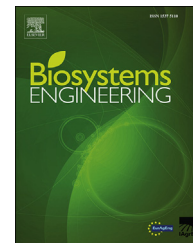




ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110**Special Issue: Proximal Soil Sensing****Research Paper****Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy**

Antonios Morellos^a, Xanthoula-Eirini Pantazi^{a,*}, Dimitrios Moshou^a,
Thomas Alexandridis^c, Rebecca Whetton^b, Georgios Tziotzios^a,
Jens Wiebensohn^d, Ralf Bill^d, Abdul M. Mouazen^b

^a Agricultural Engineering Laboratory, Aristotle University of Thessaloniki, Faculty of Agriculture, Univ. Box. 275, Thessaloniki 54124, Greece

^b Cranfield Soil and AgriFood Institute, Cranfield University, Bedfordshire MK43 0AL, United Kingdom

^c Laboratory of Remote Sensing and GIS, Aristotle University of Thessaloniki, Faculty of Agriculture, Univ. Box. 259, Thessaloniki 54124, Greece

^d Professorship for Geodesy and Geoinformatics, Faculty of Agricultural and Environmental Sciences, Rostock University, Germany

ARTICLE INFO

Article history:

Published online xxx

Keywords:

VIS-NIR spectroscopy

Data mining

Chemometrics

Soil properties

It is widely known that the visible and near infrared (VIS-NIR) spectroscopy has the potential of estimating soil total nitrogen (TN), organic carbon (OC) and moisture content (MC) due to the direct spectral responses these properties have in the near infrared (NIR) region. However, improving the prediction accuracy requires advanced modelling techniques, particularly when measurement is planned for fresh (wet and un-processed) soil samples. The aim of this work is to compare the predictive performance of two linear multivariate and two machine learning methods for TN, OC and MC. The two multivariate methods investigated included principal component regression (PCR) and partial least squares regression (PLSR), whereas the machine learning methods included least squares support vector machines (LS-SVM), and Cubist. A mobile, fibre type, VIS-NIR spectrophotometer was utilised to collect soil spectra (305–2200 nm) in diffuse reflectance mode from 140 wet soil samples collected from one field in Germany. The results indicate that machine learning methods are capable of tackling non-linear problems in the dataset. LS-SVMs and the Cubist method out-performed the linear multivariate methods for the prediction of all three soil properties studied. LS-SVM provided the best prediction for MC (root mean square error of prediction (RMSEP) = 0.457% and residual prediction deviation (RPD) = 2.24) and OC (RMSEP = 0.062% and RPD = 2.20), whereas the Cubist method provided the best prediction for TN (RMSEP = 0.071 and RPD = 1.96).

© 2016 IAGrE. Published by Elsevier Ltd. All rights reserved.

* Corresponding author. Tel.: +30 2310998264; fax: +30 2130998729.

E-mail address: renepantazi@gmail.com (X.-E. Pantazi).

<http://dx.doi.org/10.1016/j.biosystemseng.2016.04.018>

1537-5110/© 2016 IAGrE. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Soil is a heterogeneous natural resource, the processes and mechanisms of which are complex and difficult to understand. Laboratory analysis has been the main key to better understand the soil system and to assess its quality and functions (Viscarra Rossel, Walvoort, McBratney, Janik, & Skjemstad, 2006). Accurate information on soil at regional and national scale is essential, since it enables improved soil management according to land potential (Odeh & McBratney, 2000). Spatial assessment of soil properties allows researchers to understand the dynamics of ecosystems (Hively et al., 2011). Understanding the soil properties and how these affect agriculture can lead to the implementation of sustainable agricultural and environmental management (Viscarra Rossel, Cattle, Ortega, & Fouad, 2009). In precision agriculture, the scale of soil information required for land and crop management is much smaller, and normally rely on proximal soil sensing (Kuang et al., 2012) to allow data collection at high sampling resolution. However, the traditional laboratory methods for soil analysis are not able to fulfil the requirement of high sampling resolution, since they are tedious, time consuming, expensive and require expert laboratory operator.

One of the most common proximal soil sensing techniques is visible (VIS) and near infrared (NIR) spectroscopy used to estimate soil properties and can be considered as a complementary to chemical laboratory analysis methods. It can be adopted for laboratory and field (both portable and on-line) measurements. Detailed information about accuracy and performance under different application conditions has been provided in an intensive review by Kuang et al. (2012). There is increasing interest in VIS-NIR analysis techniques as they are non-destructive, fast, cost-effective and more importantly allow for high sampling resolution (Tekin, Kuang, & Mouazen, 2013; Viscarra Rossel & Hicks, 2015), which is particularly necessary for the implementation of variable rate farm inputs (e.g., fertilisers) in precision agriculture.

Soil mapping and classification has historically been performed through various methods, including statistical techniques such as principal components regression (PCR) (Chang, Laird, Mausbach, & Hurburgh., 2001; Islam, Singh, & McBratney, 2003; Mouazen, Kuang, De Baerdemaeker, & Ramon, 2010), partial least squares regression (PLSR) (McCarty, Reeves, Reeves, Follett, & Kimble, 2002; Mouazen, Kuang, et al., 2010) and also the use of machine learning techniques such as different types of artificial neural networks, decision trees and support vector machines (SVM) (Brown, Shepherd, Walsh, Mays, & Reinisch, 2006; Kuang, Tekin, & Mouazen, 2015; Mouazen, Kuang, et al., 2010; Vasques, Grunwald, & Sickman, 2008; Viscarra Rossel & Behrens, 2010). Stevens, Nocita, Tóth, Montanarella, and van Wesemael (2013) used support vector machines and Cubist to predict organic carbon (OC). Cubist is able to make very efficient spectral variable selection and the rule structure is transparent to the user regarding the association of the spectra with soil properties, allowing useful conclusions to be made about this relationship. However, these authors have used processed (dried, grinded and sieved) soil samples in their analysis. Since processed soil samples have different

physical conditions from fresh samples under field spectroscopy analyses, calibration models need to be developed with fresh samples. Then the performance of advanced data mining techniques needs to be evaluated for improved prediction capability of the VIS-NIR spectroscopy of studied soil properties for field spectroscopy application.

The aim of this paper is to compare the performance of four different regression methods for the prediction of total nitrogen (TN), OC and moisture content (MC) in fresh (wet and unprocessed) soil samples by means of a portable VIS-NIR spectrophotometer which is designed for field applications. These include two linear multivariate methods (e.g., principal components regression (PCR) and partial least squares regression (PLSR)), and two machine learning (e.g., least squares-support vector machines (LS-SVM) and the Cubist).

2. Materials and methods

2.1. Soil sampling and chemical analyses

A total of 140 soil samples were collected from the top soil layer (0–20 cm) of an arable field with an area of 31 ha in Premslin, Germany (Fig. 1) during August 2013, after harvest of winter wheat. The soil type according to the Food and Agriculture Organization (FAO) is a Luvisol. The soil samples were analysed in the soil laboratory of Cranfield University for TN, OC and MC. Soil OC and TN were measured by a TrusSpecCNS spectrometer (LECO Corporation, St. Joseph, MI, USA), using the Dumas combustion method. Soil MC was determined by oven drying of the soil samples at 105 °C for 24 h.

Figure 2 displays the histograms of the measured parameters. Although the ranges are rather narrow, they fall within the broader ranges of typical agricultural fields of the area.

2.2. Optical soil measurements

The preparation of soil samples for optical measurements was carried out, as described by Kuang et al. (2015). Fresh (wet and non-processed) soil samples were put into glass containers and mixed well, after large stones and plant residue had been removed (Mouazen, Karoui, Deckers, De Baerdemaeker, & Ramon, 2007). The optical measurements were taken from the smooth surface of soil samples, in order to achieve a higher signal to noise ratio (Mouazen, Karoui, et al., 2007). The soil samples were scanned by the AgroSpec portable VIS-NIR spectrophotometer (Tec5 Technology for Spectroscopy, Germany) that provides spectral measurements in the range between 305 and 2200 nm. A 100% white reference was measured before scanning, and this was repeated every 30 min. The 100% white reference was made from lime material to ensure 100% of light is reflected back. A total of 10 scans were collected from each glass container and these were averaged in one spectrum. The spectra from 305 to 370 nm and from 2150 to 2200 nm at the fringe of the active range of the spectrophotometer showed an excessive noisy pattern and were removed from further analysis.

Download English Version:

<https://daneshyari.com/en/article/5471965>

Download Persian Version:

<https://daneshyari.com/article/5471965>

[Daneshyari.com](https://daneshyari.com)