Geoderma 268 (2016) 165–175

Contents lists available at ScienceDirect

Geoderma

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

Pitfalls in the use of middle-infrared spectroscopy: representativeness and ranking criteria for the estimation of soil properties

Bernard Ludwig [a,*], Deborah Linsler [a], Heinrich Höper [b], Harald Schmidt [c], Hans-Peter Piepho [d], Michael Vohland [e]

[a] *Department of Environmental Chemistry, University of Kassel, Nordbahnhofstr. 1a, 37213 Witzenhausen, Germany*
[b] *Landesamt für Bergbau, Energie und Geologie, Stilleweg 2, 30655 Hanover, Germany*
[c] *Foundation Ecology & Agriculture (SÖL), Weinstraße Süd 51, 67089 Bad Dürkheim, Germany*
[d] *Institute of Crop Science, University of Hohenheim, Biostatistics Unit, Fruwirthstr. 23, 70599, Stuttgart, Germany*
[e] *Geoinformatics and Remote Sensing, Institute for Geography, Leipzig University, Johannisallee 19a, 04103 Leipzig, Germany*

ARTICLE INFO

Article history:
Received 12 August 2015
Received in revised form 8 December 2015
Accepted 11 January 2016
Available online 5 February 2016

Keywords:
Calibration set selection
Carbon fractions
Multiple linear regression
Nitrogen fractions
Outlier detection
Partial least squares regression

ABSTRACT

Middle-infrared spectroscopy (MIRS) is an established method for estimating the contents of soil organic carbon (SOC) and total soil nitrogen (N). However, obtained estimation accuracies vary between studies and only few studies are available that deal with C and N fractions. Objectives were to determine estimation accuracies for contents of SOC, microbial biomass C (C_{mic}) and C and N fractions for two samples of surface soils using different software packages (with different data treatments) and to discuss the usefulness and limitations of MIRS for a quantitative assessment of soil properties. Eighty-four surface soils were collected from arable sites from eight German states; their middle infrared spectra were recorded and their physical, chemical and biological properties determined. Estimates of SOC contents were obtained with WinISI software in cross-validations with and without removal of spectral ($H > 10$) outliers and units with large deviations between measured and estimated values ($T > 2.5$). Sample I (all 84 soils) consisted of soils from different horizons (partly with a substantial fraction of tangle of roots) and comprised of pseudo-replicates (different managements, but same mineralogy); for this ill-defined sample WinISI achieved an apparently excellent estimation accuracy when suspected outliers were removed. We suggest that T outliers should not be removed from samples in soil infrared studies except for preliminary evaluations. In contrast, for the consistently defined subset sample II (i.e., soils were taken from Ap and M-Ap horizons from 51 German arable sites with typical SOC contents and without pseudo-replicates) without outliers only a good estimation accuracy was reached. This indicates that besides a search for optimum estimation accuracies, equal attention should be paid to the representativeness of the sample for a specific population, an appropriate handling of suspected outliers and the generalizability of the MIRS results. With respect to accuracies, we obtained good results, approximative quantitative results or accuracies with the potential to discriminate between high and low values for all C and N fractions (except for light-fraction N) and also C_{mic}. An estimation of these properties without infrared data using the contents of SOC, N, pH, sand, silt and clay in multiple linear regressions was generally slightly less successful than the MIRS estimates using OPUS. However, when we created artificial spectra based solely on the measured pH, contents of SOC and N and texture data – without any real underlying infrared data – and then used them for a PLS regression in OPUS, the performance was similar to the MIRS estimates, with a slight difference for passive C and N. Overall, our study indicates that MIRS is a useful method for an estimation of the spectrally active main constituents SOC and N and possibly for passive C and N. However, there is not much benefit of using MIRS to obtain a spectral assessment for those properties, where approaches without infrared data (either multiple linear or PLS regressions using pH, SOC, N and texture data) give estimates of similar accuracy, which was the case for of C_{mic}, light-fraction C and N, mineral-associated C and N or intermediate C and N for the dataset investigated here.

1. Introduction

The application of infrared spectroscopy (near infrared (NIR) range: 800–2500 nm, visible and NIR (visNIR) range: 400–2500 nm, middle infrared (MIR) range: 2500–25,000 nm) has been suggested to be useful

---

* Corresponding author.
E-mail address: bludwig@uni-kassel.de (B. Ludwig).

http://dx.doi.org/10.1016/j.geoderma.2016.01.010
0016-7061/

in estimating a large number of physical, chemical and even biological soil properties (for reviews see e.g., Kuang et al., 2012 and Soriano-Disla et al., 2014). The accuracies of estimation will generally be less than those obtained by direct wet-chemical analysis, but may still be sufficient for many research goals, and spectroscopic approaches may be especially important for studies which require large data sets, such as for long-term monitoring of soil properties.

Infrared spectroscopy is well-suited for those properties which directly absorb in the infrared region such as soil organic carbon (SOC) and N (Janik and Skjemstad, 1995; Reeves, 2010; Bellon-Maurel et al., 2010) because of its sensitivity to C–H, C–O and C–N groups that dominate organic matter (Soriano-Disla et al., 2014). Also for pH, the usefulness of infrared spectroscopy is established with the absorptions probably related to contents of organic acids and carbonates (Reeves, 2010). Soriano-Disla et al. (2014) summarized for clay and sand a suitability of MIR spectroscopy (MIRS) and NIR spectroscopy (NIRS) due to the absorptions of quartz and various types of clay in the infrared region, whereas for silt greater uncertainties were reported, as the exact nature of soil components associated with this large range of particle size and their spectral behavior is not clear.

The estimation of biological properties such as microbial biomass C (C_{mic}) and N (N_{mic}) and enzyme activities by infrared spectroscopy is discussed controversially. Soriano-Disla et al. (2014) summarized for C_{mic} that results are likely due to a strong correlation between the biological properties and the quantity and quality of soil organic matter; they cited studies suggesting that correlations with C and N may only partially explain the infrared PLS regression results for biological properties, and proposed differences in macronutrient availability and specific signature wavelengths directly related to C_{mic} to be also responsible for the modeled values. However, Ludwig et al. (2015) did not see a marked benefit of MIRS for an estimation of C_{mic} and N_{mic} for a sample of 422 soils, since an estimation of these properties using cross-validated multiple linear regressions with pH and the contents of C, N, P, S, sand and clay as explanatory variables was as accurate (C_{mic}: $r^2 = 0.77$, N_{mic}: $r^2 = 0.72$) as an estimation using MIR data (entire set, C_{mic}: $r^2 = 0.78$, N_{mic}: $r^2 = 0.73$). Vohland et al. (2014) found MIR key wavelengths for C_{mic} that largely overlapped with those of OC and thus also suggested indirect correlations to be the reason for a spectral assessment of C_{mic}. However, indirect estimations by infrared spectroscopy are problematic since they have less predictive value and are per se not transferable to other samples.

Less information is available for the use of infrared spectroscopy for estimations of C and N contents in density fractions. Luce et al. (2014) used visNIR spectroscopy (visNIRS) for an estimation of particulate organic matter (POM) N (POMN) and light fraction N (LF-N). They found for a sample of 200 soils (including pseudo-replicates) that the validation statistics for the randomly selected validation subset (49 units) indicated a potential of visNIRS for these properties (LF-N: $r^2 = 0.84$, POMN: $r^2 = 0.88$) and concluded that their data suggested the presence of a spectral basis for estimations of these properties by visNIRS. Similarly, Bornemann et al. (2010) reported a usefulness for MIRS for an estimation of differently sized POM fractions due to specific absorptions of lignin, cellulose, aliphatic C–H groups, aromatic moieties and carboxylic groups. In contrast, we hypothesize that a spectral basis may be of minor importance and that LF-N and light fraction C (LF-C) can be as accurately estimated due to their relationships to main (spectrally active) properties such as C, N, pH and texture as due to infrared spectra.

For the estimation of passive C and N (turnover time > 100 years, v. Lützow et al., 2008), there may be a potential for infrared spectroscopy, since, e.g., Rumpel et al. (2001) studied a sample of soils with different amounts of lignite-derived C and reported that many of the frequencies in the MIR spectra identified for the humic and compost substances were similar to those for organic matter of lignite-containing mine soils, but their relative proportions were different and the peaks in the spectra region from 1700 to 1200 cm^{-1} were better resolved. Such a potential was also suggested by the MIR studies by Janik et al. (2007) who investigated 177 Australian soils with UV photo-oxidation and CPMAS ^{13}C NMR spectroscopy, by Michel et al. (2009) who studied mixtures of different chars (lignite, anthracite, charcoal, or a mixture of the three coals) and forest-floor Oa material and by Bornemann et al. (2008) who investigated a sample of soils characterized for SOC and benzene polycarboxylic acids as specific black C markers. A problem, however, with the estimation of passive C is that different stabilization processes exist, e.g., production of charcoal by fire or interaction of SOC with mineral surfaces (v. Lützow et al., 2008), which should result in different spectral characteristics. Thus, the potential of MIRS for a direct assessment may be also limited.

All these examples show that the focus in spectroscopic studies should not be solely placed on estimation accuracies of the spectroscopic approaches, at least for spectrally non-active constituents which may be assessed due to their correlations to active components.

Estimation accuracies are typically expressed in terms of performance to deviation value (RPD, ratio of standard deviation of laboratory results and standard error of cross-validation (SECV)) and r^2 values. The accuracies of estimations for different properties may be ranked according to their RPD and r^2 values ranging from excellent estimation (RPD > 3.0 and $r^2 > 0.91$) to no usable estimation (RPD < 1.5 and $r^2 < 0.50$) as suggested by Saeys et al. (2005). We believe that the ranking system by Saeys et al. (2005) is very appropriate for soil analyses, but that several misunderstandings regarding this system may have occurred in some studies. Firstly, the interpretation of RPD based on a skewed sample or a sample with outliers is problematic as pointed out by Bellon-Maurel et al. (2010) and Gobrecht et al. (2014). Thus, studies should at least report the required descriptive statistics of the properties analyzed and interpret their RPD data accordingly. Secondly, many studies aimed primarily to achieve as high RPD values as possible, disregarding other aspects. For instance, Stenberg et al. (2010) summarized for visNIRS studies that in the literature, suspected pseudo-independent validation samples are fairly common. Thirdly, many different scales of interest are possible for NIR and MIR studies. For instance, replicates used may stem solely from one field for a field scale study or replicates may stem from different fields for a regional scale study and RPD values may vary markedly between field and regional scale studies. The underlying specific population should be stated at least and in case of a regional study, pseudoreplication (several replicates from one field) should either be avoided or at least discussed critically. Overall, we think that an exclusive focus on RPD is problematic, since other aspects such as the representativeness of the sample for a specific population, an appropriate handling of suspected outliers and a discussion of the generalizability of the MIRS results depending on the scale are equally or even more important.

With the above findings in mind, we set ourselves the following objectives: (i) we wanted to demonstrate that it is possible to achieve apparently excellent estimation accuracies using MIRS for samples from ill-defined populations and that the representativeness of the sample for a specific population and an appropriate handling of suspected outliers are equally important and (ii) we studied, whether MIRS is useful for a direct estimation of C_{mic} and C and N fractions in soils. For objective (ii), we compared the statistical analyses using the laboratory data of the main physical and chemical properties with those using MIR data.

2. Material and methods

2.1. Soils

2.1.1. Sample I (n = 84 soils)

Surface soils (24 soils: 0–20 cm, 60 soils: 0–10 cm) were taken at different locations across Germany on arable land with different soil types and textures. (i) Thirty-one soils were taken from different long-term monitored areas in Lower Saxony. (ii) Additionally, thirty-one