width and depth) may

* Corresponding author.
E-mail address: vasat@af.czu.cz (R. Vašát).

http://dx.doi.org/10.1016/j.geoderma.2017.05.018
Received 20 September 2016; Accepted 10 May 2017
Available online 18 May 2017

lead to further improved efficiency of that methodology.

In a few cases, other environmental ancillary predictor variables were incorporated into the multivariate regression to improve the predictive performance of spectroscopic models. Brown et al. (2006) suggested the use of sand content and soil pH, which, in their opinion, are simple and inexpensive to obtain. By combining these two variables with vis–NIR spectra they achieved an improved boosted regression trees (BRT) and PLSR prediction. Stevens et al. (2013) tested two particle fractions, sand and clay content (considered by the authors as readily available and unlikely to change significantly over time), incorporated into the multivariate calibration in two ways. Firstly attached to the first few principal component scores (that were explaining > 99% of the variation of the spectra) and used with PLSR, BRT, RF, radial-basis SVM regression, multivariate adaptive regression splines and Cubist regression, and secondly attached to the small set of optimal spectral bands as selected with recursive feature elimination. By using sand content their predictive models for both grassland and woodland areas were noticeably improved, however, by using clay content, only grassland models were improved. Most recently, Peng et al. (2015) combined laboratory and remotely sensed spectra, as well as other ancillary environmental variables (including terrain characteristics such as elevation and landscape type maps, and soil maps), and by employing Cubist regression they improved the accuracy of SOC prediction, while emphasizing particularly the role of plant available water (derived from soil texture maps).

Apparently, the few attempts to improve the spectroscopic models performance were mainly based on incorporating other soil variables, i.e. soil texture fractions. The information about soil texture, however, may not be always available as it requires time- and labor consuming laboratory analyses. On the other hand, remote sensing (RS) data (satellite and/or airborne) are becoming more and more accessible. One of the widely available RS data is the digital elevation model (DEM), which comprises information about the terrains' surface. Using terrain attributes to enhance the predictive models based on vis–NIR DRS could be promising as many researchers found a mild to moderate relationship of SOC to elevation (Adhikari et al., 2014; Florinsky et al., 2002), slope (Adhikari et al., 2014; Florinsky et al., 2002; Moore et al., 1993; Xiong et al., 2014), topographic (or SAGA) wetness index (Adhikari et al., 2014; Florinsky et al., 2002; Moore et al., 1993), aspect, vertical curvature, horizontal curvature and catchment area (Florinsky et al., 2002).

In this study, we therefore attempted to further explore the potential of terrain attributes as an auxiliary source of information in vis–NIR DRS modeling, in order to improve the prediction accuracy of SOC. For this purpose we tested two predictive approaches that differ in the nature of spectroscopic predictor variables, which were five AF parameters (according to Vašát et al., 2015b) for Approach A, and reflectance (or normalized reflectance) values at each wavelength in Approach B.

2. Materials and methods

2.1. Study area and soil sampling

The study area (100 ha) is located in the southeast of the Czech Republic, and so it belongs to the most fertile areas of the country. For centuries it has been intensively farmed by cultivating crops such as wheat and sweet corn. Nowadays, due to significant slope and inadequate field management practices, the land is strongly affected by water erosion, with distinct erosion furrows occurring mainly in the steepest parts of the area. Soils were developed on loess parent material rich in carbonates. Originally, the area was covered uniformly with a single soil unit classified as Haplic Chernozem. However, due to intensive machinery cultivation and consequent strong water erosion, depending on the terrain conditions, three other soil units were later developed. These were classified as Regosols (degraded Chernozem), colluvial Chernozem and Colluvial soil. Regosols occur at the most sloping parts of the area and are characterized by the eroded humic topsoil horizon (that was largely washed away, partly even down to the parent material). Soils classified as Colluvial Chernozem occur at depressions in the upper and middle parts, and are enriched slightly in the eroded material coming from the upper parts. The lowermost parts of the area are covered with Colluvial soils that are highly enriched in the transported material (mainly the eroded topsoil humic horizon, but partly also mixed with the loess). The original Haplic Chernozem remained at the top flat parts and gently sloping parts, however, it is partly eroded too (Zádorová et al., 2013). All the soil units were classified according to World Reference Base for Soil Resources (IUSS Working Group WRB, 2014). For more detailed characterization of the area see e.g. Zádorová et al. (2011a) or Jakšík et al. (2015). One hundred and seven topsoil samples (< 20 cm) were collected following a regular sampling design with a varying spacing (Fig. 1) in the spring 2014.

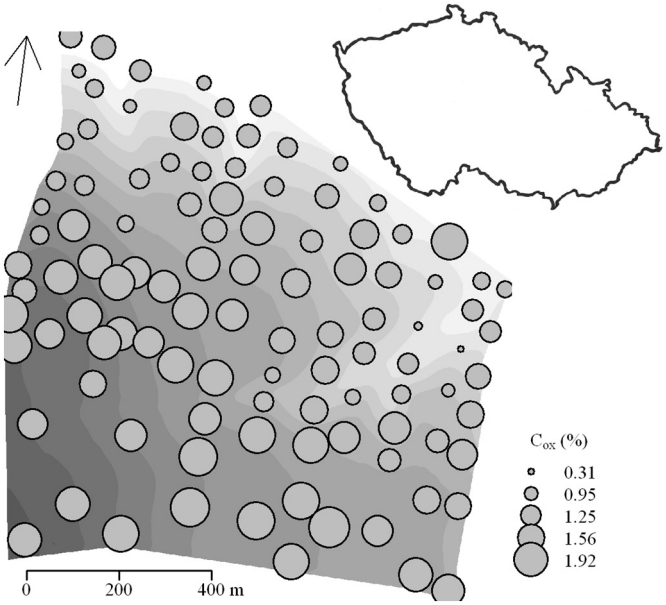

Fig. 1. Spatial pattern of 107 topsoil samples (< 20 cm) with the content of C_{ox} differentiated by circle sizes. Digital Elevation Model (DEM) in the background.

2.2. Soil oxidizable carbon laboratory determination

Due to the nature of the analytical method (wet oxidation) used to determine the content of SOC, we refer to it more precisely as soil oxidizable carbon (C_{ox}) hereinafter. The point is that not all the organic matter may be completely oxidized by the oxidant during the oxidation phase of the procedure (Skjemstad and Baldock, 2008).

The samples were air-dried at room temperature, ground and mixed thoroughly using mortar and pestle, and finally sieved to a particle fraction ≤ 0.25 mm. The C_{ox} measurement was then carried out in two sub-steps according to the dichromate redox titration method (Skjemstad and Baldock, 2008). First, the samples were oxidized with $K_2Cr_2O_7$, and then the solution was potentiometrically titrated with ferrous ammonium sulphate.

The histogram (Fig. 2) of C_{ox} content (%) indicated a statistical distribution close to normal, and the summary statistics was as follows: mean = 1.24, median = 1.1, standard deviation = 0.37, minimum = 0.31 and maximum = 1.92. The spatial distribution of C_{ox} content (%), as expressed by the different circle sizes (Fig. 1), implies that the content of C_{ox} varied across the area as a result of the character of the terrain. Generally, the content was higher at the top flat and gently sloping parts and lower at steep parts of the area.