Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Geoderma 136 (2006) 235–244
GEODERMA
www.elsevier.com/locate/geoderma

High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures

Thomas Selige [a,*], Jürgen Böhner [b], Urs Schmidhalter [a]

[a] *Chair of Plant Nutrition, Department of Plant Sciences, Technical University of Munich, Am Hochanger 2, D-85350 Freising, Germany*
[b] *Department of Physical Geography, University Göttingen, Goldschmidtstrasse 5, D-37077 Göttingen, Germany*

Received 12 July 2004; received in revised form 9 December 2005; accepted 19 March 2006
Available online 24 May 2006

Abstract

The spatial variability of within field topsoil texture and organic matter was studied using airborne hyperspectral imagery so as to develop improved fine-scale soil mapping procedures. Two important topsoil variables for precision farming applications, soil organic matter and soil texture, were found to be correlated with spectral properties of the airborne HyMap scanner. The percentage sand, clay, organic carbon and total nitrogen content could be predicted quantitatively and simultaneously by a multivariate calibration approach using either partial least-square regression (PLSR) or multiple linear regression (MLR). The different topsoil parameters are determined simultaneously from the spectral signature contained in the single hyperspectral image, since the various variables were represented by varying combinations of wavebands across the spectra. The methodology proposed provides a means of simultaneously estimating topsoil organic matter and texture in a rapid and non-destructive manner, whilst avoiding the spatial accuracy problems associated with spatial interpolation. The use of high spatial resolution and hyperspectral remotely sensed data in the manner proposed in this paper can also be used to monitor and better understand the influence of management and land use practices on soil organic matter composition and content.

Keywords: Hyperspectral airborne data; Topsoil mapping; Partial least-square regression; Multiple linear regression; Soil texture; Soil organic matter

1. Introduction

Topsoils frequently show a fine tessellated pattern and heterogeneity across fields indicated by e.g. color, roughness, infiltration, erosion and surface sealing phenomena. These characteristics cause differences in crop germination, nutrient and water uptake and thus markedly influence crop growth processes. This has implications for the pattern and the spatial extent of appropriate land use management practices and soil conservation strategies including site specific management in precision agriculture systems (Runge and Hons, 1998; Bullock and Bullock, 2000). To optimize crop growth, the various cropping practices, including soil tillage, seed bed preparation, fertilization and herbicide use must be adapted to the local topsoil properties. However, lack of high spatial resolution topsoil data is a serious limitation to the establishment of sub-field soil and crop management.

There is at present no effective way to map fine-scale soil heterogeneity so as to derive site specific data about

* Corresponding author.
E-mail address: selige@wzw.tum.de (T. Selige).

0016-7061/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.geoderma.2006.03.050

topsoil physical/chemical characteristics. Several authors established relationships between soil spectral reflectance data and organic matter characteristics (Dalal and Henry, 1986; Ben-Dor et al., 1997; Ingleby and Crowe, 2000; Reeves et al., 2002; Udelhoven et al., 2003) and soil texture (Al-Abbas et al., 1972; Ben-Dor and Banin, 1995; Thomasson et al., 2001; Ben-Dor et al., 2002; Cozzolino and Morón, 2003). Soil attributes, soil texture and soil organic matter, all play an interdependent and decisive role in assessing topsoil characteristics e.g. soil aggregation, aggregate stability and resistance to water and wind erosion (Neemann, 1991). As a consequence it would be an advantage to be able to map both sets of physical characteristics from the one set of image data.

The aim of the work reported here was to develop a method of mapping fine-scale topsoil organic and texture parameters from a combination of field and hyperspectral image data. The work investigated the use of both multiple linear regression (MLR) and partial least-square regression (PLSR) to construct the model necessary to estimate the soil physical/chemical variables from the image data. Field data were combined with the image data for the construction of independent models derived using both MLR and PLSR. This innovative approach to digital soil mapping achieved gain from the simultaneous estimation of a suite of topsoil parameters. Additionally, the use of high spatial resolution remotely sensed data provides a more detailed pattern recognition of the soil's heterogeneity. Generally in soil mapping a soil data set is available that is restricted to only few sampled locations. The approach reported here avoids the attendant problems of accurate and reliable spatial prediction of spatial interpolation that has to deal with a small soil data set that does not uncover the fine tessellated pattern of soils. Moreover, as a methodology that is applicable over large areas at high spatial resolution this approach can be expected to support change detection research by mapping of past and future land use effects on soil organic matter accumulation and decomposition and contributes to sustainable and site adapted soil management.

2. Material and methods

2.1. Geographical survey, geology and terrain

The East German study area *Wulfen* (11°55′E, 51°49′ N) is located in the federal state Sachsen–Anhalt, and stretches over 20 km from the district town *Köthen* in the south to the Elbe river in the north. It is characterized by three soilscapes. The southern part is a slightly undulated tertiary plain at 70 m altitude that is covered by a thin Loess layer up to 1.2 m deep. The northern part is the alluvial plain (glacial valley) of the river Elbe at 50 m altitude that served as the origin for the Aeolian deposit in the south. Both landscapes are separated by late Pleistocene terminal moraines that form a transition zone of low, gently undulating hills, with elevations ranging from 50 to 90 m above sea level. Situated in the rain shadow of the *Harz* mountains the region is distinctly dry with 430 mm mean annual precipitation and 9 °C mean annual temperature and has a marked negative annual climatic water balance, so that a natural surface drainage system has not evolved in the area. There are a few man-made drainage channels designed primarily to remove temporary stagnant subsoil water in spring rather than to drain off surface waters or to prevent flooding. The mean size of the agricultural fields in the study area is about 45 ha and height differences rarely exceed more than 3 m within a field.

The predominant soil type of the Loess covered Tertiary plain is Chernozem in conjunction with Cambisols and Luvisols. The alluvial plain is characterized by coarse sand to fine sand, loamy and clayey sediments. The predominant soil types are Mollic Gleysols, Fluvisols and Planosols. In the moraine belt Cambisols, Regosols, and Kolluvisols evolved mainly from fine sandy and gravelly parent material. The highly diverse soil properties within the fields of the landscapes results in fine-scale pattern of soil texture and organic matter and challenge the future application of modern site specific management practices.

2.2. Data sets

The remotely sensed data was acquired using the HyMap™ scanner (Integrated Spectronics Pty Ltd, Australia), installed on the DLR Cessna Caravan aircraft. This scanner records spectra from 420 to 2480 nm wavelength, in 128 wavebands with full width half maximum (FWHM) bands of 15 and 20 nm for the 420–1803 nm range and for the 1949–2480 nm range, respectively (Cocks et al., 1998). The imagery was acquired with 6 m pixel size at nadir and a swath width of 30° from 12:30 to 13:15 h on 19th May 1999. Flight position and aircraft motion were recorded by an Inertial Navigation System with differential GPS. The data was atmospherically and geometrically corrected using the ATCOR procedures (Richter and Schläpfer, 2002) and a rectification procedure (Schläpfer and Richter, 2002).

To provide a representative soil data set for calibration purposes, we did not focus on extended data sampling but rather designed the sample selection procedure. The sample sites were picked out across the testsite's landscapes according to different combinations of soil forming geo-factors and represent largely the typical range of soil texture and soil organic matter in the arable soils of the region. A total of 72 locations were sampled on 12 bare soil fields covering roughly 700 ha distributed across the study area of 200 sq. km. Spatial coordinates of the locations were measured using Differential GPS. These locations were selected so as to cover the range of SOM and soil texture conditions that exist in the study site. From these samples, a subset of 60 samples was selected by an algorithm providing random subset selection. The selected samples were used as the calibration data for multivariate regression modeling. This data set was also used for cross-validation purposes (one-leave-out cross-validation). The remaining subset of 12 samples was used for stand-alone testset validation purposes to test the robustness of the different models. All 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 (LECO FP-2000). The C_{org} and N_t data range from 0.7% to 3.85% and from 0.07% to 0.37%, respectively. 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. The texture data ranges from 16% to 84% (sand), from 5% to 75% (silt) and from 7% to 26% (clay). Effects of soil surface moisture and roughness were excluded from this study by selecting bare soil fields after seed bed preparation and organizing the flight campaign after a period of dry weather to ensure continuous dry soil surface conditions.

The complexity of soil forming processes gives rise to multi-collinearity and auto-correlation of most soil parameters. In particular C_{org} and N_t are highly correlated in arable soils expressed by the limited range of C/N index for the most parts of arable land. The close relation between the two parameters in the data set used is given in Fig. 1. Consequently the C/N ratio just ranges from 8 to 12. Despite of the close connection, both C_{org} and N_t differences are in fact quite relevant since they reflect on one hand differences of organic matter composition and related soil properties e.g. nutrient mineralization, and on the other hand differences of past land use practices that either accumulated or decomposed SOM. Due to the potential occurrence of auto-correlation it was investigated to what extent it has to take notice of a spatial trend in the data. A

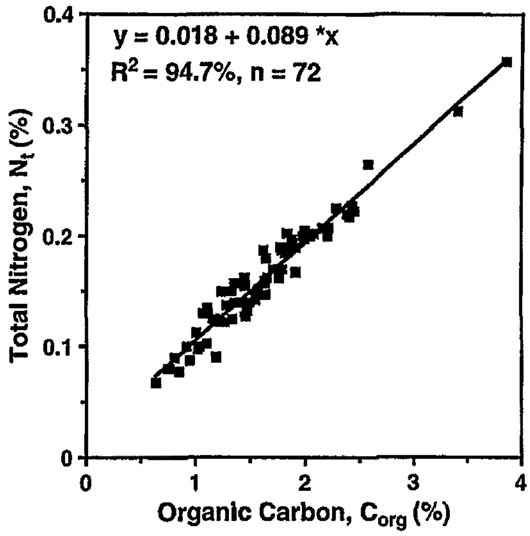

Fig. 1. Relation between organic carbon (C_{org}) and total nitrogen (N_t) in the data set used.

semivariogram analysis performed on the residual values reveals that there are no spatial correlations of the regression residues.

2.3. Multivariate regression modeling

In this paper, the focus is on the development of calibrated spectral models for the large scale soil mapping of C_{org}, N_t and the sand and clay contents. The complexity of soil formation and attendant multi-collinearity and auto-correlations between the various soil parameters, as were found for example between C_{org} and N_t, has led to the use of multivariate calibration techniques. Two multivariate regression techniques were used to develop the models to estimate soil parameters from the hyperspectral image data. Multi-variate calibration was performed using multiple linear regression (MLR) and partial least-square regression (PLSR). Both techniques allow the simultaneous quantitative determination of several soil parameters from individual spectral signatures. With the aim to develop a method of fine-scale topsoil mapping, one sub-goal of this study was to test the applicability of multivariate regression techniques in hyperspectral remote sensing rather than to compare the statistical level of model fit. Consequently a simpler (MLR) and a more complex numerical technique (PLSR) were chosen for investigation.

To optimize the PLSR models and to ensure the robustness against variability of natural factors, a recorded spectral signature should be considered as a whole (Dardenne, 1996). Similarly, it would be usual to include all of the possibly occurring variations of natural factor combinations in the model construction so as to achieve a robust model algorithm (Schenk and Westerhaus, 1991). PLSR reduced the reflectance spectra to a few relevant factors and regressed them to the soil value of a given sample (Martens and Næs, 1989). 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. A more detailed overview of PLSR model may be found in Hastie et al. (2003).

As distinct from PLSR, MLR performs a regression model selecting and combining those wavebands from the spectrum that explain the soil values at best. The general result of the MLR procedure is a prediction equation of the following kind:

$$C_p = B_0 + B_1 R_{\lambda_1} + B_2 R_{\lambda_2} + \ldots \ldots B_n R_{\lambda_n}$$

where C_p stands for the predicted soil compound value, B_0 is a constant coefficient for the current population, $B_1 - B_n$ are coefficients for each wavelength reading, R is the reflectance or its manipulation by the different spectra pre-treatment techniques and λ stands for the full width half maximum mean wavelength of the particular spectral channel of the HyMap sensor. The numerical methods are discussed in detail for MLR in Martens and Næs (1989) and for PLSR in Næs et al. (2002).

From both regression procedures, models were derived enabling prediction of the soil value from the spectra of samples with unknown soil value. Various wavelength regions and data pre-treatments were analyzed using an optimization routine to find the best calibration algorithm. To reduce noise and offset effects from the data, several spectra pre-treatment methods were utilized including vector normalization, min–max normalization, derivates (Savitzky and Golay, 1964) and multiple scatter correction (Geladi et al., 1985). Since several data pre-treatments result in similar error, one selected those leading to the inclusion of the fewest factors in the regression model (Næs et al., 2002). For each of the soil parameters, the model algorithm with the lowest root mean square error of cross-validation (RMSECV) was chosen as statistically the best. For this, cross-validation was computed as a one-leave-out cross-validation.

Finally, the optimal calibration model for each soil parameter was used to predict the particular parameter from the HyMap spectra for each HyMap image pixel of bare soils resulting in a map showing the distribution of the topsoil parameter. For independent prediction of the respective parameter, it was desirable similarly to find calibration algorithms for the different parameters that depend on different wavebands. The spectrum of each of the sample sites was extracted from the image data. For this, we applied a seeded region growing algorithm to identify the region with the spectrally most similar neighboring pixels (Adams and Bischof, 1994). The geographical coordinates of the sampled locations were used as seed coordinates for the seeded region growing algorithm. The algorithm's implementation guaranteed both to minimize the spectral Euclidean distance as a measure of the region's spectral similarity and to select at least 8 pixels to represent each sample location. From the spectra of these pixels, the mean spectrum of each site was calculated. The potential impact of smoothing spectra through averaging on the calibration results is a more robust calibration model through noise reduction in the spectral data. The channel at 1949 nm was excluded from the spectra due to insufficient signal to noise ratio at this water induced absorption band.

3. Results and discussion

3.1. Calibration procedure by PLSR

PLSR reduces the whole reflectance spectra to a few relevant factors and regresses them to the measured parameter of a given sample. While doing so, a calibration design is recommended that covers the whole range of possible values (Brereton, 2000). It is also reported that a certain redundancy within the spectra is useful to stabilize PLSR models against noise (Martens and Næs, 1989). When this is done, PLSR models are considered to be more robust than MLR calibration models. This might be relevant in particular for SOM parameters, since SOM has numerous broad overlapping absorption features located throughout the spectra and consequently influences the overall shape of the reflectance curve (Baumgartner et al., 1985). Thus, multivariate calibration was performed with PLSR first.

Table 1
Partial least-square regression (PLSR) and multiple linear regression (MLR) models statistics for the different topsoil parameters

Parameter	PLSR			MLR		
	Factors	R^2	RMSECV	Wavelength (nm)	R^2	RMSECV
C_{org}	7	0.90	0.29	800, 830, 1194, 1322	0.86	0.22
N_t	7	0.92	0.03	1194, 2115, 2185, 2220	0.87	0.02
Sand	9	0.95	9.7	2202, 2238, 2322, 2371	0.87	12.9
Clay	5	0.71	4.2	902, 950, 998, 1165	0.65	3.8

Table 2
Prediction equations of the best MLR regression models of the different soil compounds (C_p) and selected HyMap sensor wavebands (channels)

C_p	MLR prediction equation	HyMap channels
C_{org}	$3.5688-0.0318\,R_{800}+0.0362\,R_{830}-0.0173\,R_{1194}+0.0122\,R_{1322}$	26, 28, 52, 62
N_t	$0.3691-0.0001529\,R_{1194}-0.0007059\,R_{2115}+0.002087\,R_{2185}+0.001208\,R_{2220}$	52, 106, 110, 112
Sand	$3.3-0.68\,R_{2202}+1.15\,R_{2238}-0.76\,R_{2322}+0.3\,R_{2371}$	111, 113, 118, 121
Clay	$19.48-0.19\,R_{902}+0.14\,R_{950}+0.14\,R_{998}-0.09\,R_{1165}$	34, 37, 40, 51

PLSR models were derived enabling prediction of the C_{org} content from the spectra of samples with unknown C_{org}. The same calibration procedure was also employed to derive prediction models for the N_t content as well as prediction models for the sand and clay contents. All wavelength regions of the spectra and different data pre-treatments were analyzed using an optimization routine to find the best calibration algorithm. The selected PLSR models use both 7 factors for C_{org} and N_t, 9 factors for sand content and 5 factors for clay content. Overall the min–max normalization gave the best results for the different data pre-treatments. The PLSR model statistics are shown in Table 1.

3.2. Calibration procedure by MLR

The PLSR procedure reduces the spectra to a few factors and as such it does not support the idea of identifying the most significant individual wavebands. The MLR procedure, by contrast, does enable the selection of the most significant wavebands. Due to the increase of calculation time with the number of regression variables in MLR, the selection of algorithms was limited to regression models with a maximum of 4 spectral variables in this study.

Thus, all possible subsets of regression models with at least 1 and up to any combination of 2, 3 and 4

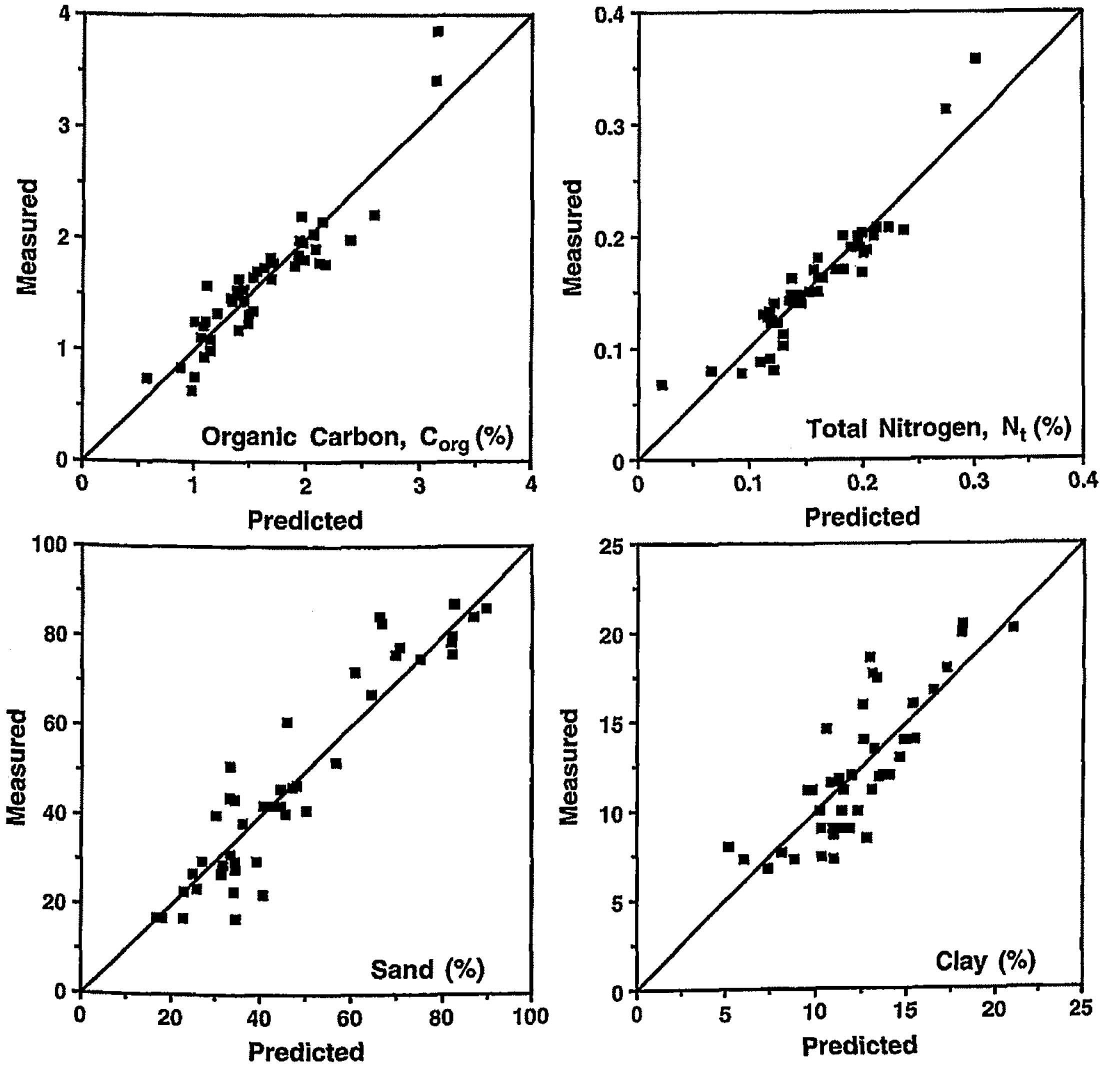

Fig. 2. Predicted soil parameters values from calibrated models versus measured calibration data set values. Solid line represents 1:1 agreement between the methods.

spectral variables were calculated for each of the selected soil parameters. The algorithm with the lowest RMSECV was chosen as statistically the best for each soil variable. As was found with the PLSR calibration, the min–max normalization gave the best results for the different data pre-treatments. The prediction equations of the best MLR regression models are listed in Table 2. MLR model statistics are compiled in Table 1.

The calibration model output is shown versus the measured reference values for C_{org} and for N_t in Fig. 2. Both models are in the same range of quality expressed by R^2 and RMSECV. Both the C_{org} and the N_t models use different channels, with the exception of HyMap's channel 52 at 1194 nm. This underlines the independence of the C_{org} and the N_t prediction in the spectral domain.

Fig. 2 also shows the calibration model output versus the measured reference values for sand and for clay in Fig. 2, as well. The sand model fits well to the calibration data as expressed by R^2 and RMSECV, whereas the clay model is characterized by a much weaker regression fit. Both texture MLR models also used different wavebands. Thus all four MLR models are independent in their parameter prediction from the spectral domain.

Only a few studies have been found in the literature that investigated the combination of hyperspectral remote sensing and regression modeling procedures to map soil parameters. Ben-Dor et al. (2002) investigated the MLR calibration procedure to quantify organic carbon using the DAIS-7915 airborne scanner, while Galvao et al. (2001) used the AVIRIS sensor and principle component regression for this purpose. Deronde et al. (2004) reported about the unsupervised classification of 8 coastal sand types with different mineralogical and organic composition and granulometric properties using the CASI airborne sensor. The different sensor instrumentation and numerical methods, the different regions and data pools make it difficult to compare the different procedures and the statistical quality.

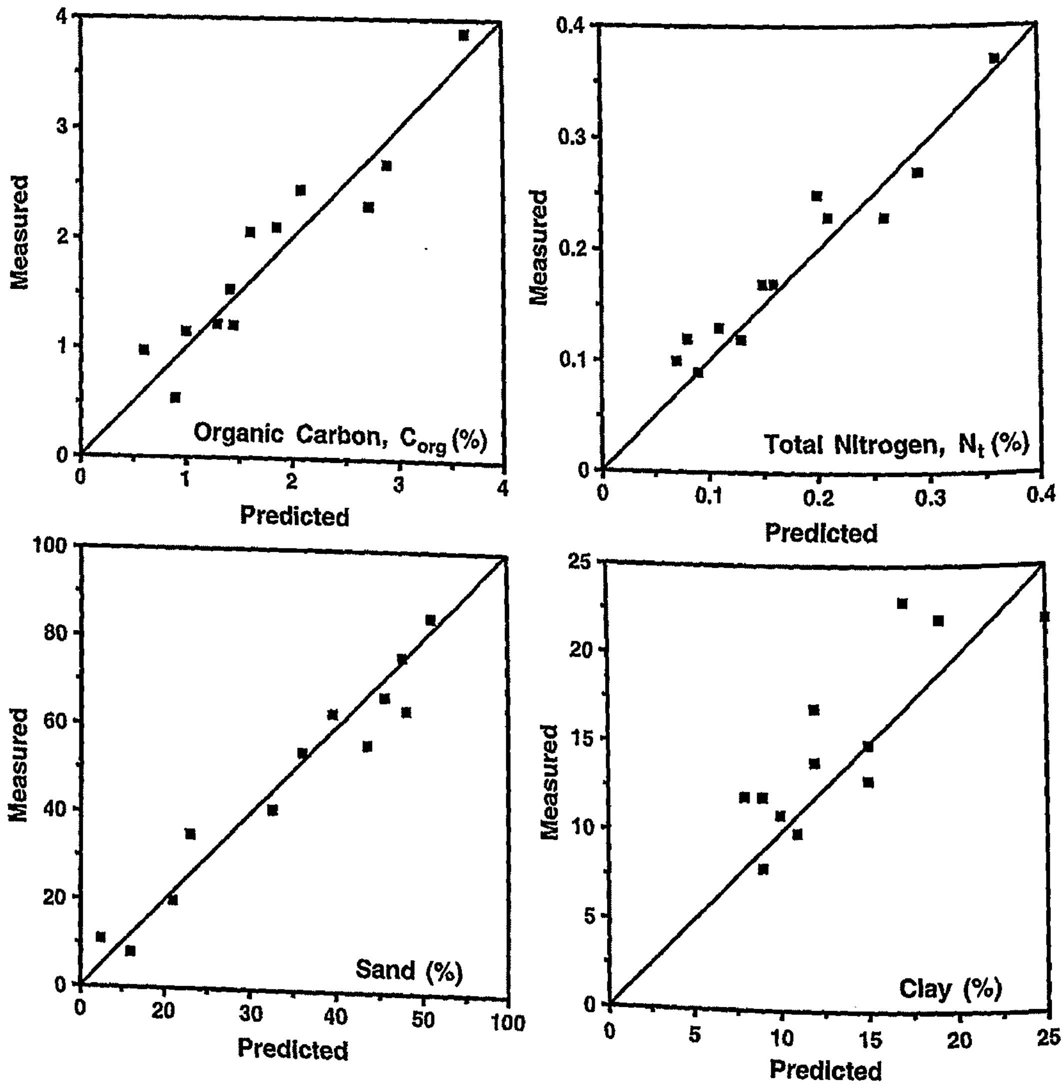

Fig. 3. Predicted soil parameters values from calibrated models versus measured testset validation data values. Solid line represents 1:1 agreement between the methods.

3.3. Validation

A stand-alone validation was performed to test the robustness of the different models. This validation test using 12 independent field samples (neither used in the calibration procedures nor in the one-leave-out cross-validation) indicated a close relation of $R^2=0.89$*** (*** = significant level at $P<0.001$) and $R^2=0.91$***, respectively between the C_{org} and N_t content predicted from the HyMap data using the MLR models and the C_{org} and N_t content measured by dry combustion method as standard reference value (Fig. 3). The validation test gave relations of $R^2=0.94$*** and $R^2=0.64$***, respectively between the sand and clay content predicted from the HyMap data using the MLR models and the sand and clay content measured by sieve and pipette analysis as standard reference value (Fig. 3).

With both calibration techniques, PLSR and MLR, we achieved similar results for R^2 and RMSECV. Some MLR models are characterized by a slightly better RMSECV. This is an unexpected result as there is generally strong multi-collinearity and auto-correlation between soil and spectral data. These characteristics have been found to cause a problem when using MLR but not when using PLSR (Næs et al., 2002) and as a consequence it was expected that PLSR would give a statistically better result than MLR. Further investigation is required of this issue, for which an extended data set over that used in this study would be necessary. The relatively weak regression model for clay content might be attributed to the relatively narrow range of data values. The robustness test based on the 12 independent samples gave similar results as the one-leave-out cross-validation concerning the regression coefficients and standard errors. Since this data set was completely independent from the data set used in calibration and cross-validation, the small statistical differences confirm the general applicability of the models. But of course, results from such a small data set can only be an approximate estimate of the models robustness and transferability. A more extended number of samples will be required for future studies.

3.4. Application

The optimal calibration models were used to calculate the C_{org}, N_t, sand and clay contents from the HyMap spectra of each HyMap image pixel. Both the PLSR and the MLR calibration procedures led to models which are characterized by comparable results in terms of R^2 and RMSECV. As the MLR models were more suitable for simple grid operations, they were used to calculate C_{org}, N_t, sand and clay maps of reference fields. Resulting topsoil maps show the distributions of C_{org}, N_t, sand and clay contents across an 88 ha-sized reference field as an example. The C_{org} values (Fig. 4) range from 1.2% (light color) to 2.5% (dark color). The N_t values (Fig. 5) range from 0.13% (light color) to 0.24% (dark color). The sand values (Fig. 6) range from 10% (light color) to 50% (dark color). The clay values (Fig. 7) range from 5% (light color) to 20% (dark color). As silt content supplements sand and clay to 100% (silt=100-sand–clay), a silt map was just calculated from the sand and clay map. A silt model was not attempted to do so far.

In the south-westerly part of the N_t map appears an area with distinctly higher values than in the surrounding. But a similar effect does not appear in the C_{org} map. After studying historical cadastral information it was found that this area matched ancient boundaries in the land register. Prior to land consolidation during the 1950s and 1960s, the area consisted of a number of fields that were independently owned and managed by different farmers. Enquiries of the local community revealed the opinion these fields had received significantly more dung than was usually applied, prior to consolidation. If these opinions are correct, then one possible explanation for the high N_t values in the area is that the historical record of high application of dung and manure is still recorded in the hyperspectral data and can be revealed using multivariate analytical techniques even after the subsequent 40 years of uniform management practice.

To investigate this effect in greater detail at this locality, an additional set of 16 samples arranged in 8 pairs across the ancient boundary was taken in the immediate vicinity of the ancient field boundaries and was analyzed for C_{org} and N_t by dry combustion method. The data confirmed that the N_t values along the former field boundaries differ by about 0.015% N_t for C_{org} values lower than 2%. At higher C_{org} values the N_t differences across the boundary were not significant. In contrast to this no significant differences were found for the C_{org} values at all data pairs accordingly to the visible pattern in the C_{org} map. It is concluded from these results that the organic matter within the sub-field has risen by the application of dung and manure and has developed a different composition with higher N_t content. But if a higher C_{org} content exist in the soil (here at around 2%) a higher organic matter content did not emerge from the higher application of dung and manure.

Past management effects were also observed on fields in other locations within the study area. In many of these cases the observed differences were within the

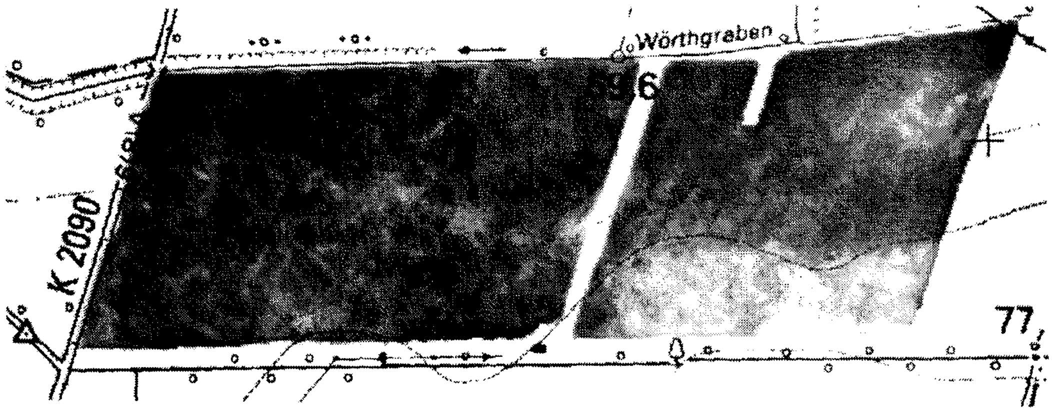

Fig. 4. Spatial distribution of the predicted organic carbon content (C_{org}) across the 88 hectare reference field 'Pfingstbreite' (range from 1.2%=light color to 2.5%=dark color).

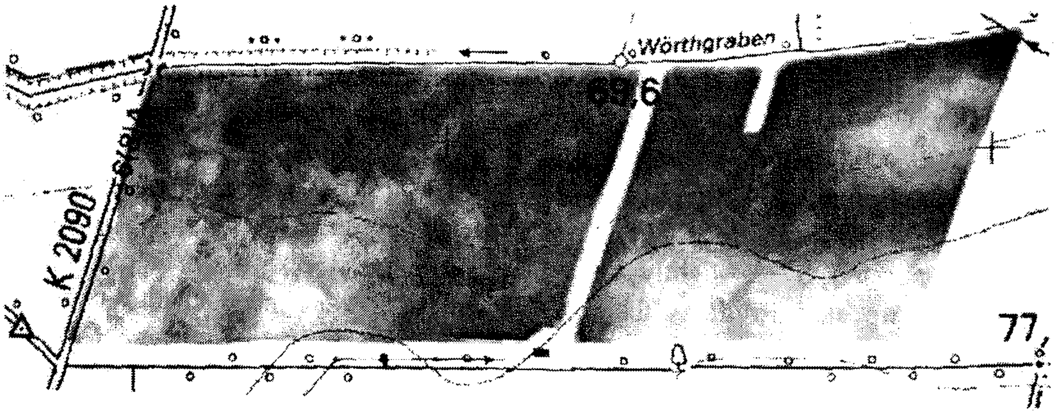

Fig. 5. Spatial distribution of the predicted total nitrogen content (N_t) across the 88 hectare reference field 'Pfingstbreite' (range from 0.13%=light color to 0.24%=dark color).

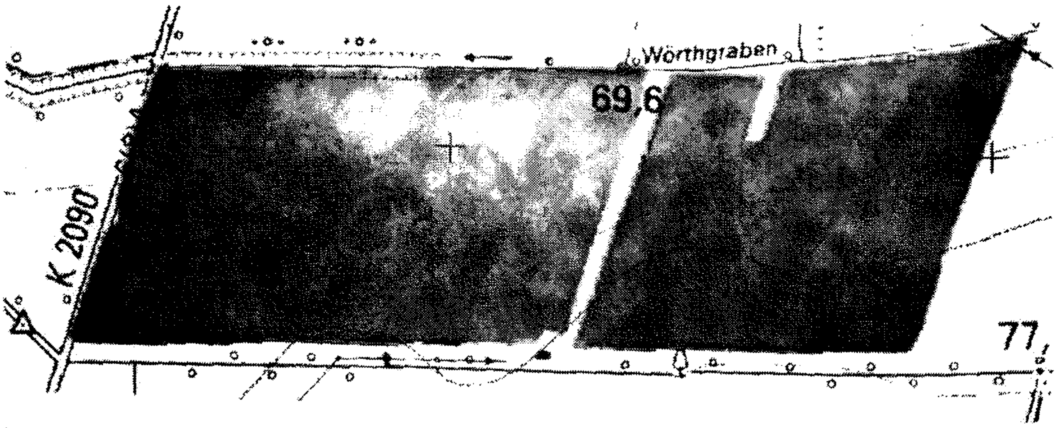

Fig. 6. Spatial distribution of the predicted sand content across the 88 hectare reference field 'Pfingstbreite' (range from 10%=light color to 50%=dark color).

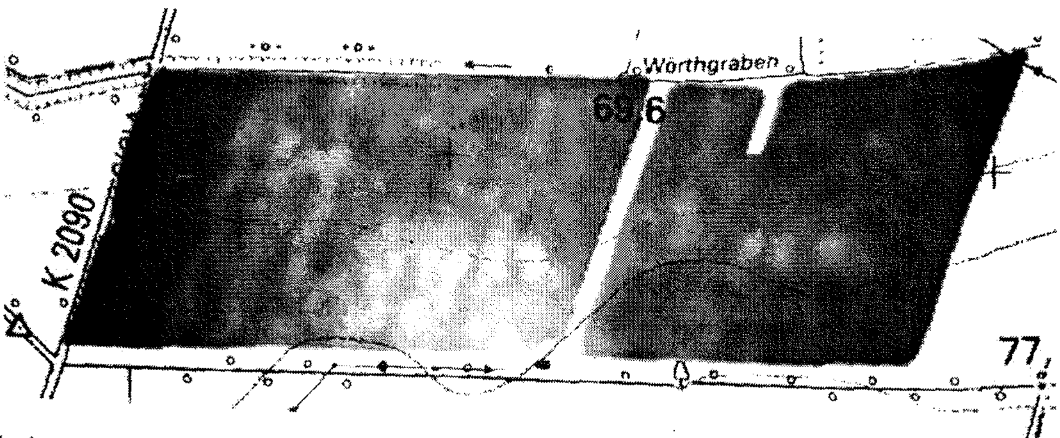

Fig. 7. Spatial distribution of the predicted clay content across the 88 hectare reference field 'Pfingstbreite' (range from 5%=light color to 20%=dark color).

accuracy tolerances of the model (as expressed by the RMSCV) and as such care must be taken with the conclusions that can be drawn. However these methods, and the statistical confidence that can be placed on these results, are all based on a pixel by pixel analysis, so that spatial correlation is not taken into account. Thus even when the results are small in the instance of individual pixels, when they form patterns that match other patterns in the landscape, then they become important evidence of historical activities. As a consequence the method has the potential to be a tool to understand the historical development of land using practices in an area, as well as providing the fine resolution soil mapping required for site specific soil management.

4. Conclusion

This is supposed to be the first study that employed the HyMap sensor and multivariate regression modeling for indirect topsoil parameter measurement. This remote sensing approach shows the potential benefits of using image data with carefully located in situ field data in digital soil mapping. The results also indicate that soil mapping procedures must be adapted to the soil parameter of interest and that multivariate calibration techniques allow calibration modeling as a generic procedure. With the HyMap spectrometer, the wavebands that are relevant to the mapping of the different soil parameters can be recorded and used in adapted calibration models to simultaneously predict a suite of different soil parameters. The method proposed provides a means of simultaneously estimating topsoil SOM and texture in an extensive, rapid and non-destructive manner, whilst avoiding the spatial accuracy problems associated with interpolation. The use of remotely sensed data in the manner proposed in this paper can also be used to monitor and better understand the influence of management and land use practices on SOM composition and content. In precision agriculture it can be used to establish the precise spatial locations of specific management practices, as a pre-requisite to much of the modeling and estimation that needs to be conducted for variable rate applications.

The preparation of spatially accurate, high resolution soil maps for site specific management is still a matter that is to be resolved. It has been shown in this work that hyperspectral remotely sensed data, combined with field data in model calibration based on Partial Least-Square Regression or Multiple Linear Regression has the potential to contribute significantly to the goal of fine-scale soil mapping. With this simultaneous, rapid, non-destructive determination of topsoil SOM and texture over extended areas, maps of secondary soil attributes, as aggregate stability or erodibility of soil surface, can be calculated. Such soil attributes are necessary to be considered for the appropriate application of different site specific management activities as variable rate seeding and fertilization. The results presented in this study lead us to the suggestion to better amalgamate soil sciences with remote sensing and multivariate calibration techniques to improve precision agriculture applications and SOM monitoring studies in future.

Acknowledgment

This paper draws on work from research projects funded by the German Federal Ministry of Education and Research and the Bavarian Research Foundation.

References

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16 (6), 641–647.

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

Baumgartner, M.F., Silva, L.F., Biehl, L.L., Stoner, E., 1985. Reflectance properties of soils. Adv. Agron. 38, 1–44.

Ben-Dor, E., Banin, A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Sci. Soc. Am. J. 59, 364–372.

Ben-Dor, E., Inbar, Y., Chen, Y., 1997. The reflectance spectra of organic matter in the visible, near infrared and short wave infrared region (400–2500 nm) during a control decomposition process. Remote Sens. Environ. 61, 1–15.

Ben-Dor, E., Patkin, K., Banin, A., Karnieli, K., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data — a case study over clayley soils in Israel. Int. J. Remote Sens. 23 (6), 1043–1062.

Brereton, R.G., 2000. Introduction to multivariate calibration in analytical chemistry. Analyst 125, 2125–2154.

Bullock, D.S., Bullock, D.G., 2000. From agronomic research to farm management guidelines: a primer on the economics of information and precision technology. Prec. Agric. 2, 71–101.

Cocks, T., Jenssen, R., Stewart, A., Wilson, I., Shields, T., 1998. The HyMap airborne hyperspectral sensor: the system, calibration and performance. In: Schaepman, M., Schläpfer, D., Itten, K.I. (Eds.), Proc. 1st EARSeL Workshop on Imaging Spectroscopy. 6–8 October 1998, Zurich, EARSeL, Paris, pp. 37–42.

Cozzolino, D., Morón, A., 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. J. Agric. Sci. 140, 65–71.

Dalal, R.C., Henry, R.J., 1986. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. Soil Sci. Soc. Am. J. 50, 120–123.

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

Deronde, B., Houthuys, R., Sterckx, S., Debruyn, W., Fransaer, D., 2004. Sand dynamics along the Belgian coast based on airborne hyperspectral data and lidar data. EARSeL Proc. 3 (1), 26–33.

Galvao, L.S., Pizarro, M.A., Epiphanio, J.C.N., 2001. Variations in reflectance of tropical soils: spectral–chemical composition relationships from AVIRIS data. Remote Sens. Environ. 75, 245–255.

Geladi, P., McDougall, D., Martens, H., 1985. Linearisation and scatter correction for near infrared reflectance spectra of meat. Appl. Spectrosc. 39, 491–500.

Hastie, T., Tibshirani, R., Friedman, J., 2003. The Elements of statistical Learning: Data Mining, Inference, and Prediction. Springer, New York.

Ingleby, H.R., Crowe, T.G., 2000. Reflectance models for predicting organic carbon in Saskatchewan soils. Can. Agric. Eng. 42 (2), 57–63.

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

Næs, T., Isaksson, T., Fearn, T., 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). Geol. Jahrb., Reihe F, Bodenkd. 25 (131 pp.).

Reeves III, J., McCarty, G., Mimmo, T., 2002. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. Environ. Pollut. 116, 277–284.

Richter, R., Schläpfer, D., 2002. Geo-atmospheric processing of airborne imaging spectrometry data: Part 2. Atmospheric/topographic correction. Int. J. Remote Sens. 23, 2631–2649.

Runge, E.C.A., Hons, F.H., 1998. Precision agriculture: development of a hierarchy of variables influencing crop yields. In: Robert, P.C. (Ed.), Proc. 4th Int. Conf. On Precision Agriculture. American Society of Agronomy, Madison, WI, pp. 143–158.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least-squares procedures. Anal. Chem. 36, 1627–1639.

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

Schläpfer, D., Richter, R., 2002. Geo-atmospheric processing of airborne imaging spectrometry data: Part 1. Parametric orthorectification. Int. J. Remote Sens. 23, 2609–2630.

Thomasson, J.A., Sui, R., Atkins, D.C., 2001. Soil Reflectance Sensing for determining soil properties in precision agriculture. ASAE Paper, vol. 001044. ASAE, St. Joseph, Michigan.

Udelhoven, T., Emmerling, C., Jarmer, T., 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: a feasibility study. Plant Soil 251, 319–329.