Geoderma 158 (2010) 398–411

Contents lists available at ScienceDirect

Geoderma

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

Integration of mid-infrared spectroscopy and geostatistics in the assessment of soil spatial variability at landscape level

Juan Guillermo Cobo [a,b], Gerd Dercon [a,1], Tsitsi Yekeye [c], Lazarus Chapungu [c], Chengetai Kadzere [b], Amon Murwira [c], Robert Delve [b,2], Georg Cadisch [a,*]

[a] University of Hohenheim, Institute of Plant Production and Agroecology in the Tropics and Subtropics, 70593 Stuttgart, Germany
[b] Tropical Soil Biology and Fertility Institute of the International Center for Tropical Agriculture (TSBF-CIAT), MP 228, Harare, Zimbabwe
[c] University of Zimbabwe, Department of Geography and Environmental Science, MP 167, Harare, Zimbabwe

ARTICLE INFO

Article history:
Received 7 January 2010
Received in revised form 16 May 2010
Accepted 21 June 2010
Available online 17 July 2010

Keywords:
DRIFT
MIRS
Sampling designs
Soil fertility
Spatial patterns
Zimbabwe

ABSTRACT

Knowledge of soil spatial variability is important in natural resource management, interpolation and soil sampling design, but requires a considerable amount of geo-referenced data. In this study, mid-infrared spectroscopy in combination with spatial analyses tools is being proposed to facilitate landscape evaluation and monitoring. Mid-infrared spectroscopy (MIRS) and geostatistics were integrated for evaluating soil spatial structures of three land settlement schemes in Zimbabwe (i.e. communal area, old resettlement and new resettlement; on loamy-sand, sandy-loam and clay soils, respectively). A nested non-aligned design with hierarchical grids of 750, 150 and 30 m resulted in 432 sampling points across all three villages (730–1360 ha). At each point, a composite topsoil sample was taken and analyzed by MIRS. Conventional laboratory analyses on 25–38% of the samples were used for the prediction of concentration values on the remaining samples through the application of MIRS–partial least squares regression models. These models were successful ($R^2 \geq 0.89$) for sand, clay, pH, total C and N, exchangeable Ca, Mg and effective CEC; but not for silt, available P and exchangeable K and Al ($R^2 \leq 0.82$). Minimum sample sizes required to accurately estimate the mean of each soil property in each village were calculated. With regard to locations, fewer samples were needed in the new resettlement area than in the other two areas (e.g. 66 versus 133–473 samples for estimating soil C at 10% error, respectively); regarding parameters, less samples were needed for estimating pH and sand (i.e. 3–52 versus 27–504 samples for the remaining properties, at same error margin). Spatial analyses of soil properties in each village were assessed by constructing standardized isotropic semivariograms, which were usually well described by spherical models. Spatial autocorrelation of most variables was displayed over ranges of 250–695 m. Nugget-to-sill ratios showed that, in general, spatial dependence of soil properties was: new resettlement > old resettlement > communal area; which was potentially attributed to both intrinsic (e.g. texture) and extrinsic (e.g. management) factors. As a new approach, geostatistical analysis was performed using MIRS data directly, after principal component analyses, where the first three components explained 70% of the overall variability. Semivariograms based on these components showed that spatial dependence per village was similar to overall dependence identified from individual soil properties in each area. In fact, the first component (explaining 49% of variation) related well with all soil properties of reference samples (absolute correlation values of 0.55–0.96). This showed that MIRS data could be directly linked to geostatistics for a broad and quick evaluation of soil spatial variability. It is concluded that integrating MIRS with geostatistical analyses is a cost-effective promising approach, i.e. for soil fertility and carbon sequestration assessments, mapping and monitoring at landscape level.

1. Introduction

Soil properties are inherently variable in nature mainly due to pedogenetical factors (e.g. parental material, vegetation, climate), but heterogeneity can be also induced by farmers' management (Dercon et al., 2003; Samake et al., 2005; Yemefack et al., 2005; Giller et al., 2006; Wei et al., 2008). Soil spatial variability can occur over multiple spatial scales, ranging from micro-level (millimeters), to plot level (meters), up to the landscape (kilometers) (Garten et al., 2007). Thus,

* Corresponding author. Tel.: +49 711 459 22438; fax: +49 711 459 22304.
E-mail address: georg.cadisch@uni-hohenheim.de (G. Cadisch).
[1] Present address: Soil and Water Management and Crop Nutrition Subprogramme, Joint FAO/IAEA Division of Nuclear Techniques in Food and Agriculture, Department of Nuclear Sciences and Applications, International Atomic Energy Agency, IAEA, Wagramerstrasse 5, A-1400, Vienna, Austria.
[2] Present address: Catholic Relief Services, P.O. Box 49675-00100, Nairobi, Kenya.

0016-7061/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2010.06.013

soil spatial variability is a function of the different driving factors and spatial scale (in terms of size and resolution), but also of the specific soil property (or process) under evaluation and the spatial domain (location), among others factors (Lin et al., 2005). Recognizing spatial patterns in soils is important as this knowledge can be used for enhancing natural resource management (e.g. Liu et al., 2004; Borůvka et al., 2007; Wang et al., 2009), predicting soil properties at unsampled locations (e.g. Wei et al., 2008; Liu et al., 2009) and improving sampling designs in future agro-ecological studies (e.g. Yan and Cai, 2008; Rossi et al., 2009). In fact, the identification of spatial patterns is the first step to understanding processes in natural and/or managed systems, which are usually characterized by spatial structures due to spatial autocorrelation: i.e. where closer observations are more likely to be similar than by random chance (Fortin et al., 2002). Conventional statistical analyses are not appropriate to identify spatial patterns, as these analyses require the assumption of independence among samples, which is violated when autocorrelated (spatially dependent) data are present (Fortin et al., 2002; Liebhold and Gurevitch, 2002). Thus, since 1950s, alternative methods, so-called spatial statistics, have been developed for dealing with spatial autocorrelation (Fortin et al., 2002). Today several methods for spatial analyses exist (e.g. Geostatistics, Mantel tests, Moran's I, Fractal analyses), while the reasons for the different studies carried out to date on spatial assessments are also diverse (e.g. hypotheses testing, spatial estimation, uncertainty assessment, stochastic simulation, modeling) (Goovaerts, 1999; Liebhold and Gurevitch, 2002). However, a common characteristic is that all methods intent to capture and quantify in one way or another underlying spatial patterns of a specific spatial domain (Liebhold and Gurevitch, 2002; Olea, 2006).

Geostatistics is one of the most used and powerful approaches for evaluating spatial variability of natural resources such as soils (Sauer et al., 2006). However, construction of stable semivariograms (the main tool on which geostatistics is based) requires considerable amount of geo-referenced data (Davidson and Csillag, 2003). Infrared spectroscopy (IRS) has been suggested as a viable option to facilitate access to the extensive soil data required (Shepherd and Walsh, 2007; Cécillon et al., 2009). IRS is able to detect the different molecular vibrations due to the stretching and binding of the different compounds of a sample when illuminated by an infrared beam in the near, NIRS (0.7–2.5 μm), or mid, MIRS (2.5–25 μm) ranges. The result of the measurements is summarized in one spectrum (e.g. wavelength versus absorbance), which is later related by multivariate calibration to known concentration values of the properties of interest (e.g. carbon content, texture) from reference samples. Thus, a mathematical model is created and used later for the prediction of concentration values of these properties in other samples from which IRS data is also available (Conzen, 2003). IRS measurements are therefore not destructive, take few minutes, and one spectra can be related to multiple physical, chemical and biological soil properties (Janik et al., 1998; McBratney et al., 2006). Hence the technique is more rapid and cheaper than conventional laboratory analysis, especially when a large number of samples must be analyzed (Viscarra-Rossel et al., 2006). IRS has the additional advantage that spectral information can be used as an integrative measure of soil quality, and therefore employed as a screening tool of soil conditions (Shepherd and Walsh, 2007). The few existing initiatives in this regard are, however, limited to NIRS. For example, a visible-NIRS (VNIRS) soil fertility index based on ten common soil properties has been developed and applied in Madagascar (Vågen et al., 2006); ordinal logistic regression and classification trees were used to discriminate soil ecological conditions by using biogeochemical data and VNIRS in the USA (Cohen et al., 2006); and in Kenya, Awiti et al. (2008) developed an odds logistic model based on principal components from NIRS for soil fertility classification. Nevertheless, despite its multiple applications, to date IRS has not been widely used, especially for wide-scale purposes and in developing countries (Shepherd and Walsh, 2007).

African regions are usually characterized by food insecurity and poverty, which have been extensively attributed to low soil fertility and soil mining (Sanchez and Leakey, 1997; Vitousek et al., 2009). Therefore, to boost land productivity in the continent, there is an increasing need to develop and apply reliable indicators of land quality at different spatial scales (Cobo et al., 2010). In fact, Shepherd and Walsh (2007) proposed that the successful "combination of infrared spectroscopy and geographic positioning systems will provide one of the most powerful modern tools for agricultural and environmental monitoring and analysis" in the next decade. The present study aims to contribute to this goal, and follows up a study from Cobo et al. (2009), in which three villages as typical cases of three settlement schemes in north-east Zimbabwe (i.e. communal area, old resettlement and new resettlement) were evaluated to determine specific cropping strategies, soil fertility investments and land management practices at each site. The assessment, however, was done at plot and farm level, and did not take into account spatial structures of soil properties. Hence, the same three villages of Cobo et al. (2009) were systematically sampled, soils characterized by MIRS, and data subsequently analyzed using conventional statistics and geostatistics tools. The main objectives of this study were: i) to evaluate advantages and disadvantages of using MIRS and geostatistics in the assessment of spatial variability of soils, ii) to test if MIRS can be directly integrated with geostatistics for landscape analyses, and iii) to present recommendations for guiding future sampling designs.

2. Materials and methods

2.1. Description of study sites

The study sites consisted of three villages, selected as typical cases of three small-holder settlement schemes, in the districts of Bindura and Shamva, north-east Zimbabwe (Table 1). The first village, Kanyera, is located in a communal area, covers 730 ha, and is mainly characterized with loamy-sand soils of low fertility. The second village, Chomutomora, is located in an old resettlement area (from 1987), covers 780 ha and mostly presents sandy-loam soils of low quality. The third village, Hereford farm, is located in a new resettlement area (from 2002), covers 1360 ha and is predominantly characterized by clay soils of relatively higher fertility. All villages are located in natural region II, which covers a region with altitudes of 1000 to 1800 m a.s.l. and unimodal rainfall (April to October) with 750–1000 mm per annum (FAO, 2006). Maize (*Zea maiz* L.) is the main crop planted in the three areas, and farmers have free access to communal grazing areas and woodlands. A full description of the sites' selection and characteristics is provided in Cobo et al. (2009).

Table 1
Main characteristics of the villages under study.

Village name	Settlement type	Settlement time	Location (district, ward)	Dominant soil type[a]	Mean soil textural class	Village area (ha)
Kanyera	Communal area	1948	Shamva, 6	Chromic Luvisols	Loamy sand	730
Chomutomora	Old resettlement	1987	Shamva, 15	Chromic Luvisols	Sandy loam	780
Hereford Farm	New resettlement	2002	Bindura, 8	Rhodic Ferrasols	Clay	1360

[a] According to FAO soil classification.