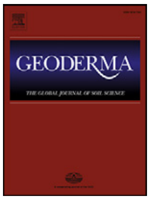
Geoderma 214–215 (2014) 114–125

Contents lists available at ScienceDirect

Geoderma

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



Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale

Guillaume Debaene *, Jacek Niedźwiecki, Alicja Pecio, Anna Żurek

Institute of Soil Science and Plant Cultivation — State Research Institute, ul. Czartoryskich 8, 24-100 Puławy, Poland

ARTICLE INFO

Article history:
Received 21 March 2012
Received in revised form 20 August 2013
Accepted 11 September 2013
Available online 11 October 2013

Keywords:
Precision agriculture
Soil texture
Visible and near infrared spectroscopy
Soil organic carbon
Soil pH
K-means clustering

ABSTRACT

Precision agriculture (PA) is a management method that measures and manages within-field variability. Previously, PA has required expansive and time consuming measurement of soil physical and chemical properties. In this paper we use a new and more rapid method of data collection based on Visible and Near-Infrared Spectroscopy (VIS–NIRS) in the 400–2200 nm spectral range to predict soil organic carbon (SOC), plant available [Mg, P, K], pH and texture at the farm scale. The experimental work was done at the experimental Station at Baborówko (52.583778°N, 16.647353°E) in Poland. The focus of the paper was to look at the effect of the number of samples on the calibration. Different calibration schemes using PLS regression with calibration datasets of different sizes were applied. The best predictions were obtained using K-means clustering for calibration sample selection. Using this scheme and 79 calibration samples, satisfactory results were obtained predicting SOC ($r^2 = 0.63$; RMSEP = 0.13%) and soil texture (e.g. clay, $r^2 = 0.71$; RMSEP = 0.36%). The use of the entire dataset did not improve significantly the prediction ability ($r^2 = 0.72$; RMSEP = 0.12% for SOC and $r^2 = 0.73$; RMSEP = 0.32% for clay). Reasonable results were obtained for available Mg content ($r^2 = 0.53$; RMSEP = 1.54 mg.100 g^{-1}) and pH ($r^2 = 0.52$; RMSEP = 0.34 pH unit). Available [P, K] gave unsatisfactory results ($r^2 < 0.5$ for both; RMSEP 6.27 and 3.31 mg.100 g^{-1} respectively). The maps (SOC and pH) generated with the K-means clustering scheme were compared with those obtained with reference data. The results show that the VIS–NIRS method is suitable to adequately predict SOC and texture using 1.5 samples per ha (79 samples). The method can also be useful as a rough screening for pH and available Mg thereby significantly reducing the cost of mapping.

1. Introduction

Site-specific management (SSM) is a precision agriculture practice that can provide positive economic and environmental impacts on modern crop production. It is based on the management of crop at a spatial scale more limited in scale than that of the entire field (Franzen et al., 2002; Mzuku et al., 2005). Visible and Near-Infrared Spectroscopy (VIS–NIRS) is a method to access high-resolution soil data, which is needed to achieve SSM. VIS–NIRS is a rapid and non-destructive analytical technique that correlates diffusely-reflected, near-infrared radiation with the chemical and physical properties of materials (Chang and Laird, 2002). Furthermore, quantitative predictions of primary (e.g. total N, total C) and secondary (e.g. cation exchange capacity, clay content) soil attributes can be very accurate, as Viscarra Rossel and McBratney (2008) have shown in their review. Information on spatial variability is lost when samples are bulked to obtain a representative sample from a definite area. In this case the acquisition of samples from e.g. a sampling grid and analysing them with a VIS–NIRS system can provide insightful information about that variability. The problem is to decide how many samples are required to capture variability. Transition from a standard soil testing approach to the adoption of VIS–NIRS method for precision agriculture requires a shift in mentality as underlined by Van Vuuren et al. (2006) but is also a technical challenge (Auernhammer, 2001).

VIS–NIRS has been used in agriculture for assessing grain, fertilisers and soil qualities (Ben-Dor and Banin, 1995; Faraji et al., 2004; Mohan et al., 2005). VIS–NIRS is now a dominant analytical technique which is used globally for food, feed, and grain analyses and for quality assessment purposes (Chang et al., 2001). Nonetheless, the complex nature of the soil matrix is still a challenge for the VIS–NIRS community (Stenberg et al., 2010). Soil properties that have been predicted using VIS–NIRS include soil moisture, SOC content, electrical conductivity (EC), cation exchange capacity (CEC), soil acidity, and some macro and micro elements (Chang et al., 2001; Dunn et al., 2002; Sørensen and Dalsgaard, 2005; Velasquez et al., 2005; Viscarra Rossel et al., 2006).

Abbreviations: PA, precision agriculture; SSM, site-specific management; VIS–NIRS, visible and near infrared reflection spectroscopy; SOC, soil organic carbon; PLS, partial least square; PCA, principal component analysis; CV, cross-validation; RPD, ratio of performance to deviation; RMSEP, root mean square error of prediction; LOOCV, leave-one-out cross validation; NN, natural neighbour.

* Corresponding author. Tel.: +48 81 886 34 21x396; fax: +48 81 886 45 47.
E-mail addresses: gdebaene@iung.pulawy.pl (G. Debaene), jacekn@iung.pulawy.pl (J. Niedźwiecki), alap@iung.pulawy.pl (A. Pecio), azurek@iung.pulawy.pl (A. Żurek).

0016-7061/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geoderma.2013.09.022

Absorption in the near-infrared spectral region (780–2500 nm) is dominated by molecules that contain strong bonds between light atoms. Specifically, these are molecules that contain C – H, N – H or O – H bonds. This makes the near infrared region particularly useful for measuring forms of carbon, nitrogen and moisture content and eventually all properties related to those. The size, the possibility of analysing wet samples and the use of fibre optics make a VIS–NIRS spectrophotometer suitable for field work (Christy, 2008) and are advantages over a mid-infrared (MIR) spectrometer.

Recent research seems to prove that local calibration (at a field or farm scale) still has an important role to play since it provided better estimation when predicting (SOC and clay) over national and reduced national libraries (Wetterlind and Stenberg, 2010). For that reason, the objective of this work was to test the VIS–NIRS method for determining various soil chemical (gravimetric SOC content, soil acidity, soil available [Mg, K, P]) and physical (soil texture) properties of topsoil collected on a grid sampling field in western Poland. Particular emphasis was given to SOC and clay content because of their importance not only in precision agriculture but also in the context of climate change as soil is one of the largest pools in the global carbon system and regular losses of greenhouse gases are measured in soils (e.g. Loubet et al., 2011). Moreover, clay content is often positively correlated with SOC content (Homann et al., 2007). Since increasing sampling density is often too expensive and time-consuming for a farmer to implement (Wetterlind et al., 2008), we decided to focus our attention on the influence of the number of samples required for calibration and on the effect of that number on the root mean square error of prediction (RMSEP) and on map accuracy. With that aim in mind, four methods to select representative calibration samples were tested on 199 validation samples. The sample selection schemes investigated were based on (1) random sample selection, (2) analyte concentration, (3) K-means clustering, and (4) PCA scores. These schemes were applied to calibration datasets of different size (20 to 199 samples). The mapping potential of the VIS–NIRS method was tested on SOC and pH.

2. Materials and methods

2.1. Sample collection and preparation

More than four hundred (403) soil samples were collected from a regular soil sampling grid of resolution of 35 m (Fig. 1) in the experimental station of the Institute of Soil Science and Plant Cultivation (IUNG) in Baborówko near Poznań (Wielkopolska, Poland). The station conducts its field trials on an area of 53.6 ha on three adjacent fields

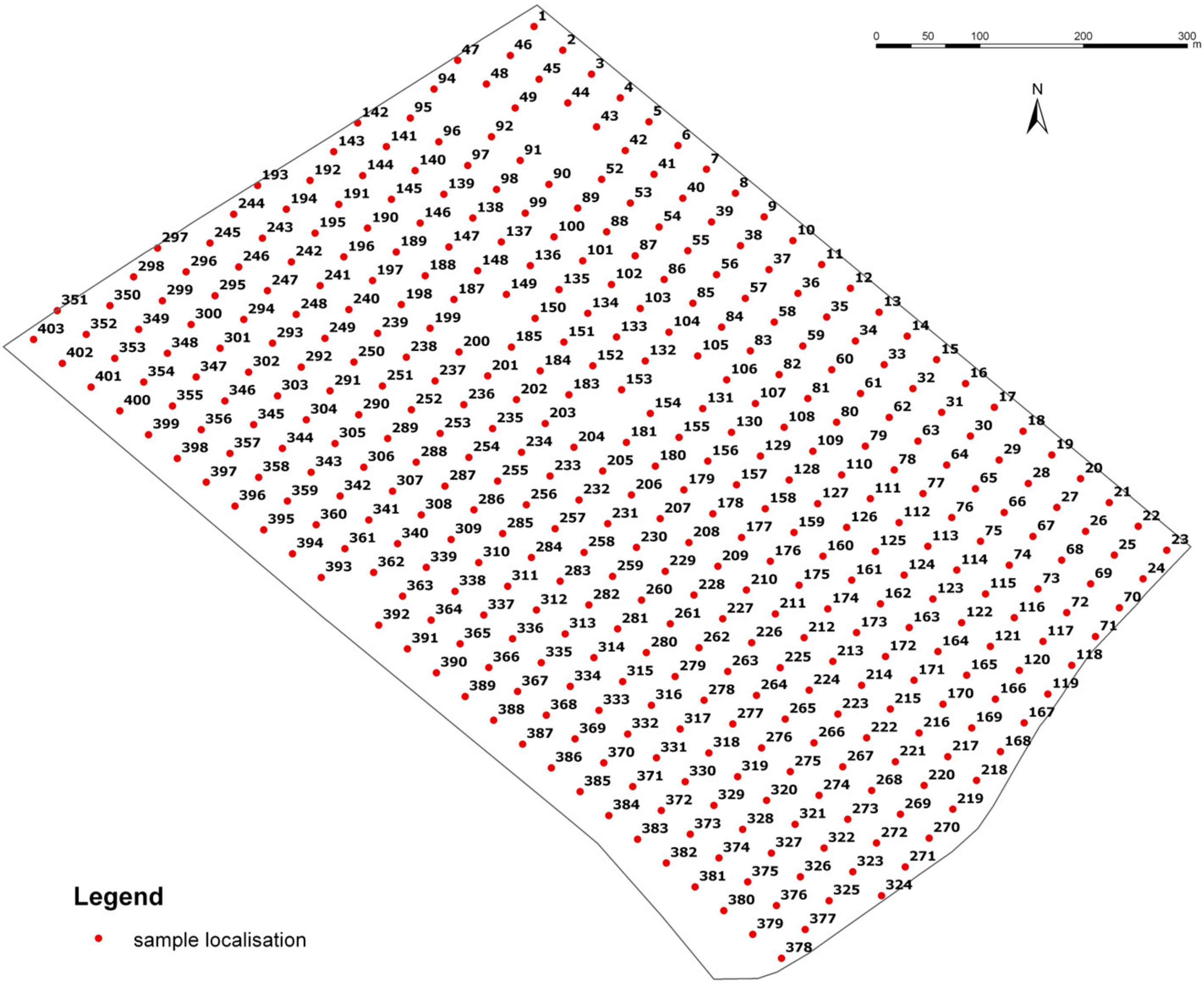


Fig. 1. Baborówko field with sample locations.