Geoderma 263 (2016) 244–253

Contents lists available at ScienceDirect

Geoderma

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

Fuzzy clustering of Vis–NIR spectra for the objective recognition of soil morphological horizons in soil profiles

Mario Fajardo *, Alex McBratney, Brett Whelan

University of Sydney, Soil Security Laboratory, Biomedical Building, 1 Central Avenue, Australian Technology Park, Eveleigh, NSW 2015, Australia

ARTICLE INFO

Article history:
Received 18 December 2014
Received in revised form 11 May 2015
Accepted 15 May 2015
Available online 6 June 2015

Keywords:
Near infrared spectroscopy
Pedology
Soil horizons
Fuzzy clustering
Digital gradient
Digital soil morphometrics

ABSTRACT

In the past decades the use of Vis–NIR spectra information applied to soil science studies has seen an exponential growth, specially in predicting commonly used soil properties. We used the ability of Vis–NIR for detecting physico-chemical characteristics along with fuzzy clustering techniques to discriminate spectrally homogeneous zones in soil cores and applied a DG to define its boundaries i.e., SPD hor. We tested this methodology in 59 air dried soil cores varying between 85 and 130 cm depth from the HWCPID, NSW, Australia. We observed that SPD hor had great similarity with traditional horizons. The SPD hor were more homogeneous in terms of Vis–NIR spectral variability and also offered more information about the relationship between the different spectral classes. Because of the intrinsic characteristics of the methodology it can be easily applicable with or in conjunction with other proximal sensing devices which can add further detail when recognizing morphological soil horizons.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

It has been almost 100 years since Professor Curtis Marbut stated that soil studies would not thrive as a science until a generally accepted classification system was developed, suggesting with this, the use of soil horizons as a key element of it (Hartemink & Minasny, 2014; Bockheim et al., 2005). Diagnostic soil horizons have been commonly accepted since then, however it is of common knowledge in the soil science community that the identification of soil horizons and their boundaries could in many situations be inaccurate or biased due to varying description criteria. Furthermore, to classify a diagnostic horizon could require additional laboratory analysis (Weindorf et al., 2012) and given analytical procedures may change in time, this could eventually lead to biased observations (Ciampalini et al., 2013) and in the end misleading interpretations.

For these reasons there is a general challenge in homogenizing soil description criteria and a considerable amount of resources exclusively assigned for this purpose worldwide e.g., Soil Taxonomy, World Reference Base for Soil Resources, Australian Soil Classification, and German Soil Classification, among others (Ad-Hoc-AG, 2005; CSIRO, 2009; Jahn et al., 2006; Schoeneberger, 2002).

As noted by Hartemink & Minasny (2014), soil science is witnessing a historic moment, where a vast amount of new technologies are replacing or complementing the new soil science toolbox. Among these new technologies, Vis–NIR stands as one of the most widely used in both remote sensing and proximal sensing.

One of the biggest advantages in using Vis–NIR is that it can easily capture a great part of the physico-chemical variability of the sample which can be used later when comparing between different types of materials.

The objective of the present work is to use Vis–NIR to recognize different materials in soil profiles and to apply a methodology for detecting their relative patterns in depth, to finally establish in a quantitative way, boundaries between homogeneous groups of those materials i.e., soil horizons. Previous studies have used quantitative approaches to distinguish between different soil materials and/or soil horizons (Weindorf et al., 2012; Ben-Dor et al., 2008; Grunwald et al., 2001; Rooney & Lowery, 2000). The main contribution of the present work resides in the creation of a semi-automated soil morphological description procedure where the final SPD hor are comparable with others through their membership to global spectral classes which themselves work as a basic example of a classification system.

2. Materials and methods

2.1. Study area

The area of study was located approximately 140 km north of Sydney in the HWCPID in the lower Hunter Valley (Fig. 1). Geologically the area is situated in the Sydney basin, a depositional area formed by both Permian and Triassic materials with thick successions of mainly

* Corresponding author.
E-mail address: mario.fajardopedraza@sydney.edu.au (M. Fajardo).

http://dx.doi.org/10.1016/j.geoderma.2015.05.010
0016-7061/© 2015 Elsevier B.V. All rights reserved.

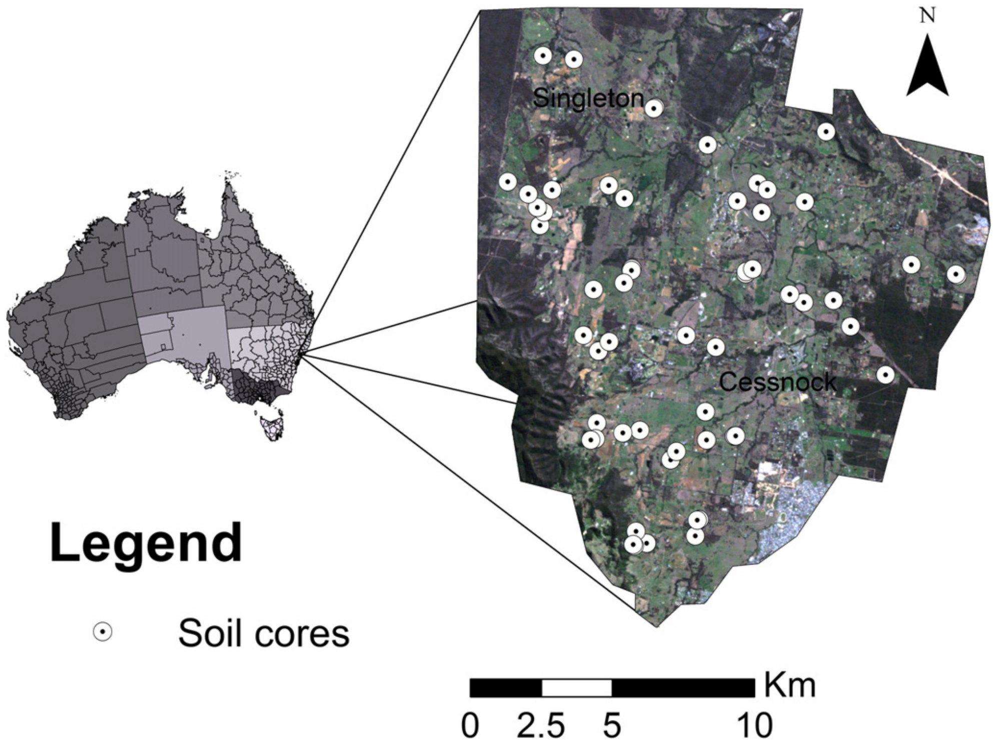

Fig. 1. Lower Hunter valley study area and sample locations, NSW, Australia.

siliciclastic rocks demonstrating a rhythmic pattern of sedimentation followed by uncommon volcanic units and carbonate rocks in a few areas (Percival & NSW, 2012). The dominant soil types according to the Australian Soil Classification (Isbell, 2002) are Red and Brown Dermosols (depending in the base saturation value, equivalent to some Udults, Udalfs and Udepts in Soil Taxonomy) and on some hill summits Red Calcarosols (equivalent to some Typic Calciudepts in Soil Taxonomy) (Odgers et al., 2011; Staff, 1999).

2.2. Sampling design

The dataset consisted of 59 soil cores varying between 85 and 130 cm depth taken 50 m away from a previous soil survey which followed a Latin hypercube sampling design where compound topographic index, parent material and normalized difference vegetation index were used as environmental variates (Minasny & McBratney, 2006) in order the maximize the variability of the samples, the cores were air dried and scanned with an ASD Agrispec 350-2500 spectrometer using a Spectralon panel as a reference, every 2 cm resulting in a datasets of 3190 separate soil scans, additionally the soil cores were morphologically described following CSIRO (2009) specifications (Fig. 1).

2.3. Processing of Vis–NIR spectra for SPD hor detection

The following treatments were employed on the dataset in the order below:

- ◇ Step correction between Vis–NIR sensors overlap in bands 1000 nm and 1800 nm.
- ◇ Selection of spectral region between 500 nm and 2450 nm.
- ◇ Conversion to absorbance from raw reflectance data.
- ◇ A second order Savitzky–Golay filter with a smoothing window of 11 bands to each spectrum.
- ◇ Based on the fact that soil spectral features change smoothly with depth we used a running median smoother on each wavelength of the spectrum depth-wise using a smoothing filter described in Hardle & Steiger (1995) and implemented by Martin Maechler in R language (R Core Team, 2013). The smoother basically works as a moving window of variable size throughout the series of numbers i.e., the values of each band through the soil profile. The selected size of the window was 10 cm (5 observations every 2 cm) after considering the observed spectral variation in the sampled soil cores and the different windows size tested.
- ◇ Selection of every 10th band in order to reduce correlation between variables and high dimensionality.
- ◇ Standard normal variate transformation of the spectra.
- ◇ Outlier detection using a Mahalanobis distance criterion (Filzmoser et al., 2005), cores with more than 10 outliers were excluded from the following analyses.

A principal component analysis was performed to each processed spectrum and the first 5 components were used (>95% of variance explained) for the next stage of fuzzy classification.

2.3.1. Fuzzy clustering

A fuzzy clustering algorithm (Maechler et al., 2014) was performed on the entire dataset. The algorithm aims to minimize the objective function, (Eq. (1)):

$$\sum_{k=1}^{c} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} u_{ik}^{r} u_{jk}^{r} d(i,j)}{2\sum_{j=1}^{n} u_{jk}^{r}} \tag{1}$$

where u_{ik} and u_{jk} are the memberships of samples i and j to class k, n is the number of observations, c is the total number of classes, r is the membership exponent and $d(i, j)$ is the dissimilarity between observations i and j. Note that if r tends to 1 it gives increasingly crisper clusterings whereas r tends to infinite it leads to complete fuzziness as specified in Maechler et al. (2014).

Download English Version:

https://daneshyari.com/en/article/6408486

Download Persian Version:

https://daneshyari.com/article/6408486

Daneshyari.com