Stafford, J., Werner, A. (eds): Precision Agriculture.
Wageningen Academic Publishers, ECPA Berlin, 15.-19.6.03
S. 633-638

Spatial detection of topsoil properties using hyperspectral sensing

T. Selige, L. Nätscher and U. Schmidhalter
Chair of Plant Nutrition, Department of Plant Sciences, Technical University Munich, Am Hochanger 2, 85350 Freising, Germany
selige@wzw.tum.de

Abstract

The spatial variability of topsoil texture and organic matter across fields was studied using field-spectroscopy and airborne hyperspectral imagery with the aim of improving fine-scale soil mapping procedures. Two important topsoil parameters for precision farming applications, organic matter and clay content, were correlated with spectral properties. Both parameters can be determined simultaneously from a single spectral signature since organic carbon largely responds to wavebands in the visible range and clay responds to wavebands in the Near Infrared. Because of cross-correlations, one has to consider iron oxides and high amounts of coarse sand in order to infer clay content from the spectral signature. The composition of the organic matter should be considered in order to infer the organic matter content from the spectral signature. It is shown that the clay and organic matter content can be predicted quantitatively and simultaneously using Partial Least Squares Regression by a multivariate calibration approach.

Keywords: clay content, field-spectroscopy, organic matter content, Partial Least Squares Regression

Introduction

Frequently, significant heterogeneity across fields can be found in topsoils which cause differences in crop germination, nutrient and water uptake and thus markedly influence crop growth. Topsoil heterogeneity also causes patterns of e.g. erosion, surface sealing, nutrient mineralisation, and water balance. It has, therefore, implications for site specific management practices and strategies. For optimizing crop growth, soil tillage, seed bed preparation, fertilization and herbicide use, in particular, must be adapted to local topsoil properties.
However, there is still a serious lack of site specific data about physico-chemical topsoil characteristics for precise and spatially variable management. Several authors have established relationships between soil reflectance and organic matter (Dalal & Henry, 1986, Udelhoven et al., 2001) and soil texture (Al-Abbas et al., 1972). Both soil parameters play significant role in assessing topsoil characteristics e.g. aggregation and resistance to erosion (Neemann, 1991).
Thus, a study was conducted to identify, quantify and locate topsoil properties using a novel approach combining multivariate calibration techniques, field-spectroscopy and hyperspectral remote sensing. The aim of this work was to derive a quantitative determination of clay content and organic matter content (OMC) from single spectral signatures. This would potentially allow mapping soil texture and organic matter simultaneously by remote or proximal sensing.

Material and methods

The area of investigation was located in Sachsen-Anhalt, Germany. It was characterised by two soilscapes, a slightly undulated tertiary plain at 70 m altitude and the alluvial plain of the river Elbe at 50 m altitude. With 450 mm mean annual precipitation and 9 °C mean annual temperature, the region has a negative water balance during the vegetation period. The tertiary plain is covered by

Aeolian sediments mixed with different portions of tertiary sand and clay. The predominant soil type is Chernozem in conjunction with Cambisols. The alluvial plain is characterised by coarse sand to fine sand, loamy and clayey sediments. The predominant soil types are Mollic Gleysols and Fluvisols. The fine-scale pattern of soil texture and organic matter within the fields of both landscapes results in highly diverse soil properties and virtually forces the application of site specific management.

To represent a relevant spectrum of soil texture and OMC that is frequently found with arable soil, the study was based on 29 topsoil samples for regression analysis and on 72 topsoil samples for multivariate calibration using Partial Least Squares Regression. The samples were passed through a 2 mm sieve and were air dried. The soils were analyzed for total amount of organic carbon (C_{org}) and the total amount of nitrogen (N_t) by dry combustion using an elemental analyzer. The factor 1.724 (Nelson & Sommers, 1982) was used to estimate the OMC from the C_{org} values. The min-max range of the soil data is given in Table 1. The sample sites were selected according to different combinations of soil forming geo-factors and represent largely the typical range of texture and OMC in arable soil.

Based on a digital terrain model an adapted topographic wetness index (aTWI) was calculated (Böhner et al., 2002) to characterize differences in site-specific dry and wet conditions of the organic matter formation which was expected to cause differences in organic matter composition. For a subset of 29 samples, the organic matter composition was additionally characterized by aliphatic and aromatic compounds using extraction methods (Schnitzer, 1982). The particle size distribution was analyzed using sieve analysis for the sand fractions and the coarse silt fraction and pipette analysis for the fine fractions of silt and clay. Iron oxides were determined by measuring the ditionite solubility (Fe_d) and the oxalate solubility (Fe_o) (del Campillo and Torrent, 1992).

Topsoil reflectance (330-2500 nm) was measured in the field using a GER 3700 field-spectrometer and a lambertian Spectralon reference panel of known reflectivity. A 50x50 cm plot was used as field of view. Due to insufficient spectral signal-to-noise ratio, two ranges (1350-1420 nm and 1780-1980 nm) of the continuous spectral signatures were excluded from further analysis. The airborne HyMap sensor was used at an early May flight campaign for recording hyperspectral images (420-2480 nm, 128 channels) of bare soil fields in the study area. In order to simulate HyMap spectra, the field-spectrometer spectra were resampled according to the position and range of the HyMap spectrometer wavebands.

In the first data analysis step, the OMC and the clay content were correlated to each waveband in order to find relevant wavebands. In a second step, Partial Least Squares Regression (PLSR) was applied to develop and calibrate an inverse model that establishes a quantitative relationship between the spectra of the field-spectrometer measurements and the soil parameters (Næs et al., 2002). PLRS models were calibrated separately for OMC and for clay content. To ensure that the PLSR model can be used after the calibration procedure concurrently to the HyMap data, the Piecewise Direct Standardization (Wang et al., 1993) was applied to both spectral data sets. Several pre-treatment techniques were applied to the data set including vector normalization, min-max normalization and first derivative (Martens & Næs, 1989). Finally, the optimal calibration model was used to predict the OMC from the HyMap spectra for each HyMap image pixel of bare soils resulting in a map showing the distribution of the topsoil OMC.

Table 1. The min-max ranges of soil parameters.

	Sand (%)	Silt (%)	Clay (%)	OMC (%)	N_t (%)	C/N Ratio	Fe_o (%)	Fe_d (%)
Min.	16	5	7	0.7	0.05	8.2	0.06	0.14
Max	88	72	35	5.2	0.26	12.4	0.48	0.98

Results

Most of the measured soil parameters were not independent. Different relations with significant correlations were found. A close correlation was found between clay and iron oxides. Clay and the oxalate soluble iron oxides Fe_o showed a linear correlation (r=0.90***). Clay and ditionite soluble iron oxides were also linearly correlated (r=0.94***).

The regression analysis of the spectral wavebands also revealed a strong spectral response for the clay content. Individual wavebands showed a significant correlation with the clay content but only in the spectral range >2300 nm. The best relation to clay was represented by a non-linear regression with the wavebands at 2427-2436 nm (Figure 1; R^2=0.78***). The relation was independent of the OMC but was linked to iron oxides (R^2=0.82***). The analysis of the data scattering suggested that the spectral response to the clay content was independent of the organic matter but affected by the sand fraction. Higher reflection values than expected were related to samples that were characterized by a high amount of coarse sand.

The OMC correlated most strongly in the range of visible wavebands and the shorter Near Infrared. The best non-linear regression was found for the wavebands 344-357 nm (Figure 2; R^2=0.68***). Similar and just marginally weaker coefficients were found for many of the individual wavebands up to 920 nm, which was comparable with results of Al-Abbas et al. (1972).

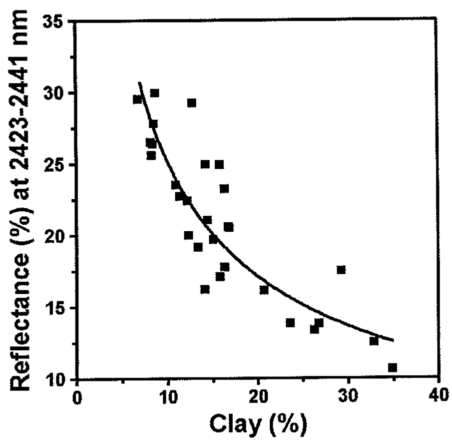

Figure 1. Relationship between clay content and spectral reflection at 2423-2441 nm.

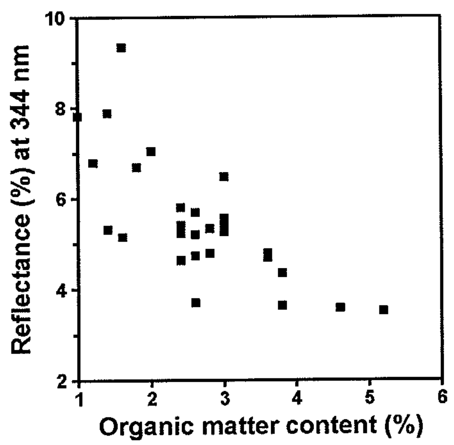

Figure 2. Relationship between the organic carbon content and the spectral reflection at 344 nm.

The data scattering of the OMC was compared with the visible soil color. It became obvious, that the organic matter composition plays a significant role in the absorption. This seemed to be connected to the origin of organic matter due to site specific transformation processes (Ben-Dor & Banin, 1995). The data scattering could be explained after separating the dataset in two sub-classes 'dry' and 'wet' conditions of organic matter formation. Both data sub-classes also fitted to the aliphatic and aromatic composition of the organic matter. It was assumed that the effect of organic matter formation can be linked to terrain features which are related to water accumulation. Thus, the aTWI was calculated as a metric value to describe the organic matter decomposition. The samples which were characterized by higher values of aTWI, respectively aromatic compounds, showed lower reflectance values. As a result, the composition of the organic matter must be considered for more precise characterisation of the organic matter and for a better understanding of the spectral response mechanism.

The univariate regression analysis indicates that even single wavebands are correlating with either the clay content or the OMC. Since clay and organic matter respond to different spectral wavebands, it is possible to extract both parameters quantitatively from a single signature. Unfortunately, the results are strongly affected by other soil parameters due to the high variability of uncertain combinations of geo-factorial effects. The organic matter therefore cannot simply be described by its content in relation to the spectral properties. The composition of the organic matter has to be included in the analysis. The spectral response of clay was affected by iron oxides and high amounts of coarse sand. Both parameters had to be considered to precisely derive the clay content using univariate analysis techniques.

Complexity and auto-correlations between the soil parameters in the above-mentioned results led to the use of multivariate calibration techniques, particularly PLSR. Multivariate calibration was performed with PLSR, which reduced the reflectance spectra to a few relevant factors and regressed them to the OMC of a given sample (Martens & Næs, 1989). From this regression, a model was derived enabling prediction of the OMC from the spectra of samples with unknown OMC. The same calibration procedure was also employed to derive a prediction model for the clay content. Various wavelength regions and data pre-treatments were analysed using an optimization routine to find the best calibration algorithm. The algorithms with the lowest Root Mean Square Error of Cross Validation (RMSECV) were chosen as statistically the best. Since several data pre-treatments result in similar error, those were chosen that had the lowest number of factors included in the regression model (Næs et al., 2002). The selected models use 5 factors for clay content and 7 factors for OMC. The results of cross validation for the different models and data pre-treatments are listed in Table 2.

Overall, the vector normalization and the min-max normalization gave the best results. The application of the first derivative led to less comparable results, possibly due to noise and the relatively broad wavebands that were used after resampling the field-spectrometer spectra into simulated HyMap spectra. But any pre-treatment produced better models than the use of the raw data.

The study indicates that multivariate calibration techniques, here represented by PLSR, can enable the determination of the organic matter and the soil texture more precisely when using optimized spectral waveband combinations than using only the best fitting to single wavebands. To achieve a calibration algorithm that is robust against variability of natural factors, more wavebands, namely the whole spectra, should be considered (Dardenne, 1996). Concurrently, most of the possibly occurring variations of natural factor combinations should be represented in the calibration set to achieve a robust calibration algorithm (Schenk and Westerhaus, 1991). The PLSR algorithm will automatically give high weights to the decisive wavelength regions and low or zero weights to uninformative wavelengths provided that the spectral and natural variability included in the calibration set is high enough.

The optimal calibration model was found with the pre-treatment of vector normalization. Thus, this algorithm was used to calculate the OMC from the HyMap spectra for each HyMap image pixel

Table 2. PLSR estimates of the organic matter content (OMC) and the clay content of topsoils from the resampled field-spectrometer signatures.

Soil parameter	Data Preprocessing	Field-spectrometer	
		R^2	RMSECV
OMC	None	74.2	0.44
OMC	Vector normalisation	89.6	0.29
OMC	Min-max normalisation	88.1	0.30
OMC	First derivative	82.3	0.36
Clay content	None	84.3	8.6
Clay content	Vector normalisation	92.3	4.2
Clay content	Min-max normalisation	91.5	4.4
Clay content	First derivative	89.7	6.5

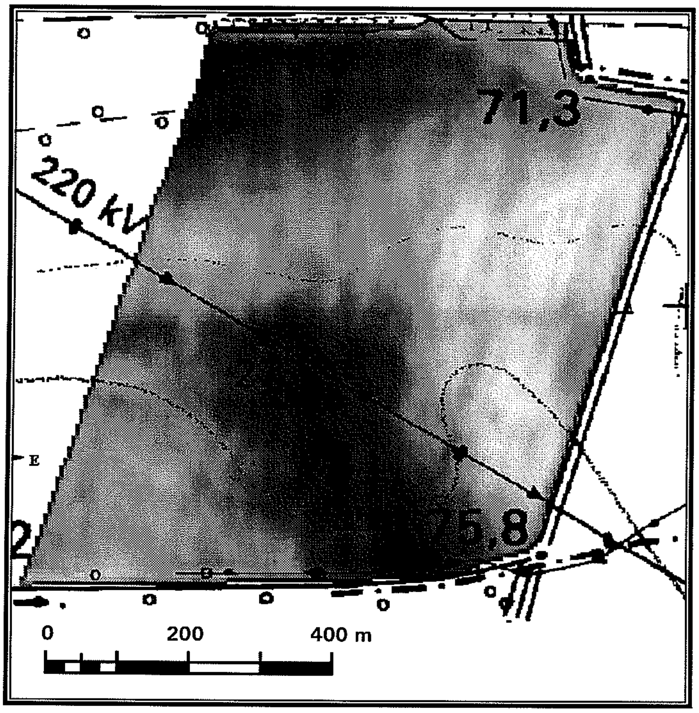

Figure 3. Spatial distribution of organic matter in the reference field 'Finkenherd'.

of the bare soil reference field "Finkenherd". The resulting map shows the distribution of the topsoil OMC on this 45 ha field (Figure 3). The OMC values range from 0.5 (light) to 4.0 % (dark). A first validation test, merely based on seven samples, indicated for this field a close correlation of r = 0.89*** between the OMC predicted from the Hymap data and the OMC measured by dry combustion method. The high variability of OMC across the entire field (45 ha) illustrates the demand for variable tillage and fertilizer applications to provide optimal conditions for site-specific crop growth.

Conclusion

How to process precise and spatially highly resolved soil maps is still a critical question for site specific management. It could be shown that hyperspectral remote and field-spectroscopic proximal sensing combined with inverse model calibration based on Partial Least Squares Regression has the potential to contribute significantly to this purpose. Simultaneous and rapid, nondestructive determination of OMC and clay content in topsoils is possible by using multivariate analysis of spectral data. The achievements presented in this study are likely to be good and lead to the recommendation for further investigations to amalgamate soil sciences with remote sensing and multivariate calibration techniques respectively modeling towards a fine-scale topsoil mapping procedure for precision agriculture.

Acknowledgment

This study was partly funded by the German Federal Ministry of Education and Research (BMBF)

References

Al-Abbas, A. H., Swain, P. H., and Baumgartner, M. F. 1972. Relating organic matter and clay content to the multispectral radiance of soils. Soil Science 114 477-485.

Ben-Dor, E. and Banin, A. 1995. Near-Infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of America Journal 59 364-372.

Böhner, J., Conrad, O., Gross, J., Köthe, R., Ringeler, A. and Selige, T. 2002. Soil regionalisation by means of terrain analysis and process parameterisation. In: Soil Classification 2001, edited by E. Micheli, F. Nachtergaele, R. Jones, L. Montanarella, Publication of European Union, EUR 20398 EN.

del Campillo, M. C. and Torrent, J. 1992. A rapid acid-oxalate extraction procedure for the determination of active Fe-oxide forms in calcareous soils. Zeitschrift für Pflanzenernährung und Bodenkunde 155 437-440.

Dalal, R. C., Henry, R. J. 1986. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. Soil Science Society of America Journal 50 120-123.

Dardenne, P. 1996. Stability of NIR spectroscopy equations. NIR news 7 (5) 8-9.

Martens, H. and Næs, T. 1989. Multivariate calibration. John Wiley & Sons, Chichester, UK.

Næs, T., Isaksson, T. Fearn, T. and Davies, T. 2002. A User-Friendly Guide to Multivariate Calibration and Classification. NIR-Publication, Chichester, UK.

Neemann, W. 1991. Bestimmung des Bodenerodierbarkeitsfaktors für winderosionsgefährdete Böden Norddeutschlands. (Determination of soil erodibility factors for wind-erosion endangered soils in Northern Germany.) Geologisches Jahrbuch, Reihe F 25, 131 pp.

Nelson, D. W., and Sommers, L. E. 1982. Total carbon, organic carbon, and organic matter. In: Methods of soil analysis - Chemical and microbiological properties, second edition. Agronomy No. 9, Part 2, edited by A.L. Page, R.H. Miller and D.R. Keeney, American Society of Agronomy, Madison, Wisconsin, USA, 539-579.

Schenk, J. S. and Westerhaus, M. O. 1991. Population definition, sample selection, and calibration procedures for near infrared spectroscopy, Crop Science 31 469-474.

Schnitzer, M. 1982. Organic matter characterization. In: Methods of soil analysis - Chemical and microbiological properties, second edition. Agronomy No. 9, Part 2, edited by A.L. Page, R.H. Miller and D.R. Keeney, American Society of Agronomy, Madison, Wisconsin, USA, 581-593.

Udelhoven, T., Jarmer, T. and Hill, J. 2001. Spectral reflectance properties of soils as support for hyperspectral applications in precision farming. In: Proceedings of the international workshop of spectroscopy applications in precision farming, Weihenstephan, Germany, 80-84.

Wang, Y., Lysaght, M. J. and Kowalski, B. R. 1993. Improvement of multivariate calibration through instrument standardization. Analytical Chemistry 64 562-564.