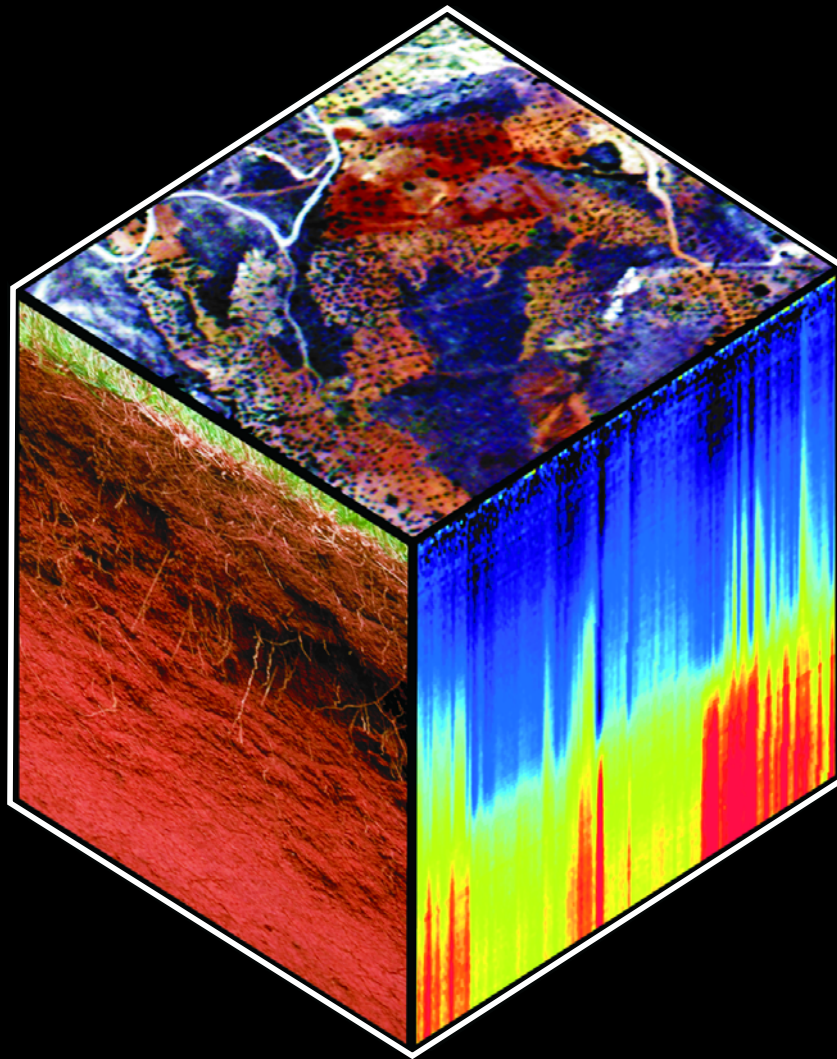# THE INFLUENCE OF VEGETATION ON THE SPECTROSCOPIC ESTIMATION OF SOIL PROPERTIES

Harm Bartholomeus

The influence of vegetation cover on the spectroscopic estimation of soil properties

Harm Bartholomeus

Thesis committee

Thesis supervisor
Prof. Dr. sc. nat M.E. Schaepman
Professor of Geo-Information Science with special attention for Remote Sensing
Wageningen University, The Netherlands
Professor of Remote Sensing
University of Zurich, Switzerland

Thesis co-supervisor
Dr. ir. L. Kooistra
Assistant Professor Laboratory of Geo-Information Science and Remote Sensing
Wageningen University, The Netherlands

Other members
Prof. Dr. L. Brussaard   Wageningen University - The Netherlands
Prof. Dr. rer. nat. J. Hill   Universität Trier - Germany
Prof. Dr. S. De Jong   Utrecht University - The Netherlands
Dr. R. Viscarra Rossel   CSIRO Canberra - Australia

The influence of vegetation cover on the spectroscopic estimation of soil properties

Harm Bartholomeus

Thesis
submitted in partial fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Doctorate Board
to be defended in public
on Wednesday 30 September 2009
at 1.30 PM in the Aula.

Table of contents

Voorwoord

Eigenlijk wilde ik helemaal geen promotie onderzoek doen… Maar zoals dat gaat lopen dingen af en toe anders dan je ooit gepland of bedacht hebt. Werken in het onderwijs beviel goed, maar ik wilde niet de rest van mijn leven voor klassen met pubers doorbrengen. Onderzoek trok me toch ook wel, maar vier jaar lang in je eentje aan een proefschrift werken? Nee, dank je… Maar met een beetje geluk hoef je niet te kiezen. De perfecte baan bij het CGI kwam precies op tijd. Promotiewensen kwamen nou niet direct op het lijstje van gevraagde kwalificaties voor. "Ach, dat regel ik wel als ik er eenmaal zit", dacht ik bij mezelf. Zoals je aan dit boek kunt zien is het met dat promotieonderzoek toch wel goed gekomen. Promoveren blijkt toch niet zo'n eenzame onderneming te zijn als ik vreesde. Daarom wil ik een hoop mensen bedanken voor hun rol in dit onderzoek.

Allereerst moet ik de mensen bedanken die aan de basis hebben gestaan van dit werk, al is hun uiteindelijke rol daarin beperkt gebleven. Gerrit was degene die er, zodra ik bij het CGI begon, op hamerde dat ik ook onderzoek moest doen. Mijn interesses sloten goed aan bij zijn onderzoeksveld en door samen met hem afstudeerders te begeleiden heb ik een hoop geleerd en heb ik bovenal de hoofdrichting van mijn onderzoek bepaald. Jan en Arnold, toen ik tijdens de eerstvolgende "contractbesprekingen" aangaf dat ik het onderzoek wat ik al deed in een promotie wilden laten eindigen waren jullie direct enthousiast. Bedankt voor het vertrouwen dat jullie me daarmee gaven, al zal Jan zich vast wel eens achter zijn oren gekrabd hebben als hij me weer vroeg naar mijn einddatum… 2007, 2008…uhm, 2009 dan maar? Zeker in de eerste jaren van het onderzoek was het vooral iets wat ik voor mijn gevoel erbij deed, naast het onderwijs. Gelukkig was daar die eerste jaren altijd nog mijn kamergenoot Joep, die me maar wat graag wees op mijn traag vorderende onderzoek. Dat genoegen was overigens geheel wederzijds, het was prettig om de dagen in C310 door te brengen met iemand met dezelfde belangstelling voor onderwijs, reizen, muziek en sport (actief dan wel passief).

Na de eerste tijd wat te hebben "aangerommeld" kwam de vaart er pas in op het moment dat Michael mijn promotor werd. De manier van begeleiden kan waarschijnlijk het best omschreven worden met "veelvuldige korte overleg momenten bij de koffiemachine". Dat werkte prima en als ik een keer een wat groter probleem had stond je deur altijd open. In de afrondingsfase werd de communicatie door je vertrek naar Zurich wat lastiger, maar uiteindelijk is het toch allemaal goed gekomen. Lammert, ik had me geen betere copromotor kunnen wensen. Iets minder koffiemachine, wat meer "regular meetings", waarbij je mijn teksten altijd weer had weten te voorzien van een hoop rode inkt, zelfs in de derde versie ☹. Is het dan nooit goed? Nee, het kan altijd beter, er blijven altijd vragen, niet te snel tevreden

zijn! Goede, soms wat frustrerende, lessen voor een wat gemakzuchtig uitgevallen persoon als ik.

Daarnaast wil ik bij deze alle collega's bij het CGI bedanken met wie ik in de afgelopen jaren heb samengewerkt. Jullie zijn met teveel om persoonlijk te noemen, maar dankzij jullie is het CGI zo'n leuke plek om te werken.

Aan de werkkant wil ik afsluiten met het bedanken van de vele studenten met wie ik in de loop der jaren heb gewerkt. Dit zijn ten eerste alle studenten die ik in verschillende cursussen remote sensing kennis en enthousiasme heb proberen bij te brengen. Deze cursussen vormen voor mij een welkome afwisseling met het onderzoek. Daarnaast heb ik in de loop der jaren ook een groot aantal afstudeerders begeleidt. Hun onderzoek en de discussies die we samen gehad hebben, hebben in een grote mate bijgedragen aan dit werk. Jullie vragen, waarop ik niet altijd meteen een antwoord wist, dwongen me om verder te kijken en dieper te graven. Bedankt voor jullie vragen, blijf ze stellen…

Naast het werk heb ik nog een ander leven, wat voor mij minstens zo belangrijk is. Een goede balans tussen inspanning en ontspanning zorgt ervoor dat ik overal lol in blijf houden. Michiel, Arnaud, Lisette, Elwin en Timo, we gaan de komende jaren nog veel empirisch onderzoek doen naar de grenzen van ons kunnen… 24 uur? Ja, lukt… 5 dagen? Oké, proberen we gewoon… Jullie zijn gek, maar wel gezellig… Daarnaast mag ik gelukkig ook normale mensen tot mijn vrienden rekenen. De frequentie waarmee we elkaar zien en de gespreksonderwerpen zijn de laatste jaren flink veranderd. Uitgewaaierd over het hele land en zelfs héél ver daarbuiten hebben we het tegenwoordig over huizen kopen, gezinsauto's, carrièremogelijkheden en de aankomende of opgroeiende kinderen. Ik kijk nu al uit naar onze gesprekken over pensioenregelingen en de nieuwste modellen rollators.

Koen en Ruud, alvast bedankt voor het beantwoorden van alle bedrijfseconomische en echohydrologische vragen tijdens de verdediging. Pap en mam, jullie hebben me altijd aangemoedigd en de mogelijkheden gegeven om dat te studeren wat ik wilde. Ik realiseer me nu dat ik daarmee jarenlang in een luxepositie heb verkeerd.

Het laatste bedankje is natuurlijk voor Baukje, Gijs en Art. Hoewel het aantal avonden dat ik thuis heb zitten werken beperkt is gebleven, was ik er de laatste maanden met mijn hoofd toch niet altijd 100 % bij. Gelukkig riepen jullie me steeds weer snel bij de les. Eigenlijk is het ook veel leuker om eindeloos met de trein en blokken te spelen, boekjes te lezen en liedjes te zingen…

Harm

Introduction

1. Introduction

1.1 Background

Soils and its resources are of great importance for the production of fiber and food. Recently, the soil functioning as carbon pool has been added to the list of soil functions (Bouma and Droogers 2007). Therefore, conservation of soil resources is high on the political agenda. Within the EU soil thematic strategy, the loss of organic matter and erosion are mentioned as some of the major threats to the soil resource (Van-Camp et al. 2004). One of the recommendations in Van-Camp et al. (2004) focuses on the spatial and temporal changes of soil processes and parameters. To determine soil quality as a resource, there is a need for regular monitoring of its chemical and physical properties. To map and monitor relevant chemical and physical soil components adequately, comprehensive and extensive datasets of high quality are needed, which is necessary to cover variation in time and place.

Quantitative estimation of the exact amount, spatial distribution and temporal change of soil properties is still challenging. Conventionally, soil samples are analyzed by means of soil extraction procedures which are accurate due to new technologies that increase sensitivity but which are also time consuming, cost-inefficient and destructive (Janik et al. 1998; Liang 2004). Spatially, soil samples are sampled according to a specific sampling scheme and spatial interpolation techniques are used to prepare continuous maps, but for accurate interpolation intensive sampling is required (Garten Jr. and Wullschleger 1999; Smith 2004). Altogether, this does not provide sufficient accuracy and an adequate spatial resolution (Barnes et al. 2003), because the sampling density required to map the high spatial variability and slow temporal changes in soil properties can be very high.

Therefore, it is widely acknowledged that there is a need for analytical techniques that allow rapid sampling and instant determination of soil properties at field and regional level. To get accurate, relevant and correct data, this routine should be able to deal with variability in time and place (Malley et al. 1999; McKenzie et al. 2000; Viscarra Rossel and McBratney 1998).

1.2 VNIR spectroscopy of soil properties

Visible and Near InfraRed (VNIR) spectroscopy, which implies the measurements and analysis of reflected radiance in the range of 350-2500 nm (Brown et al. 2006), has proven to be a useful tool to quantify various soil characteristics simultaneously in extended

geographical and application areas (Escadafal 1993; Farrand 1997; Palacios-Orueta et al. 1999; Warell 2003). Mathematical models, ranging from simple linear regression models to more complex multivariate statistical models, are used to link the measured reflected radiance to actual soil properties. In general, these models are empirical. In the calibration stage, prediction equations are developed using training sets which represent the sample population. The models are validated by predicting the properties for independent samples.

There are three types of VNIR spectroscopy techniques, used for different spatial scales and in different environments: laboratory spectroscopy, portable or field spectroscopy and imaging spectroscopy. Ground based sensors (usually point spectrometers) have demonstrated their capability to accurately determine soil properties in the laboratory (Ben-Dor et al. 1997; Reeves III et al. 2000; Sørensen and Dalsgaard 2005; Sudduth and Hummel 1993; Viscarra Rossel et al. 2006), or directly in the field with a portable spectrometer (Barnes et al. 2003). In the case of laboratory spectroscopy, illumination conditions and pre-treatments (e.g. drying, grounding) can be controlled, which yields most stable model calibrations. Field spectroscopy has been used for rapid *in-situ* monitoring of soils (Kooistra et al. 2003b; Udelhoven et al. 2003). In the usual setup, the sensor measures only the reflectance of the top-soil (≈50 µm), although devices are developed to measure the reflectance at different depths within an auger hole (Veristech 2009). When dedicated to the site-specific management of inputs in precision agriculture the sensor is often mounted on a tractor (Mouazen et al. 2007). In the field situation the sun is usually used as a light source, and illumination conditions can vary. Furthermore, no pre-treatment is applied to the samples. Regional studies most often rely on imaging spectroscopy, mainly based on airborne imaging spectrometers (Ben-Dor et al. 2002; Demattê et al. 2004).

Soil properties can spectrally be measured with a reasonable accuracy level, depending on the type of instrument used and the environmental conditions. The accuracy is generally lower than for most routine laboratory analysis methods, but the large number of samples that can be produced outweighs the slight loss of precision compared to traditional chemical analyses. However, certain complications of VNIR spectroscopy can be recognized. Since physical and chemical soil components interact in a complex way (Barnes et al. 2003; Gomez et al. 2008), optically active background variables can contribute to the reflection and absorption characteristic. Hence, significant correlations between reflection characteristics at specific wavelengths are not per definition causal due to other, unidentified, interacting chemical and physical soil components. As a result, transferability of reflectance based models between data sources is low and calibrations are only site-specific (Udelhoven et al. 2003).

As outlined in the previous sections, it can be concluded that (multivariate) statistical models combined with spectral measurements give promising results for the estimation of

soil properties (Stevens et al. 2008; Udelhoven et al. 2003; Viscarra Rossel et al. 2006). However, the found relations are often site-specific and, due to the complex interaction between soil components and the use of statistical analysis, it often remains unclear if the model relations between absorption features and soil properties are causal or not. Next to that, the often used statistical models (like Partial Least Squares Regression (PLSR)) are not directly transferable to sensors with other spectral characteristics, but require a recalibration. Indices, that are based on physically known absorption features, can be more appropriate for the use with different sensors, as long as the requirements for specific spectral bands are met.

1.3 Quantitative remote sensing and scaling

The role and importance of remote sensing for the monitoring of soil properties (carbon in specific) has been discussed by Post et al. (2001), who indicated that it can be especially useful in regions lacking detailed (in situ) geographical information. Imaging spectroscopy offers the possibility to resolve the limitation on the availability of appropriate information on soil and land resources, due to its potential advantages like speed and efficiency (McBratney et al. 2006), contiguous spectral coverage and continuous spatial coverage. The need for relatively high spectral resolution data to determine soil properties using reflectance spectra has been extensively discussed in Ben-Dor et al. (1999). Presently, technical improvements allow mapping of soil properties using higher spectral and spatial resolution imaging spectrometers with better signal-to-noise ratios (Demattê et al. 2004). The application of imaging spectroscopy for soil property estimation was limited to airborne configurations, but recently Gomez et al. (2008) used a space-borne imaging spectrometer (Hyperion) to quantify soil organic carbon (SOC). However, imaging spectroscopy suffers some specific drawbacks, due to the larger spatial scale at which it is operated. Going from point spectroscopy to imaging spectroscopy is not only a transfer between spatial scales, but also needs to deal with problems like data having a low signal-to-noise level, contamination of the atmosphere, large datasets and spectral mixing problems (Ben-Dor et al. 2008). Due to vegetation cover, which causes spectral mixing, the applicability of imaging spectroscopy to derive soil properties in partially vegetated areas is limited. Fractional vegetation cover has a large influence on spectral reflectance and limits accurate quantification of soil properties. Siegal and Goetz (1977) reported that mineral absorption features can be obscured by as little as 10 percent green vegetation and that its presence may severely hinder or limit soil and lithological discrimination. Murphy and Wadge (1994) concluded that fractional vegetation cover complicates the identification of different soil/rock types and even makes identification of specific soil/rock types impossible, due to masking of absorption features. A traditional solution is to mask out the areas with high vegetation cover, using vegetation indices with often case-specific threshold values (e.g. Wester et al. (1990)). Despite these

attempts, soil information for vegetated areas is lacking, and quantitative estimates of the errors due to the presence of vegetation are usually not given. In temperate regions, agricultural fields are covered with crops for the largest part of the year and vegetation shows large variation in time and space. This limits image acquisition for bare soil studies in the agro-ecosystem to only a few weeks. By removing the influence of vegetation on spectral soil reflectance, the flight-window in which airborne spectrometers can be used for soil investigations, can be extended. This will increase the possibilities for more frequent and accurate quantitative mapping of soil properties over large areas.

Earlier studies have developed methods for quantitative estimation of soil properties in partially vegetated areas. However, for these cases, vegetation was mostly used as proxy to estimate soil properties. Asner et al. (2003) related SOC and nitrogen field observations to fractional cover data for photosynthetic and non-photosynthetic vegetation and were able to show the trends in these soil properties at an ecosystem level. Kooistra et al. (2003a) used vegetation development as a proxy to estimate SOC and Zn in floodplains. For discrimination of different soil units, Schmidtlein (2004) used the spectral characterization of plant functional types.

Gomez et al. (2008), who performed a SOC study based on space-borne imaging spectrometer data (Hyperion), assumed that the drop of accuracy with this dataset compared to field spectroscopy data, was partly due to the Hyperion spatial resolution of 30 m. They suggested spectral unmixing as a possible technique, needed to extract the soil spectra from the mixed spectroscopic data.

Spectral unmixing techniques have played a central role in the analysis of remote sensing images with fractional vegetation cover over the past decades (Asner and Heidebrecht 2002; Garciá-Haro et al. 1996; Garcia and Ustin 2001; Sohn and McCoy 1997). Spectral unmixing algorithms are used to estimate the abundance of endmembers within a pixel and are well able to estimate the amount of vegetation in mixed pixels. Some advanced spectral unmixing techniques have been proposed to tackle the problem of vegetation influence in the case of soil classification. Luo et al. (2005) eliminated the vegetation effect with fully constrained spectral unmixing techniques, which increased the overall accuracy of their soil classification with 18 percent. Bierwirth (1990) removed the vegetation effect from reflectance data of geological materials by extrapolating the determined amounts of non-vegetated materials. He found that with careful modeling and depending on the instrument sensitivity, only a small signal from geological materials may be required to obtain spatially meaningful geological information in strongly vegetated areas.

1.4 Objectives

The main objective of this thesis is to quantify the robustness of VNIR spectroscopy based soil property models and the influence of vegetation on these models. Furthermore, it is intended to develop a methodology for quantitative mapping of soil properties in fractionally vegetation covered agricultural fields.

To achieve the objectives the following questions are addressed:

A. Which scaling issues have to be considered when using VNIR spectroscopy for estimation of soil properties?
B. What is the potential of VNIR spectroscopy for the prediction of soil properties and what is the stability of the calibrations?
C. How sensitive are models to variation in soil type and can models be used to predict soil properties for soil types that are not included in the model calibration?
D. What is the influence of fractional vegetation cover on the estimation of soil properties?
E. Can we determine soil properties from a mixed soil and vegetation signal?

1.5 Outline

The core of this thesis (Chapters 2-6) is based on a series of five peer-reviewed papers. Each chapter in this thesis is a step in answering the research questions mentioned above and is introduced separately by stating its research goals and by outlining its relationship with other relevant work.

Chapter 2 describes the issues of scaling and one of the technical solutions to overcome the spectral mixing problems that arise when the step from point spectroscopy to imaging spectroscopy is made: spectral (un)mixing. This technique will be used in the other chapters to deal with spatial scaling issues (Question A).

Chapter 3 aims to investigate the potential of different VNIR spectroscopy approaches for the monitoring of SOC in croplands at the regional scale by evaluating i) the loss of accuracy by using such techniques outside the controlled conditions of the laboratory, ii) the stability of the calibrations and iii) the capacity of such techniques to process a larger number of samples than conventional techniques allow (Question B).

Chapter 4 discusses the development of index based SOC prediction models and their robustness towards variation in soil type. Given the difficulties to transfer PLSR models from one sensor configuration to another, SOC prediction models based on spectral indices

related to the biochemical composition of SOC are developed. The robustness of these models towards inclusion of independent soil types (not included in the calibration phase) is an important measure for their usability (Question C).

Chapter 5 shows the effect of fractional vegetation cover on the reflectance of mixed pixels and the influence this has on the retrieval of iron content. Vegetation influence is simulated to investigate the magnitude of the error introduced by fractional vegetation cover. Finally, a location-specific solution is proposed to improve the mapping of iron oxides for the fractionally covered areas (Question D and E).

Chapter 6 describes a more generic method to improve the estimation of soil properties in agricultural fields with fractional vegetation cover. A spectral unmixing based approach is presented, to eliminate vegetation influence from the spectral reflectance of mixed pixels. This procedure is named Residual Spectral Unmixing (RSU), which returns so-called residual soil spectra. These are used in combination with PLSR to create a SOC-prediction model. In this way, imaging spectrometer data can be used to map soil properties of fractionally vegetation covered fields (Question D and E).

Finally, chapter 7 concludes this thesis with a summary, discussion of the main findings and suggestions for further work.

References

Asner, G.P., Borghi, C.E., & Ojeda, R.A. (2003). Desertification in Central Argentina: Changes in Ecosystem Carbon and Nitrogen from Imaging Spectroscopy. *Ecological Applications, 13*, 629-648

Asner, G.P., & Heidebrecht, K.B. (2002). Spectral unmixing of vegetation, soil and dry carbon in arid regions: Comparing multi-spectral and hyperspectral observations. *Int. J. Remote Sens., 23*, 400-410

Barnes, E.W., Sudduth, K.A., Hummel, J.W., Lesch, S.M., Corwin, D.L., Yang, C., Daughtry, C.S.T., & Bausch, W.C. (2003). Remote- and ground-based sensor techniques to map soil properties. *Photogramm. Eng. Remote Sens., 69*, 619-630

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) during a Controlled Decomposition Process. *Remote Sensing of Environment, 61*, 1-15

Ben-Dor, E., Irons, J. R. and Epema, G. F., 1999. Soil Reflectance. In: Rencz, A. N. (Eds.). Remote Sensing for the Earth Sciences: Manual of Remote Sensing. John Wiley & Sons, Inc. , New-York, pp. 111-188

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

Ben-Dor, E., Taylor, R.G., Hill, J., Demattê, J.A.M., Whiting, M.L., Chabrillat, S., & Sommer, S. (2008). Imaging Spectrometry for Soil Applications. In, *Advances in Agronomy* (pp. 321-392)

Bierwirth, P.N. (1990). Mineral mapping and vegetation removal via data-calibrated pixel unmixing, using multispectral images. *International Journal of Remote Sensing, 11*, 1999-2017

Bouma, J., & Droogers, P. (2007). Translating soil science into environmental policy: A case study on implementing the EU soil protection strategy to The Netherlands. *Environmental Science & Policy, 10*, 454-463

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., & Reinsch, T.G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma, 132*, 273-290

Demattê, J.A.M., Campos, R.C., Alves, M.C., Fiorio, P.R., & Nanni, M.R. (2004). Visible-NIR reflectance: a new approach on soil evaluation. *Geoderma, 121*, 95-112

Escadafal, R. (1993). Remote Sensing of soil color: principles and applications. *Remote Sensing Rev., 7*, 261-279

Farrand, W.H. (1997). Identification and mapping of ferric oxide and oxyhydroxide minerals in imaging spectrometer data of Summitville, Colorado, USA and the surrounding San Juan Mountains. *International Journal of Remote Sensing, 18*, 1543-1552

Garciá-Haro, F.J., Gilabert, M.A., & Melia, J. (1996). Linear spectral mixture modelling to estimate vegetation amount from optical spectral data. *International Journal of Remote Sensing, 17*, 3373-3400

Garcia, M., & Ustin, S.L. (2001). Detection of interannual vegetation responses to climatic variability using AVIRIS data in a Coastal Savanna in California. *IEEE Transactions on Geoscience and Remote Sensing, 39*, 1480-1490

Garten Jr., C.T., & Wullschleger, S.D. (1999). Soil Carbon Inventories under a Bioenergy Crop (Switchgrass): Measurement Limitations. *J. Environ. Qual., 28*, 1359-1365

Gomez, C., Viscarra Rossel, R.A., & McBratney, A.B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma, 146*, 403-411

Janik, L.J., Merry, R.H., & and Skjemstad, J.O. (1998). Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture, 38*, 681-696

Kooistra, L., Leuven, R.S.E.W., Wehrens, R., Nienhuis, P.H., & Buydens, L.M.C. (2003a). A comparison of methods to relate grass reflectance to soil metal contamination. *International Journal of Remote Sensing, 24*, 4995-5010

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R., & Buydens, L.M.C. (2003b). The potential of field spectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta, 484*, 198-200

Liang, S. (2004). Quantitative Remote Sensing of Land Surfaces. Hoboken, NJ: Wiley & Sons

Luo, H., Ye, H., Ke, Y., Pan, J., Gong, J., & Chen, X. (2005). Removing vegetation using unsupervised fully constrained least squares linear spectral mixture analysis method in soils surveying by remote sensing. In, *Proceedings of SPIE - The International Society for Optical Engineering,* 90-98

Malley, D.F., Yesmin, L., Wray, D., & Edwards, S. (1999). Application of near-infrared spectroscopy in analysis of soil mineral nutrients. *Communications in Soil Science and Plant Analysis, 30*, 999-1012

McBratney, A.B., Minasny, B., & Viscarra Rossel, R. (2006). Spectral soil analysis and inference systems: a powerful combination for solving the soil data crisis. *Geoderma, 136*, 272-278

McKenzie, N.J., Cresswell, H.P., Ryan, P.J., & Grundy, M. (2000). Contemporary land resource survey requires improvements in direct soil measurement. *Communications in Soil Science and Plant Analysis, 31*, 1553-1569

Mouazen, A.M., Maleki, M.R., De Baerdemaeker, J., & Ramon, H. (2007). On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil & Tillage Research, 93*, 13-27

Murphy, R.J., & Wadge, G. (1994). The effects of vegetation on the ability to map soils using imaging spectrometer data. *International Journal of Remote Sensing, 15*, 63-86

Palacios-Orueta, A., Pinzón, J.E., Ustin, S.L., & Roberts, D.A. (1999). Remote sensing of soils in the Santa Monica Mountains: II. Hierarchical foreground and background analysis. *Remote Sensing of Environment, 68*, 138-151

Post, W.M., Izaurralde, R.C., Mann, L.K., & Bliss, N. (2001). Monitoring and verifying changes of organic carbon in soil. *Climatic Change, 51*, 73-99

Reeves III, J.B., McCarty, G.W., & Meisinger, J.J. (2000). Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *J. Near Infrared Spectroscopy, 8*, 161-170

Schmidtlein, S. (2004). Coarse-scale substrate mapping using plant functional response types. *Erdkunde, 58*, 137-151

Siegal, B.S., & Goetz, A.F.H. (1977). Effect of vegetation on rock and soil type discrimination. *Photogrammetric Engineering and Remote Sensing, 43*, 191-196

Smith, P. (2004). Monitoring and verification of soil carbon changes under Article 3.4 of the Kyoto Protocol. *Soil Use and Management, 20*, 264-270

Sohn, Y., & McCoy, R.M. (1997). Mapping desert shrub rangeland using spectral unmixing and modeling spectral mixtures with TM data. *Photogrammetric Engineering and Remote Sensing, 63*, 707-716

Sørensen, L.K., & Dalsgaard, S. (2005). Determination of clay and other soil properties by Near Infrared Spectroscopy. *Soil Science Society America Journal, 69*, 159-167

Stevens, A., Wesemael, B.v., Bartholomeus, H., Rosillon, D., Tychon, B., & Ben-Dor, E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma, 144 (1-2)*, 395-404

Sudduth, K.A., & Hummel, J.W. (1993). Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Transactions of the ASAE, 36*, 1571-1582

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 and Soil, 251*, 319-329

Van-Camp, L., Bujarrabal, B., Gentile, A.R., Jones, R.J.A., Montanarella, L., Olazabal, C., & Selvaradjou, S.K. (2004). Reports of the Technical Working Groups Established under the Thematic Strategy for Soil protection. In O.f.O.P.o.t.E. Communities (Ed.), *EUR 21319 EN/6* (p. 872). Luxembourg

Veristech (2009). http://www.veristech.com/pdf_files/Optical_8thintlconf.pdf.

Viscarra Rossel, R.A., & McBratney, A.B. (1998). Soil chemical analytical accuracy and costs: Implications from precision agriculture. *Australian Journal of Experimental Agriculture, 38*, 765-775

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., & Skjemstad, J.O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma, 131*, 59-75

Warell, J. (2003). Properties of the Hermean regolith: III. Disk-resolved vis-NIR reflectance spectra and implications for the abundance of iron. *Icarus, 161*, 199-222

Wester, K., Lundén, B., & Bax, G. (1990). Analytically processed Landsat TM images for visual geological interpretation in the northern Scandinavian Caledonides. *ISPRS Journal of Photogrammetry and Remote Sensing, 45*, 442-460

Chapter 2

Scaling dimensions in spectroscopy of soil and vegetation

Weighted Total Contribution: 60%

Section 2.1: Concept: 20%, Writing 20%

Section 2.2: Review and writing: 100%

Section 2.3: Writing: 20%

2.1 Introduction

2.1.1 Definitions and classifications of scale

Quantitative research methods in environmental sciences are based on measurements of environmental variables. Each measurement of any quantitative parameter is expressed in the units of the measured variable. These units determine the level of the object observation, which is called scale. However, not only quantitative measurements, but also a qualitative description is related to a specific level of detail. For example, interactions of an elementary radiation element, photon, with the cellular structures inside a leaf (spatial dimension in micrometer) will be different from its behaviour inside a plant canopy between clumps of leaves, small twigs, branches, and trunks (spatial dimension in meter). This phenomenon of coexistence and interactions of differently sized objects makes scale an inseparable and basic component of any scientific domain. The essential effort to connect different scales in environmental sciences and the importance to define their common preferable spatial scales has been identified by Marceau (1999). The term scale, however, has a worldwide meaning with several disconnected definitions. Key concepts related to scale are used in different ways across disciplines and scholars, which makes the comparison and communication among researchers and research results across subfields and disciplines more difficult (Lam et al. 2004; Schneider 1994). After struggling with the confusion created by different uses of the same word, Gibson et al. (2000) presented definitions for the basic terms related to the concept of scales (Table 1.1). For a standardized lexicon, Quattrochi (1993) defined scale as ''the integral of space and time over which a measurement is made''. In the geo-statistical sciences the word support is defined as ''an n-dimensional volume within which linear average values of a regionalized variable may be computed'' (Olea 1991). The term scale is most commonly used in relation to the absolute or relative scale of space (Meentemeyer 1989) (Table 1.1). Cao and Lam (1997) introduced four scale concepts of the spatial-temporal domain: (1) cartographic scale, (2) scale of spatial extent, (3) scale of action, and (4) spatial resolution. First, the cartographic or map scale refers to the ratio of a distance on a map against the corresponding distance on the ground. A large-scale map covers a small area with a high detail, where a small-scale map covers a larger area with less detailed information. The geographical or observational scale, which refers to the size or spatial extent of the study, takes the opposite perspective. A geographic large-scale study covers a large area of interest as opposite to a geographic small-scale study covering a small area (Cohen et al. 2003). Third, the operational scale refers to the level at which observed processes operate in the environment (Turner et al. 2003). This scale, called also scale of action, represents a level at which a certain process phenomenon is best observed. Finally, the fourth meaning is the resolution of the measurement scale. Spatial resolution refers to the smallest distinguishable parts of an object, for instance a pixel size of the raster map or

remote sensing image (Dungan 2001). Scale is a frequent and general feature such that a "science of scale" was proposed (Goodchild and Quattrochi 1997). It has been concluded that such a science should address and answer the following interrelated issues: invariance of scale, the ability of scale change, measures of the impact of scale change, scale as a parameter of process models, and an implementation of multi-scale approaches.

Table 2.1 Definitions of the basic terms related to the concept of scale according to Gibson et al. (2000)

Term	Definition
Scale	The spatial, temporal, quantitative, or analytical dimensions used to measure and study any phenomenon
Extent	The size of the spatial, temporal, quantitative, or analytical dimensions of a scale
Resolution (grain)	The precision used in the measurement
Absolute scale	The distance, time, or quantity measured on an objectively calibrated measurement device
Relative scale	A transformation of an absolute scale to one that describes the functional relationship of one object or process to another (e.g. the relative distance between two locations based on the time required by an organism to move between them)

2.1.2 Imaging spectroscopy—science of increasing resolution

Spectroscopy is the scientific branch of physics concerned with the production, transmission, measurements, and interpretation of electromagnetic spectra (Swain and Davis 1978). Spectral properties can be measured using numerical spectroradiometers or imaging spectrometers in the laboratory or field, as well as from an airborne or satellite platform. Imaging spectroscopy is an imaging technique to record for every pixel in the image a separate electromagnetic spectrum (Van Der Meer and De Jong 2000). Imaging spectroscopy is an inseparable part of the passive optical remote sensing. It has many names including imaging spectrometry, hyperspectral remote sensing, or ultraspectral imaging (Kumar et al. 2001). Actually, imaging spectroscopy brought a new and expanded perception of the term scale in remote sensing mainly in the spectral domain. Imaging spectroscopy data, in contrast to multispectral data, contain a high number of narrow spectral bands. Moreover, they can be acquired from several oblique viewing angles and with a higher revisit frequency by means of modern satellite systems (e.g. spectrometer CHRIS on satellite PROBA).

2.1.3 Scale definitions and parameters in imaging spectroscopy

An imaging spectroscopy measurement of the Earth surface reflectance R is predominantly a function f_R defined by spatial, spectral, directional, and temporal scale (Baret, personal communication):

$$R = f_R(x, y; \lambda; \Omega_v; \Omega_s; t) \tag{2.1}$$

where x, y are spatial coordinates, λ the wavelength of the electromagnetic spectra, Ω_v, Ω_s describes the angular viewing geometry of the sun–object–sensor system, and t is the time frequency of the observation. A definition of remote sensing spectral image scale must consider all of these four dimensions. Adjusting the scale definition of Quattrochi (1993) we can state that the spectroscopy scale of optical imaging data is "the combination of space, electromagnetic wavelengths, their directions, and time intervals over which a spectrometric measurement is made". A more precise definition may be reached by extension of the term support (Olea 1991). The original specification of support includes the geometrical shape, size, and orientation of the volume. The volume of spectroscopy support should enclose in addition to the spatial and geometrical content, a spectro-directional component and the time intervals between successive observations.

Traditional parameters describing the scale of remote sensing image data are resolution (grain) and extent. Consequently, considering the four-dimensional spectroscopy scale scheme, spatial resolution is equal to the elementary pixel size of a remotely sensed image and spatial extent corresponds to the total area covered within an image swath. These spatial parameters are functions of the digital matrix of the spectral sensor and the 'instantaneous field of view' (IFOV) given by the optical system, flight altitude, and its flight velocity, respectively (Forshaw et al. 1983). Spectral resolution is described by Lillesand and Kiefer (1994) as "the ability to discriminate fine spectral differences". The spectral resolution of a sensor is often described by the 'full-width-half-maximum' (FWHM) of the instrument response to a monochromatic source (Liang 2004). The spectral extent, also named spectral range, is the difference between the minimum and maximum wavelengths in which measurements are made ($\lambda_{max} - \lambda_{min}$). A new parameter, spectral sampling, has to be introduced to describe the number and position of the spectral channels. Spectral sampling interval is the spacing between sample points in the spectrum (Liang 2004). As shown in Fig. 2.1, the sampling interval is independent of the spectral resolution, which implies that there can be overlap between consecutive bands. This is usually the case in imaging spectroscopy instruments, since their aim is to derive a contiguous spectrum where oversampling reduces the amount of incoming noise but at the cost of information redundancy. Spectral and spatial resolutions of the multi- and/or hyperspectral images are in an inverse relationship due to the technical constrains on the sensor side. There is usually a trade off between high spectral and low spatial resolution or vice versa, because of the limited extent and minimal element size of a 'charged-coupled device' (CCD) array recording the spectral image. Lower spatial resolution caused by the binning of the spatial array columns allows a narrowing of the FWHM and, subsequently, increasing spectral resolution. Conversely, spectral binning of the wavelengths widens the FWHM and gives opportunity to increase the spatial resolution.
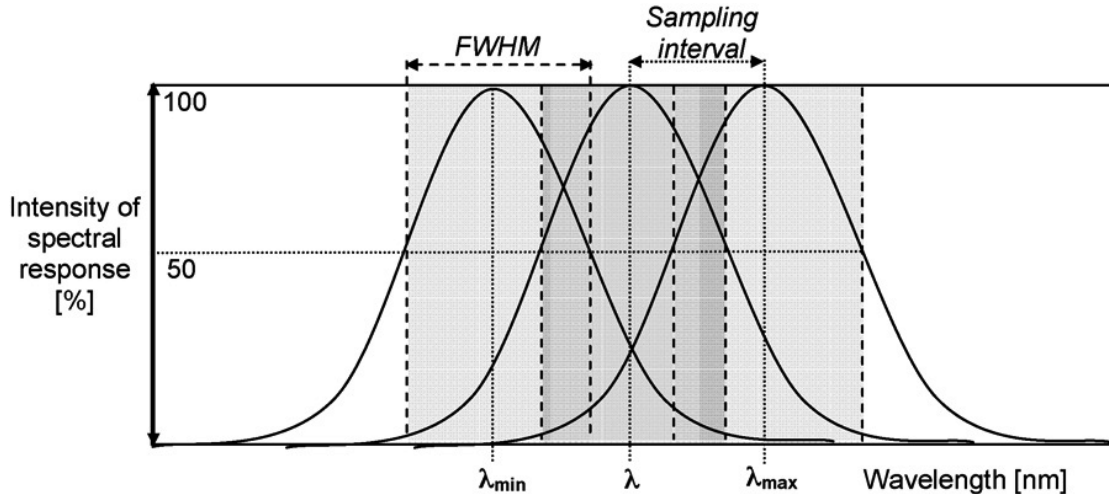
Fig. 2.1. Spectral resolution ('full-width-half-maximum'—FWHM), extent ($\lambda_{max} - \lambda_{min}$) and sampling interval of imaging spectroscopy data.

More traditional multispectral satellite instruments operate on high spatial resolution and lower spectral resolution with a small temporal sampling interval (e.g. LANDSAT 7 ETM+ or SPOT 5 HRG). New imaging spectroscopy satellite sensors acquire data with coarser spatial resolution and higher spectral sampling interval and resolution (e.g. medium resolution imaging spectrometer (MERIS) on the ENVISAT satellite), as well as at both high spatial and spectral resolutions (e.g. the Hyperion sensor on board of the satellite EO-1). Reflectance by the Earth's surface and scattering by atmospheric particles and gases have a strong directional behaviour. This phenomenon is scientifically described by the concept of the bi-directional reflectance distribution function (BRDF). The BRDF is a conceptual quantity that describes the reflectance of a target as a function of the independent variables describing viewing and illumination angles and variables determining the geometrical and optical properties of the observed target (Deering 1989; Liang 2004; Myneni and Ross 1991; Nicodemus et al. 1977). The BRDF describes the scattering of a parallel beam of incident light from one direction in the hemisphere into another direction in the hemisphere (Schaepman-Strub et al. 2004). The incident and viewing directions are each defined by the zenith and azimuth angles of illumination (in nature sun zenith θ_s and azimuth ϕ_s angle) and view of sensor (viewing zenith θ_v and azimuth ϕ_v angle). Difference of the viewing and illumination azimuth angle is called the relative azimuth angle ($\phi = \phi_v - \phi_s$) (Schönermark et al. 2004). Then the BRDF [sr^{-1}] can be expressed as a reflectance function f_{BRDF} of source illumination projected solid angle Ω_s, viewing projected solid angle Ω_v, and wavelength λ:

$$BRDF = f_{BRDF}(\Omega_s, \Omega_v, \lambda) = \frac{dL_v(\theta_s, \phi_s; \theta_v, \phi_v; \lambda)}{dE_s(\theta_s, \phi_s; \lambda)} \qquad (2.2)$$

where L_v is reflected radiance and E_s is incident solar irradiance. The unitless bi-directional reflectance factor (BRF) is proportional to the BRDF according to the relation: BRF = $\pi \times$ BRDF. Finally, the reflectance acquired under illumination of the ambient hemispherical sky is called the hemispherical-directional reflectance factor (HDRF). The HDRF is physically defined in the same way as the BRF, except that the HDRF includes illumination coming from the entire hemisphere (Schaepman-Strub et al. 2004). Any outdoor HDRF measurement depends not only on the scattering optical properties of the observed object, but also on atmospheric conditions, the surrounding of the object, the topography, and wavelength. The spectral field measurements of the HDRF are often performed by a goniometer device. In order to obtain a high accuracy, goniometer reflectance measurements are usually performed at a local scale and on a specific vegetation or soil (snow) surface. Still, natural spatial patterns are not uniformly distributed within the space and their expanse covers the whole range from micro to macro scale (e.g. cell structure, leaves, branches, trees, forest). Details of ground directional reflectance measurements are given in Bruegge et al. (2004).

Accordingly to the BRDF concept, the directional resolution of the angular spectral image data is represented by the IFOV given by the parameters of the optical set, size of the CCD array basic element, tilt, motion speed, and altitude of a sensor. Directional extent is specified by the interval between minimal viewing direction $(\theta_v \min, \phi_v \min)$ and maximal viewing direction $(\theta_v \max, \phi_v \max)$ aside (maximal–minimal oblique viewing angles). Finally, directional sampling is expressed by the total number of viewing directions and their angular position within the hemispherical space. Presently, only a few real multi-angular imaging spectroscopy satellite sensors are operational.

Examples of successful missions are the Compact High Resolution Imaging Spectrometer (CHRIS) sensor on board the PROBA satellite, providing five angular images in 63 spectral bands (NADIR, ±36 º, ±55 º), or the Multiangle Imaging SpectroRadiometer (MISR) on the NASA EOS Terra platform consisting of nine cameras capturing four VIS/NIR spectral bands in nine backwards and forwards along track viewing directions. Developments in remote sensing technology and radiative transfer modelling indicate that angular signatures can be exploited to provide not only improved accuracies relative to single-angle approaches but also unique diagnostic information about the Earth's atmosphere and surface, e.g. identification of atmospheric aerosol, cloud, or surface vegetation type (Diner et al. 1999), capitalizing on both the geometric aspects of the technique as well as the radiometric variations in signal with angle.

Temporal scale must be considered, since geochemical or geophysical constituents of the surface (e.g. concentration of various chemical compounds, or water content) exhibit specific spectral features that vary over time. Monitoring the variation over time becomes more

important when ecosystems and their reactions to climate effects are observed. Some authors traditionally refer to temporal resolution as the image frequency which depends on the revisit time of a sensor; in other words, how often is an image acquired over a specific location on Earth (Franklin 2001). However, strictly following the foregoing concept of the imaging spectroscopy spatial and spectro-directional scale, we propose the sensor revisit time to be called the temporal sampling interval rather than temporal resolution. Since image spatial resolution is given by size of the smallest CCD array element then similarly the temporal resolution should be defined as the shortest time span needed to integrate the reflected radiative information by the CCD array into the image. Nevertheless, this parameter is commonly called an integration and/or dwell time of a sensor. Temporal extent is taken as the time interval between the last and first observation of the same location ($t\max$–$t\min$), which can be several years for a given satellite platform. Perhaps the most important temporal characteristic, revisit frequency of the satellite driven by the orbit parameters and viewing extent varies from mission to mission. Among low revisiting frequency satellite platforms are LANDSAT 7, and EO-1 (both 16 days) or SPOT (26 days). Examples of the high frequency revisiting sensors are the moderate resolution imaging spectroradiometer (MODIS) aboard the Terra (EOS AM) and Aqua (EOS PM) satellites viewing the entire Earth's surface every one to two days or the Medium Resolution Imaging Spectrometer (MERIS) with the revisit frequency of three days. Note that the theoretical temporal sampling interval is usually higher than the practical one, due to cloudiness that can cover the location of interest during the time of the sensor overpass.

The increasing availability of remote sensing sensors provides the possibility of choosing the systems that are best adapted for specific research interests. Various factors such as cost, availability at a certain time and place, sensor characteristics (spatial, spectral, temporal, directional resolution) and, of course, specific research interests determine the final decision. Sensor characteristics and research interests are strongly related and this is where scale considerations play a major role. The choice of the appropriate scale for every dimension in a particular application depends on several factors and is a function of the type of environment and the kind of information desired (Woodcock and Strahler 1987).

2.1.4 Scaling up and down-scaling techniques for soil and vegetation surfaces

Transfer of data content from one scale to another one is called scaling. According to Dungan (2001), scaling when applied in remote sensing and GIS is a procedure that changes the size of a measurement unit. Basically, scaling can be performed by means of two approaches: bottom-up and top-down. The bottom-up approach up-scales information from smaller to larger observational scales, while the top-down approach down-scales, in other word decomposes, information at a certain geographical scale into its constituents at smaller scales (Marceau and Hay 1999). The capability to process and present geographic information "up" and "down" from local, regional, to global scales has been advocated as a

solution to understanding the global systems of both natural (e.g. global climate change) and societal (e.g. global economy) processes and the relationships between the two (Lam et al. 2004). But, as mentioned by Jarvis (1995), scaling represents a scientific challenge because of the non-linear nature between processes and variables, and heterogeneity of characteristics determining the rates of processes. Much literature has been published on scaling in environmental research, with the vast majority concerning spatial scaling. Since the principles of spatial scaling in imaging spectroscopy do not differ from scaling in other research fields, this topic was not discussed deeply in the original paper by Malenovsky et al., (2007). Several books have been written on this subject, so for further reading we direct the reader to the following references: Tate and Atkinson (2001), Cao and Lam (1997), and Goodchild and Quattrochi (1997).

Several techniques to perform up- or down-scaling in spectroscopy are available. Radiative transfer modeling, spectral unmixing, and data-fusion are widely investigated and accepted in the research field of spectroscopy. Spectral (un)mixing as a technique to perform up- or down-scaling in spectroscopy is described in detail. The fact that technique is widely investigated and well accepted does not mean that this technique is easy to apply or results in standard products. Expert knowledge of the technique and the physical processes specific for the study-area are needed for valuable use. Therefore, the described limitations of the technique should be taken seriously; nonetheless spectral unmixing may be useful for those who deal with spatial scaling problems in spectroscopy. The technique is illustrated with a case study in the field of environmental research, specifically vegetation and soil studies. Other scaling techniques are described in the full original paper by Malenovsky et al. (2007)

2.2 Spectral (un)mixing

2.2.1 Spectral mixing

Each pixel in a remote sensing image is treated as a homogenous area with a single reflectance value. However, homogeneity is a rare phenomenon at the Earth surface, so this one value is a combined reflectance of all objects present within the pixel. The spatial distribution of the objects within this pixel is lost, but much of the spectral information is preserved in the spectral signature. When more materials or objects are present within a pixel, the measured reflectance is a result of the fractions in which the materials occur and the spectral characteristics of each material. For each spectral band the reflectance value of a linear mixture is determined by the following equation:

$$R = R_1 f_1 + R_2 f_2 + ... + R_n f_n \tag{2.3}$$

where R is the reflectance value of the pixel, R_n the reflectance of object n in the specific wavelength, and f_n is the fraction in which object n occurs. The measured spectrum of the pixel is a weighted average of the abundances in the pixel. Hapke (1981) developed the theory of mixture modeling, to estimate the spectral response pattern of a mixed object with known pure components (endmembers).

The spectral signature of a mixed pixel contains spectral information of all endmembers present in the represented area. Spectral mixing can be considered as a linear process, when the mixing scale is macroscopic such that photons interact with one material instead of with several materials (Singer and McCord 1979). This is in general the case when the surface consists of snow (Painter et al. 2003) or bare soil but not always true when spectral mixing of vegetation endmembers occurs. Only a part of the radiation is directly reflected (dependent on the wavelength), the rest is absorbed or transmitted to other layers, from which it can be reflected to the sensor again. As a result, the reflected radiation interacts with more than one object, resulting in a non-linear spectral mixing (Huete 1988; Ray and Murray 1996; Roberts et al. 1993). Mixing of spectral signals can be used for spatial upscaling of any kind of spectral information. The object optical and structural properties determine whether the mixing process is linear or non-linear.

2.2.2 Spectral unmixing

Smith et al. (1985) reversed the mixing process in order to estimate the mixing components from the mixed spectral signal and the spectral response of the endmembers: Spectral unmixing (also named Spectral Mixture Analysis or SMA). SMA is used to find the fractions (abundances) of a number of endmembers that best explain the recorded mixed pixel reflectance spectrum. SMA is a physically based model that provides quantitative estimates of the distribution of materials within the image scene (Tompkins et al. 1997). The mixing and unmixing process is illustrated in Fig. 2.2. Spectral unmixing offers information about image components with a spatial extent smaller than the ground instantaneous field of view, and therefore can be considered as a spatial down-scaling procedure.

The linear unmixing model assumes that a surface within a pixel is made up of a limited number of endmembers. The mathematical notation of the linear unmixing procedure is as follows (Adams et al. 1986):

$$R_b = \sum_{i=1}^{n} f_i R_{i,b} + \varepsilon_b \qquad \text{under constraint} \qquad \sum_{i=1}^{n} f_i = 1 \qquad (2.4)$$

Where R_b is the reflectance of the pixel being examined in band b, f_i the fraction of endmember i, $R_{i,b}$ the reflectance value of endmember i in band b, ε_b the residual error associated with band b, n is the number of spectral endmembers. According to these formula

the value of a pixel must equal the sum of the values of the endmembers times their abundance. The fit of the model can be assessed by the root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum_{b=1}^{M} \varepsilon_b^2}{M}} \tag{2.5}$$

where M is the number of bands used in the spectral unmixing. Higher values of RMSE indicate regions that could contain lacking endmembers. These equations show that the maximum amount of endmembers can never be more than the number of bands minus one. Having more endmembers than spectral bands results in more unknowns than equations, which makes it impossible to determine the fractions (Settle and Drake 1993). Therefore, when applying linear spectral unmixing the maximum amount of endmembers is always one less than the number of spectral bands available, making imaging spectroscopy of great use. Spectral unmixing is particularly useful with imaging spectroscopy data, where the number of useful bands is much higher than the number of model endmembers, and solutions to the basic SMA equations are overdetermined (Okin et al. 2001). The inversion of the spectral mixing process can be done in different ways. The most frequently used method is the least squares fitting approach (Shimabukuro and Smith 1991; Strang 1988). Through the deconvolution, spectral unmixing seeks the abundances of the endmembers that (statistically) best explain the observed spectral characteristics of a pixel. This analysis assumes that endmembers are orthogonal and thus uncorrelated (Van Der Meer and De Jong 2000). The matrix inversion introduces a problem: when endmembers are not strictly orthogonal, the matrix cannot be inverted without error, thus giving rise to doubtful results. This problem occurs when endmembers are linearly scaled versions of each other (degenerate), since these can not be separated by the least-squares approach. This is especially the case when applying spectral unmixing techniques to low spectral resolution image data, where most spectra are largely similar. The problem can be avoided by working on orthogonally transformed data, such as minimum noise fractions (MNF) or maximum autocorrelation factors (MAF) (Nielsen 2001; Van Der Meer and De Jong 2000). Other techniques used to perform the unmixing include a regularization method (Settle and Drake 1993), minimum variance methods (Manolakis et al. 2000), singular value decomposition (Tompkins et al. 1997), and variable endmember methods (Ramsey and Christensen 1998; Roberts et al. 1998). Non-linear mixing is likely to lead to an overestimation of green vegetation during the unmixing process and further complicates the analysis if the spectra used in a spectral library are collected under conditions where the substrate in the image differs from what was present when the library spectra were collected (Okin et al. 2001). The used constraints are only meaningful when a real world situation is described. From statistical viewpoint it is not necessary to use these constraints and values outside the bounds have some physical meaning (Adams et al. 1993; Sabol Jr et al. 1992). Fractions greater than one and deviations in the 'negative' direction from other endmembers can be
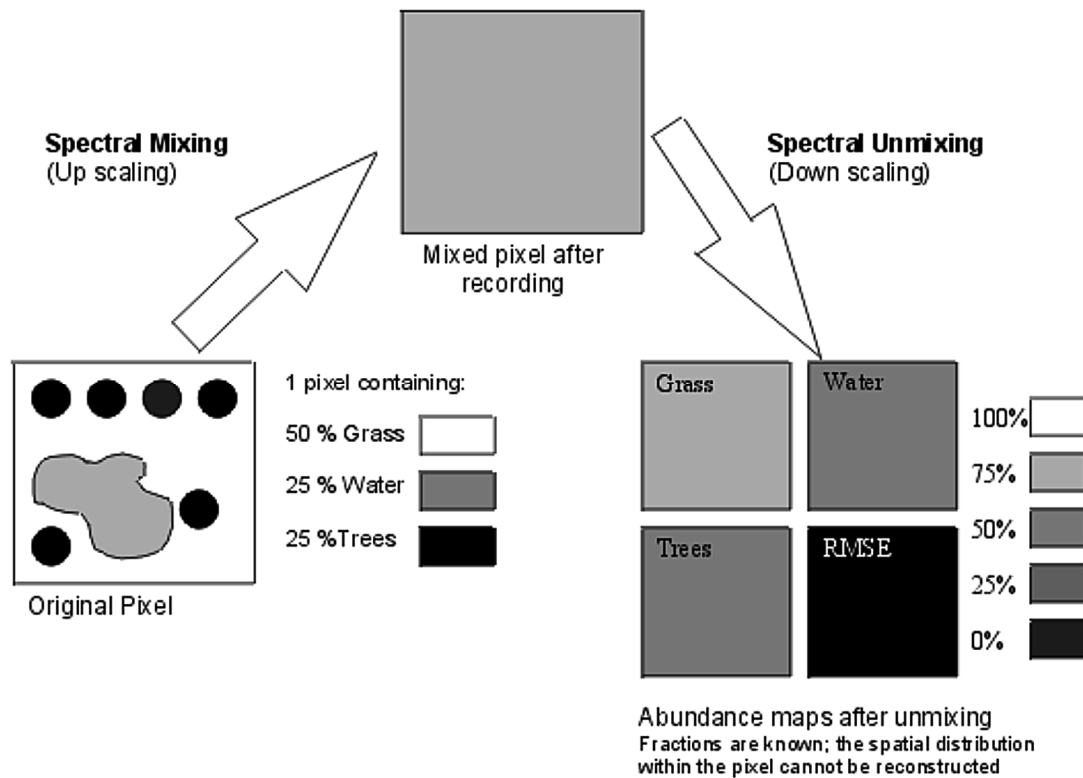
Fig. 2.2. The results of the spectral mixing and unmixing process. A heterogeneous surface element recorded in a single pixel results in a mixed pixel. This can be unmixed into abundance images, which show the fractional cover of the surface elements.

expected if the library reflectance values are derived from an 'average' material response. Negative fractions can occur due to poor endmember selection, an improper number of endmembers or because of spectral ambiguity (Gross and Schott 1998). Negative abundance factors can be, among other factors, the result of correlation of the endmembers. Van Der Meer and De Jong (2000) showed that decorrelation of the endmembers improves the abundance estimates. Marden and Manolakis (2002) observed that enforcing the constraints led to performance degradation in most cases they investigated.

2.2.3. Endmembers selection

An essential step in the unmixing process is the definition of the appropriate set of endmembers, which should be chosen to represent surface components (Tompkins et al. 1997). In practice, the number of endmembers and their composition is unknown and a selection is made of a number of endmembers supported by the dimensionality of the data. The estimation of these endmembers is not straightforward since they vary upon the scale and purpose of the study (Milton and Emery 1995). Endmember selection can be done in three ways: (1) derived from a spectral library (Boardman 1990), (2) derived from field measurements, or (3) extracted from the image.

2.2.3.1. Spectral libraries

Spectral libraries are collections of laboratory spectra of a wide variety of materials. Users should be aware that the spectra in these libraries can be collected for different purposes and are measured using various types of instruments (portable non-imaging spectrometers) of often unknown quality and precision. Spectral libraries can also contain directional spectra that have been modelled with radiative transfer approaches for surfaces such as snow or soil (Painter et al. 2003). Wavelength shifts, unreliable instrument calibration and significant degrees of noise are not uncommon to these data. Furthermore, spectral measurements in the laboratory can vary in terms of directionality. Directional-hemispherical measurements (e.g. with an integrating sphere) should not be directly mixed up with bi-directional measurements. The quality of spectra taken from spectral libraries should always be checked using known absorption features of, e.g. minerals or oxygen (760 nm). Since laboratory spectra are collected without atmospheric influence they cannot be directly combined with TOA airborne or spaceborne imagery. Hence, spectral libraries can be very useful for the interpretation of spectroscopic imagery but should be applied with care.

2.2.3.2. Field measurements

A second method is the use of field spectra. Using a field spectrometer the spectra of each endmember may be approximated. When applying this method good correction for the atmospheric influence is essential because this influence will be present within the images and not in the field spectra.

2.2.3.3. Image derived endmembers

The third possibility is to derive the endmembers directly from the purest pixels in the image. The advantage is that the selected endmembers were collected under similar atmospheric conditions and spectral/radiometric biases. An image-derived endmember can either be a single pixel or the average of several pixels in a homogeneous area. When there is no field knowledge and the spatial locations of pure materials are not known a variety of techniques can be applied to identify endmembers directly from the imagery. The first step of automated procedures often consists of a compression of the data with a minimum noise fraction (MNF) rotation (Green et al. 1988) or principal component analysis (PCA) (Smith et al. 1985). A common approach for determining the number of endmembers (e.g. Smith et al. 1985) is to view the spectral data as points in the space of the complete set of spectral bands and to find the PCA eigenvectors or directions accounting for most of the variance in the data. The remaining variance is equal to the instrumental error. Since sampling a continuous reflectance spectrum in many narrow, contiguous spectral bands results in a high covariance between the bands, the number of useful eigenvectors will be less than the number of bands in the imaging spectroscopy datasets (Bateson and Curtiss 1996). Another algorithm for finding these pure pixels is the Pixel Purity Index (PPI) (Boardman et al. 1995), which yields the best results when applied on compressed dataset (PCA-transformed or

MNF). Numerous other techniques for automated image endmember selection are published. Some examples are the vertex component analysis (VCA), described by Nascimento et al. (2004) and endmember average RMSE (EAR), published by Dennison and Roberts (2003).

2.2.4. Advanced unmixing techniques

In addition to the above-described linear spectral unmixing approach a wide variety of advanced unmixing techniques are published. Iterative spectral unmixing (ISU) (Van Der Meer 1999) uses the RMS error image to automatically select additional endmembers and reposition of existing ones. The Multiple Endmember Spectral Mixture Analysis (MESMA) accounts for spatial heterogeneity in a geographic approach that allows the number of endmembers and the endmembers themselves to vary on a pixel-by-pixel basis (Roberts et al. 1998). Unmixing based on stepwise regression (Gross and Schott 1996) solves the same problem by adaptively selecting a spatially varying set of endmembers and solves the fractions of the selected endmembers within each pixel. Zhang et al (2002) introduced a Derivative Spectral Unmixing (DSU) model, which makes it possible to estimate the fraction of an endmember characterized by one or more diagnostic absorption features, despite having only general knowledge of the spectral shapes of the remaining endmembers. Changing the constraints to describe the situation in a specific study area can help to improve unmixing results (Vikhamar and Solberg 2003). Ultimately, these algorithms share unmixing at the core but also a significant increase in computational expense.

2.2.5. Unmixing example

Linear unmixing of a reflective optics system imaging spectrometer (ROSIS) image of an agricultural area just south of Álora (Spain) was studied in this case study. The research goal was the quantification of the soil iron content in olive fields. The objective of the spectral unmixing was to estimate the fractional vegetation cover, since the accuracy of the quantification of the soil iron content is highly influenced by vegetation. The iron bearing soils were mainly formed on Tortonian Molasse deposits. The area is covered with olive trees and bare soils, with shrub cover in some locations. Senescent grass is present in many places but generally in low abundance. The first step in the analysis was the MNF transformation (Fig. 2.3). The importance of using uncorrelated data for unmixing was emphasized before. The PPI algorithm was run on the MNF transformed images, resulting in the image shown in Fig. 2.4, where the pixels with a high PPI-score are highlighted. From this image the locations of endmembers were identified and selected using field investigations.

As a last step, a non-constrained linear spectral unmixing was performed resulting in the fractional images displayed in Fig. 2.5. The result of the spectral mixture analysis is a series of abundance images and an image depicting the RMSE. The fractional images are scaled between 0 and 1, where a value of 1 implies 100% abundance of the specific endmember. ROSIS measures in 115 spectral bands, of which the first 15 bands were not used due to

striping effects. The last image of Fig. 2.5 shows the RMSE of the unmixing process. Some features can still be recognized, indicating that these features were not selected as endmembers. Man-made structures like the road and a house can be recognized in the RMSE image.

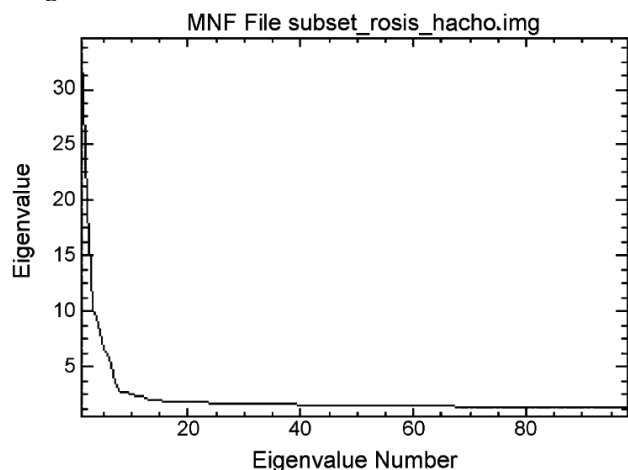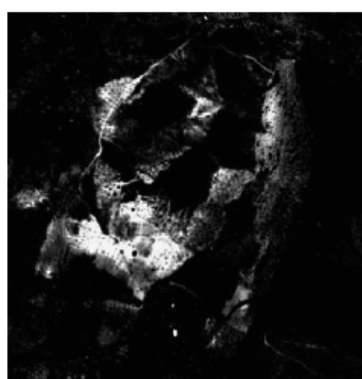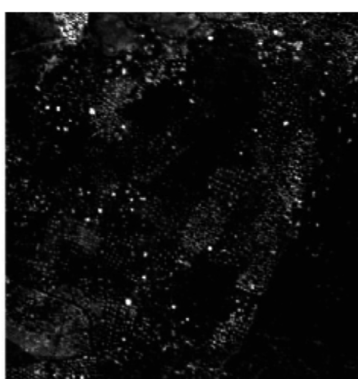
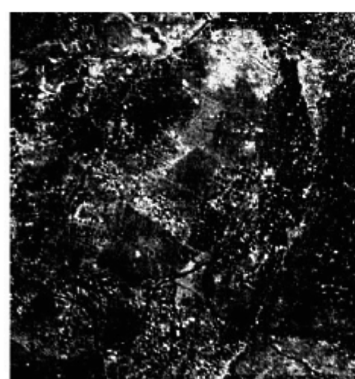Fig. 2.3. Eigenvalues of the ROSIS bands after MNF transformation.

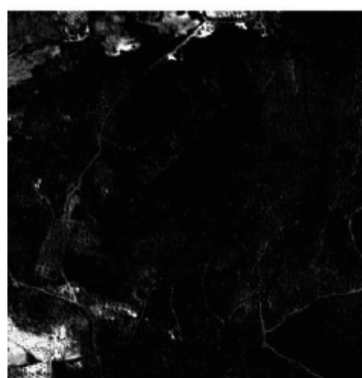Fig. 2.4. PPI image of the MNF transformed ROSIS image.
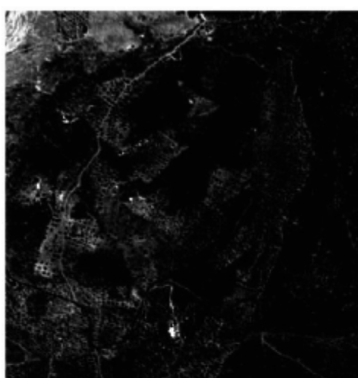

EM 1: Iron bearing soil

EM 2: Olive trees

EM 3: Shrubs

EM 4: Soil type 2

EM 5: Senescent grass

RMSE

Fig. 2.5. Fractional cover maps, showing the abundances of the five endmembers (white colour indicates a high abundance; black colour indicates a low abundance).

2.2.6. Other applications

Spectral unmixing of imaging spectroscopy data is done in all fields of environmental research. Painter et al. (2003) applied a multiple endmember SMA to AVIRIS data acquired over snow-covered areas in the Sierra Nevada to estimate the fractional snow covered area and the grain size of the fractional snow cover. Metternicht and Zinck (2003) described the use of spectral unmixing to determine soil salinity using the DAIS-7915 sensor. The grain size of sediments in an intertidal estuarine zone, using Daedalus 1268 Airborne Thematic Mapper data, was investigated by Rainey et al. (2003). Lelong et al. (1998) determined the fractions of two wheat endmembers which were interpreted in terms of crop vitality relative to stress presence. Vikhamar and Solberg (2003) based their SnowFrac-algorithm on spectral unmixing and endmember constraints to estimate the snow-cover fraction of a pixel. Unmixing techniques can also be applied on a multispectral dataset, but the limited amount of spectral bands limits the amount of endmembers that can be used in the analysis. Furthermore, the higher correlation between endmembers, due to broad wavelength ranges, should decrease the accuracy of multi-spectral unmixing techniques. To improve the amount of input data multi-temporal approaches are used. Multi-spectral recordings of several moments in time are combined and treated as a single image, where the bands represent the reflectance at a specific moment in, e.g. the phenological cycle (Lobell et al. 2002; Shoshany and Svoray 2002). Imaging spectroscopy offers valuable input data for spectral down-scaling using spectral unmixing. The large number of spectral bands used in spectroscopy offers the opportunity to increase the number of endmembers in the analysis. Furthermore, spectroscopy reveals reflectance differences in narrow wavelength ranges.

2.3 General discussion and conclusions

Traditional understanding of the term scale in remote sensing is defined by two parameters: grid resolution (grain), and spatial extent. This review paper shows that the scale of the imaging spectroscopy data is not a simple spatial function but it must be defined as a complex four dimensional function of the space, wavelengths of electromagnetic spectra, angular geometrical vectors, and time. Consequently, this new idea, proposed for imaging spectroscopy data by Baret (personal communication), is developed into the concept of the spatial, spectro-directional and temporal scaling dimensions of imaging spectroscopy. Each of the dimensions is described by means of the resolution and extent parameter as summarized in Table 2.2. As a third parameter the sampling interval was defined to express the entire complexity of the spectro-directional and temporal scale. Changes in magnitudes and subsequently units of these parameters are associated with a process of scaling. The statistic-empirical or physical method decreasing spatial, spectro-directional, and temporal spectroscopic resolution is termed upscaling, while the opposite procedure increasing this parameter is called down-scaling.

Table 2.2 Parameters of the four dimensional concept of scale for spectroscopic data

Scale dimension	Resolution	Extent	Sampling interval
Spatial	Image Pixel-size	$x, y_{max} - x, y_{min}$	-
Spectral	FWHM	$\lambda_{max} - \lambda_{min}$	Spacing between band centres
Directional	IFOV[a]	$\Omega_{max} - \Omega_{min}$	Spacing between hemispherical directions
Temporal	Integrating time	$t_{max} - t_{min}$	Spacing between sensor visits

[a] IFOV is applicable for imaging spectrometers. FOV (field-of-view) of the optical set applies in case of numeric spectroradiometers.

The literature review carried out in the context of this study supports the conclusion that spatial scaling is the most often investigated procedure within the scope of remote sensing to present, probably due to long term accessibility of coarse satellite and detailed airborne images. Availability of hyperspectral images and frequently revisiting satellites induced several studies considering the spectral and temporal scaling. Yet, only a few studies have been conducted on directional scaling thus far. Use of the directional spectral information is currently growing in remote sensing research, so one may expect rapid progress in directional methods supporting development of a new multi-angle optical sensor.

One of the major concerns of the remote sensing community is how to scale and generalize information collected on the local level up through regional to the global level. Remote sensing data have a significant potential to generate inputs for global as well as regional and local eco-physiological models of natural ecosystem processes (e.g. models of carbon fluxes) (Chen et al. 2003; Wang et al. 2004). Yet, methodological questions remain that require attention and careful treatment. How should we properly combine several remote sensing sources at different scales insuring accuracy and further applicability of the results? Consequently, how should we appropriately validate accuracy and uncertainties of the coarse resolution satellite products against the ground-based measurements (Chen et al. 2002; Tian et al. 2002)? The answers and potential solutions are available in the multi-source approach proposed in Malenovsky et al. (2007). This concept combines remote sensing and ground data of various sources and resolutions (spaceborne and airborne images, field measurements) by means of several scaling techniques (statistical and empirical methods, physical models, etc.) including specific data corrections and calibration mechanisms, as well as successive validation feedbacks. This way the multi-source approach improves the reliability of the remote sensing products, which subsequently becomes suitable input for ecosystem process models.

References

Adams, J.B., Smith, M.O., & Gillespie, A.R. (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, 145-166

Adams, J.B., Smith, M.O., & Johnson, P.E. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research, 91*, 8098-8112

Bateson, A., & Curtiss, B. (1996). A method for manual endmember selection and spectral unmixing. *Remote Sensing of Environment, 55*, 229-243

Boardman, J.W. (1990). Inversion of high spectral resolution data. Proceedings of SPIE - The International Society for Optical Engineering, 1298, 222-233

Boardman, J.W., Kruse, F.A., & Green, R.O. (1995). Mapping target signatures via partial unmixing of AVIRIS data. *Summaries, Fifth JPL Airborne Earth Science Workshop, 1*, 23-26

Bruegge, C.J., Abdou, W.A., Painter, T.H., Paden, B.E., Dozier, J., Schaepman, M., Strub, G., Beisl, U., Itten, K.I., Demircan, A., Geiger, B., Helmlinger, M.C., & Martonchik, J. (2004). Field measurements of bi-directional reflectance. *Reflection Properties of Vegetation and Soil With a BRDF Database*, 195-224

Cao, C., & Lam, N.S. (1997). Understanding the scale and resolution effects in remote sensing and GIS. *Scale in Remote Sensing and GIS*, 57-72

Chen, J.M., Liu, J., Leblanc, S.G., Lacaze, R., & Roujean, J.L. (2003). Multi-angular optical remote sensing for assessing vegetation structure and carbon absorption. *Remote Sensing of Environment, 84*, 516-525

Chen, J.M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S.G., White, H.P., Hall, R.J., Peddle, D.R., King, D.J., Trofymow, J.A., Swift, E., Van Der Sanden, J., & Pellikka, P.K.E. (2002). Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment, 80*, 165-184

Cohen, W.B., Maiersperger, T.K., Yang, Z., Gower, S.T., Turner, D.P., Ritts, W.D., Berterretche, M., & Running, S.W. (2003). Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: A quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment, 88*, 233-255

Deering, D.W. (1989). Field measurements of bidirectional reflectance. *Theory and Applications of Optical Remote Sensing*, 14-65

Dennison, P.E., & Roberts, D.A. (2003). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment, 87*, 123-135

Diner, D.J., Asner, G.P., Davies, R., Knyazikhin, Y., Muller, J.P., Nolin, A.W., Pinty, B., Schaaf, C.B., & Stroeve, J. (1999). New Directions in Earth Observing: Scientific Applications of Multiangle Remote Sensing. *Bulletin of the American Meteorological Society, 80*, 2209-2228

Dungan, J. (2001). Scaling up and scaling down: The relevance of the support effect on remote sensing of vegetation. *Modelling Scale in Geographical Information Science*, 221-235

Forshaw, M.R.B., Haskell, A., Miller, P.F., Stanley, D.J., & Townshend, J.R.G. (1983). Spatial resolution of remotely sensed imagery. A review paper. *International Journal of Remote Sensing, 4*, 497-520

Franklin, S.E. (2001). Remote Sensing for Sustainable Forest Management. Lewis Publishers, Boca Raton.

Gibson, C.C., Ostrom, E., & Ahn, T.K. (2000). The concept of scale and the human dimensions of global change: A survey. *Ecological Economics, 32*, 217-239

Goodchild, M.F., & Quattrochi, D.A. (1997). Scale, multiscaling, remote sensing, and GIS. *Scale in Remote Sensing and GIS*, 1-11

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). Transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing, 26*, 65-74

Gross, H.N., & Schott, J.R. (1996). Application of spatial resolution enhancement and spectral mixture analysis to hyperspectral images. *Proceedings of SPIE: Hyperspectral Remote Sensing & Applications, 2821*

Gross, H.N., & Schott, J.R. (1998). Application of spectral mixture analysis and image fusion techniques for image sharpening. *Remote Sensing of Environment, 63*, 85-94

Hapke, B. (1981). Bidirectional reflectance spectroscopy. 1. Theory. *J. Geophys. Res., 86*, 3039-3054

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Rem. Sens. Environ., 25*, 89-105

Jarvis, P.G. (1995). Scaling up. *Plant Cell Environ*, 1079-1089

Kumar, L., Schmidt, K.S., Dury, S., & Skidmore, A.K. (2001). Imaging spectrometry and vegetation science. *Imaging Spectrometry*, 111-155

Lam, N.S.N., Catts, D., McMaster, R., Quattrochi, D., & Brown, D. (2004). Scale. *A Research Agenda for Geographic Information Science*, 93-128

Lelong, C.C.D., Pinet, P.C., & Poilvé, H. (1998). Hyperspectral imaging and stress mapping in agriculture: A case study on wheat in Beauce (France). *Remote Sensing of Environment, 66*, 179-191

Liang, S. (2004). Quantitative Remote Sensing of Land Surfaces, John Wiley & Sons, Hoboken, New Jersey.

Lillesand, T.M., & Kiefer, R.W. 1994. Remote Sensing and Image Interpretation. John Wiley & Sons, Inc., New York.

Lobell, D.B., Asner, G.P., Law, B.E., & Treuhaft, R.N. (2002). View angle effects on canopy reflectance and spectral mixture analysis of coniferous forests using AVIRIS. *International Journal of Remote Sensing, 23*, 2247-2262

Malenovsky, Z., Bartholomeus, H.M., Acerbi-Junior, F.W., Schopfer, J.T., Painter, T.H., Epema, G.F., & Bregt, A.K. (2007). Scaling dimensions in spectroscopy of soil and vegetation. *International Journal of Applied Earth Observation and Geoinformation, 9*, 137-164

Manolakis, D.G., Ingle, V.K., & Kogon, S.M. (2000). *Statistical and Adaptive Signal Processing*

Marceau, D.J. (1999). The scale issue in social and natural sciences. *Canadian Journal of Remote Sensing, 25*, 347-356

Marceau, D.J., & Hay, G.J. (1999). Remote Sensing Contributions to the Scale Issue. *Canadian Journal of Remote Sensing, 25*, 357-366

Marden, D.B., & Manolakis, D. (2002). Constrained detectors for hyperspectral remote sensing applications: Theory versus practice. *Proceedings of SPIE - The International Society for Optical Engineering, 4816*, 211-221

Meentemeyer, V. (1989). Geographical perspectives of space, time, and scale. *Landscape Ecology, 3*, 163-173

Metternicht, G.I., & Zinck, J.A. (2003). Remote sensing of soil salinity: Potentials and constraints. *Remote Sensing of Environment, 85*, 1-20

Milton, E.J., & Emery, D.R. (1995). The identification of reference endmembers using high spatial resolution multispectral images. *Proc. RSS95 Remote Sensing in Action, Remote Sensing Society, 95*, 579-586

Myneni, R.B., & Ross, J. (1991). Photon-Vegetation Interactions: Applications in Optical Remote Sensing and Plant Ecology. Springer-Verlag, Berlin.

Nascimento, J.M.P., & Dias, J.M.B. (2004). Fast Unsupervised Extraction of Endmembers Spectra from Hyperspectral Data. *Proceedings of SPIE - The International Society for Optical Engineering, 5239*, 314-321

Nicodemus, F.E., Richmond, J.C., Hsia, J.J., Ginsberg, I.W., & Limperis, T. (1977). Geometrical considerations and nomencalture for reflectance. *Natl Bur Stand (US) Monogr*, 1-52

Nielsen, A.A. (2001). Spectral mixture analysis: Linear and semi-parametric full and iterated partial unmixing in multi- and hyperspectral image data. *Journal of Mathematical Imaging and Vision, 15*, 17-37

Okin, G.S., Roberts, D.A., Murray, B., & Okin, W.J. (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment, 77*, 212-225

Olea, R.A. (1991). Geostatistical Glossary and Multilingual Dictionary: Members of the 1984–1989 IAMG Committee on Geostatistics. Oxford University Press, New York.

Painter, T.H., Dozier, J., Roberts, D.A., Davis, R.E., & Green, R.O. (2003). Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data. *Remote Sensing of Environment, 85*, 64-77

Quattrochi, D.A. (1993). The need for a lexicon of scale terms in integrating remote sensing data with geographic information systems. *Journal of Geography, 92*, 106-212

Rainey, M.P., Tyler, A.N., Gilvear, D.J., Bryant, R.G., & McDonald, P. (2003). Mapping intertidal estuarine sediment grain size distributions through airborne remote sensing. *Remote Sensing of Environment, 86*, 480-490

Ramsey, M.S., & Christensen, P.R. (1998). Mineral abundance determination: Quantitative deconvolution of thermal emission spectra. *Journal of Geophysical Research B: Solid Earth, 103*, 577-596

Ray, T.W., & Murray, B.C. (1996). Nonlinear spectral mixing in desert vegetation. *Remote Sensing of Environment, 55*, 59-64

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R.O. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment, 65*, 267-279

Roberts, D.A., Smith, M.O., & Adams, J.B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment, 44*, 255-269

Sabol Jr, D.E., Adams, J.B., & Smith, M.O. (1992). Quantitative subpixel spectral detection of targets in multispectral images. *Journal of Geophysical Research, 97*, 2659-2672

Schaepman-Strub, G., Schaepman, M.E., Painter, T., Dangel, S., Martonchik, J.V., & Verstraete, M. (2004). Review of reflectance nomenclature used in optical remote sensing with quantitative comparisons. *Spectrodirectional Reflectance Analysis and Definition for the Estimation of Vegetation Variables*, 73-110

Schneider, D.C. (1994). Quantitative Ecology: Spatial and Temporal Scaling. Academic Press, San Diego.

Schönermark, M.v., Geiger, B., Roser, H.P., 2004. Reflection Properties of Vegetation and Soil. Wissenschaft und Technik Verlag, Berlin

Settle, J.J., & Drake, N.A. (1993). Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing, 14*, 1159-1177

Shimabukuro, Y.E., & Smith, J.A. (1991). The least-squares mixing models to generate fraction images derived from remote sensing multispectral data. *IEEE Transactions on Geoscience and Remote Sensing, 29*, 16-20

Shoshany, M., & Svoray, T. (2002). Multidate adaptive unmixing and its application to analysis of ecosystem transitions along a climatic gradient. *Remote Sensing of Environment, 82*, 5-20

Singer, R.B., & McCord, T.B. (1979). Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. *Proc. 10th Lunar Planet. Sci. Conf.*, 1835-1848

Smith, M.O., Johnson, P.E., & Adams, J.B. (1985). Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. *J.Geophys. Res.,* 90, 797–804.

Strang, G. (1988). Linear Algebra and Its Applications. Harcourt Brace Jovanovich.

Swain, P.H., & Davis, S.M. (1978). Remote sensing: the quantitative approach. McGraw-Hill International Book Company.

Tate, N.J., & Atkinson, P.M. (2001). Modelling Scale in Geographical Information Science, John Wiley & Sons, Ltd., New York.

Tian, Y., Woodcock, C.E., Wang, Y., Privette, J.L., Shabanov, N.V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B., Häme, T., Andersson, K., Ozdogan, M., Knyazikhin, Y., & Myneni, R.B. (2002). Multiscale analysis and validation of the MODIS LAI product I. Uncertainty assessment. *Remote Sensing of Environment, 83*, 414-430

Tompkins, S., Mustard, J.F., Pieters, C.M., & Forsyth, D.W. (1997). Optimization of endmembers for spectral mixture analysis. *Remote Sensing of Environment, 59*, 472-489

Turner, D.P., Ritts, W.D., Cohen, W.B., Gower, S.T., Zhao, M., Running, S.W., Wofsy, S.C., Urbanski, S., Dunn, A.L., & Munger, J.W. (2003). Scaling Gross Primary Production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation. *Remote Sensing of Environment, 88*, 256-270

Van Der Meer, F. (1999). Iterative spectral unmixing (ISU). *International Journal of Remote Sensing, 20*, 3431-3436

Van Der Meer, F., & De Jong, S.M. (2000). Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members. *International Journal of Remote Sensing, 21*, 2781-2797

Vikhamar, D., & Solberg, R. (2003). Snow-cover mapping in forests by constrained linear spectral unmixing of MODIS data. *Remote Sensing of Environment, 88*, 309-323

Wang, Q., Tenhunen, J., Falge, E., Bernhofer, C., Granier, A., & Vesala, T. (2004). Simulation and scaling of temporal variation in gross primary production for coniferous and deciduous temperate forests. *Global Change Biology, 10*, 37-51

Woodcock, C.E., & Strahler, A.H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment, 21*, 311-332

Zhang, B., Zhang, X., Liu, L., Zheng, L., & Tong, Q. (2002). Spectral unmixing and image classification supported by spatial knowledge. *Proceedings of SPIE - The International Society for Optical Engineering, 4897*, 279-283

Chapter 3

Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils

Antoine Stevens, Bas van Wesemael, Harm Bartholomeus, Damien Rosillon,
Bernard Tychon, Eyal Ben-Dor

Abstract

The temporal evolution in Soil Organic Carbon (SOC) content is often used in estimations of greenhouse gas fluxes and is an important indicator of soil quality. Regional estimates of SOC changes can only be obtained by analyzing very large number of samples over large areas due to the strong spatial variability in SOC contents. Visible and Near Infrared Spectroscopy (VNIRS) provides an alternative to chemical analyses. The benefits of this technique include a reduction of the sampling processing time, an increase of the number of samples that can be analyzed within time and budget constraints and hence an improvement of the detection of small changes in SOC stocks for a given area. Carbon contents are predicted from spectra through Partial Least Square Regressions (PLSR). The performance of three different instrumental settings (laboratory, field and airborne spectroscopy) has been assessed and their relative advantages for soil monitoring studies have been outlined using the concept of Minimal Detectable Difference. It appears that ground-based spectrometers give Root Mean Square Errors of Cross-Validation similar to the limit of repeatability of a routine SOC analytical technique such as the Walkley and Black method (± 1 g C kg^{-1}). The airborne spectrometer, despite its greater potential to cover large areas during a single flight campaign, has some difficulties to reach such values due to a lower Signal-to-Noise Ratio. Because of its statistical nature, the method and its potential rely on the stability of the calibrations obtained. It appears that calibrations are currently site-specific due to variation in soil type and surface condition. However, it is shown that PLSR can take into account both soil and spectral variation caused by different measuring campaigns and study areas. Further research is needed to develop regional spectral libraries in order to be able to use VNIRS as a robust analytical technique for precisely determining the SOC content and its spatial variation.

3.1 Introduction

In the context of global environmental change, the estimation of carbon fluxes between soils and the atmosphere has been the object of a growing number of studies (Ryan and Law, 2005). This has been motivated notably by the possibility to sequester CO_2 into soils by increasing the Soil Organic Carbon (SOC) stocks (Lal, 2004) and by the role of SOC in maintaining soil quality. Within the EU soil thematic strategy, the decline of organic matter is mentioned as one of the major threats to the soil resource (Van-Camp et al., 2004). Even if a number of studies have already demonstrated the impact of specific management practices or land use changes on SOC stocks (*e.g.* Johnson and Curtis, 2001, Guo and Gifford, 2002; West and Post, 2002), several difficulties in estimating SOC stocks and their temporal evolution remain challenging (Post et al., 2001). One of them is linked to the spatial variability of SOC that masks its slow accumulation or depletion. Even at the field scale, the sampling density required to detect a change in SOC content can be very high.

Hence, there is an urgent need to develop a cost effective monitoring system that would allow to calculate the emissions of CO_2 from individual fields, follow the evolution of SOC concentration in the topsoil as an indicator of Good Agri-Environmental Conditions and estimate the nitrogen that can be mineralized from the soil organic matter in order to reduce mineral fertiliser use. Current methods of soil analysis are too expensive and time consuming to meet the amount of data required for statistical inference in soil monitoring. New analytical techniques would allow rapid sampling and instant determination of SOC values, at the field and regional level.

Visible and Near InfraRed Spectroscopy (VNIRS) can be a suitable technique to rapidly quantify various soil characteristics simultaneously (Malley et al., 2004). There are three types of VNIRS techniques operating at different spatial scales and in different environments: (1) Laboratory Spectroscopy (LS); (2) Portable field Spectroscopy (PS) and (3) Remote Spectroscopy (RS). LS and PS rely on ground-based sensors (usually point spectrometry) and RS on air- or space- borne sensors (usually image spectrometry).

Numerous studies used laboratory spectrometers to analyse SOC content (*e.g.* McCarty and Reeves III, 2001; McCarty et al., 2002; Martin et al., 2002) and show notably their relevance for SOC inventories. Furthermore, using the NIRS concept, Ben-Dor et al. (1997) were able to show that the decomposition process of soil organic matter can be assessed solely from spectroscopy in the laboratory. Portable Spectroscopy has been used for rapid *in-situ* monitoring of soils (Kooistra et al., 2003; Udelhoven et al., 2003), or, when the sensor is mounted on a tractor, it is dedicated to the site-specific management of inputs in precision agriculture (e.g. Mouazen et al., 2007). Remote Spectroscopy has mainly been used for the

mapping of surface soil properties of agricultural fields with high spectral and spatial resolutions (Ben-Dor et al., 2002; Selige et al., 2006; Stevens et al., 2006).

SOC can be spectrally measured with a reasonable accuracy level, depending on the type of instrument and environmental conditions (LS, PS or RS), with Root Mean Square Errors (RMSE) ranging from 1 to 15 g C kg^{-1} (Stevens et al., 2006). This accuracy is lower than that of most routine laboratory SOC analysis. For instance, the Walkley and Black method (Walkley and Black, 1934) reaches maximum values of 1-2 g C kg^{-1} (Colinet, 2005). However, the large number of samples that can be produced outweighs the slight loss of precision compared to traditional chemical analyses.

This paper aims to investigate the potential of different VNIRS approaches for the monitoring of SOC in croplands at the regional by evaluating i) the loss of accuracy by using such techniques outside the controlled conditions of the laboratory, ii) the stability of the calibrations and iii) the capacity of such techniques to process a larger number of samples than conventional techniques would not allow. Spectral measurements from LS, PS and RS were collected during a field campaign in 2005 and SOC contents were predicted using Partial Least Square Regressions (PLSR). The stability of the calibration of PS techniques was addressed by means of PS spectral datasets from two field campaigns (in 2003 and 2005) and three study areas. Finally, the loss of accuracy will be weighed against the potential of VNIRS techniques to high density datasets for regional SOC change monitoring using the concept of Minimal Detectable Difference.

3.2 Materials and Methods

3.2.1 Study Sites
The area selected for the 2005 campaign, from which LS, PS and RS data were collected, is located in the Belgian Lorraine (Tintigny; 49°43' N 5°27' E and 49°39' N 5°32' E). This site has a mean altitude of 350 m.a.s.l with a rather flat topography and a mean temperature of 8.5 °C and an annual precipitation of 1013 mm. The Tintigny area (52.5 km²) is characterized by diverse soil types, from sandy to clayey and overall the Fe$_2$O$_3$ concentration ranges from 8.8 to 24.2 g kg^{-1} (based on 30 field samples, using a dithionite extraction). Most of the selected fields are on sandy-loam or loamy-sand soils (Haplic and Gleyic Luvisol, FAO-ISRIC-ISSS, 1998). The two other study areas in the Belgian Ardennes (Ortho) and the Belgian Lorraine (Attert), from which additional PS spectra have been measured in September 2003, are fully described in Stevens et al. (2006). Soils in the Ortho area site are mainly loamy acid brown soils (Cambisols, FAO-ISRIC-ISSS, 1998) while the Attert test site has similar pedological characteristics as the Tintigny area.

3.2.2 Spectral Measurements

The time window dedicated to spectral measurements (end of spring), constrained by other projects using the same remote sensor, introduced non-ideal circumstances since bare fields, needed for soil sensing, are extremely rare in the region during spring. In order to circumvent the problem, we cleared 81 areas of 7.5 by 7.5 m in 8 maize fields. The size of these bare plots corresponds to an area of 3 by 3 pixels of the sensor. Furthermore, two fields of approx. 0.25 ha each were ploughed and left bare by the farmers within which 36 similar plots have been delimited, resulting in a total of 117 plots.

3.2.2.1 Remote Spectroscopy

The remote sensor AHS-160 was mounted onboard a CASA aircraft and flown on June 20, 2005, at an altitude of 1000 m.a.s.l. This sensor provides a total of 80 bands covering the spectrum in the Visible (VIS: 400-750 nm), Near InfraRed (NIR: 750-2500 nm), Mid Infrared (MIR: 3300-5400 nm and 8200-12700 nm) region, with a Instantaneous Field Of View of 2.5 mrad and a Field Of View of 90° enabling pixel size of 2.6 m and swath of 2000 m. The acquired data cube was first ortho-rectified, then corrected for atmospheric interferences by using the MODTRAN-4 radiative transfer code (Berk et al., 1999) embedded in a modified version of ATCOR-4 (The, 2000; Richter, 2005). For each plot, 9 pixels centred on the middle of the plot were extracted using the ENVI software (ITT Visual Information Solutions, Boulder, CO) and averaged to give the plot representative spectrum that was further used for the statistical analysis. Some pixels, especially those located at the edge of the plot, were influenced by the surrounding vegetation and were omitted from this averaging. This was done by removing those pixels with NDVI values greater than 0.3. In order to evaluate the quality of RS spectra, a Signal-to-Noise Ratio (SNR) was estimated similarly to that in Ben-Dor and Levin (2000; eq. 3.1).

$$SNR = AV/SD \tag{3.1}$$

where AV is the mean signal of an homogeneous target and SD is the Standard Deviation of the same pixels values.

3.2.2.2 Portable Spectroscopy

PS measurements were taken with a FieldSpec Pro FR spectrometer (Analytical Spectral Devices Inc., Boulder Co) during two field campaigns in 2003 and 2005 using a similar measurement protocol. This instrument is characterized by a Full Width Half Maximum of 3 nm for the 350-1000 nm region and 10 nm for the 1000-2500 nm region and by Field Of View of 25° that from a measurement height of 1 m gave a spot size of 0.45 m at nadir. Since measurements were taken the same day as the flight (under perfectly clear sky and between 11 am and 5 pm), light conditions (stability and intensity) were optimal. Each spectrum consisted of 9 individual spectra taken across each experimental plot and measured against

Spectralon white reference under similar geometry and level of radiation. The 9 spectra were averaged to give a plot representative spectrum. The plot of the reflectance at 780 nm against the reflectance at 670 nm of the mean spectra revealed that 7 mean spectra were influenced by the presence of residual vegetation. These spectra are not a part of the "soil line" and were therefore removed from the analysis. A strong atmospheric noise due to water vapour affected wavelengths between 1340-1430 nm, 1810-1970 nm and beyond 2400 nm. These were removed.

3.2.2.3 Laboratory Spectroscopy

Air-dried and sieved soil samples in small cups were measured with a FieldSpec Pro FR spectrometer that was hooked to a contact probe with a built in halogen bulb for illumination and Spectralon as a white reference. Nine spectra were produced per sample from which the mean spectrum of the plot in question was calculated. As the measurement conditions of the contact probe are ideal, the entire spectral range (350-2500 nm) was further used.

3.2.3 Soil Analysis

In each plot, 9 sub-samples were collected to a depth of 5 cm, mixed and a representative sample was taken for carbon and spectral analysis in the laboratory. All soil samples were air dried (30 ˚C) and sieved to pass a 2 mm mesh. The SOC content was determined using the Walkley and Black method (Walkley and Black, 1934). Moreover, 15 soil samples were taken randomly in all of the selected plots, up to a maximum depth of 1 cm and put into a hermetic plastic bag in order to determine the soil moisture content upon drying the sample at 105°C for 24 hours. SOC and soil moisture statistics of each field of the Tintigny area are given in Table 3.1.

Table 3.1. Summary statistics of field samples (in g kg^{-1})

Field	N	Carbon content				Moisture content				Main texture[a]
		Mean	SD	Min	Max	Mean	SD	Min	Max	
G1	21	11.2	1.3	9.2	14.6	16.3	9.8	6.7	44.0	Loamy-sand
G2	15	15.6	1.2	13.3	17.7	5.3	2.5	3.0	12.2	silt/silt-loam
P1	15	13.6	1.1	11.5	15.9	6.1	2.6	3.4	12.7	sandy-loam
P2	15	16.6	2.3	13.0	22.1	17.4	6.5	8.7	28.3	sandy-loam
P3	9	13.3	1.8	10.7	15.9	21.8	7.5	12.1	35.8	silt-loam/silt
P4	6	9.8	2.1	5.9	11.9	22.4	9.8	6.2	35.5	Loamy-sand/sandy-loam
P5	8	13.3	3.7	10.5	22.1	30.4	39.1	10.1	126.8	Loamy-sand/sandy-loam
P6	9	15.6	2.1	13.3	19.8	19.7	10.8	5.5	38.9	silt-loam/loam/sandy-loam
P7	10	12.6	1.2	11.3	15.5	14.1	5.4	5.7	23.0	Loam/clay-laom/silt-loam
P8	9	10.8	1.5	8.6	12.7	12.8	5.8	4.7	21.8	Loamy-sand/sandy-loam
TOTAL	**117**	**13.4**	**2.7**	**5.9**	**22.1**	**15.2**	**13.6**	**3.0**	**126.8**	-

[a]USDA textural class, translated from the Belgian soil map

3.3 Statistical Analysis

3.3.1 Signal pre-processing

Before the quantitative statistical analysis reflectance is converted into "absorbance" (log [1/reflectance]). Noise reduction was achieved through standard pre-treatments like differentiation and smoothing. Each pre-treatment or combination of pre-treatments was related separately to carbon content. These pre-treatments consisted of (Reeves and Delwiche, 2003): (i) 1st and 2nd derivatives, (ii) 1st and 2nd gap derivatives (differentiation is done between points x_{i-n} and x_{i+n} with $n > 1$), (iii) Savitzky-Golay smoothing and differentiation algorithms (Savitzky and Golay, 1964), (iv) mean centering and variance scaling, and (v) skipping every n data points (allow to reduce the number of spectral bands).

3.3.2 Multivariate calibration and validation with Partial Least Square Regression

Each spectral pre-treatment was related to the carbon content by PLSR as implemented in *SAS* (*SAS* Institute Inc., Cary, NC). The maximum number of Latent Variables (LV) was set to 10. The optimal number of LV was determined by *leave-one-out* cross-validation.

Two different approaches, depending on the number of samples available, were used to validate the model and evaluate its prediction error: (i) test set validation and (ii) cross-validation. In the test set validation procedure a third of the data was randomly selected and used to validate the model by confronting measured and predicted values. The performance of the model was measured by the Root Mean Square Error of Prediction (RMSEP; eq. 3.2):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

(3.2)

Where $\hat{y}_i$ and yi are the predicted and the observed values of the sample i in the test set of n samples. In the cross-validation procedure the dataset is divided in several groups, each one being used alternatively as validation set. We used the *leave-one-out* cross-validation procedure, during which each sample was validated individually on the basis of the calibration set compiled from the remaining dataset. Then, the difference between predicted and observed value of each individual pass was determined to calculate the Root Mean Square Error of Cross-Validation (RMSECV). For a small dataset, this method has the advantage to estimate the 'true' prediction error more accurately (Ben-Dor and Banin, 1990; Martens and Dardenne, 1998). The RMSEP overestimates the true prediction uncertainty since it includes measurement errors in the reference values. Therefore, a simple correction for this error leads to the corrected RMSEP (RMSEP$_{corr}$; Faber et al., 2004; eq. 3.3).

$$RMSEP_{corr} = \sqrt{RMSEP^2 - RMSEL^2}$$

(3.3)

where RMSEL is the Standard Error of Laboratory analyses i.e. the Standard Deviation (SD) of differences between duplicate samples of the reference method. A corrected RMSECV (RMSECV$_{corr}$) has been calculated according to the same rule. RMSEL of the Walkley and Black method has been fixed to 0.1 g C kg^{-1} (Colinet et al., 2005). This latter value is a rough estimate, because RMSEL can vary with the range of SOC content measured (usually lower for lower SOC content). The Ratio of Performance to Deviation (RPD) was computed in order to interpret the prediction ability of each pre-treatment (Chang and Laird, 2002). RPD is the ratio between the SD of the reference method against that of the RMSEP or RMSECV. Chang and Laird (2002) defined three classes of RPD: category A (RPD > 2) are models that can accurately predict the property in question, category B (RPD = 1.4~2) is an intermediate class which regroups models that can be possibly improved, and models falling in category C (RPD < 1.4) have no prediction ability.

In order to construct a robust calibration model, an automatic outlier removal procedure was implemented in *SAS*, which is close to the procedure developed by Koshoubu et al. (2001). The quality of the calibration was also evaluated by checking the shape of the X-weights of the LV's. Noisy features may indicate overfitting and thus that the calibration may not be suitable for future prediction. If so, the PLS model was rejected. The best pre-treatment able to predict un-sampled sites with the highest precision was selected on the basis of its RPD value and the percentage of X (wavelengths) and Y (carbon content) variation explained by the model.

3.3.3 Assessing the stability of calibrations

The problem of calibration stability can be described as the shifting, warping and scaling of the spectral predictors arising from using spectra measured under conditions/instruments different from the ones used during the model calibration (Marx and Eilers, 2002). The PS dataset of the 2005 field campaign was joined with PS spectra collected in 2003 in the Ardennes and Lorraine regions in order to test the stability over time and across different physiographic zones of the calibrations obtained by PLSR. The resulting dataset contained 201 samples with a range of carbon contents, texture and field surface conditions. Three different tests were carried out: (i) predict samples of the Lorraine region taken during the 2003 field campaign using the regression obtained from samples of the same area but from a different field campaign (2005), (ii) validate the entire PS dataset using a random validation set and (iii) cross-validate the entire PS dataset.

3.3.4 SOC monitoring and Minimum Detectable Difference

The number of samples needed for monitoring and the detection of significant change in SOC content depends on the variability within the study area. Power analysis can be used to calculate the Minimal Detectable Difference (MDD) between two measuring campaigns as a function of sample number and hence, designing a future sampling strategy. Here, we use

this methodology to assess the benefits of spectroscopic techniques. MDD has been calculated for each field on the basis of the variability in SOC and the number of samples collected during the field campaigns. MDD is given by (Garten and Wullschleger, 1999; eq. 3.4):

$$MDD = \frac{2 \cdot \Phi \cdot s}{\sqrt{n}} \qquad\qquad (3.4)$$

where n is the minimum sample size, s is the estimated standard deviation and Φ is a tabulated critical value. Minimum sample size required to achieve a specified precision (MDD) has been calculated as well from eq. 3.4.

3.4. Results and Discussion

3.4.1 Spectral Data Quality and Comparison

Several pixels of homogeneous light (gravel quarry), grey (parking) and dark (water body) areas were used to calculate a mean Signal to Noise Ratio (SNR) of the RS sensor, according to eq. 3.1. Generally, the SNR of the RS sensor was low in the 1900-2500 nm region (mean: 5.2) where its spectral resolution is high while the SNR was high in the 430-1030 nm region (mean: 39.8) where its spectral resolution is low. The low SNR in the 1900-2500 nm region is attributed to a degradation of the optics during the flight campaigns and atmospheric absorption. Mean SNR of LS and PS across the entire spectral range were respectively 64.8 and 16.7. The difference between LS and PS is probably due to a lower stability of incoming light and variation in viewing/illumination geometry under field conditions. LS and PS spectra were resampled to fit the resolution of the RS sensor and plotted in order to visually compare the spectra retrieved by the three techniques (Fig. 3.1). Maximum reflectance varied from 15 % to 55 %. Fig. 3.1a shows spectra of the same experimental plot as measured by the three sensors. Small absorption features can be seen around 500 nm, 700 nm and 2300 nm with a more distinct one at 2200 nm associated with combination mode of OH (v_{OH} and δ_{OH}) in the clay lattice (Ben-Dor and Banin, 1990). Classical absorption at 1400 nm and 1900 nm due to the soil water are masked by the poor spectral resolution of the RS sensor at these wavelengths.

Above all, LS, PS and RS spectra showed large difference in absolute reflectance (Fig. 3.1a). The mean reflectance of the PS is much lower than LS. This is probably caused by shadowing due to soil roughness that adds a micro shade domain to the measured reflectance. This reduction seems not linear across the spectrum since the visible part is less affected by this reduction than the NIR region. Nevertheless, the shape of the spectrum obtained by PS seems similar to the shape of the spectrum of LS and, when looking at the first derivative of

the reflectance, the LS and PS yield more or less the same spectrum (Fig.3.1b). The RS differs strongly from LS and PS in the 1900-2500 nm, while having a closer match in the 430-1400 nm region. The strong divergence as well as the low SNR in the 1900-2500 nm region indicate that the atmospheric correction of the RS sensor was not adequate or the radiometric calibration of this instrument is questionable. The MIR window of the RS sensor suffered the same problem. The 1900-2500 nm and the MIR region were thus considered useless and, as such, removed from the analysis.
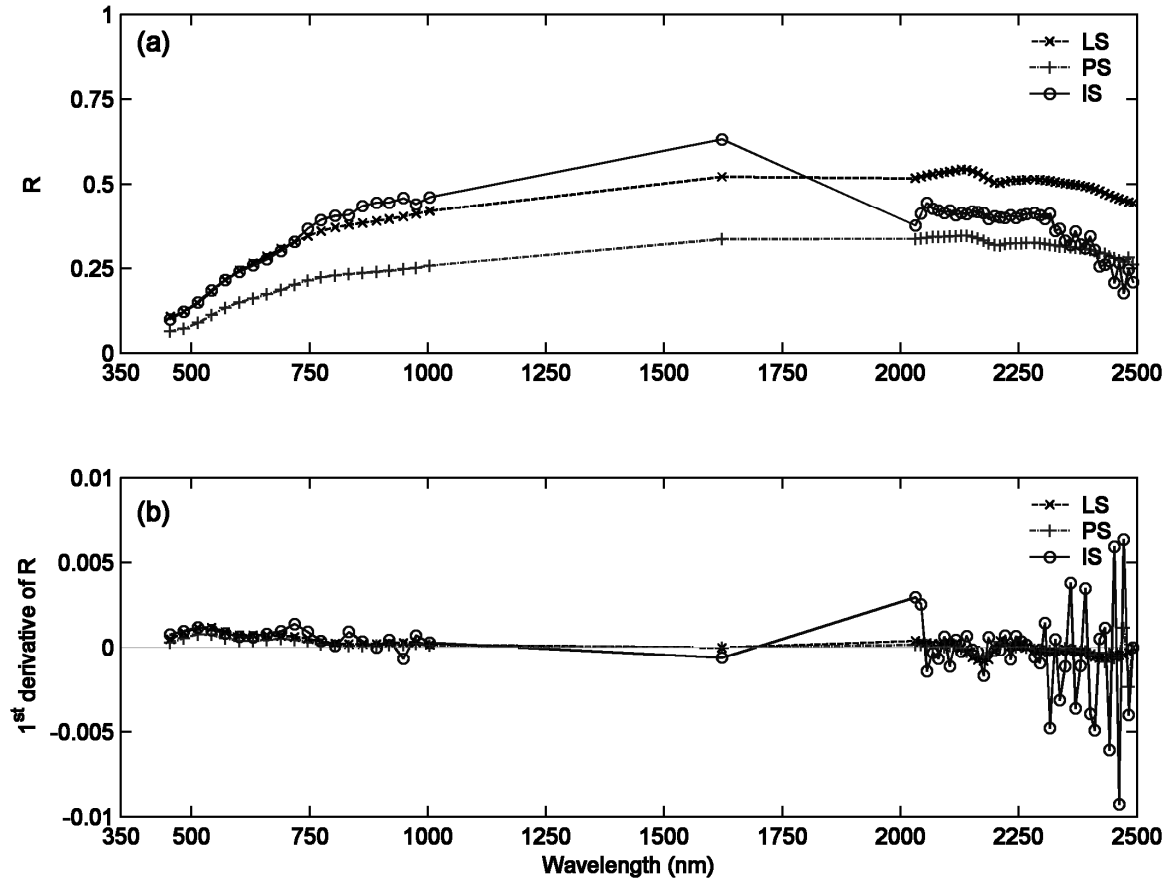

Fig. 3.1. Comparison between (a) the reflectance and (b) the 1st derivative of the reflectance, for three spectra of the same plot, as measured by RS (solid line), PS (dashed line) and LS (dash-dotted line). PS and LS spectra have been resampled to the AHS (RS) configuration.

3.4.2 Model Calibration and Validation

3.4.2.1 Comparing the predictive ability of LS, PS and RS

Spectral data collected during the 2005 field campaign were used to compare the performance of the different spectroscopic techniques (LS, PS and RS). These three datasets contain more or less 100 samples each, so that cross-validation was considered more appropriate to estimate future predictive ability than test set validation. The results show that RPD decreased from ground-based (LS and PS) to remote measurements (RS; Table 3.2).

This decrease of predictive ability is undoubtedly due to difference in sensor characteristics, an increase in environment-induced variation and in uncontrolled measuring/experimental conditions (e.g. SNR, light source quality etc.). RPD values ranged from 1.47 (RS) to a maximum of 2.11 (PS). According to the classification of Chang and Laird (2002), we can state that ground-based spectroscopy (either LS and PS) can be a reliable technique to measure SOC content. While PS and LS have similar RMSECV (1.2 g C kg⁻¹), PS has a better RPD than LS due to a slightly higher standard deviation in the calibration set. This result is counter-intuitive, since LS is measuring spectra under optimal conditions and thus is supposed to produce more accurate predictions. Two reasons can potentially explain the slightly higher accuracy for PS. First, one can assume that a large part of the noise has been successfully removed by pre-treatments. Secondly, we presume that spectral variation caused by disturbing factors not encountered in the laboratory was effectively handled. These disturbing factors can be: (i) soil moisture content, (ii) soil roughness and (iii) vegetation cover. Since soil moisture was very low across samples and showed little variation (Table 3.1), no signal processing was necessary to remove the moisture effect. Similarly, soil roughness was generally low because measurements were taken on fields in seedbed conditions. The problem of vegetation cover was resolved by removing samples different from the "soil line". In summary, we can state that the performance of PS is equivalent to LS when measuring SOC under specific surface conditions and subject to appropriate signal pre-treatments. Compared to ground-based sensors, RS has a RPD value close to the one of a very bad model (RPD close to 1.4), according to the classification of Chang and Laird (2002) and thus produces less accurate measurements. However, the removal of the 1900-2500 nm region before the PLSR may cause some discrepancies when comparing the results of RS (430-1400 nm) with LS and PS (350-2500 nm).

The RMSECV of RS is 1.7 g C kg⁻¹, yielding a RMSECV$_{corr}$ twice as high as the one of ground-based spectroscopy (1.4 g C kg⁻¹). The RMSECV$_{corr}$ constitutes a limit beyond which it is impossible to reliably differentiate two samples having a smaller difference in concentration of the substance being analyzed. Those values are comparable to the limit of repeatability (RMSE of replicate samples) of 1 g C kg⁻¹ of the Walkley-Black method (Colinet et al., 2005). This would mean that ground-based spectroscopy is at least as good as a standard analytical method. However, it should be noted that the samples have a relatively low mean (13.4 g C kg⁻¹) and standard deviation (2.7 g C kg⁻¹) so that this statement is only true for this particular type of soils and range of SOC content.

3.4.2.2 Testing calibration stability with PS datasets

Since the PS dataset of 2003 and 2005 field campaigns contained 201 samples, a true test set validation could be used rather than a cross-validation. Location of the field campaigns and SOC statistics of the combined dataset are given in Table 3.3.

Table 3.2. Predictive statistics of the best model for each sensing technique.

Sensor type	X var[a]	Y var[a]	LV number[b]	N	Standard Deviation	RMSECV[c]	Corrected RMSECV[c]	RPD[d]	Outlier	Treatment	Window size[e]
						$\longrightarrow$ g C kg^{-1} $\longrightarrow$					
IS	98.36	52.74	2	110	2.5	1.7	1.4	1.47	3	No treatment	0
PS	72.68	76.97	3	99	2.6	1.2	0.7	2.11	1	Savitzky-Golay 1st derivative 2th order polynomial	65
LS	99.96	75.30	8	117	2.5	1.2	0.7	2.03	5	Savitzky-Golay smoothing 5th order polynomial	65

[a] Total predictor (X) or response (Y) variation (%) explained by the model; [b] Smallest number of PLS factors determined by the cross-validation procedure, [c] Root Mean Square Error of Cross-Validation, [d] Ratio of Performance to Deviation; [e] Window size of the best pre-treatment in number of spectral bands.

As a preliminary approach, we tried to build a stable calibration curve for the same physiographic region. Samples taken during the 2005 field campaign in Lorraine were used to calibrate a model, which was validated with spectral data measured in 2003 in the same region. RPD in the calibration set is high (4.26) while RPD in the validation set is low reaching a value of 1.09 (Table 3.4). It indicates a very bad model and poor replication of a calibration trough time on the same area. The calibration made on 2005 samples was obviously not representative of spectra measured in 2003. Nevertheless, the range of SOC content measured in 2005 was within the range of SOC content measured in 2003 (Table 3.3) and soil types were similar. Moreover, even if light conditions were different between the two campaigns, spectral calibration through white reference measurements should ensure a constant spectral shape. This first test shows that the methodology is very sensitive to small changes in the predictors and that a calibration is still needed before each measurement campaign.

Table 3.3. Summary statistics of SOC (g C kg^{-1}) for the different field campaigns in the PS dataset

Field campaign	N	Mean	SD	Min	Max	Region
Attert 2003	37	13.3	4.9	5.7	22.8	Lorraine
Ortho 2003	65	26.9	3.5	19.9	37.3	Ardenne
Tintigny 2005	99	13.2	2.7	5.9	22.1	Lorraine
All campaigns	201	17.7	7.3	5.7	37.3	-

In a second experiment, a calibration was carried out using the dataset from the different campaigns randomly leaving out a portion (1/3) that was used for validation (Table 3.4, second line). RPD in the calibration set was considered satisfactory since it reached a value of 2.70. However, in the validation set, RPD is falling to 1.21 with a RMSEP$_{corr}$ of 6.96 g C kg^{-1}.

Table 3.4. Predictive statistics of the best model for PS under different validation settings.

Validation type	X var[a]	Y var[a]	LV[b]	N	SD[c]	RMSECV[d]	Corrected RMSECV[d]	RPD[e]	Outlier	Bias	N	SD[c]	RMSEP[d]	Corrected RMSEP[d]	RPD[e]	Bias	Treatment	Window size[f]
				Calibration							Validation							
Validation of 2005 samples against 2003 samples in Lorraine	79.55	79.30	4	99	2.59	1.17	0.61	4.26	5	-0.26	37	4.85	4.56	4.45	1.09	-0.45	Savitzky-Golay 1st derivative 5th order polynomial	129
Random test set validation of all field spectral data	99.74	84.81	5	134	6.90	2.67	2.48	2.70	7	0.26	67	7.79	7.03	6.96	1.21	1.21	Savitzky-Golay smoothing 5th order polynomial	5
Cross-validation of all field spectral data	99.75	87.50	5	201	7.12	2.50	2.29	2.84	17	0.00	-	-	-	-	-	-	No treatment	0

[a] Total predictor (X) or response (Y) variation (%) explained by the model; [b] Smallest number of PLS factors determined by the cross-validation procedure, [c] Standard Deviation in the calibration or validation set, [d] Root Mean Square Error of Cross-validation or Validation, [e] ratio of Performance to Deviation in the calibration or validation set, [f] Window size of the best pre-treatment in spectral bands unit.

This decrease is due to a few outliers, clearly visible in Fig.2 and increasing the latter statistic. For the same reason, the bias in the validation set was unexpectedly large (1.21 g C kg^{-1}). Either these samples are spectral outliers and are influenced by other factors than the dependent variable (like roughness), or they are not comparable to the group of samples used in the calibration in terms of SOC concentration. In our case, the five outliers are samples from one particular field, for which all other samples in the calibration set have been removed during the outlier detection phase. The processing of all the pre-treatments by PLSR confirmed that spectra from this field were systematically removed from the calibration set during the outlier detection phase (results not shown). A preliminary Discriminant Analysis could help to identify these kinds of problematic spectra. As a matter of fact, PLSR was unable to predict samples falling outside of the calibration set. The development of a broader spectral library for soils of southern Belgium should be able to settle this difficulty.

In a third experimental test, we analyzed the entire PS dataset using cross-validation. This gave better results (RPD = 2.84, RMSECV$_{corr}$ = 2.29 g C kg^{-1} , 17 outliers) indicating that the statistical procedure can effectively predict samples when they are within the same range as the calibration set. Scores are linear combinations of the predictors (X-score) or response (Y-score). Plotting X-scores by Y-scores can be used to visualize samples that are similar in spectral shape (X-score axis) and SOC concentration range (Y-score axis). It appears that spectra of the three study areas (Tintigny (A), Attert (B), Ortho (C); Table 3) can be discriminated on the basis of their scores, defining groups of similar characteristics (Fig.3). Tintigny spectra generally have low X-scores while Attert and Ortho spectra generally have high X-scores, this pattern corresponds to the two field campaigns (A: 2005 and B+C: 2003).
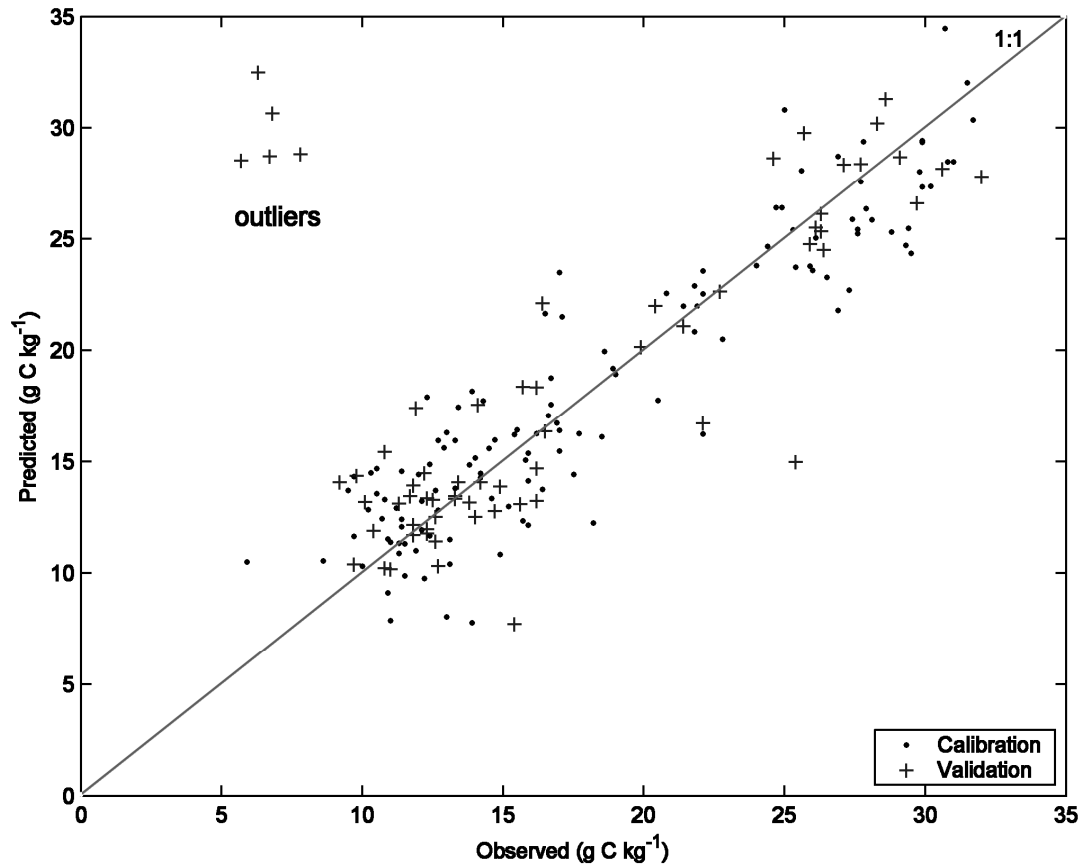
Fig.3.2. Plot of predicted *vs* observed organic carbon as obtained after a random test set validation on field spectroscopic spectra from different study areas and field campaigns (calibration: dots; validation: cross)

Along the Y-axis, there are two groups again (A+B with low Y-scores and C with high Y-cores) differing in their carbon content range (see Table 3.3). These observations are encouraging because the model constructed by PLSR, reaching relatively high RPD, appears to be able to take into account this double variation in spectral shape and carbon content.

This fact is particularly important when assessing the stability of the calibration at the broader scale and between measurement campaigns. Indeed, the power of the technique would be greatly enhanced if the technique shows a high stability across time and space. Still, RMSECV$_{corr}$ is higher than the one obtained by the model restricted to only one study area (Table 3.2). This difference can be due to a larger range of SOC contents but also possibly to the variability in field conditions across study areas. For instance, some fields in the 2003 campaign were not in seedbed conditions while others were covered by vegetation residues. Therefore, we need to define surface conditions characteristics required to measure SOC in the field with acceptable accuracy (like low variation in moisture content of the soil surface, low roughness and absence of vegetation).

Currently, calibrations are often applied to limited areas with homogeneous soils or geology (*e.g.* Kooistra et al., 2003; Udelhoven et al., 2003) and attempts to predict samples from different locations and studies have shown relatively low validation results (*e.g.* Brown et al., 2005; Stevens et al., 2006) due to varying soil types or soil surface conditions. Brown et al. (2006) produced a generic model for several thousands of soils from all over the world but failed to predict C at an acceptable accuracy for most applications (RMSE: 9 g C kg^{-1}). However, such a global spectral library can be used to classify measured spectra (e.g. based on scores). Then, only a portion of the library may be used for prediction.
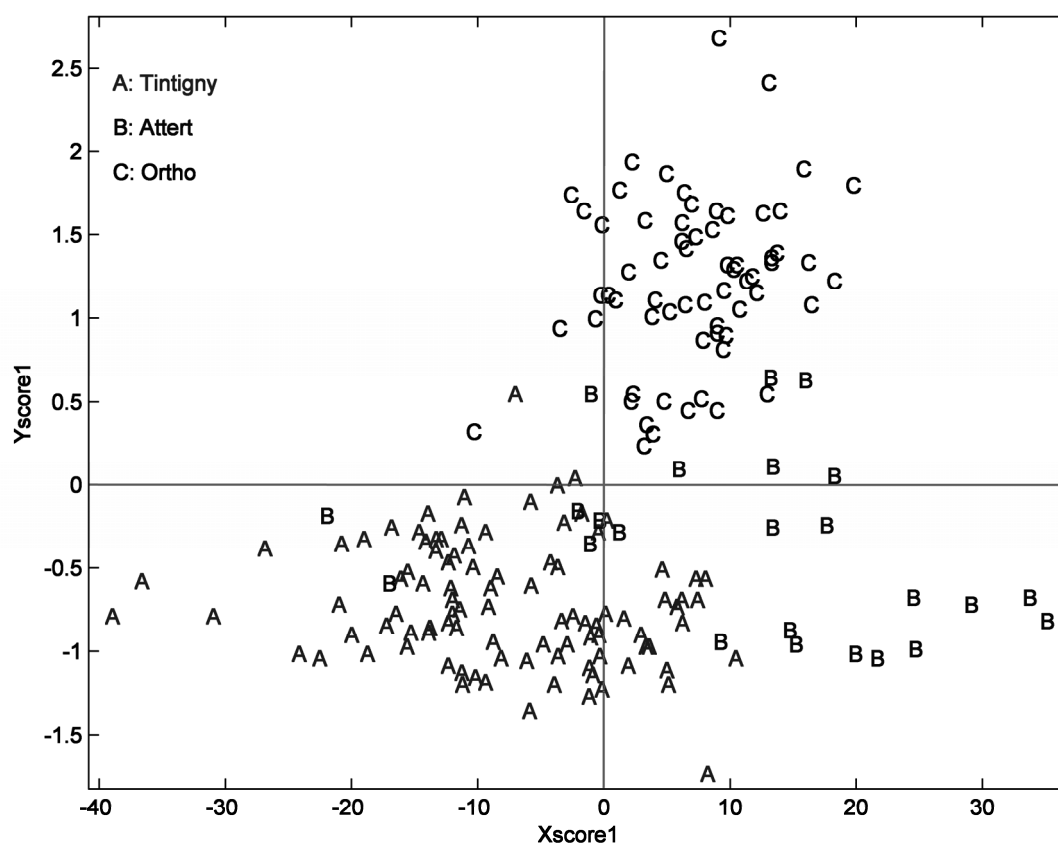

Fig.3.3. Plot of Y-Score *vs* X-Score of the first Latent Variable as obtained after a cross-validation on the joined PS dataset. Samples are labeled according to their origin: "A" for Tintigny 2005, "B" for Attert 2003, and "C" for Ortho 2003.

3.4.3 Monitoring SOC Change in croplands

Due to the large spatial variability in SOC content, it is difficult detect a significant change in SOC stocks of croplands sampled at two dates. The Minimum Detectable Difference (MDD, eq. 3.4) decreases non-linearly with sample size and for a given sample size increases with the variance. Considering that generally SOC stock changes as a result of management in European agricultural soils are lower than 2 t C ha^{-1} y^{-1} (Freibauer et al., 2004), the maximum

change in SOC stock to be expected over a reasonable time span (3–5 years) would not be higher than 5 t C ha^{-1}. This threshold corresponds to a concentration of 1.2 g C kg^{-1} (considering a mean bulk density of 1.3 g cm^{-3} and a depth of 30 cm). At the field scale, the number of samples required to detect such a change, assuming normality, ranges from 9 to 156 samples (Table 3.5). With the number of samples collected in each field during the measuring campaigns, one can detect changes varying between 1.2 and 4.5 g C kg^{-1} (Table 5). These calculations show the need for high sampling densities, even at the field scale. To detect a similar change in mean SOC content in the three study areas, the number of samples required would be sensibly higher (n = 486; Table 3.5). Hence, at regional scale, the need for high sampling densities is even more important because the effect of local management practices on SOC content is likely to be obscured by spatial heterogeneity generated by drivers acting at broader scale (e.g. climate, geology; Holmes et al., 2004). Increasing the sampling density will result in a smaller detection limit of SOC stock change.

Table 3.5. SOC statistics per field and Minimal Detectable Difference (MDD).

Study area	Field	Mean	SD	N[a]	MDD	N required[b]
		——— g C kg^{-1} ———			g C kg^{-1}	
Attert	A	19.0	2.7	10	3.5	83
	C	6.8	0.8	9	1.2	9
	F	12.0	1.0	10	1.3	12
	H	15.2	1.7	8	2.6	38
Ortho	L	29.7	3.7	11	4.5	156
	P	23.3	3.3	11	4.0	121
	R	25.6	3.0	11	3.7	103
	S	27.1	1.7	11	2.0	32
	T	28.1	3.3	11	4.0	122
	U	27.6	2.6	10	3.4	81
Tintigny	G1	11.2	1.4	20	1.2	19
	G2	15.6	1.2	15	1.2	14
	P1	13.9	1.2	9	1.6	17
	P2	16.4	2.3	14	2.4	58
	P3	13.3	1.8	9	2.5	39
	P4	9.8	2.1	6	3.8	61
	P5	12.1	1.0	7	1.7	14
	P7	12.6	1.2	10	1.6	18
	P8	11.1	1.4	8	2.0	23
	ALL	17.7	7.3	201	1.9	486

[a]Number of samples collected in the field; [b]Number of samples in each field required to detect a stock change of 5 t C ha^{-1}.

Conventional sampling strategies are often too time consuming and expensive to provide such large amounts of samples. LS, PS and RS may be an attractive alternative to estimate SOC changes with higher confidence at various scale. In this respect, users willing to choose one of these techniques for a particular application have to weigh the accuracy against the

area to be covered. While LS has the advantage to allow stable calibration trough time, it still requires sample preparation (collecting sample, sieving, crushing, etc...). This study shows that PS can reach accuracy comparable to LS but it requires a calibration before each campaign. PS can be used to rapidly scan the soil surface of individual fields and estimate with confidence their mean SOC content. RS might be a practical way to spatially evaluate soil carbon and a useful vehicle to track the SOC content change on large scales. RS can provide a large amount of samples (up to 1600 pixels/ha) but still shows unacceptable level of accuracy. Similar to PS, a calibration before each campaign would be also necessary due to difficulties in spectral calibration.

3.5. Conclusions

Spatial variability and slow temporal change in SOC stocks at various scales reduce our capacity to detect changes within a short time span. This problem can be resolved by using a high sampling density that can only be achieved by means of more rapid analytical methods like VNIRS. Soil spectra can be measured with different instruments and settings (laboratory, field, airborne), each having its comparative advantages and limitations.

The accuracy of field spectroscopy is equivalent to laboratory spectroscopy when measuring SOC under specific surface conditions (low variation in moisture content of the soil surface, low roughness, absence of vegetation) and appropriate pre-treatments able to extract information from noisy spectra. The potential of this technique is high, since it requires very little prior sample manipulation. Our calibrations showed that it is possible to reach accuracies comparable to standard analytical method (Walkley-Black). These two techniques can thus be potentially used for monitoring studies where their speed is a valuable advantage. Airborne techniques, such as RS, appear, for the time being, not able to predict SOC with an acceptable accuracy due to their low SNR and problem to achieve true spectral information. Nevertheless, the greater potential lies in this technique and more efforts have to be put in spectrum calibration.

The use of different datasets from different study areas and field campaigns showed that calibrations are currently site-specific and partly fail to predict, under a proper test set validation procedure, samples belonging to another study area or falling outside of the range of the calibration set. The development of a regional calibration, valid for soils of the same physiographic region is thus one of the first research priorities. The potential benefit of such techniques for soil monitoring studies has been evaluated using the concept of the minimum detectable difference (MDD). The limit of accuracy of the method is sufficient to capture changes in SOC stocks at the field and regional scale. Even for RS, having still rather low precision level, the large number of samples that can be taken could improve estimates of

mean SOC stocks compared to more precise chemical analyses of an inherently smaller number of samples.

Acknowledgements

References

Ben-Dor, E. and Banin, A., 1990. Diffuse reflectance spectra of smectite minerals in the near infrared and their relation to chemical composition. *Sciences Géologiques Bulletin, 43*: 117-128.

Ben-Dor, E., Inbar, Y. and Chen, Y., 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400-2500 nm) during controlled decomposition process. *Remote Sensing and Environment, 61*: 1-15.

Ben-Dor, E. and Levin, N., 2000. Determination of surface reflectance from raw hyperspectral without simultaneous ground data measurments: a case study of the GER 63-channel sensor data acquired over Naan, Israel. *International Journal of Remote Sensing, 21*: 2053-2074.

Ben-Dor, E., Patkin, K., Banin, A. and Karnieli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data - a case study over clayey soils in Israel. *International Journal of Remote Sensing, 23*: 1043-1062.

Berk, A., Anderson, G.P., Acharya, P.K., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W. and Alder-Golden, S.M., 1999. Modtran4 User's Manual. Technical report, Hanscom, USA.

Brown, D.J., Bricklemyer, R.S. and Miller, P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma, 129*: 251-267.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Mays, M.D. and Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma, 132*: 273-290.

Chang, C. and Laird, D.A., 2002. Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Science, 167*: 110-116.

Colinet, G., Laroche, J., Toussaint, B., Goffaux, M., Martinez, M. and Oger, R., 2005. Base de données sols de REQUASUD, 2ème synthèse. Technical report, ASBL REQUASUD, Gembloux, Belgium. (in French)

Faber, N.M., Schreutelkamp, F.H. and Vedder, H.W., 2004. Estimation of prediction uncertainty for a multivariate calibration model. *Spectroscopy Europe 16*:17-20.

FAO-ISRIC-ISSS, 1998. World Reference Base for Soil Ressources. *World Soil Resources Report 84*, FAO, Rome, Italy.

Freibauer, A., Rounsevell, M.D.A., Smith, P. and Verhagen, J., 2004. Carbon sequestration in the agricultural soils of Europe. *Geoderma, 122*: 1-23.

Garten, C.T. and Wullschleger, S.D., 1999. Soil carbon inventories under a bioenergy crop (Switchgrass): measurement limitations. *Journal of Environmental Quality, 28*: 1359-1365.

Guo, L.B. and Gifford, R.M., 2002. Soil carbon stocks and land use change: a meta analysis. *Global Change Biology, 8*: 345-360.

Holmes, K.W., Roberts, D.A., Sweeney, S., Numata, I., Matricardi, E., Biggs, T.W., Batista, G. and Chadwick O.A, 2004. Soil databases and the problem of establishing regional biogeochemical trends. *Global Change Biology, 10*: 796-814.

Johnson, D.W. and Curtis, P.S., 2001. Effects of forest management on soil C and N storage: meta analysis. *Forest Ecology and Management, 140*: 227-238.

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R. and Buydens, L.M.C., 2003. The potential of field scpectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta, 484*: 198-200.

Koshoubu, J., Iwata, T. and Minami, S., 2001. Elimination of the uninformative calibration sample subset in the modified UVE (Uninformative Variable Elimination)-PLS(Partial Least Squares) method. *Analytical Sciences, 17*: 319-322.

Lal, R., 2004. Soil carbon sequestration to mitigate climate change. *Geoderma, 123*: 1-22.

Malley, D., Martin, P. and Ben Dor, E., 2004. Application in analysis of soils. In: Craig. R.; Windham R.; Workman J. (Editors), *Near Infrared Spectroscopy in Agriculture. A Three Society Monograph* (ASA, SSSA, CSSA) Madison, WI, pp. 729-784.

Martens, H.A. and Dardenne, P., 1998. Validation and verification of regression in small data sets. *Chemometrics and Intelligent Laboratory Systems, 44*: 99-121.

Martin, P.D., Malley, D.F., Manning, G. and Fuller, L., 2002. Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Canadian Journal of Soil Science, 82*: 413-422.

Marx, B.D. and Eilers, H.C., 2002. Multivariate calibration stability: a comparison of methods. Journal of *Chemometrics, 16*: 129-140.

McCarty, G.W. and Reeves III, J.B., 2001. Development of rapid instrumental methods for measuring soil organic carbon BookTI - Assement methods for soil carbon. In: R. Lal (Editors), Lewis, Boca Raton, pp. 371-380.

McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F. and Kimble, J.M., 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society America Journal, 66*: 640-646.

Mouazen, A.M., Maleki, M.R., De Baerdemaeker, J. and Ramon, H., 2007. On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil and Tillage Research, 93*: 13-27.

Post, W.M., Izaurralde, R.C., Mann, L.K. and Bliss, N., 2001. Monitoring and verifying changes of organic carbon in soil. *Climatic Change, 51*: 73-99.

Reeves III, J.B. and Delwiche, S.R., 2003. SAS partial least squares regression for analysis of spectroscopic data. *Journal of Near Infrared Spectroscopy, 11*: 415-431.

Richter, R., 2005. Atmospheric/Topographic correction for airborne imagery. Technical report, Wesseling, Germany.

Ryan, M.G. and Law, B.E., 2005. Interpreting, measuring, and modeling soil respiration. *Biogeochemistry, 73*: 3-27.

Savitzky, A. and Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*: 1627-1638.

Selige, T., Böhner, J. and Schmidhalter U., 2006. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma, 136*: 235-244.

Smith, P., 2005. An overview of the permanence of soil organic carbon stocks: influence of direct human-induced, indirect and natural effects. *European Journal of Soil Science, 56*: 673-680.

Stevens, A., van Wesemael, B., Vanderschrick, G., Touré, S. and Tychon, B., 2006. Detection of carbon stock change in agricultural soils using spectroscopic techniques. *Soil Science Society America Journal, 70*: 844-850.

The, T.H., 2000. Applying atmospheric correction coefficients for tidal and inland waters with the use of look-up tables. Technical Consultancy Report for VITO, Mol, Belgium.

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

Van-Camp, L., Bujarrabal, B., Gentile, A., Jones, R.J.A., Montanarella, L., Olazabal, C. and Selvaradjou, S., 2004. Reports of the Technical Working Groups Established under the Thematic Strategy for Soil Protection. EUR 21319 EN/3. Office for Official Publications of the European Communities, Luxembourg. pp. 872.

Walkley, A. and Black, I.A., 1934. An estimation of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science, 37*: 29-37.

West, T.O. and Post, W.M., 2002. Soil organic carbon sequestration rates by tillage and crop rotation: a global data analysis. *Soil Science Society of America Journal, 66*: 1930-1946.

Chapter 4

Spectral reflectance based indices for soil organic carbon quantification

Harm Bartholomeus, Michael Schaepman, Lammert Kooistra,
Antoine Stevens, Willem Hoogmoed, Otto Spaargaren

Abstract

We investigated 40 samples from nine different soil types, originating from several climatic zones and a large variety in SOC content (0.06 – 45.1%). Spectral measurements for all soil samples were performed in a controlled laboratory environment. We tested the performance of several spectral indices which have been developed to detect biochemical constituents (e.g., cellulose, lignin) for their ability to retrieve SOC, and compared it to PLS. Good relations were found for indices based on the visible part of the spectrum (R^2 = 0.80) and for the absorption features related to cellulose (around 2100 nm) (R^2 = 0.81). The best index based relations were compared to the results for PLS (R^2 = 0.87). Cross validation was used to evaluate the predictive capacity of the spectral indices. The results demonstrate that it is feasible to use spectral indices derived from laboratory measurements to predict SOC in various soil types. However, a large variance in SOC is required for the calibration of the prediction model, since extrapolation beyond the SOC range in the training dataset results in large errors. PLS proves to be much less sensitive towards extrapolation of the model beyond the mineralogy and SOC levels used during the calibration.

4.1. Introduction

Soil Organic Carbon (SOC) represents a significant fraction of the total amount of carbon involved in the global carbon cycle. However, a quantitative estimation of the exact amount, the spatial distribution and the temporal change is still difficult and requires an intensive sampling strategy (Garten Jr. and Wullschleger, 1999; Smith, 2004). Estimates for the size of the carbon reservoir in mineral soils range from 1115 to 2200 Pg in a metre soil profile (Post et al., 1982; Esweran et al., 1993; Batjes, 1996) and the Intergovernmental Panel on Climate Change (IPCC) estimates the total size on 1750 ± 250Pg (IPCC, 2001). The role and importance of remote sensing for the monitoring of soil carbon has been discussed by Post et al. (2001), who indicated that it can be especially useful when used in regions lacking detailed (in situ) geographical information. Spectroscopy is one of the techniques that can be used for a better estimation of the amount and spatial distribution of SOC and has demonstrated its capability to accurately determine SOC contents in the laboratory (Sudduth and Hummel, 1993; Ben-Dor et al., 1997; Reeves III et al., 2000; Sørensen and Dalsgaard, 2005), directly in the field with a portable spectrometer (Barnes et al., 2003) or from airborne hyperspectral sensors (Ben-Dor et al., 2002). Imaging spectrometers offer the possibility to resolve the soil data problem that regional studies face, due to its advantages like speed and efficiency (McBratney et al., 2006), and contiguous spectral coverage. Estimation of SOC based on remote sensing techniques is mostly based on the overall decrease in reflectance in the Visible and Near-Infrared part of the spectrum (Irons et al., 1989; Chen et al., 2000; Kooistra et al., 2003). In recent years, multivariate techniques like Partial Least Squares (PLS) regression have been commonly used to build SOC prediction models (Fidencio et al., 2002b; Udelhoven et al., 2003; Madari et al., 2006; Stevens et al., 2006; Viscarra Rossel et al., 2006). PLS is easy to use and often used in different fields of application (Martens and Naes, 1984). However, an important drawback is the complexity of the transfer of prediction models from one sensor to another. Sensor characteristics like wavelength position, bandwidth or number of bands, often differ from one remote sensing instrument to another, which requires new model calibrations for each sensor. Furthermore, it remains a question whether PLS based models are predicting SOC based upon the absorptions of organic matter and carbonates, or whether they are built indirectly on SOC correlations with other soil constituents like iron oxides and clay minerals (Brown et al., 2005).

Considering the difficulties to transfer PLS models from one sensor to another, we discuss the development of SOC prediction models based on spectral indices related to the biochemical composition of SOC. Such models have the advantage that the transferability of spectral indices to various sensors is less complicated than for models based on multivariate techniques. Spectral indices related to organic carbon have – to our best knowledge – not been presented, although organic carbon contains biochemical constituents for which

absorption features are known. PLS will be used as reference technique to which the performance of the in this study developed indices is compared.

The biochemical composition of SOC depends on the source material and its decomposition stage (Ben-Dor et al., 1997). Biochemical constituents present in the vegetation can also be present within the soil, because SOC includes (next to microbes and animals) living biomass from plants and vegetation remnants. Soil organic matter therefore contains biochemical constituents like chlorophyll, oil, cellulose, pectin, starch, lignin and humic acids (Beyer et al., 2001), which influence the reflectance in the Visible (VIS, 400- 700 nm), Near-Infrared (NIR, 700-1400 nm) and Short Wave Infrared (SWIR, 1400- 3000 nm) region (Ben-Dor et al., 1997) of the electromagnetic spectrum. Many of these biochemical constituents decompose fast (e.g. chlorophyll), others have a longer turnover time (e.g. cellulose, lignin and starch) but low levels of these elements can be found in the soil, primarily in the upper horizons (Beyer et al., 2001). Elvidge (1990) described the spectral response of individual biochemical components in dry plant materials, and currently significant advances have been made in assessing the plant pigment (Ustin et al., 2006) and non-pigment system from space (Kokaly et al., 2006). Starch, cellulose and several other biochemical constituents show absorption features in the SWIR region (2000-2500 nm), which are often close to each other and show significant overlap (Curran, 1989). Some wavelengths have also been related to amount of SOC in the soil (Dalal and Henry, 1986; Morra et al., 1991; Sudduth and Hummel, 1991; Henderson et al., 1992; Shepherd and Walsh, 2002) and it was proven that it is possible to explain the variation in soil cellulose concentrations by means of Near Infrared Spectroscopy (Hartmann and Appel, 2006). Published prediction models are often derived for datasets with a limited variance in SOC concentration and a limited number of soil types (Kooistra et al., 2003). Positive results for an effective calibration with a large number of heterogeneous samples were achieved by Kemper et al. (2005), Fidencio et al. (2002a) and Brown et al. (2006). A widely applicable SOC prediction model should be based on a wide range of SOC-values and soils with different mineralogical background, since soil mineralogy is one of the main factors causing differences in soil reflectance (Hartmann and Appel, 2006). Although a greater variability in the training phase of a statistical model may lead to an improved robustness of the model and an increased ability to characterize a diverse range of samples, it may also result in a decrease in prediction accuracy (McCarty et al., 2002).

In this paper we investigate the relation between spectral indices based on absorption features of biochemical constituents and the SOC content. We show that the amount of SOC can be detected with reflectance spectroscopy based on cellulose, starch and lignin influenced wavelengths. The relation is investigated for a small dataset that contains samples from nine soil types from different geographic locations and represents a large variance in SOC content, mineralogical background and grain size. Finally, we discuss the

accuracy and robustness of the spectral reflectance based SOC prediction models in comparison with PLS, and the opportunities for application in remote sensing studies.

4.2. Methodology

A selection of 40 soil samples, originating from a) ISRIC's World Soil Reference Collection (ISRIC, 2006) b) the CARBIS field campaign (Stevens et al., 2008) and c) the MIES field campaign (Van der Molen, 2005) were used for analysis. The samples represent nine soil types which originate from different source materials, climatic zones and show a large variance in SOC content (0.06 to 45.1 %). The samples originate from surface as well as sub-surface horizons (Table 4.1). All soil samples were sieved over a 200 μm sieve, keeping the smaller fraction, and analysed for SOC content using the Walkley Black method (Walkley and Black, 1934). The samples were divided into two sets of 20 samples, which are both used for calibration as well as for validation. The sets were constructed in such a way that both sets contain an equal distribution of samples of all nine soil types and practically the entire range of used SOC-values is included. The SOC-values in both sets show a non-normal distribution which is strongly skewed towards the lower SOC-values, which corresponds to the worldwide SOC distribution function. To allow regression analysis and calculation of R^2-values it was converted to a normal distribution by calculating the double square-root of the SOC-content (Osborne and Waters, 2002). This will be referred to as $SOC^{\frac{1}{4}}$. An overview of used samples, the different soil types, the SOC and $SOC^{\frac{1}{4}}$ content and the grain size distribution is given in Table 4.1.

Spectral measurements were performed on soil samples that were air-dried at room temperature (20-25°C) to reduce the effect of soil-moisture on reflectance (Ben-Dor et al., 1999), using an ASD Fieldspec Pro FR in laboratory setup (Analytical Spectral Devices, Inc.). The ISRIC samples were measured with a viewing angle of 30° of nadir, and a 1° aperture angle at a distance of 40 cm. For the measurement a collimated tungsten quartz halogen lamp (Lowel pro-lamp, Lowel-Light Manufacturer Inc., New Your, NY) was used for illumination at a distance of 55 cm in nadir position, resulting in the final measurements being measured under bi-conical conditions (Schaepman-Strub et al., 2006). This setup results in a ground projected field of view (FOV) of approximately 0.38 cm^2. The chemical analysis of the ISRIC samples was carried out in 1993 and 1997. The spectral measurements were made in 2003. It was assumed that the concentration of organic carbon had not significantly changed during the storage period. The reflectance of the CARBIS and MIES samples was measured with an ASD Contact Probe, which results in a FOV of 3.14 cm^2. Also the contact probe uses a collimated illumination source resulting in the same geometrical-optical measurement properties as above. Even though the probe has a smaller illumination angle (12°), the sieving at the given grain size, minimized the potential directional effects

Table 4.1: Overview of selected soil samples: Soil unit = type of soil according to FAO or Local classification; SOC = Soil Organic Carbon Content in %; SOC$^{1/4}$ = the SOC to the power 0.25; Setup = measurement method with either the ASD and lamp or the ASD with the Contact Probe (ASD + CP).

Set 1								
Soil type	SOC	SOC$^{1/4}$	Depth	Sand	Silt	Clay	Setup	Country
Andosol[1]	0.06	0.49	90-135	95.8	2.6		Lamp	Costa Rica
Andosol[1]	1.55	1.12	25-80	67.8	27.3	4.9	Lamp	Costa Rica
Andosol[1]	6.42	1.59	0-25	58.8	28.3	12.8	Lamp	Costa Rica
Cambisol[1]	0.21	0.68	50-75	12.8	34.7	52.5	Lamp	Nicaragua
Cambisol[1]	0.99	1.00	5-40	9.7	19.3	71.1	Lamp	Nicaragua
Chernozem[1]	0.22	0.68	100-120	82.2	14.5	3.3	Lamp	Hungary
Chernozem[1]	1.42	1.09	25-40	66	21.8	12.4	Lamp	Hungary
Fluvisol[1]	0.17	0.64	140-160	74	14.8	11.2	Lamp	Nicaragua
Histosol[1]	22.90	2.19	0-5				CP	The Netherlands
Histosol[1]	45.10	2.59	0-5				CP	The Netherlands
Luvisol[1]	0.75	0.93	15-40	15.4	33.6	51.2	Lamp	Nicaragua
Luvisol[1]	2.98	1.31	0-15	7.1	33.5	59.4	Lamp	Nicaragua
Sand on Peat[2]	4.00	1.41	0-5				CP	The Netherlands
Sand on Peat[2]	10.70	1.81	0-5				CP	The Netherlands
Sand on Peat[2]	16.00	2.00	0-5				CP	The Netherlands
Sandy Loam[2]	0.92	0.98	0-5				CP	Belgium
Sandy Loam[2]	1.54	1.11	0-5				CP	Belgium
Sandy Loam[2]	2.21	1.22	0-5				CP	Belgium
Ferralsol[1]	0.94	0.98	15-20	6.8	29.4	63.8	Lamp	Costa Rica
Ferralsol[1]	3.96	1.41	0-8	8.7	37	54.3	Lamp	Costa Rica
Average	6.15	1.26						
STDEV	10.93	0.55						
Skewness	2.82	0.91						

[1]Soil Unit according to FAO Classification, [2]Soil Unit according to description of the analyzed surface sample, no entire profile description done.

Table 4.1 - continuous on next page.

present under these conditions. The chemical analysis of these soil samples was carried out shortly after sampling and spectral measurements.

The PLS regression method (Geladi and Kowalski, 1986) was used to establish relations between the spectra and SOC$^{1/4}$ content. For the PLS regression, reflectance spectra were converted to absorbance (log 1/Reflectance) before the models were developed. Since PLS is a widely adopted method in laboratory studies it was taken as the reference analysis, to which the performance of the spectral indices developed within this study is compared.

Table 4.1 - Continued

Set 2

Soil type	SOC	SOC[¼]	Depth	Sand	Silt	Clay	Setup	Country
Andosol[1]	0.20	0.67	100-115	89.2	8.1	2.7	Lamp	Costa Rica
Andosol[1]	1.76	1.15	20-70	71.3	28.8		Lamp	Costa Rica
Andosol[1]	3.99	1.41	7-20	63.1	35.8		Lamp	Costa Rica
Cambisol[1]	0.57	0.87	40-50	13.5	23.8	62.7	Lamp	Nicaragua
Cambisol[1]	2.88	1.30	0-5	12.1	22.9	65.0	Lamp	Nicaragua
Chernozem[1]	0.30	0.74	70-90	77.7	17.6	4.7	Lamp	Hungary
Chernozem[1]	2.06	1.20	0-25	61.3	24.5	14.1	Lamp	Hungary
Fluvisol[1]	0.32	0.75	60-70	61.3	23	15.5	Lamp	Nicaragua
Histosol[1]	23.80	2.21	0-5				CP	The Netherlands
Histosol[1]	40.20	2.52	0-5				CP	The Netherlands
Luvisol[1]	1.08	1.02	25-43	4.8	28.1	67.1	Lamp	Nicaragua
Luvisol[1]	2.21	1.22	15-33	5.7	33.7	60.5	Lamp	Nicaragua
Sand on Peat[2]	4.00	1.41	0-5				CP	The Netherlands
Sand on Peat[2]	10.70	1.81	0-5				CP	The Netherlands
Sand on Peat[2]	16.10	2.00	0-5				CP	The Netherlands
Sandy Loam[2]	0.95	0.99	0-5				CP	Belgium
Sandy Loam[2]	1.54	1.11	0-5				CP	Belgium
Sandy Loam[2]	2.21	1.22	0-5				CP	Belgium
Ferralsol[1]	0.73	0.92	23-110	7.5	27.1	65.5	Lamp	Costa Rica
Ferralsol[1]	3.84	1.40	0-15	6.4	24.3	69.3	Lamp	Costa Rica
Average	5.97	1.30						
STDEV	10.05	0.50						
Skewness	2.61	1.10						

[1]Soil Unit according to FAO Classification, [2]Soil Unit according to description of the analyzed surface sample, no entire profile description done.

As a first step, the direct relation between reflectance values and SOC[¼] content was investigated for each wavelength individually. Next, spectral indices based on the summed reflectance in the VIS and the slope for several ranges in the VIS were calculated. In addition, we investigated the usability of spectral indices based on wavelengths influenced by SOC related biochemical constituents. For this the reflectance from 1600-1800 nm (related to lignin) and 2000-2300 nm (related to cellulose, starch and lignin) was studied. First, the

general reflectance pattern was investigated visually for presence of absorption features that vary in location, shape or size. Next, based on the reflectance pattern, spectral indices were developed. The first type of index uses the area of the absorption feature, which was defined as the sum of the total reflectance minus the continuum removed function (Grove et al., 1992). The continuum removed function is usually defined as a convex hull, but can also be defined as a straight line (often named baseline) if a single absorption feature is investigated (Kokaly and Clark, 1999). The border values for the baseline were systematically varied in order to find an optimal relation (highest R^2-value) with SOC[14]. For each individual wavelength the difference between the baseline and the measured value was calculated and those differences are summed. This sum is a measure for the size of the absorption feature and expresses the amount of absorption. The second type of index uses the slope of the spectral signature corresponding to the higher wavelength part of the absorption feature. In this case, a straight line was plotted between the reflectance value at the start of the absorption feature and the reflectance value at the centre of the absorption feature, where the relation of the slope (change in reflectance per nanometer) of this line and the SOC[14] was studied. Although the start and the position of the centre of an absorption feature can vary with the amount of absorption, and may therefore vary per sample, we chose to use the same start and centre of the absorption feature for all samples. Absorption features of pure biochemical constituents may overlap in the SWIR, making the assignment of the different components to a specific absorption feature difficult (Elvidge, 1990). No actual quantification of biochemical constituents in the soil samples by means of chemical analysis was available. Only wavelength regions influenced by cellulose, starch and lignin, as described in literature, were studied.

Both linear and curvilinear regression functions were tested to relate the indices with SOC[14]. The quality of the fit was assessed using the R^2-value with a confidence level of 0.95 and results were cross-validated by applying the found relations for Set 1 on Set 2 and vice versa. This means that both sets are used as calibration and validation set. Quality of the SOC content prediction is expressed in Standard Error of Calibration (SEC) and Standard Error of Performance (SEP), which are both given in percent SOC. The SEC is related to the Standard Deviation (STDEV) of the training set and the SEP is related to the STDEV of the validation set. The ratio of STDEV in the validation set to the SEC or SEP is referred to as the Ratio of Performance to Deviation (RPD) and can be used as an index of model accuracy. Chang et al. (2001) defined three classes of RPD: category A (RPD>2) are models that can predict accurately the property in question, category B (RPD = 1.4 – 2) is an intermediate class which regroups models that can be possibly improved, and models falling in category C (RPD < 1.4) have no prediction capability.

To test the robustness of the prediction model to variation in soil type, both sets were combined and the calibration was repeated, each time leaving all samples of one soil type out. The soil type which was left out of the calibration phase was used for validation, to test whether the prediction model is applicable on soil types which are not used for calibration of the model. Robustness was in this study defined as how sensitive the model reacts to inclusion of soil types with a mineralogy that was not used during calibration.

4.3. Results and Discussion

For individual wavelengths, the highest correlation between $SOC^{1/4}$ and reflectance is found in the VIS, with a maximum around 600 nm (Figure 4.1). It is well know that higher levels of organic material in the soil lead to darker soils, and therefore a lower reflectance. Scatter plots, with the $SOC^{1/4}$ plotted against the reflectance for individual wavelengths (figures not shown), expose that the relation between $SOC^{1/4}$ and reflectance is not linear. Both linear and curvilinear relations are reported in literature but whether a relation is linear depends on the range of SOC-values considered (Schreier, 1977; Baumgardner et al., 1985). In our case the relation can be made linear by calculating the inverse of the reflectance (1/reflectance), which reveals highest correlation between 640 and 690 nm ($R^2_{set\ 1} = 0.74$; $R^2_{set\ 2} = 0.76$). This spectral region corresponds with the results reported by Vinogradov (1982).
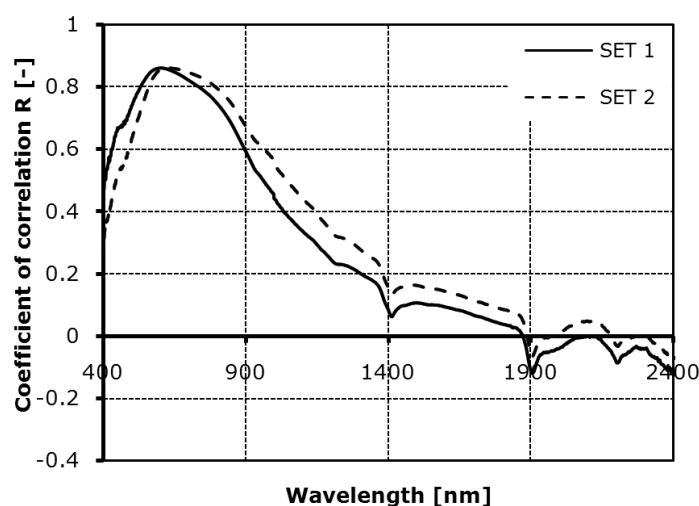

Figure 4.1: Coefficient of correlation (R) between $SOC^{1/4}$ and 1/Reflectance.

4.3.1 Partial Least Squares Regression

For the PLS analysis the wavelengths contributing most to the prediction model are found in the VIS (maximum at 650 nm), around the water absorption features (1400 nm and 1900 nm), and in the SWIR region (2212 nm). The optimal balance between model complexity and

predictive power was found for 2 latent variables. Calibrations developed using PLS resulted in high R^2 values ($R^2_{set\ 1}$ = 0.87; $R^2_{set\ 2}$ = 0.76) and a low SEC and SEP (Table 4.2). For set 1 the SEP was calculated two times, and both values are given in table 4.2. The first value represents the SEP for the entire set of 20 samples. Because this implies an extrapolation outside the minimum and maximum SOC$^{\text{¼}}$ range of Set 2, the SEP was calculated also for the interpolated range of SOC$^{\text{¼}}$ content only. This decreases the SEP from 4.21 to 2.41, but even without exclusion of this sample the PLS reaches a RPD > 2 for both datasets, resulting in a category A model. Research by Kooistra et al. (2001) showed that pre-processing on the spectral data can have a large influence on the prediction accuracy of PLS, but the effect it has depends on the property that is examined and the composition of the dataset.

4.3.2 Indices in the VIS wavelengths

The reflectance in the VIS and NIR of three samples with a large range in SOC content is shown in Figure 4.2. Besides the overall decrease in reflectance with an increase in SOC, also the shape of the spectral signature varies, resulting in a change from convex to concave when the SOC content decreases, and a flatter spectral profile at higher SOC contents.
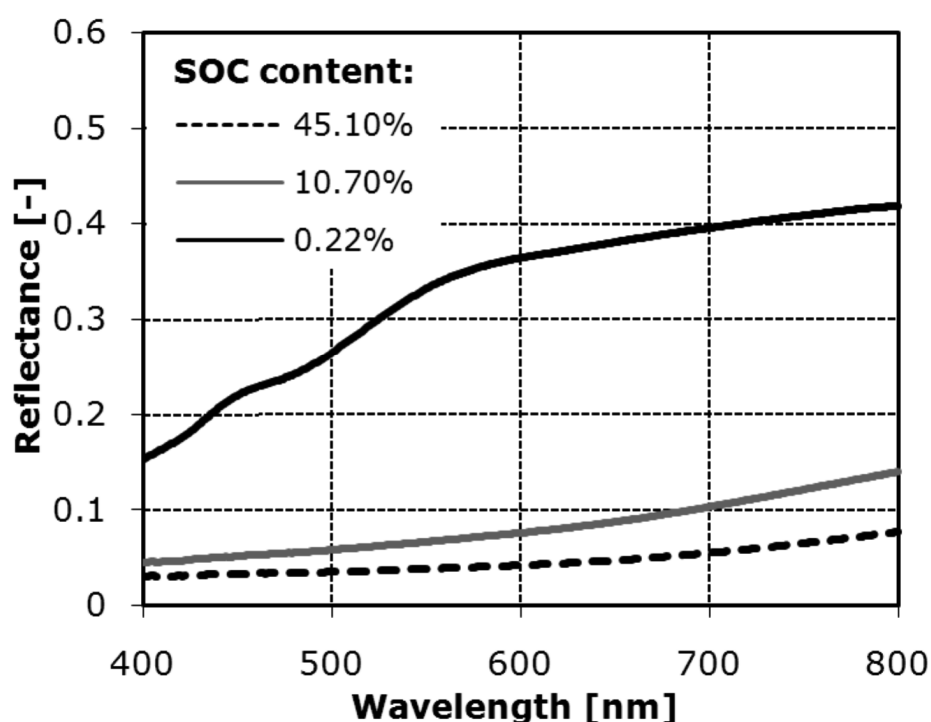

Figure 4.2: Reflectance in the Visible and Near-Infrared part of the spectrum for three samples with a large range in SOC content.
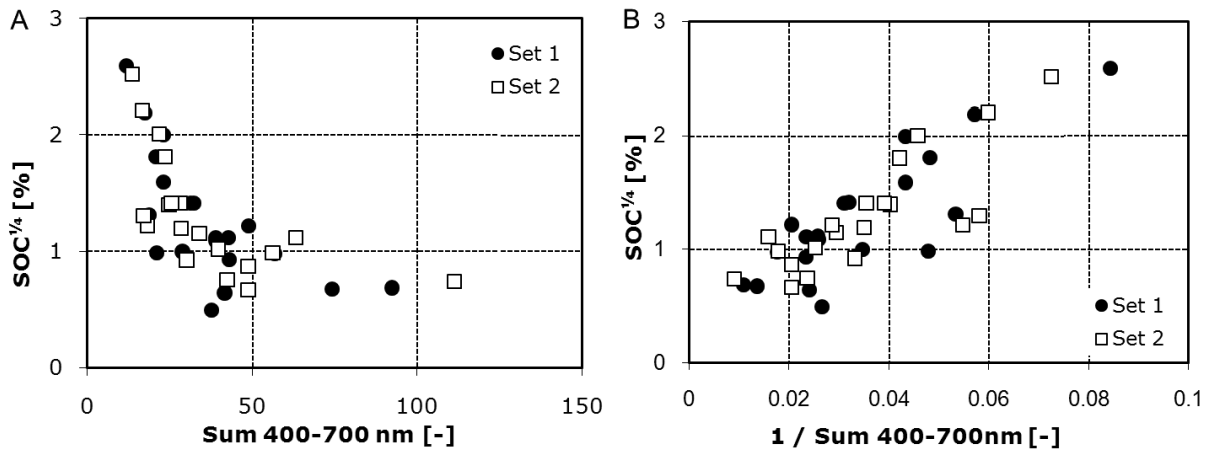
Figure 4.3: SOC$^{1/4}$ content plotted against 'sum 400-700' (graph A) and the SOC$^{1/4}$ content plotted against '1/ sum 400-700' (graph B).

Based on the VIS wavelength region, two indices were calculated. The first index is based on the overall decrease in reflectance between 400 and 700 nm. Figure 4.3A shows that the relation between the 'Sum 400-700 nm' and SOC$^{1/4}$ can be considered linear by dividing the SOC$^{1/4}$ range into two parts: from 0 - 1.5 and >1.5. These values correspond to the SOC range from 0 - 5 % and >5 % respectively. A combination of samples with high and low SOC content results in a non-linear relation but the relation can be made linear by calculating '1/Sum 400-700 nm'. Figure 4.3B shows that this results in a moderately high correlation coefficient ($R^2_{set\ 1}$ = 0.69; $R^2_{set\ 2}$ = 0.68).

The slope in the VIS varies with the SOC$^{1/4}$ content (Figure 4.4) and the correlation of this slope with SOC$^{1/4}$ was tested for several spectral ranges (400-700 nm, 400-600 nm, 500-700 nm, 500-600 nm), of which '1/Slope 400 - 600 nm' appeared to correlate best ($R^2_{set\ 1}$ = 0.71; $R^2_{set\ 2}$ = 0.80). Validation shows that this results in a category A model if the sample with the highest SOC content of set 1 is excluded, to prevent extrapolation of the model beyond the range for which it was calibrated. Ben-Dor et al. (1997) showed that this slope is influenced by the decomposition stage of the organic material, which especially effects the surface samples that may contain a reasonable amount of fresh organic material. The datasets contain fresh samples as well as soil samples that were stored for a longer period before spectral measurements were done, but these samples do not show any structural deviation from the fresh samples. Another complicating factor for the use of these indices is the strong influence of iron oxides in this spectral range (Ben-Dor et al., 1999), which can be used for quantification of dithionite extractable iron (Bartholomeus et al., 2007). A deviation in prediction accuracy can therefore be expected in case of large variations in iron oxide concentrations.
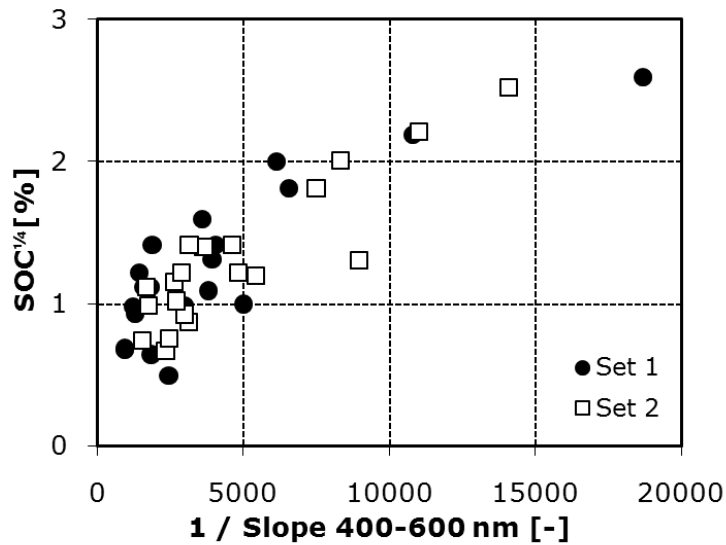
Figure 4.4: SOC$^{1/4}$ content plotted against '1/Slope 400-600 nm'.

4.3.3 NIR and SWIR wavelengths

In the spectral range from 1600 to 1800 nm the overall level of reflectance varies, but there is hardly any variance in the pattern of the reflectance curve (Figure 4.5). Therefore, it is not possible to derive an index based on this spectral region, which in vegetation studies is used as an indication for lignin content (Serrano et al., 2002). A slight variance in slope between 1600 and 1800 nm is visible, but this does not show a relationship with SOC content (R^2 = 0.09). Dalal and Henry (1986) reported the reflectance at 1744 nm from Australian soils to be sensitive to organic matter, but this could not be supported by our dataset (R^2 = 0.1).
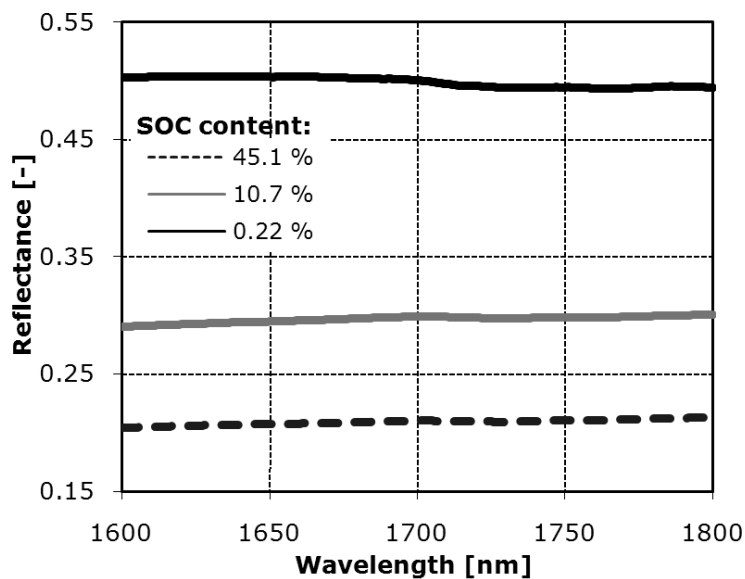

Figure 4.5: Spectral signature (1600 – 1800 nm) of three soils with a large range in SOC content.

Figure 4.6 depicts the spectral response from 2000 to 2300 nm. Besides the decrease in overall reflectance, the reflectance pattern also varies; the spectral profile flattens when the SOC[¼] content increases. In contrast to vegetation, where the presence of starch and cellulose results in a clear absorption feature at 2100 nm (Curran, 1989; Elvidge, 1990), the presence of SOC[¼] does not lead to a clear absorption feature in soil reflectance spectra. Although the reflectance does not show a dip as such the flattening of the spectra can be interpreted and analysed as an absorption feature.
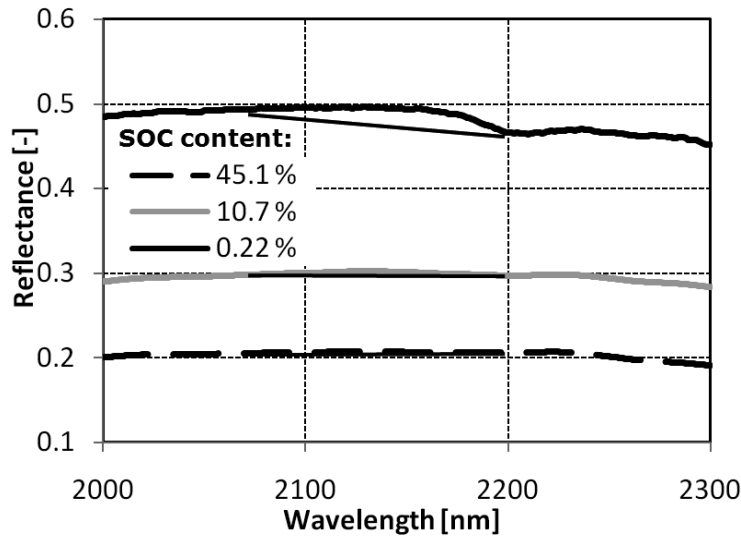

Figure 4.6: Spectral signature (2000 - 2300 nm) of three soils with a large range in SOC content. This spectral region is influenced by the presence of several biochemical constituents, e.g. cellulose, lignin and starch. The straight thin black line between 2050 and 2200 indicates the baseline to which the spectra were normalized. The area in this spectral range is a sum of the differences between the reflectance value and the baseline, for all wavelengths.

The changes of the spectral profile in the SWIR are quantified in two ways. First, the total decrease in reflectance, compared to the continuum removed values was calculated, which yielded the best results between 2050 and 2200 nm. This results in a value that describes the area of the absorption feature, and shows a negative relation with SOC[¼] (Figure 4.7A). More SOC leads to a higher amount of cellulose, lignin and starch, which results in more absorption and a flatter spectral signature, shown by a lower value for the area of the absorption feature. To create a linear relation, again the inverse of the area is taken (Figure 4.7B) and correlated with SOC[¼] ($R^2_{set\ 1}$ = 0.55; $R^2_{set\ 2}$ = 0.81). According to the RPD this index will be classified as a category B model.

As a second index, the slope of the spectral signature is calculated, which yielded the best results for '1/Slope 2138-2209 nm' ($R^2_{set\ 1}$ = 0.49; $R^2_{set\ 2}$ = 0.79) (Figure 4.8). These two wavelengths correspond to the left border and centre wavelength of the absorption dip around 2200 nm, which is caused by the clay lattice absorption feature (Ben-Dor et al., 1999).
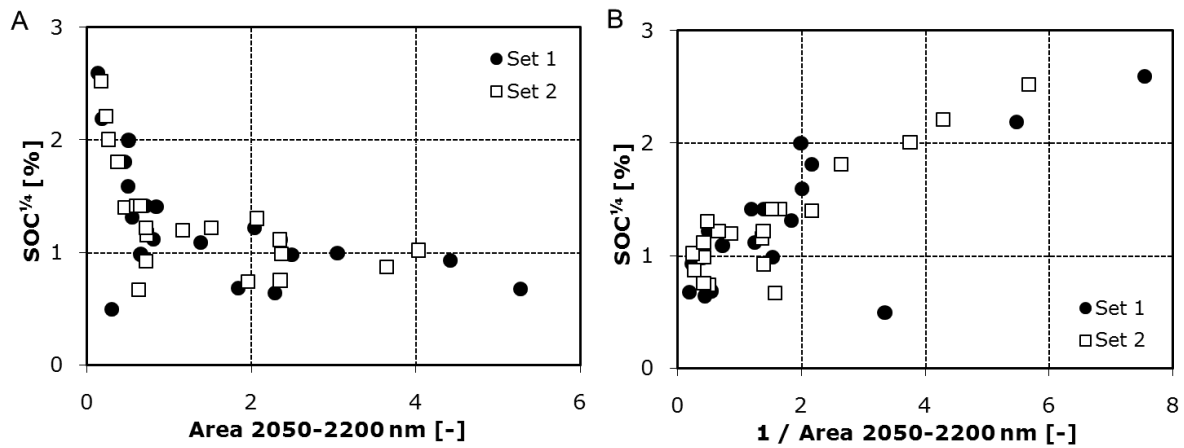
Figure 4.7: SOC$^{1/4}$ content plotted against the 'area 2050-2200 nm' (A) and the SOC$^{1/4}$ content plotted against '1/Area 2050-2200 nm' (B).

The good correlation is therefore probably not caused by absorption due to biochemical constituents as such, but by the known fact that higher amounts of SOC mask other spectral absorption features (Stoner and Baumgardner, 1980), in this case clay. Although soil samples with high clay content (up to 71%) were included in the dataset, the deviation of the predicted SOC-content shows no relation with the clay content. There is no correlation between the error in predicted SOC and clay content ($R^2_{set\ 1}$ = 0.07 and $R^2_{set\ 2}$ = 0.002)

For all indices the low R^2-values for Set 1 are caused by an Andosol sub surface sample with a low SOC content, which represents the volcanic source material on which hardly any soil forming processes took place. In general these volcanic deposits show a low reflectance and
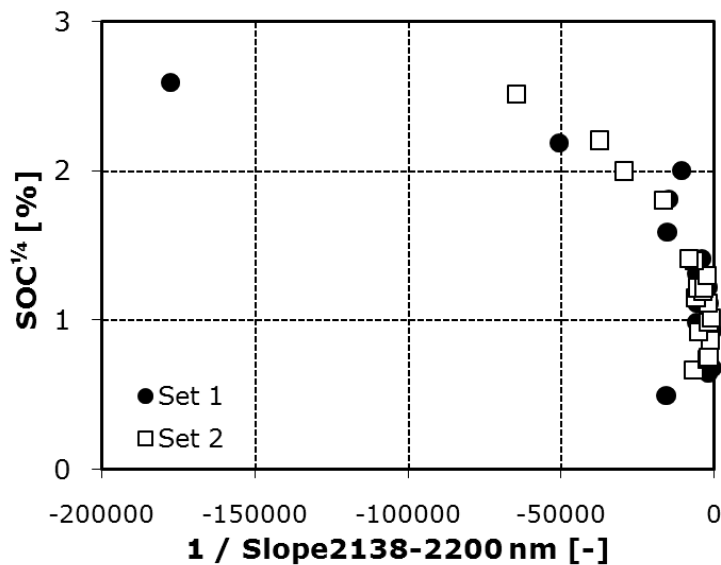

Figure 4.8: SOC$^{1/4}$ content plotted against '1/ slope 2138-2200 nm'

a feature-poor spectral signature. This confirms that mineralogical background has a large influence, which for some soils results in a less robust SOC prediction model.

The performance indicators for the evaluated relations between $SOC^{1/4}$ and spectral reflectance are summarized in Table 4.2. Because relations based on direct correlations are curvilinear, only the results of inverse relations are shown. PLS shows the overall best results, and also meets the requirements of a category A model. For the spectral indices the biochemical constituent absorption dip based index '1/Area 2050-2200 nm' ($R^2_{set\,2}$ = 0.81) and the visible wavelengths based index '1/Slope 400-600 nm' ($R^2_{set\,2}$ = 0.80), show the overall highest R^2-values. The SEC values for '1 /Area 2050-2200' are better, but SEP and RPD values are best for '1/Slope 400-600' and the RPD of this index model ($RPD_{set\,1}$ = 3.18 ; $RPD_{set\,2}$ = 2.72) is comparable to the prediction accuracy achieved with PLS. The only index that is not sensitive to extrapolation of $SOC^{1/4}$ values beyond the $SOC^{1/4}$ range used for training is '1/Sum 400-700', while '1/Slope 2138-2209 nm' appears to be very sensitive to extrapolation. These results show that the SWIR spectral region can be used for the quantification of $SOC^{1/4}$, as was also reported by e.g. Morra et al. (1991) and Henderson et al. (1992), but the predictive power is less than for PLS and VIS based indices.

Table 4.2: Performance of spectral indices towards the prediction of SOC. R^2-values are based on linear relations with $SOC^{1/4}$.

	R^2-value		SEC (%)		RPD (SEC)		SEP (%)		RPD (SEP)	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set2	Set 1	Set 2
PLS	0.87	0.76	3.18	5.23	3.43	1.92	4.07/2.41 [a]	4.84	2.68 / 4.53 [a]	2.08
1 / Sum 400-700	0.69	0.68	4.27	5.78	2.56	1.74	4.40 / 4.34 [a]	5.64	2.48 / 2.52 [a]	1.78
1 / Slope 400-600	0.71	0.80	6.05	2.92	1.81	3.44	10.61 / 3.43 [a]	3.70	1.03 / 3.18 [a]	2.72
1 / Area 2050-2200	0.55	0.81	4.00	1.63	2.73	6.17	11.73 / 5.29 [a]	5.74	0.93 / 2.06 [a]	1.75
1 / Slope 2138-2209	0.49	0.79	6.67	5.23	1.64	1.92	265.62 / 4.42 [a]	8.92	0.04 / 2.47 [a]	1.13
STDEV of SOC (%)			10.92	10.05			10.92	10.05		

[a] For Set 1 the SEP was calculated two times. The first value represents the SEP for the entire set of 20 samples. Because this implies an extrapolation outside the minimum and maximum $SOC^{1/4}$ range of Set 2, the SEP was calculated also for the interpolated range of $SOC^{1/4}$ content only. The second number represents this non-extrapolated SEP value.

4.3.4 Robustness of the models

The robustness of the models was tested in a separate procedure. The two data sets were combined and all samples belonging to single soil type were excluded from the training phase. Next, the SOC-content of the excluded soil type was predicted, and the difference between the measured and predicted SOC-content was calculated, which was repeated until all soil types were predicted. The SEP and RPD for each soil type was calculated ($RPD_{per\,soiltype}$ = $SEP_{per\,soiltype}$ / $SD_{per\,soiltype}$) to demonstrate how well the SOC content can be estimated, when there

is just a single soil type of interest. The low RPD per soil type, which reaches the requirements for a category A model only accidentally (Table 4.3) show that none of the models can be used for prediction of the SOC within a single soil type or a region with little variation in SOC. To express the potential of each model to estimate the SOC for a dataset that contains several soil types with a wide variety in SOC the RPD for all soils was calculated ($RPD_{all\ soils} = SEP_{all\ soils} / SD_{all\ soils}$). In general, PLS is more robust than the prediction models based on spectral indices (Table 4.3). Of the spectral indices, '1/Area 2050-2200' is most robust towards the removal of soil types. Using the model for prediction beyond the maximum SOC-content range used for calibration leads to extremely poor prediction results, which is shown by the large SEP for the Histosol samples (Table 4.3). This shows that a widely applicable prediction model based on laboratory measured reflectance spectra should contain a wide range of SOC values. It should be noted that the number of samples used in this exploratory study is too small to develop an operational SOC prediction model that can be applied globally. Such an operational model requires a larger number of samples and soil types in the training phase. The results obtained for the robustness of the presented models show that the presence of the specific soil type or mineralogy in the training set, determines the predictive capacity for the soil type in question.

Table 4.3: Results of the predictions of SOC values for excluded soils.

	STDEV	1 / sum 400-700		1 / slope 400-600		1 / area 2050-2200		1 / slope 2138-2200		PLS	
		SEP	RPD	SEP	RPD	SEP	RPD	SEP	RPD	SEP	RPD
Ferralsol	1.77	3.40	0.52	1.86	0.95	1.52	1.17	1.60	1.11	2.11	0.84
Luvisol	1.03	6.16	0.17	0.60	1.70	0.18	5.81	0.72	1.42	0.52	1.98
Sandy Loam	0.57	0.67	0.45	0.36	0.51	0.49	0.54	0.54	0.88	0.52	1.11
Sand on Peat	5.4	9.28	0.58	8.34	0.65	8.99	0.60	11.26	0.48	5.58	0.97
Histosol	11.33	19.95	0.57	24.38	0.46	10.78	1.05	1873.42	0.01	5.74	1.98
Fluvisol	0.11	0.62	0.17	0.66	0.16	0.55	0.19	1.14	0.09	0.07	1.45
Chernozem	0.89	0.34	2.62	1.13	0.79	0.66	1.35	1.18	0.76	0.49	1.84
Cambisol	1.19	5.38	0.22	4.78	0.25	1.00	1.19	1.17	1.01	0.38	3.16
Andosol	2.45	1.14	2.15	2.65	0.93	5.41	0.45	2.64	0.93	0.84	2.91
All soils	10.36	7.52	1.38	8.37	1.24	4.81	2.15	592.44	0.02	2.72	3.82

The results show that the prediction of SOC with spectroscopy using a highly variable dataset, with not only a wide range in SOC, but also a large variance in mineralogical background is possible with a reasonable accuracy. This gives a good perspective towards the application of these indices for remote sensing, development of on-the-go soil sensors or non-destructive laboratory analysis. PLS yields the overall best results and is most robust when it comes to predicting SOC of soil types that were not included in the calibration. However, in the search for spectral indices it was shown that '1 /Slope 400-600' is the best

spectral index to use when predictions are made for a soil that is also used in the calibration phase. It should be noted that in terms of absolute determination of SOC content the results are less accurate than reflectance based prediction models for a limited number of soil types with less variance in mineralogical background (Henderson et al., 1992; Daniel et al., 2004). The feasibility to apply the derived relations to remotely sensed images to assess SOC on a landscape or regional scale will mainly depend on the availability and quality of images. The indices that use baseline and area calculations require hyperspectral images, which are only sparsely available. The slope based indices could already be applied on remote sensing data from sensors with lower spectral resolution (e.g. ASTER). It has to be noticed that the scaling from laboratory measurements to image data is a large step, since atmospheric influences, lower signal to noise ratios and disturbing factors of the surface itself complicate the analysis.

4.4. Conclusion

The results of this study show that biochemistry based spectral indices can be used to estimate SOC for a dataset composed of nine soil types with a large variance in mineralogical composition. Prediction accuracies based on spectral indices are comparable to accuracies of PLS based predictions. However, the models based on indices become inaccurate when the SOC content is estimated for samples which have a SOC content larger than the SOC range of the calibration set, or when the models are applied to soil types that were not used for model calibration. We conclude that for the development of a spectral reflectance based SOC prediction model a larger variety of soil types, covering all SOC ranges that can occur, may be included in the training phase as well. Arriving at a worldwide operational SOC prediction model, a larger number of samples is required. However, including a-priori knowledge on expected soil associations, mineral composition pre-classification may increase the robustness of the prediction model and its applicability for extended geographical areas.

References

Barnes, E.W., Sudduth, K.A., Hummel, J.W., Lesch, S.M., Corwin, D.L., Yang, C., Daughtry, C.S.T. and Bausch, W.C., 2003. Remote- and ground-based sensor techniques to map soil properties. *Photogramm. Eng. Remote Sens., 69*: 619-630.

Bartholomeus, H.M., Epema, G.F. and Schaepman, M.E., 2007. Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy. I*nternational Journal of Applied Earth Observation and Geoinformation, 9*: 194- 203.

Batjes, N.H., 1996. Total carbon and nitrogen of the soils of the world. *European Journal of Soil Science, 47*: 151-163.

Baumgardner, M.F., Silva, L.F., Biehl, L.L. and Stoner, E.R., 1985. Reflectance properties of soils. *Adv. Agron., 38*: 1-44.

Ben-Dor, E., Inbar, Y. and Chen, Y., 1997. The Reflectance Spectra of Organic matter in the Visible Near-Infrared and Short Wave Infrared Region (400-2500) during a Controlled Decomposition Process. *Remote Sensing of Environment, 61*: 1-15.

Ben-Dor, E., Irons, J.R. and Epema, G.F., 1999. Soil Reflectance. Remote Sensing for the Earth Sciences, *Manual of Remote Sensing*. Wiley & Sons Inc. , New-York, 111-188.

Ben-Dor, E., Patkin, K., Banin, A. and Karnieli, A., 2002. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data - a case study over clayey soils in Israel. *International Journal of Remote Sensing, 23*: 1043-1062.

Beyer, L., Kahle, P., Kretschmer, H. and Wu, Q., 2001. Soil organic matter composition of man-impacted urban sites in North Germany. *J. Plant Nutr. Soil Sci., 164*: 359-364.

Brown, D.J., Bricklemyer, R.S. and Miller, P.R., 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma, 129*: 251-267.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M. and Reinsch, T.G., 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma, 132*: 273-290.

Chang, C.W., Laird, D.A., Mausbach, M.J. and Hurburgh Jr., C.R., 2001. Near-infrared reflectance spectroscopy - Principal components regression analysis of soil properties. *Soil Science Society America Journal, 65*: 480-490.

Chen, F., Kissel, D.E., West, L.T. and Adkins, W., 2000. Field-Scale Mapping of Surface Soil Organic Carbon Using Remotely Sensed Imagery. *Soil Science Society of America Journal, 64*: 746-753.

Curran, P.J., 1989. Remote Sensing of Foliar Chemistry. *Remote Sensing of Environment, 30*: 271-278.

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

Daniel, K.W., Tripathi, N.K., Honda, K. and Apasit, E., 2004. Analysis of VNIR (400-1100 nm) spectral signatures for estimation of soil organic matter in tropical soils of Thailand. *International Journal of Remote Sensing, 25(3)*: 643-652.

Elvidge, C.D., 1990. Visible and near infrared reflectance characteristics of dry plant materials. *International Journal of Remote Sensing, 11(10)*: 1775-1975.

Esweran, H., Van den Berge, E. and Reich, P., 1993. Organic carbon in soils of the world. *Soil Science Society America Journal, 57*: 192-194.

Fidencio, P.H., Poppi, R.J. and De Andrade, J.C., 2002a. Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy. *Analytica Chimica Acta, 453*: 125-134.

Fidencio, P.H., Poppi, R.J., De Andrade, J.C. and Cantarella, H., 2002b. Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. *Commun. Soil Sci. Plant Anal., 33*: 1607-1615.

Garten Jr., C.T. and Wullschleger, S.D., 1999. Soil Carbon Inventories under a Bioenergy Crop (Switchgrass): Measurement Limitations. *J. Environ. Qual. , 28*: 1359-1365.

Geladi, P. and Kowalski, B.R., 1986. Partial Least-Squares Regression: A Tutorial. *Analytica Chimica Acta, 185*: 1-17.

Grove, C.I., Hook, S.J. and Paylor, E.D., 1992. Laboratory reflectance spectra of 160 minerals 0.4 to 2.5 micrometers. *JPL publ., 92(2)*: 406.

Hartmann, H.P. and Appel, T., 2006. Calibration of near infrared spectra for measuring decomposing cellulose and green manure in soils. *Soil Biology & Biochemistry, 38*: 887-897.

Henderson, T.L., Baumgardner, M.F., Franzmeier, D.P., Stott, D.E. and Coster, D.C., 1992. High Dimensional Reflectance Analysis of Soil Organic Matter *Soil Science Society America Journal, 56:* 865-872.

IPCC, 2001. Summary for policy makers. In: J.T. Houghton et al. (Editors), Climate Change 2001: The Scientific Basis. Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge, pp. 944.

Irons, J.R., Weismiller, R.A. and Petersen, G.W., 1989. Soil Reflectance. In: G. Asrar (Editor), *Theory and Applications of Remote Sensing*. John Wiley & Sons, Washington D.C., pp. 66-106.

ISRIC, 2006. http://www.isric.org/UK/About+Soils/WDC+for+Soils/WDC+Catalogue/Reference+Soil+Samples.htm.

Kemper, T., Böttcher, K., Machwitz, M., Sommer, S. and Mehl, W., 2005. An approach to chemometric methods for soil organic matter estimation from laboratory and remote sensing data, *4th Workshop on Imaging Spectroscopy*, Warsaw, pp. April 27-29.

Kokaly, R.F., Asner, G.P., Martin, M.E., Ollinger, S.V. and Wessman, C.A., 2006. Retrieval of Quantitative and Qualitative Information about Plant Pigment Systems from High Resolution Spectroscopy, G*eoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Denver (USA).

Kokaly, R.F. and Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment, 67*: 267-287.

Kooistra, L., Wanders, J., Epema, G.F., Leuven, R.S.E.W., Wehrens, R. and Buydens, L.M.C., 2003. The potential of field spectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta, 484*: 198-200.

Kooistra, L., Wehrens, R., Leuven, R.S.E.W. and Buydens, L.M.C., 2001. Possibilities of VNIR spectroscopy for the assessment of soil contamination in river floodplains. A*nalytica Chimica Acta, 446*: 97-105.

Madari, B.E., Reeves III, J.B., Machado, P.L.O.A., Guimaraes, C.M., Torres, E. and McCarty, G.W., 2006. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma, 136*: 245-259.

Martens, H. and Naes, T., 1984. Multivariate calibration. 1. Concepts and distinctions. *Trends in analytical chemistry, 3(8)*: 204-210.

McBratney, A.B., Minasny, B. and Viscarra Rossel, R., 2006. Spectral soil analysis and inference systems: a powerful combination for solving the soil data crisis. *Geoderma, 136*: 272-278.

McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F. and Kimble, J.M., 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society America Journal, 66*: 640-646.

Morra, M.J., Hall, M.H. and Freeborn, L.L., 1991. Carbon and Nitrogen Analysis of Soil Fractions Using near-Infrared Reflectance Spectroscopy. *Soil Science Society of America Journal, 55(1)*: 288-291.

Osborne, J. and Waters, E., 2002. Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research & Evaluation, 8(2)*: 4.

Post, W.M., Emmanuel, W.R., Zinke, P.J. and Stagenberger, A.G., 1982. Soil organic pool and world life zones. *Nature, 298*: 156-159.

Post, W.M., Izaurralde, R.C., Mann, L.K. and Bliss, N., 2001. Monitoring and verifying changes of organic carbon in soil. *Climatic Change, 51*: 73-99.

Reeves III, J.B., McCarty, G.W. and Meisinger, J.J., 2000. Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *J. Near Infrared Spectroscopy, 8*: 161-170.

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S. and Martonchik, J.V., 2006. Reflectance quantities in optical remote sensing-definitions and case studies. *Remote Sensing of Environment, 103(1)*: 27-42.

Schreier, H., 1977. Quantitative predictions of chemical soil conditions from multispectral airborne, ground and laboratory measurements, *Canadian Symposium on Remote Sensing, 4th*, Quebec, Canada, 106-112.

Serrano, L., Penuelas, J. and Ustin, S.L., 2002. Remote Sensing of Nitrogen and Lignin in Mediterranean Vegetation from AVIRIS Data: Decomposing Biochemical from Structural Signals. *Remote Sensing of Environment (81)*: 355-364.

Shepherd, K.D. and Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal, 66*: 988-998.

Smith, P., 2004. Monitoring and verification of soil carbon changes under Article 3.4 of the Kyoto Protocol. *Soil Use and Management, 20*: 264-270.

Sørensen, L.K. and Dalsgaard, S., 2005. Determination of clay and other soil properties by Near Infrared Spectroscopy. *Soil Science Society America Journal, 69*: 159-167.

Stevens, A., Van Wesemael, B., Bartholomeus, H.M., Rossilon, D., Tychon, B. and Ben-Dor, E., 2008. Laboratory, Field and Airborne Spectroscopy for monitoring Organic Carbon Content in Agricultural Soils. *Geoderma, 144*: 395-404.

Stevens, A., Van Wesemael, B., Vandenschrick, G., Touré, S. and Tychon, B., 2006. Detection of Carbon Stock Change in Agricultural Soils Using Spectroscopic Techniques. *Soil Science Society of America Journal, 70*: 844-850.

Stoner, E.R. and Baumgardner, M.F., 1980. Characteristic Variations in Reflectance of Surface Soils. *Soil Science Society of America Journal, 45(6)*: 1161-1165.

Sudduth, K. and Hummel, J., 1991. Evaluation of reflectance methods for soil organic matter sensing. *Transactions of the ASAE, 34*: 1900-1909.

Sudduth, K.A. and Hummel, J.W., 1993. Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Transactions of the ASAE, 36*: 1571-1582.

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

Ustin, S., Asner, G., Gamon, J., Huemmerich, K., Jacquemoud, S., Schaepman, M.E. and Zarco-Tejada, P., 2006. Retrieval of Quantitative and Qualitative Information about Plant Pigment Systems from High Resolution Spectroscopy, *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Denver (USA). 1996-1999.

Van der Molen, M.K., 2005. On-the-go sensors for Mapping of Soil Properties, a Field study, Wageningen University, Wageningen, 65 pp.

Vinogradov, B.V., 1982. Remote Sensing of the humus content of soils. *Soviet Soil Science, 13*: 103-113.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. and Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma, 131*: 59-75.

Walkley, A. and Black, I.A., 1934. An estimation of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science, 37*: 29-37.

Chapter 5

Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy

Harm Bartholomeus, Gerrit Epema, Michael Schaepman

Abstract

The possibility of quantifying iron content in the topsoil of the slopes of the El Hacho Mountain complex in Southern Spain using imaging spectroscopy is investigated. Laboratory, field and airborne spectrometer (ROSIS) data are acquired, in combination with soil samples, which are analysed for dithionite extractable iron (Fe_d) content. Analysis of the properties of two iron related absorption features present in laboratory spectra demonstrates good relations, especially between the Standard Deviation (S.D.) of the values in an absorption feature and the Fe_d content ($R^2 = 0.67$) as well as the ratio based Redness Index ($R^2 = 0.51$). Such derived relations are less strong for the ROSIS data (R^2 for S.D. $= 0.26$ and R^2 for Redness Index $= 0.22$). The spatial distribution of iron in vegetated areas shows a strong sensitivity of these relations with the presence of vegetation. A combination of both methods shows that the overestimation of the Fe_d content with the one method is (partly) compensated by the underestimation with the other method.

5.1. Introduction

The slopes of El Hacho Mountain near Álora in Southern Spain show a large variety in iron content. El Hacho is composed of materials derived from Tertiary marls and sands, deposited on a continental slope, which are indicated with the generic term Flysch deposits (Fig. 5.1). On top of El Hacho thick cemented deposits of sand and gravel (conglomerate) occur (Buurman, 1999). Blocks of this conglomerate are now scattered over the flysch slope, where Cambisols are formed. Below these blocks unweathered flysch is exposed. Iron content is varying with the distance down slope from these blocks, due to erosion, iron deposition processes and leaching. The spatial heterogeneity of the iron distribution is further increased due to presence of gullies. Hematite is the major iron bearing mineral in the area, while some goethite may occasionally be present behind the conglomerate blocks (Buurman, 2005).

Iron (Fe) is one of the most common minerals contained in soils (Hunt, 1980). It is an indicator for the fertility of the soil, the usability of an area to cultivate specific crops and an indicator of the age of the deposits (Torrent et al., 1980; Gardner and Pye, 1981; Blount et al., 1990). Determining the spatial distribution of different types of iron with traditional fieldwork and laboratory analysis is time-consuming and expensive (Liang, 2004). Remote sensing has proven to be a useful tool to determine the presence of iron in extended areas and various research fields (Escadafal, 1993; Farrand, 1997; Palacios-Orueta et al., 1999; Warell, 2003). Quantification of the amount of iron contained in soils using spectral measurements has proven to be feasible (Torrent et al., 1983; Coyne et al., 1989; Ben-Dor and Banin, 1990), and is focusing on laboratory or field spectra and specific soil types (Dematte, 2002). Presently, technical improvements allow mapping using higher spectral and spatial resolution imaging spectrometers with better signal to noise ratios (Dematte et al., 2004). The need for relatively high spectral resolution data to determine iron content using reflectance spectra has been extensively discussed in Ben-Dor et al. (1999). We propose to use for the El Hacho area imaging spectrometer data for mapping iron quantity. In addition to the analysis of the Reflective Optics System Imaging Spectrometer (ROSIS) imagery, field spectra and laboratory spectra are acquired and analysed. The influence of iron on the spectral signature in the wavelength range covered by ROSIS (416.9 – 872.9 nm) has been discussed in literature extensively (e.g. Hunt and Salisbury, 1970; Stoner and Baumgardner, 1981; Ben-Dor et al., 1999; Bullard and White, 2002). This study focuses on the determination of the soil iron content in fractionally covered olive fields, where the ploughing, weeding and a distance of 6 meters between the trees results in a substantial amount of bare soil surface in the 2 x 2 meter ground projected instantaneous field of view (GIFOV) of the ROSIS images. Zarco-Tejada et al. (2004) performed research on olive fields too, focusing on eliminating the influence of soil by deriving plant parameters of olive trees.

However, in this study the analysis of absorption features was used to estimate iron content, using regression analysis, considering the influence of vegetated areas. Fig. 5.2 gives an impression of the study-area.
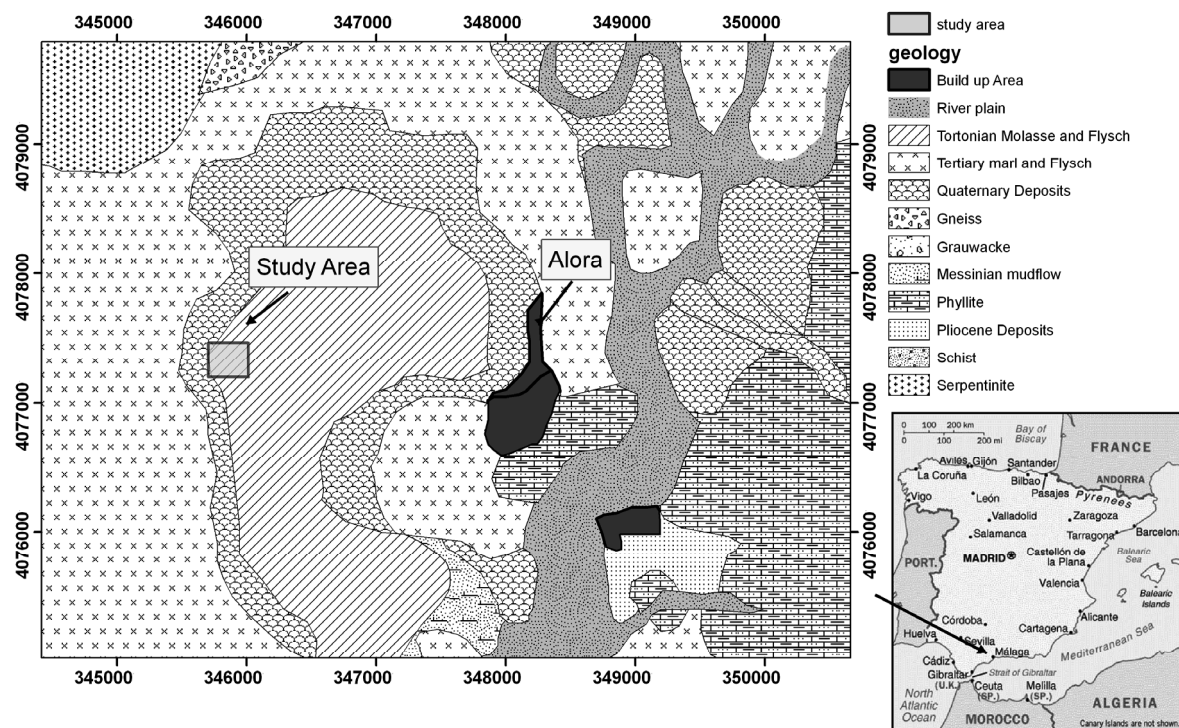

Figure 5.1: Geologic map of Álora and surroundings. The study area (corresponding with Fig.5.9) is located west of Álora and delineated with the square. Coordinates are in UTM, zone 30N, NAD27.

Figure 5.2: Impression of the study area

5.2. Methodology

5.2.1 Soil Sampling and analysis

In June 2003, 35 bare soil plots, of 2 by 2 meter size, were sampled and measured for iron content. The sample locations were positioned in two down slope transects and cover the full iron variation within these slopes (Fig. 5.3). Fractional cover of soil, rocks and litter was determined in each plot and a soil sample of the topsoil (first two centimeters) within the plot was collected for determination of iron content and laboratory spectroscopy.
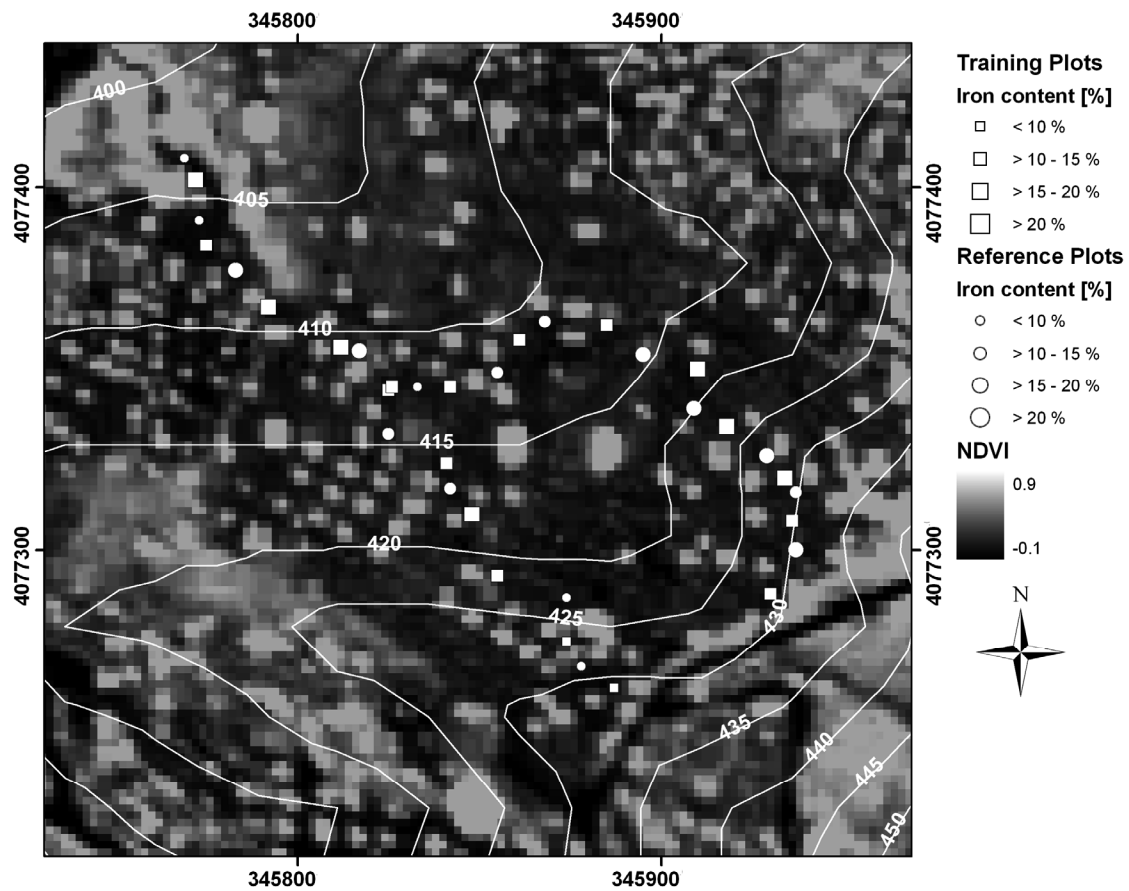

Figure 5.3: Distribution of the 35 iron sample points within the El Hacho study area. NDVI values are calculated from the ROSIS image. The isolines represent the altitude in meters. Coordinates are in UTM, zone 30N, NAD27.

The dithionite extractable iron (Fe_d) concentration of all 35 soil samples was determined using a dithionite extraction as described in Raiswell et al. (1994). An oxalate extraction was used to determine the concentration of oxalate extractable iron (Fe_o). The concentration of iron in the extraction fluids was determined with an Inductively Coupled Plasma Atomic Emission Spectrometer (ICP-AES). In literature, Fe_d is often referred to with the term free

iron; however, terminology is not uniform in this sense. Therefore, we refer to the Fe content by the name of the extraction technique instead of using uninformative phase designations, as suggested by Borggaard (1988).

5.2.2 Field and laboratory spectra

Field and laboratory spectra were acquired using an ASD Fieldspec Pro FR spectroradiometer, covering the 350 – 2500 nm wavelength region. Four reflectance measurements per plot usually were taken in the corners to estimate the full spectral variability within a plot. The aperture angle of the fore optics of the spectrometer was 25° and measurements were taken from nadir, at a distance of about 30 cm from the surface, resulting in a circular Ground projected Field of View of 13 cm diameter. The soil samples used for the laboratory spectral analysis were air-dried and sieved (<2 mm). The incidence angle of the irradiance source is set to 30° off nadir and a 3° fore optic is used at a 30 cm distance from the target, resulting in a circular Ground projected Field of View of 3 cm diameter. The sample was rotated in steps of 90°, resulting in four measurements per sample. Uncalibrated white spectralon panels were used for the field and laboratory measurements, resulting in relative reflectance values. The four measurements per plot were compared, deviating measurements were removed when necessary, and the remaining spectra were averaged. This was done for both field and laboratory spectra separately. For further analysis the relative reflectance field and laboratory spectra were limited to, and rescaled to the bandwidth of ROSIS.

5.2.3 Image data

Airborne imaging spectrometer data were collected in June 2001, during the DAISEX campaign (Berger et al., 2001). ROSIS was built for the detection of absorption features especially in coastal and inland waters (Gege et al., 1998), but it is also used for land applications (Holzwarth et al., 2003). During this overflight of El Hacho, ROSIS recorded data in 115 spectral bands, ranging from 416.9 to 872.9 nm, with a bandwidth of approximately 4 nm and a 4 nm sampling interval. The operating altitude of the aircraft was at flight level 120 (corresponding to 3510 m above ground), resulting in a pixel size of 2x2 meters. ROSIS images were calibrated and converted to reflectance values by the operator (DLR), using the PARGE/ATCOR model (Schläpfer and Richter, 2002). Quality analysis of ROSIS resulted in the exclusion of band 1 (416.9 nm) and 2 (420.9 nm), because of the low signal to noise ratio. A DEM with a resolution of 25 m in x and y direction and 1 m in z direction was used for orthorectification. The use of a handheld GPS for determination of the location of the sample plots appeared to be inaccurate. Therefore, the positioning of the plots on the ROSIS image was done visually, which can be done up to 1 to 2 pixels accurately, using clearly recognisable objects in the image like olive trees. In this way correct spectral information was achieved for the sampling points.

5.2.4 Processing of Spectral data

The plots were divided in a training set (19 plots) and a reference set (16 plots), to allow cross-validation of the results. The plots were divided in such way that both sets show a similar range in iron content variation along locations down the slope. The distribution of training and reference points is indicated in Fig. 5.3.

First, direct correlation of the reflectance values in all individual bands with the iron content was performed, for field, laboratory and ROSIS spectra. This reveals the spectral regions that are directly influenced by the presence of iron. The next step was the calculation of the Redness Index. This method uses the reflectance properties in the visible part of the spectrum and is used as indicator for iron and was among others used by Bullard and White (2002): Redness = $R_{600\text{-}700nm}$ / ($R_{600\text{-}700nm\ +}$ $R_{500\text{-}600nm\ +}$ $R_{400\text{-}500nm}$) in which R = Reflectance. Bands in these spectral ranges were averaged; corresponding band numbers are given in table 5.1.

Table 5.1: Band numbers corresponding with the wavelength ranges used in the calculation of the Redness Index.

Wavelength	Bandnumbers ROSIS and resampled ASD spectra	
Blue: 400-500 nm	Band 3 – 21	(424.9 – 496.9 nm)
Green: 500-600 nm	Band 22-46	(500.9 – 596.9 nm)
Red: 600-700 nm	Band 47-71	(600.9 – 696.9 nm)

Continuum removal (Green and Craig, 1985) between 424.9 and 872.9 nm was applied on the field, laboratory and ROSIS spectra. The minimum, average and median values were calculated of the continuum-removed spectra. Absorption feature properties used during analysis were depth, width, area and standard deviation of dips. The width of the dip is defined as the wavelength distance between the maximum values on both sides of the dip, after continuum removal. The depth of the absorption feature is the maximum distance of the spectrum to the continuum. Summing these distances over the width of the dip results in the area of the dip (Grove et al., 1992). Furthermore, the Standard Deviation (S.D.) of the continuum-removed values for each dip was calculated.

As a next step, the linear correlation between the spectral features and both types of iron was calculated on both the training and reference points. All presented R^2 values have a significance level of 0.95. Prediction of the iron content was done using the regression functions and is expressed in Standard Error of Calibration (SEC) values. Next, cross validation between the training and reference set was carried out, expressed with Standard Error of Performance (SEP) values. Predictions were done with the relations based on the corresponding spectral dataset, since the use of different reference panels does not allow

interchanging the relationships between different spectral datasets. The difference between the predicted and the actual measured iron content is expressed in mass percent iron.

Since the relations are only valid for bare soil the variance of the predictive value related to fractional vegetation cover was tested. Linear mixing of a bare soil laboratory spectrum and an olive tree field spectrum with different fractions was done to set a threshold of vegetation on the iron content prediction. Both spectra were linearly mixed representing a fractional vegetation cover up to 30 %. The iron content of the fractionally covered pixels was predicted using the regression function based on the training set of the laboratory spectra.

Prediction of the iron content with the Redness Index should underestimate the amount of Fe_d when the fractional vegetation cover increases, because of the decrease of the red reflectance and thus a lowering of the Redness index. An overestimation of the iron content is expected when the area or standard deviation of the absorption features of continuum-removed spectra is used. This is a result of the decrease of reflectance in the visible wavelengths and an increase in Near Infrared reflectance, which leads to a larger absorption feature and higher S.D. and Area. A combination of both methods may result in a better prediction of the Fe_d content; therefore this was also tested.

Finally, the technique which yields the best prediction results and is least influenced by vegetation was used to create a spatial iron distribution map. For this purpose fractional vegetation cover was determined using Spectral Mixture Analysis (SMA) (Smith et al., 1985), after which pixels with a fractional vegetation cover of more than 30 % were masked out. As a preparatory step for SMA, Minimum Noise Fraction (MNF) rotation (Green et al., 1988) was used to de-stripe the ROSIS image and smooth the spectral response. The iron content of the masked out areas was interpolated from the surrounding pixels using Inversed Distance Weighted (IDW) interpolation (Watson and Philip, 1985).

5.3. Results and Discussion

5.3.1 Soil chemical analysis

The Fe_d concentration of the samples varies between 7 and 20 %. Although these high concentrations are not common, more soils in Andalucia show this range of iron concentrations (De la Rosa et al., 1984). The amount of Fe_o is low in most samples (1.4 % at highest) and correlations with reflectance are low ($R^2_{lab} = 0.36$ at highest). Therefore only correlations with Fe_d concentration are used in the further analysis and presented in this

paper. The moisture content of the soil samples was below 1 % in all samples, so its effect on the reflectance can be neglected.

5.3.2 Spectra

The spectral curves show a number of features related to iron (Fig. 5.4). Absorption features are found from about 350 to 600 nm (further referred to as: D_{550}), and from about 850 to 970 nm (further referred to as: D_{870}). D_{550} represents the increase from blue to red and indicates that hematite is the main iron bearing mineral in the study area (Ben-Dor et al., 1999). This absorption feature is related to iron in its ferrous form (Fe^{2+}) and is the result of large absorption in the ultraviolet (UV), which is related to charge transfer transitions between iron and oxygen (Goetz, 1989). It also influences the visible part of the spectrum, since it has a strong wing to the visible wavelengths (Hunt, 1980). Besides these two dips more absorption features exist in this part of the spectrum as is extensively described by Ben-Dor et al. (1999). The effect is an overall decrease in reflectance when the iron content increases, as can be seen in Fig. 5.4.

Continuum removal shows D_{550} at the lower wavelengths. The strong absorption of iron ranges from the UV to approximately 700 nm, which corresponds with the high correlation values from the original spectra. A smaller dip, which was not visible in the reflectance spectra, is revealed around 640 nm after Continuum Removal (Fig. 5.5), which is also related to the presence of iron (Obukhov and Orlov, 1964). Because of the wide range of the absorption in the UV, this feature became part of D_{550}. The area of D_{870}, which is related to iron in its ferric form (Fe^{3+}) can not be calculated, because this absorption feature ranges beyond the spectral range used in this study. Continuum removal applied on the incomplete dip may lead to undesired results.

The selection of sample plots with the same size as an individual pixel may result in indistinctness when identifying the reflectance properties of the image. Ideally, one should take a plot size of at least 3*3 times the pixel size, but the spatial distribution of olive trees in the area does not leave enough bare soil in between the trees to match this criterion. As a result, the selected ROSIS spectra may be influenced by vegetation. For this reason, the influence of vegetation on the reliability of iron prediction is tested and discussed later on.
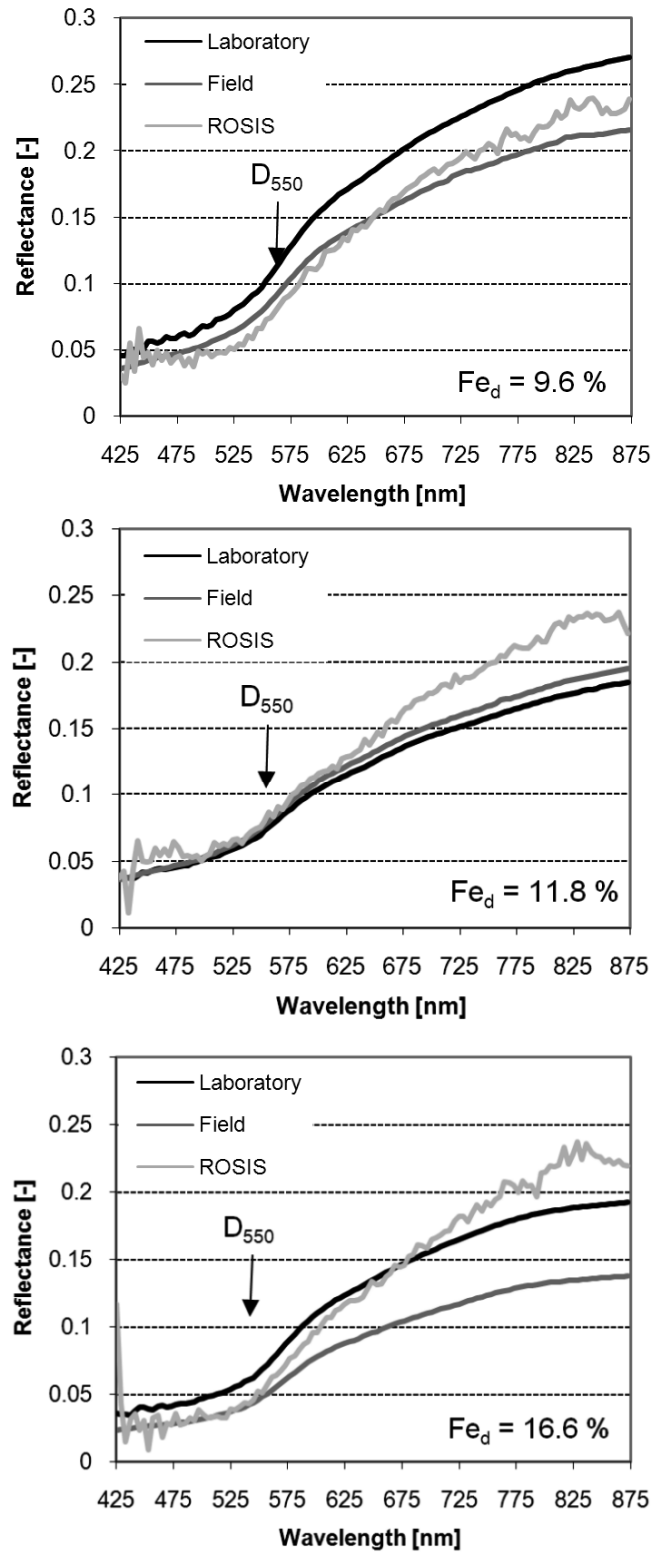
Figure 5.4: Spectral signatures from laboratory, field and ROSIS measurements with three different Fe_d contents.
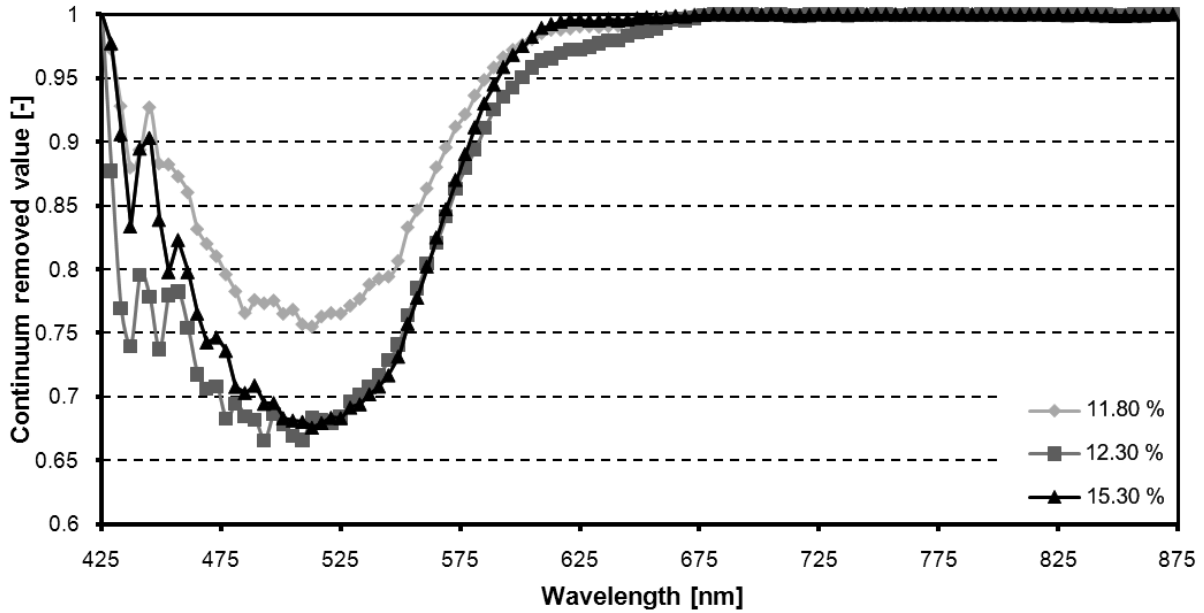
Figure 5.5: Continuum removed field spectra with three different Fe_d contents.

5.3.3 Correlation of reflectance with iron

Direct correlation of the Fe_d content with the reflectance values gives high correlations ($R^2_{lab} = 0.71$) for the laboratory spectra. Highest R^2-values occur at lower wavelengths, up to 600 nm. The field spectra show a medium correlation ($R^2_{field} = 0.48$) at lower wavelengths too. Goetz (1989) reports that this absorption is indeed due to the presence of iron. As expected, the correlations for the laboratory spectra are much higher than the correlations for field spectra (Fig. 5.6). Stable measuring conditions and samples which are dried and sieved, improve the results. Correlation with the ROSIS spectra shows comparable results ($R^2_{ROSIS} = 0.5$), but the highest values occur near 650 nm.

5.3.4 Correlation of reflectance properties with iron

The Redness Index gave medium correlating results with Fe_d. Field and laboratory spectra showed some correlation ($R^2_{lab} = 0.51$ and $R^2_{field} = 0.48$), but ratios calculated with the ROSIS spectra showed only little correlation with the iron content ($R^2_{ROSIS} = 0.22$). Correlations between Fe_d and the continuum-removed values resulted in high negative correlations for field- and laboratory spectra from the UV to 600 nm ($R^2_{lab} = 0.78$ and $R^2_{field} = 0.56$) (Fig. 5.7). This range corresponds with D_{550} again. The relation between continuum-removed ROSIS spectra and iron is again less strong ($R^2_{ROSIS} = 0.35$). Higher absorption in the beginning of the spectrum is also reflected in a positive correlation between the area of D_{550} and the Fe_d content ($R^2_{lab} = 0.65$). The width of the dip shows a positive correlation too, but the correlation is weaker ($R^2_{lab} = 0.41$).
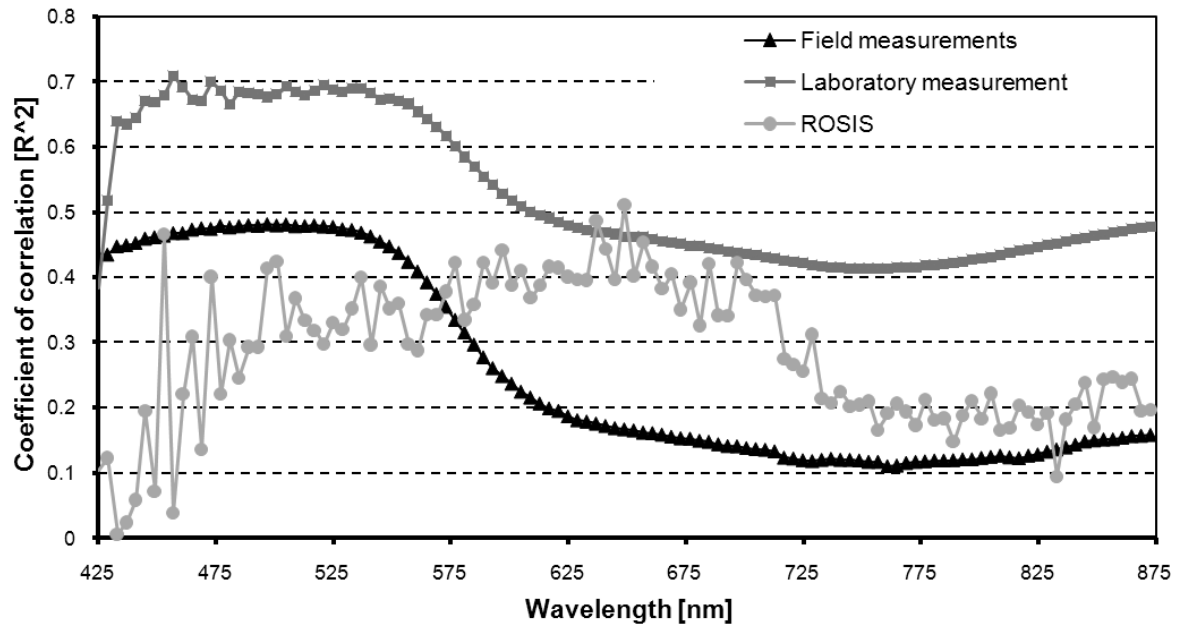
Figure 5.6: Correlation coefficients between reflectance values and Fed content.
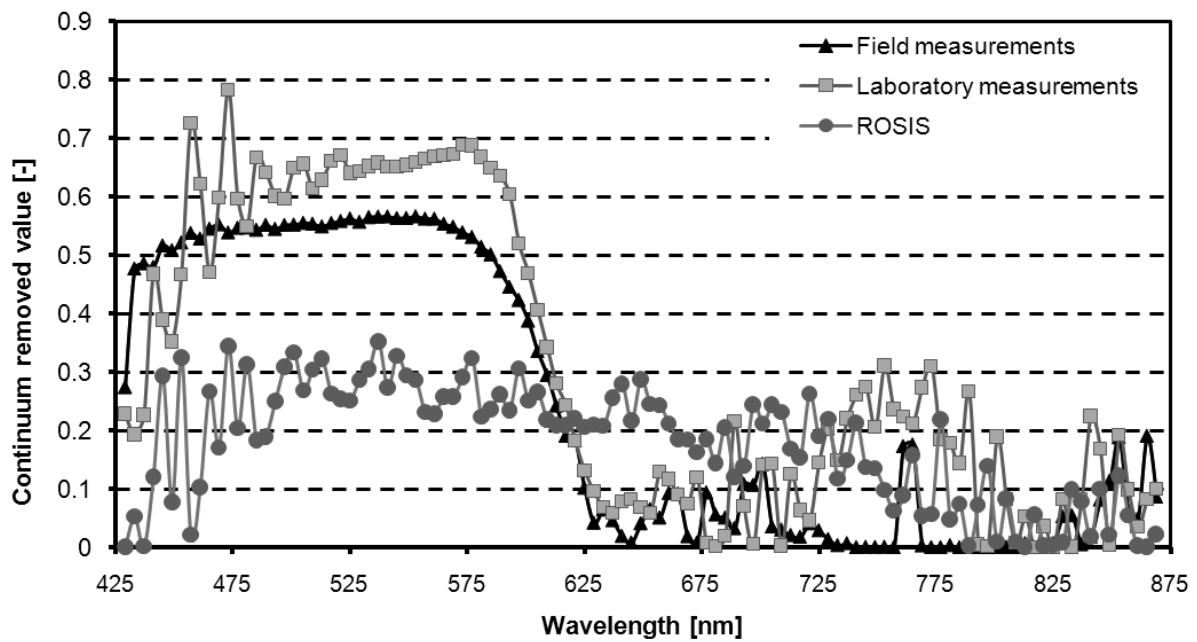

Figure 5.7: Correlation coefficients between continuum removed values and total iron content.

Calculating statistics from the full range of continuum removed values shows the most stable results over all datasets (Table 5.2). The average value after continuum removal correlated negatively with the Fe_d content ($R^2_{lab} = 0.66$). The S.D. resulted in slightly higher, but positive correlations ($R^2_{lab} = 0.67$). The S.D. of the continuum-removed values combines the effect of the width, depth and area of the dip, which are all highly correlated. When D_{550} is deeper or wider the S.D. after continuum removal increases. The area of the dip is highly

related to an increase of the width and depth, thus it has the same effect on the S.D.. An overview of the correlations is given in Table 5.2. ROSIS spectra show lower correlations than laboratory and field spectra, independent of the analysis.

Table 5.2: Correlation coefficients between absorption feature properties and Fe_d content. Values are given in R^2.

	Field Spectra		Laboratory Spectra		ROSIS Spectra	
	training set	reference set	training set	reference set	training set	reference set
Redness Index	0.27	0.51	0.48	0.44	0.22	0.10
Area of D_{550}	0.29	0.55	0.36	0.65	0.16	0.26
S.D. after cont rem.	0.29	0.55	0.40	0.67	0.19	0.26
Average after cont rem.	0.29	0.56	0.35	0.66	0.17	0.27

5.3.5 Prediction of iron content

Predictions of the iron content are made with the established regression functions. The regression functions of the training set are applied on the reference set and vice versa. SEC and SEP values for three methods, Redness Index, S.D. and Area of D_{550}, are given in Table 5.3. When laboratory or field spectra are used for Fe_d prediction, SEC and SEP values are between 2.66 and 4.95 where Fe_d varies between 7 and 20 %. Using the S.D. after continuum removal gave better results than the Redness Index, but differences are small. Fe_d prediction with ROSIS spectra gave less good results and SEC and SEP values vary between 4.38 and 11.39. The best results with this dataset are gained with S.D., giving SEP-values of 5.79 and 7.71 (Table 5.3).

The training set shows structurally lower correlation values and prediction results than the reference set. A closer inspection shows this is caused by point A03, which is located on top of one of the sliding blocks. Removing it increases R^2-values (R^2_{field} using S.D. increases from 0.29 to 0.53) and decreases the SEC and SEP of the training set (SEC_{field} using S.D. decreased from 4.9 to 3.5 and SEP_{field} using S.D. decreased from 3.5 to 2.6). Although the point is identified as an outlier there is no physical reason to remove it from the dataset. Therefore, we decided to leave it in.

Table 5.3 : SEC and SEP values of Fe_d determination for three spectral datasets. Fe_d values of the training and reference plots ranges between 7 % and 20 %.

	Field Spectra				Laboratory Spectra				ROSIS Spectra			
	training set		reference set		training set		reference set		training set		reference set	
	SEC	SEP	SEC	SEP	SEC	SEP	SEC	SEP	SEC	SEP	SEC	SEP
Redness Index	5.22	5.12	3.67	3.86	3.29	5.01	4.20	3.19	4.38	10.59	11.39	4.97
Area of D_{550}	4.96	4.81	3.42	3.57	4.18	3.92	2.78	3.21	7.26	5.73	6.41	9.08
S.D. after cont rem.	4.95	4.88	3.40	3.50	3.81	3.98	2.66	2.59	6.43	5.79	6.39	7.71

5.3.6 Effect of vegetation on iron prediction

All presented relations are based on bare soil spectra. To simulate varying fractional vegetation cover a bare soil laboratory spectrum and an olive tree spectrum measured under field conditions are linearly mixed. When using the Redness Index or the S.D. a 15% vegetation fraction (NDVI = 0.43) results in a 2.8% difference in predicted iron content. When the fractional vegetation cover is 20%, this difference is already more than 6%. This effect is much stronger when the Area of D_{550} is used to predict Fe_d. When using the Redness Index at higher vegetation cover, the iron prediction shows a structural underestimation. Vegetation effects lead to a lower reflectance in the red part of the spectrum, while the reflectance in the green wavelengths shows a relative increase. Therefore, the Redness Index will always underestimate the iron content when the regression is established on bare soil and applied to partially vegetated spectra. The S.D. after continuum removal shows a comparable deviation, but results in an overestimation of the iron content. The reflectance in the NIR increases, while the reflectance in the VIS decreases, if influence of vegetation is present. This increases the size of the absorption dip in the visible part of the spectrum. Using the area of D_{550} to predict the iron content in the partially vegetated areas is much more sensitive to fractional vegetation cover and therefore less useful in this study area.

In this simulation a combination of the Redness Index and the S.D. after continuum removal creates a valuable approach to use in partially vegetated areas. Both indices predict the iron content accurately up to 5% vegetation cover. At a higher vegetation cover the average of both iron prediction results is taken, which leads to a much lower decrease in prediction accuracy, as is shown in Fig. 5.8. The combination of both methods could allow an acceptable iron determination even in partially vegetated areas and reduces the decrease in prediction accuracy due to mixed pixels or adjacency effects.
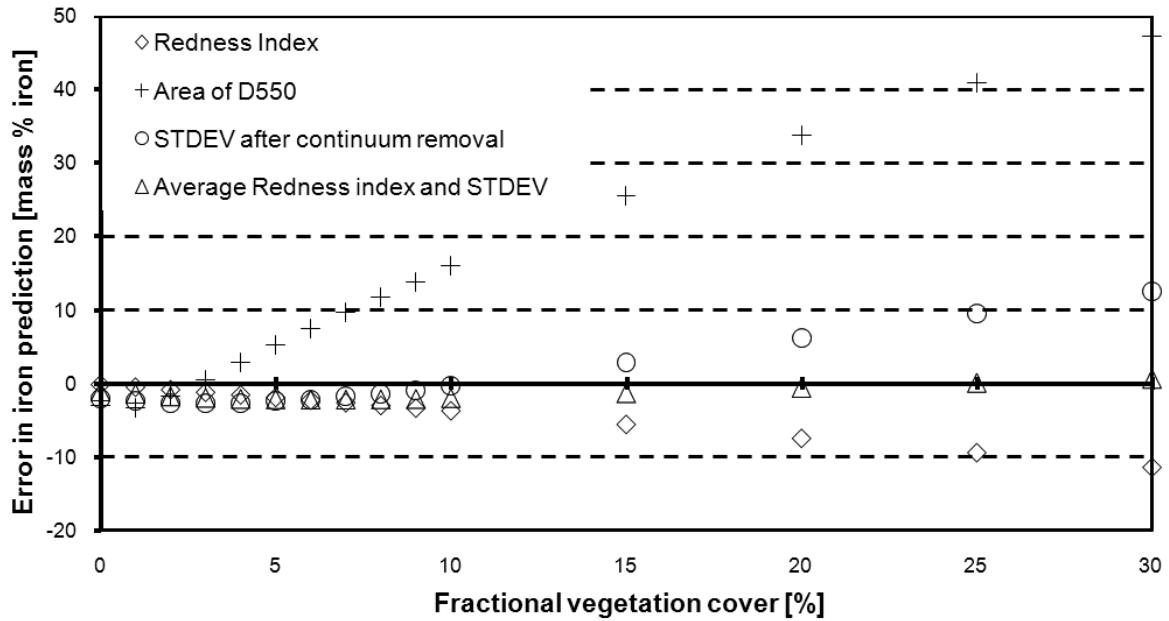
Figure 5.8: Error in Fe_d prediction in mass percent iron, related to fractional vegetation cover. The error is calculated as $Fe_{predicted} - Fe_{measured}$.

Finally a Fe_d spatial distribution map is created. First, a linear unmixing of the MNF de-striped image was done to determine fractional vegetation cover. Next all areas with a vegetation cover of >30 % were masked out. For all remaining areas the Fe_d content was calculated with the average of the S.D. after continuum removal and the Redness Index, using the regression functions based on the ROSIS spectra of the training plots. The Fe_d content of the masked out pixels is determined using IDW interpolation with power 2 and using the 12 nearest points in the calculation. The resulting Fe_d spatial distribution map is shown in Fig. 5.9.

5.4. Conclusions

The spectral range from UV to NIR, covered by ROSIS, is suitable for quantification of Fe_d content, as shown in this study. The S.D. of a continuum-removed spectrum in the VIS and NIR part of the spectrum, the area of the absorption dip around 550 nm and the Redness Index give comparable results. The area of the absorption dip around 550 nm is much more sensitive to the influence of vegetation on the spectral response than the Redness Index or the S.D. of the continuum removed spectrum. Combining the iron predictions of two methods, Redness Index and the S.D. after continuum removal, allows more accurate iron prediction in partially vegetated areas.
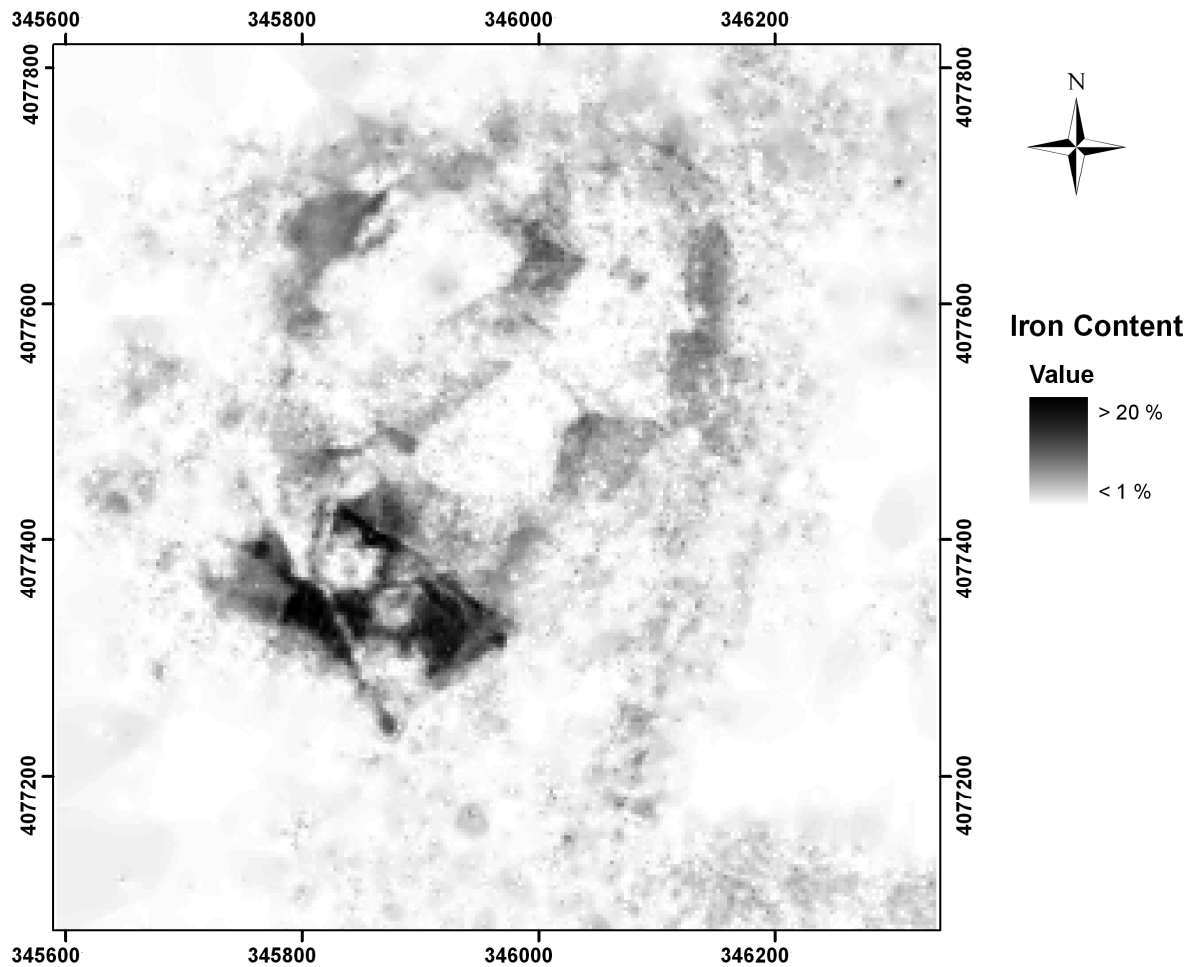
Figure 5.9: Fe_d spatial distribution map, based on ROSIS imagery. First a linear unmixing of the MNF de-striped image was done to determine the fractional vegetation cover. Next all areas with a vegetation cover of >30% were masked out. For all remaining areas the Fe_d content was calculated with the average of the S.D. after continuum removal and the Redness Index and using the regression functions based on the ROSIS spectra of the training plots. The Fe_d content of the masked out pixels is determined using inversed distance weighted interpolation.

Acknowledgements

References

Ben-Dor, E. and Banin, A., 1990. Diffuse reflectance spectra of smectite minerals in the near infrared and their relation to chemical composition. *Sci. Geol. Bull 43(2-4)*, 117-128.

Ben-Dor, E., Irons, J. R. and Epema, G. F., 1999. Soil Reflectance. In: Rencz, A. N. (Eds.). Remote Sensing for the Earth Sciences: *Manual of Remote Sensing.* John Wiley & Sons, Inc. , New-York, pp. 111-188.

Berger, M., Rast, M., Wursteisen, P., Attema, E., Moreno, J., Mueller, A., Beisl, U., Richter, R., Schaepman, M., Strub, G., Stoll, M. P., Nerry, F. and Leroy, M. E., 2001. The DAISEX campaigns in support of a future land-surface-processes mission. *ESA Bull. 105*, 101-111.

Blount, G., Smith, M. O., Adams, J. B., Greeley, R. and R., C. P., 1990. Regional Aeolian dynamics and sand mixing in the Grab Desierto: Evidence from Landsat Thematic Mapper Images. *J. Geophys. Res. 95*, 15463- 15482.

Borggaard, O. K., 1988. Phase identification by selective dissolution techniques. In: Stucki, J. W., Goodman, B. A. and Schwertmann, U. (Eds.). *Iron in soils and clay minerals*. D. Reidel Publishing Company, Dordrecht, pp. 83-98.

Bullard, J. E. and White, K., 2002. Quantifying iron oxide coatings on dune sands using spectrometric measurements: An example from the Simpson-Strzelecki Desert, *Australia. Journal of Geophysical Research B: Solid Earth 107*, 5-1.

Buurman, P., 2005. Personal Communication.

Buurman, P., 1999. Introduction to the interdisciplinary practical Sustainable Land Use in the Alora region, Spain. In. *Practical Manual Wageningen University.*

Coyne, L. M., Bishop, J. L., Sacttergood, T., Banin, A., Carle, G. and Orenberg, J., 1989. Near-infrared correlation spectroscopy: quantifying iron and surface water in series of variably cation-exchanged montmorillonite clays. In: Coyne, L. M., McKeever, S. W. S. and Blake, D. F. (Eds.). *Spectroscopic characterization of Minerals and Their Surfaces, ACS Symp. Ser. 415*. Washington, pp. 407 - 429.

De la Rosa, D., Banos, C., Mudarra, J. L., Barahona, E., Moreira, J. M., Gago, R., Puertas, J. M. and Ramos, A., 1984. *Catálogo de suelos de Andalucía. Junta de Andalucía*, Sevilla.

Dematte, J. A. M., 2002. Characterization and discrimination of soils by their reflected electromagnetic energy. P*esquisa Agropecuaria Brasileira 37*, 1445 - 1458.

Dematte, J. A. M., Campos, R. C., Fiorio, P. R., Alves, M. C. and Nanni, M. R., 2004. Visible-NIR reflectance: A new approach on soil evaluation. *Geoderma 121*, 95-112.

Escadafal, R., 1993. Remote Sensing of soil color: principles and applications. *Remote Sensing Rev. 7*, 261-279.

Farrand, W. H., 1997. Identification and mapping of ferric oxide and oxyhydroxide minerals in imaging spectrometer data of Summitville, Colorado, USA and the surrounding San Juan Mountains. *International Journal of Remote Sensing 18*, 1543-1552.

Gardner, P. and Pye, K., 1981. Nature, origin and palaeoenvironmental significance of red coastal and desert dune sands. *Prog. Phys. Geogr. 54*, 514-534.

Gege, P., Beran, D., Mooshuber, W., Schulz, J. and Piepen, H. V. d., 1998. System analysis and performance of the new version of the imaging spectrometer ROSIS. *1st EARSeL Workshop on Imaging Spectroscopy*, Remote Sensing Laboratories, University of Zurich, Switzerland, pp. 29-35.

Goetz, A. F. H., 1989. Spectral Remote Sensing in Geology. Wiley, New York.

Green, A. A., Berman, M., Switzer, P. and Craig, M. D., 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing 26*, 65-74.

Green, A. A. and Craig, M. D., 1985. Analysis of aircraft spectrometer data with logarithmic residuals. NASA-JPL Publication

Grove, C. I., Hook, S. J. and Paylor, E. D., 1992. Laboratory reflectance spectra of 160 minerals 0.4 to 2.5 micrometers. *JPL publ. 92*, 406.

Holzwarth, S., Muller, A., Habermeyer, M., Richter, R., Hausold, A., Thiesmann, S. and Strobl, P., 2003. HySens - DAIS 7915 / ROSIS Imaging Spectrometers at DLR. *Proceedings of the3 rd Earsel Workshop on Imaging Spectroscopy*, Herrsching, 13-16 may, pp. 3-14.

Hunt, G. R., 1980. Electromagnetic radiation: The communications link in remote sensing. Wiley, New York.

Hunt, G. R. and Salisbury, J. W., 1970. Visible and near infrared spectra of minerals and rocks. I. Silicate minerals. *Modern Geol. 1*, 283-300.

Liang, S., 2004. Quantitative Remote Sensing of Land Surfaces. Wiley & Sons Inc., New York.

Obukhov, A. I. and Orlov, D. C., 1964. Spectral reflectance of the major soil groups and the possibility of using diffuse reflection in soil investigations. *Sov. Soil Sci. 2*, 174-184.

Palacios-Orueta, A., Pinzón, J. E., Ustin, S. L. and Roberts, D. A., 1999. Remote sensing of soils in the Santa Monica Mountains: II. Hierarchical foreground and background analysis. *Remote Sensing of Environment 68*, 138-151.

Raiswell, T., Canfield, D. E. and Berner, R. A., 1994. A comparison of iron extraction methods for the determination of degree of pyritisation and the recognition of iron limited pyrite formation. *Chemical Geology 11*, 101-110.

Schläpfer, D. and Richter, R., 2002. Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: parametric orthorectification. *International Journal of Remote Sensing 23*, 2609-2630.

Smith, M. O., Johnstonn, P. E. and Adams, J. B., 1985. Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. *Journal of Geophysical research 90*, 797-804.

Stoner, E. R. and Baumgardner, M. F., 1981. Characteristic variations in reflectance of surface soils. *Soil Science Society of America Journal 45*, 1161-1165.

Torrent, J., Schwertmann, U., Fetcher, H. and Alferez, F., 1983. Quantitative relationships between soil color and hematite content. *Soil Science 13*, 354-358.

Torrent, J., Schwertmann, U. and Schulze, D. G., 1980. Iron oxide mineralogy of some soils of two river terrace sequences in Spain. *Geoderma 23*, 191-208.

Warell, J., 2003. Properties of the Hermean regolith: III. Disk-resolved vis-NIR reflectance spectra and implications for the abundance of iron. *Icarus 161*, 199-222.

Watson, D. F. and Philip, G. M., 1985. A Refinement of Inverse Distance Weighted Interpolation. *Geoprocessing 2*, 315-327.

Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A. and J.Agüera, 2004. Hyperspectral Indices and Model Simulation for Chlorophyll Estimation in Open-Canopy Tree Crops. *Remote Sensing of Environment 90*, 463-476.

Chapter 6

Soil Organic Carbon mapping in partially vegetated agricultural fields with imaging spectroscopy

Harm Bartholomeus, Lammert Kooistra, Antoine Stevens, Bas van Wesemael, Eyal Ben-Dor, Bernard Tychon, Martin van Leeuwen

Abstract

Soil Organic Carbon (SOC) is one of the key soil properties, but the large spatial variation makes continuous mapping a complex task. Imaging spectroscopy has proven to be a useful technique for mapping of soil properties, but the applicability decreases rapidly when fields are partially covered with vegetation. With only a few percent fractional vegetation cover the accuracy of a partial least square regression (PLSR) based SOC prediction model drops dramatically. In this paper we show that this problem can be solved with the use of spectral unmixing techniques. First, the fractional vegetation cover is determined with linear spectral unmixing, taking the illumination and observation angles into account. In a next step the influence of vegetation is filtered out from the spectral signals by a procedure termed Residual Spectral Unmixing (RSU). The residual soil spectra resulting from this procedure are used for mapping of SOC using PLSR, which could be done with accuracies comparable to studies performed on bare soil surfaces (Root Mean Standard Error of Calibration = 1.6 g/kg and Root Mean Standard Error of Prediction = 1.7 g/kg). With the presented RSU approach it is possible to filter out the influence of vegetation from the mixed spectra, and the residual soil spectra contain enough information for mapping of the SOC distribution within agricultural fields. This improves the applicability of airborne imaging spectroscopy for soil studies in temperate climates, since the use of this method can extend the flight-window.

6.1. Introduction

Soil Organic Carbon (SOC) is one of the key soil properties, because of its influence on plant growth, water holding capacity, soil structure, soil fertility while it governs many soil processes. Furthermore, the possibility to sequester CO_2 by increasing SOC stocks (Lal 2004) has resulted in a growing number of studies on the estimation of carbon fluxes between soil and atmosphere (Ryan and Law 2005). The determination of the SOC content, and in particular SOC stock in the topsoil of large units of land, using traditional methods is a time and money consuming procedure. Since the percentage of SOC may hold a large variability within a small area, it is practically impossible to sample large areas with sufficient measurements level to achieve continuous data coverage.

Imaging spectroscopy (IS) has proven to be a powerful tool to estimate the spatial distribution of soil properties (Barnes et al. 2003; Ben-Dor et al. 1999; Palacios-Orueta et al. 1999; Sørensen and Dalsgaard 2005). However, fractional vegetation cover has a large influence on spectral reflectance and it limits accurate quantification of soil properties. Siegal and Goetz (1977) reported that mineral absorption features can be obscured by as little as 10 percent green vegetation and that its presence may severely hinder or limit computer automated and photo-interpretative studies of multispectral data for soil and lithological discrimination. Murphy and Wadge (1994) studied the effect of living and non-photosynthetic vegetation on mapping of soil and rock types in arid areas, concluding that fractional vegetation cover complicates the identification of different soil/rock types and even makes identification of specific soil/rock types impossible due to masking of absorption features. A traditional solution is to mask out the areas with high vegetation cover, using vegetation indices with often case-specific threshold values (e.g. Wester et al. 1990). As a result soil, information for highly vegetated areas is lacking.

Some spectral unmixing techniques have been proposed to tackle the problem of vegetation influence in the case of soil classification. Luo et al. (2005) eliminated the vegetation effect with fully constrained spectral unmixing techniques, which increased the overall accuracy of the soil classification with 18 percent. Bierwirth (1990) removed the vegetation effect from reflectance data of geological materials by extrapolating the determined amounts of non-vegetated materials. He found that with careful modeling and depending on the instrument sensitivity, only a small spectral signal from geological materials may be required to obtain spatially meaningful geological information in strongly vegetated areas.

In this research we intend to determine the spatial distribution of Soil Organic Carbon (SOC) within a number of maize fields with varying fractional coverage. Quantitative determination of the SOC content using soil reflectance is well possible under controlled laboratory

conditions (e.g. Bartholomeus et al. 2008; Reeves III et al. 2000; Sørensen and Dalsgaard 2005; Stevens et al. 2008) or from airborne imaging spectrometers (Ben-Dor et al. 2002; Selige et al. 2006; Uno et al. 2005). However, these studies were carried out on bare soil surfaces, but the pixel size of airborne and spaceborne imaging spectrometers introduces spectral mixing problems in heterogeneous terrain, for example in areas which are partially vegetated.

Earlier studies have developed methods for quantitative estimation of soil properties in partially vegetated areas. However, for these cases vegetation was mostly used as proxy to estimate soil properties. Asner et al. (2003) related SOC and nitrogen field observations to fractional cover data for photosynthetic and non-photosynthetic vegetation and were able to show the trends in these soil properties at an ecosystem level. Kooistra et al. (2003) used vegetation development as a proxy to estimate SOC and Zn in floodplains. For discrimination of different soil units Schmidtlein (2004) used the spectral characterisation of plant functional types.

The spectral mixture problem has also been recognized by Gomez et al. (2008), who suggested that investments in spectral unmixing techniques are necessary to overcome this. Spectral unmixing techniques have played a central role in the analysis of remote sensing images with fractional vegetation cover over the past decades (Asner and Heidebrecht 2002; Garciá-Haro et al. 1996; Garcia and Ustin 2001; Sohn and McCoy 1997). In general, spectral unmixing algorithms are used to estimate the abundance of endmembers within a pixel and are well able to estimate the amount of vegetation in mixed pixels. In recent years, multivariate techniques like Partial Least Square Regression (PLSR) have been used with success to build SOC prediction models (Fidencio et al. 2002; Madari et al. 2006; Stevens et al. 2006; Udelhoven et al. 2003; Viscarra Rossel et al. 2006). This chemometric technique is now well-established for soil studies in a laboratory setup (Ben-Dor et al. 2008) and is also used for remote sensing of vegetation from imaging spectometers (LaCapra et al. 1996).

In this paper, we present a spectral unmixing based approach to eliminate vegetation influence from the spectral reflectance of mixed pixels, which results in a so-called residual soil spectrum. This procedure is named Residual Spectral Unmixing (RSU). The residual soil spectra are used in combination with PLSR to create a SOC-prediction model. The aim of this paper is 1) to discuss the influence of fractional vegetation cover on the accuracy of a PLSR model, 2) to develop a new spectral unmixing technique (RSU) to remove the vegetation spectral effects from the SOC signals and 3) to demonstrate the ability of RSU, in combination with PLSR to map the within-field variation of SOC in three partly covered maize fields using hyperspectral airborne images.

6.2. Materials and Methods

6.2.1 Study area description and data collection

The study area is located in the Belgian Lorraine region (49˚38'; 49˚43' N and 5˚27'; 5˚31' E) and is characterised by sandy-loam and loamy-sand soils (Haplic and Gleyic Luvisol, (FAO-ISRIC-ISSS 1998) with strong ferric components (Stevens et al. 2008). This site has a mean altitude of 350 m.a.s.l. with a rather flat topography, a mean annual temperature of 8.5ºC and an annual precipitation of 1013 mm.

A flight campaign with the AHS-160 sensor took place on the June 20, 2005 under clear sky and dry soil surface conditions. This sensor provides a total of 63 spectral bands covering the Visible (VIS: 430-700 nm), Near InfraRed (NIR: 700-1100 nm) and Short Wave InfraRed (SWIR: 1100-2540 nm) parts of the electromagnetic spectrum. The data were corrected for geographical, radiometric and atmospheric attenuations by the Central Data Processing Center of the Vlaamse Instelling voor Technologisch Onderzoek (CDPC-VITO) at Mol (Belgium). After geometric correction (direct georeferencing), atmospheric perturbations were removed using the MODTRAN4 radiative transfer code (Berk et al. 1999). The pixel size was 2.6 m.

A spatial subset (400*300 m) of a North-South AHS flight line was used for the estimation of vegetation cover, RSU and spatially continuous SOC prediction. This subset is located just south-east of the village of Bellefontaine and covers three fields, on which maize is grown. Vegetation was cleared for 69 plots of 7.5*7.5m and in each plot 9 soil samples were collected to a depth of 5 cm. These sub-samples were mixed and a representative sample was taken for SOC determination and spectral analysis in the laboratory. The SOC content was determined using the Walkley and Black method (Walkley and Black 1934). Observed SOC concentrations vary between 5.9 g/kg and 22.1 g/kg.

A Fieldspec Pro FR spectrometer (Analytical Spectral Devices Inc., Boulder CO) was used to collect in situ spectral reflectance of the soil surface in all plots. The soils were brought to the laboratory and air dried in the room atmosphere. Spectra of the dried soil samples were measured with the ASD Fieldspec Pro FR in the laboratory using a contact probe with a built in halogen lamp for illumination and the entire spectral range (350-2500 nm) of this population was used for further analysis. Further, the leaf reflectance of maize was measured in situ around the cleared plots, using an ASD contact probe with leafclip, that permits individual leaf measurements in constant conditions. For all Fieldspec measurements a white spectralon reference panel was used as a reference for the reflected radiation from the samples.

6.2.2 Simulation of vegetation cover influence on SOC prediction

During the time window of the flight, fields were partially covered with maize in an early phenological stage. Spectral signatures of the fractionally covered pixels show a combined spectral response, containing features from both soil and vegetation. This effect, called spectral mixing (Hapke 1981), can be found in all heterogenic pixels and can be linear or non linear (Roberts et al. 1993; Singer and McCord 1979). The maize plants were still small, reaching up to max 25 cm. Leaves were mainly erectophile, and due to the low number of leaves per plant (3-5) overlap of leaves as well within the plant as with neighboring plants was hardly present. Given the maize status and plant architecture we assumed that only linear mixing effects take place.

To quantify the influence of fractional vegetation cover on the accuracy of SOC prediction, an exploratory experiment was carried out. We simulated fractional vegetation cover by mixing a maize spectrum with bare soil spectra. Three soil samples (with SOC contents of 5.9 g/kg, 12.5 g/kg, and 22.1 g/kg) were used for mixing, and the fraction of maize was varied between 0 and 1. This was done according to the following formula:

$$R_{mix} = f_{maize} * R_{maize} + f_{soil} * R_{soil} \qquad (6.1)$$

in which R_{mix} is the reflectance of the mixed spectrum, R_{maize} is the reflectance of maize, R_{soil} is the reflectance of the soil sample, f_{maize} is the fraction of maize and f_{soil} is the fraction of soil. In these simulations f_{maize} and f_{soil} always sum up to 1. Next, the SOC content of the mixed spectra was predicted with a PLSR-model calibrated on all soil spectra measured under laboratory conditions (N=69). The calibration of this PLSR model is described in Stevens et al. (2008). The difference between observed and predicted amount of SOC is used as a measure of the error of SOC prediction in relation to the percentage of fractional maize cover. This experiment gives us an idea of the magnitude of the error caused by fractional vegetation cover.

6.2.3 Spectral unmixing

In this study, fractional maize cover is determined spatially using linear unmixing that takes into account the angular variability of endmember spectra, and the sun-sensor-target geometry of the image acquisition. The AHS sensor has a large field of view (90º), which results in a large variation in viewing geometry within a single scene. Especially for the maize fields contained in this imagery, profound bidirectional reflectance effects were encountered and needed to be resolved for.

We resolved these bidirectional reflectance effects in the maize vegetation using several sets of endmembers, based on the viewing angle. Due to the phenological stage of the maize is was not possible to derive pure maize spectra from the image, and canopy measurements in

the field were not feasible because of the small plant size. Therefore, maize endmembers were modelled using PROSAIL: which is a coupling of the PROSPECT model (Jacquemod and Baret 1990) that simulates leaf reflectance, and the SAIL model (Verhoef and Bach 1984) to simulate canopy reflectance. Throughout 100 simulation runs, the biophysical parameter values used by PROSAIL were set to encompass the natural variability of the maize fields contained in our study area. The vegetation endmembers were simulated using PROSAIL parameterized using biophysical parameters found in literature (Bach et al. 2005; Jacquemoud et al. 2000; Jacquemoud et al. 1995; Urso et al. 2004; Weiss et al. 2002). The parameters that were varied included the chlorophyll content, canopy water content, structural parameter, ratio of diffuse/ direct radiation, hotspot parameter and the leaf area index. Bare soil endmembers showed a much less apparent bidirectional reflectance effect than maize, and were for that reason taken from the image, without taken into account directionality of viewing. Bare soil pixels were selected using the Pixel Purity Index (PPI) (Boardman et al. 1995) on Minimum Noise Fraction (MNF) (Green et al. 1988) transformed image data. Before unmixing, simulated endmembers from the PROSAIL model were resampled to the spectral properties of the AHS-sensor, using both the AHS spectral band positioning and full width half maximum (FWHM).

Spectral unmixing was applied using the vegetation endmember sets to those parts of the image that were acquired under a similar viewing geometry as the endmember sets were simulated for. Every endmember set contained the 100 PROSAIL simulations varying in biophysical parametering. We then selected one unmixing result out of the 100 unmixing results for further processing, specifically the run which gave the lowest residual error (as given in the RMSE image). This is the spectral unmixing solution that is mathematically most sound as it leaves the smallest part of the spectrum unexplained. Therefore, it uses the combination of endmembers that matches the conditions in the field best. Ground truth maize cover was acquired at 14 points within the study area. The maize cover was determined by classifying digital photographs into vegetation and soil. The photographs were taken from a rod supported and nadir pointed camera at about five meters above the ground.

6.2.4 Residual spectral unmixing

Because the research is limited to maize fields, mixed pixels are composed of only two endmembers: bare soil and maize. Shadow can be a significant component as well, but due to the generally low fractional cover and good timing of the image acquisition during the day (around noon), resulting in small solar-zenith angles, this was considered to be of minor importance.

With two endmembers contributing to the spectral signature, the spectral mixture model can easily be inverted to a model that returns the residual spectrum of one of the components,

assuming homogeneity for the other. In this case, the spectral signature of bare soil in the pixel can be estimated when fractional maize cover and the spectral signature of maize are known rewriting eq. 6.1 as:

$$R_{soil} = (R_{mix} - (f_{maize} * R_{maize})) / f_{soil} \tag{6.2}$$

Applying RSU to all spectral bands returns a bare soil spectrum for the mixed pixel, which we will further name a residual soil spectrum. This results in a reflectance image of the bare soil, when applied to the entire image. Because of the computational effort we decided to use a single vegetation spectrum for residual unmixing instead of a different vegetation spectrum for each pixel. Therefore, a "general vegetation spectrum" has to be composed. In our case, this vegetation spectrum is an average of all modeled vegetation spectra used within the study area, weighted by the number of times a certain spectrum results in the lowest residual error in the unmixing model.

6.2.5 PLSR

The soil spectra of 57 plots derived from RSU were used to calibrate a carbon prediction model using PLSR. This was done using the ParLeS software of Viscarra Rossel (2008). The number of latent variables was chosen through minimization of the RMSE by leave-one-out cross-validation (RMSECV). The quality of the model was assessed through the Root Mean Square Error (RMSE) and Ratio of Peformance to Deviation (RPD). Subsequently, this PLSR model was used to create a spatially continuous SOC map based on all residual soil spectra within the three fields. Twelve plots that were left out of the calibration phase of the residual soil spectra based PLSR model were used for independent validation of the SOC map.

6.3. Results and Discussion

6.3.1 Influence of vegetation cover on SOC Prediction

The results of the spectral mixing experiment of soil and vegetation spectra shows that the PLSR based prediction of SOC is very sensitive to vegetation fraction (Figure 6.1). A small fractional maize cover already leads to an SOC overestimation of a few g/kg. Since SOC content in the study area is relatively low, this has a strong effect on the prediction accuracy. SOC is overestimated with 5 g/kg at a fractional maize cover of 0.04-0.06. This represents for the sample with the lowest SOC content (5.9 g/kg) an error of 85% of the measured value, which show that the prediction of SOC in vegetation influenced pixels is very unreliable. The differences between measured and predicted SOC content for several fractional covers is shown in Figure 6.1.
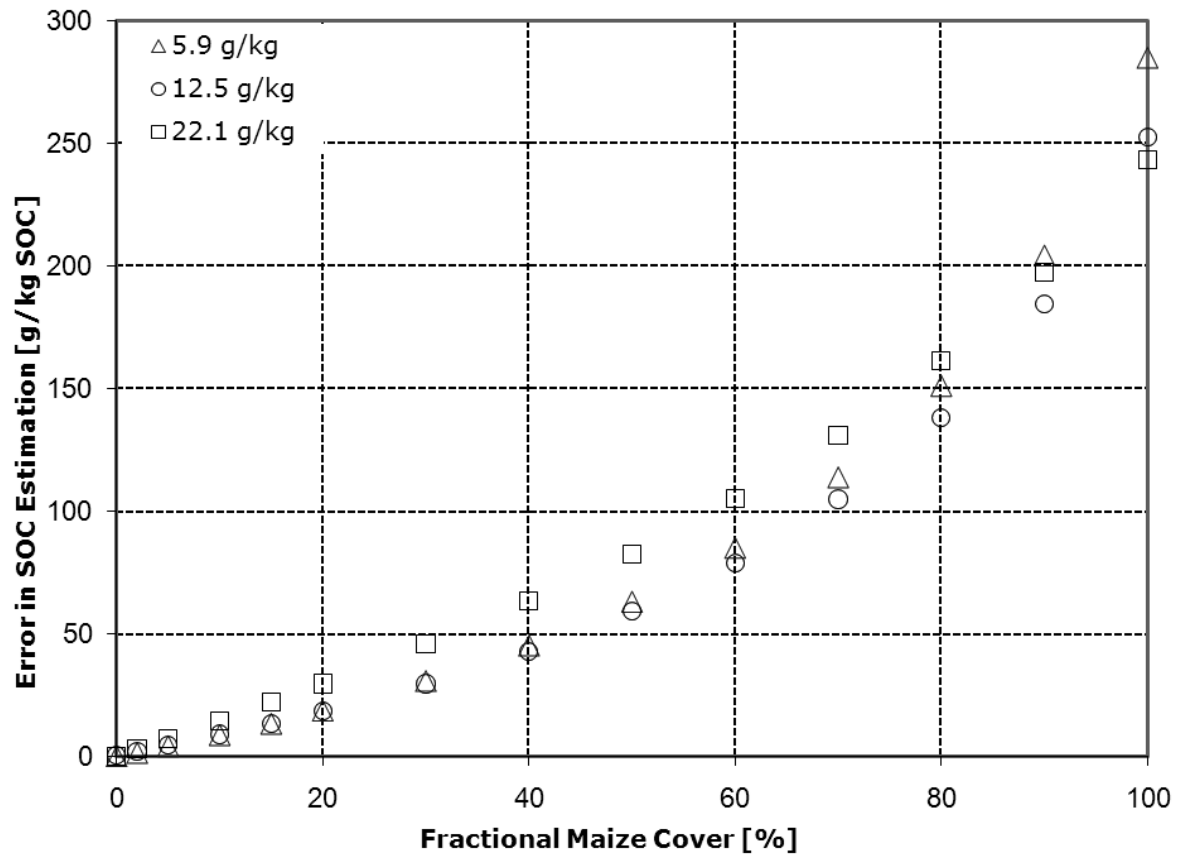
Figure 6.1. Error in soil organic carbon (SOC) estimation (g/kg) as a function of fractional maize cover (%). The different symbols indicate the three soil samples, with different SOC concentration, that were used for the analysis.

6.3.2 Spectral unmixing

The majority of fractional cover values lies between 0.12 and 0.60. Spatial patterns of fractional cover correspond well with the distribution observed in the fields (Figure 6.2). In the south-eastern part large amounts of weeds cause a higher fractional vegetation cover. A fractional cover of 0 shows up as expected at the plots where vegetation was cleared. Incidentally values above 1 are found. Although these are no valid unmixing results they can occur if the used set of endmembers is not able to reconstruct the reflectance of the pixel. This is the case for some pixels in the area with high weed coverage. Since weed has a slightly different reflectance than maize, errors in the unmixing results can be expected in this area. This is also shown when the unmixing results are compared with the fractional cover observed in the field. Most plots are accurately estimated (deviation <0.1), but larger deviations (deviations up to 0.25) show up in the more densely vegetated areas. Because there are only few pixels that show values above 1, the results are considered to be valid and used as they are. We chose not to constrain the unmixing results to sum to 1, since this forces the unmixing algorithm to come up with "correct" values without giving any measures about the quality of the unmixing procedure.
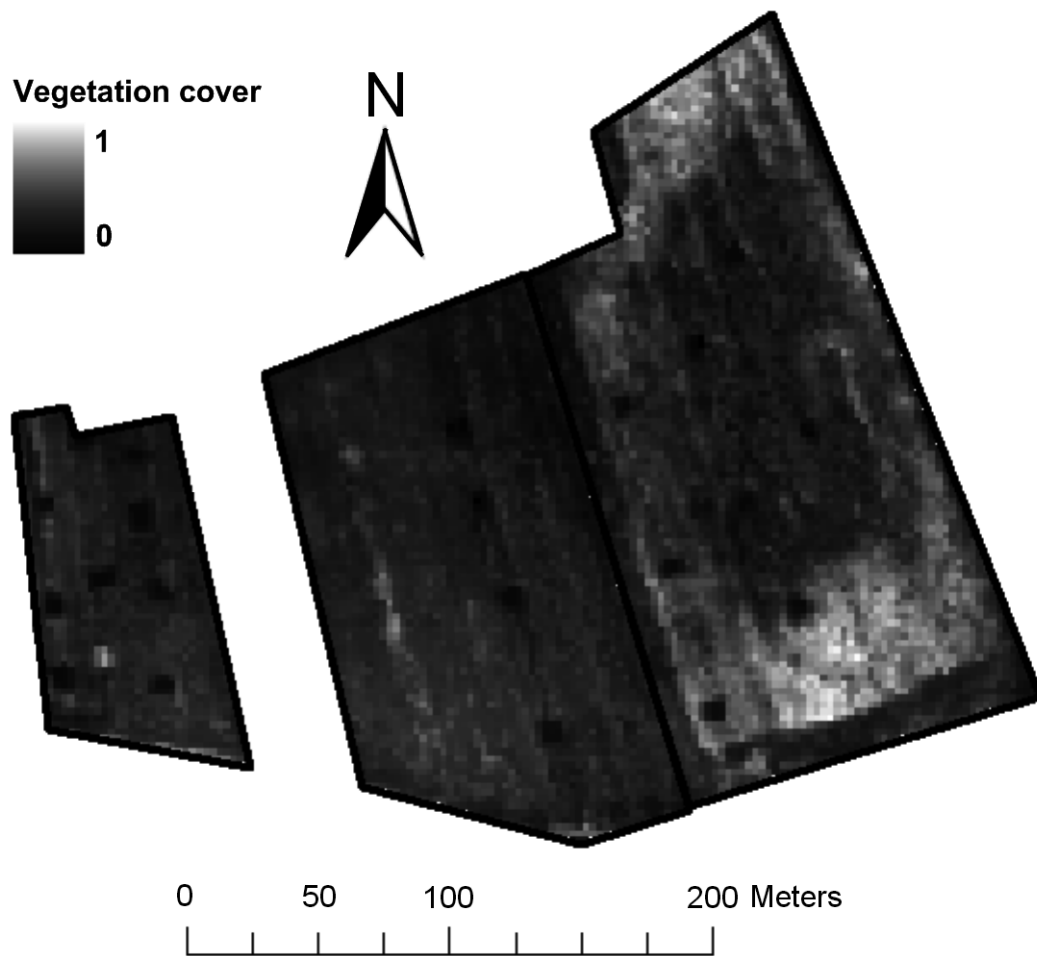
Figure 6.2. Fractional vegetation cover of the three studied fields

6.3.3. Residual spectral unmixing

Figure 6.3a shows the spectrum of the maize endmember used for RSU and a bare soil spectrum derived from the image. RSU results are shown in Figure 6.3 b-d, where the original and residual soil spectra of three locations with varying vegetation cover are shown. It can clearly be seen that the shape of the original mixed signatures is influenced by vegetation reflectance, with a more pronounced red edge as fractional cover increases. After RSU, the residual spectra all show a spectrum similar to the bare soil spectrum as shown in Figure 6.3a. Absorption features in the visible part of the spectrum, such as the absorption feature due to iron (around 0.7 µm), can be observed in the residual soil spectra.
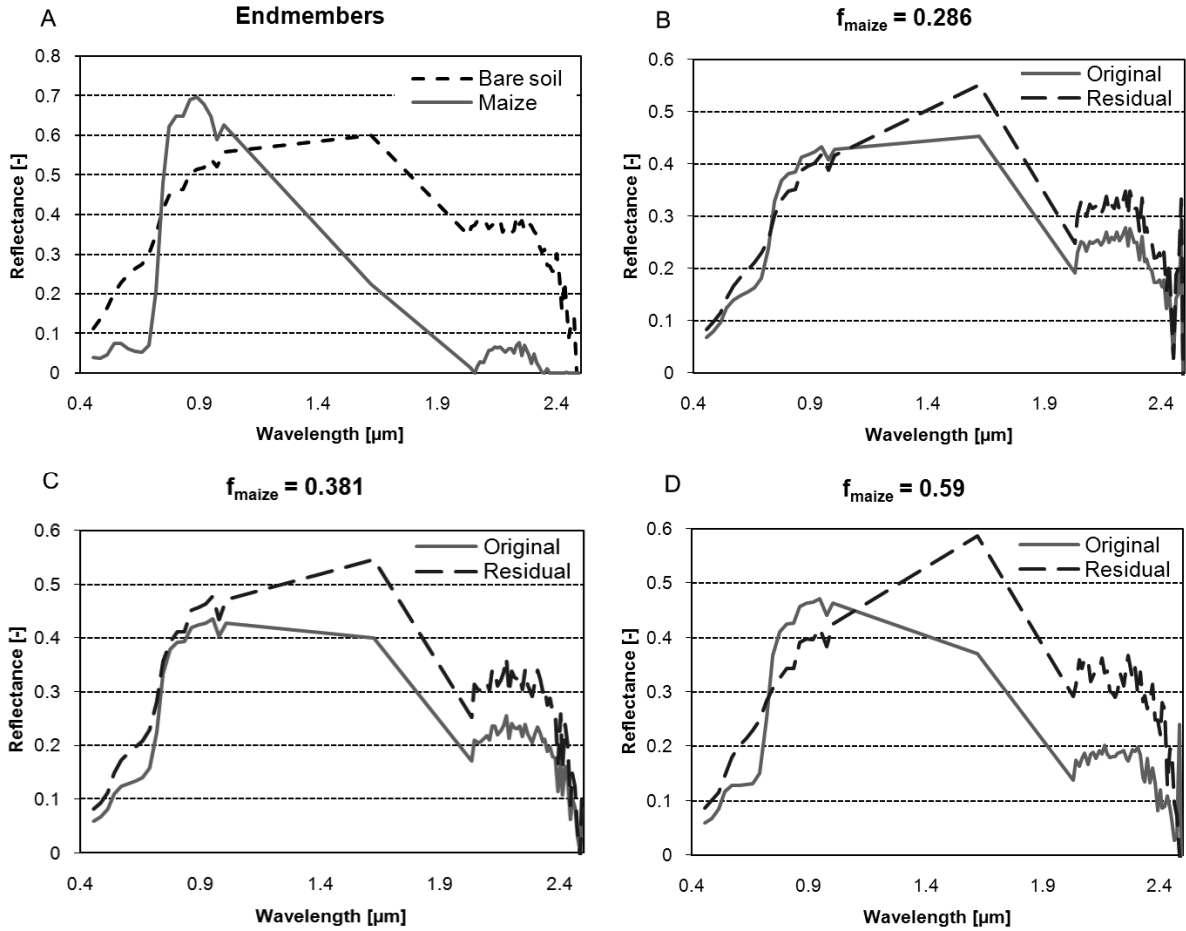
Figure 6.3: Reflectance before (grey lines) and after residual unmixing (dashed lines) of plots showing varying fractional vegetation cover. Endmembers are given in graph (a).

Inversion of the linear spectral mixing model is quite straightforward, but some difficulties may appear. While the use of the fractional cover from the unmixing model showing the lowest unresolved part intends to minimize the error, spectral unmixing always returns an error factor. This implies that the residual soil spectrum always contains deviations, introduced by the unresolved part of the unmixing model. Errors are introduced due to the nature of the unmixing algorithm, the variability within a single vegetation type (because of e.g. biochemical differences, varying leaf angle distribution), shadow effects and the variability in the soil spectrum. The difference between the image derived and field estimated fractional vegetation cover is often more than 0.05 in fractional cover, especially in the parts of the image where the fractional vegetation cover is high. Because of these inaccuracies it is not possible to use $1-f_{maize}$ instead of f_{soil} in equation 6.2. In our specific case this may result in negative residual soil reflectance values.

6.3.4 PLSR prediction of SOC

To predict the spatial distribution of SOC on the fields we used PLSR. For calibration of the PLSR model a set of 57 residual soil spectra was used. We selected not only residual spectra originating from the three fields of interest, but also a number of samples from neighboring bare soil fields. This choice was made for practical reasons; in total only 22 plots were sampled on the fields of interest and 12 of these plots had to be used for independent validation of the final carbon distribution map as well. This would have left too few points for proper calibration of the PLSR model. Some of the additional plots from the neighboring fields had no fractional vegetation cover, but these bare soil samples show no deviations in the PLSR calibration, when compared to the soil spectra obtained by RSU. Due to the poor signal to noise ratio in the shortwave infrared, we used only the first 30 spectral bands of AHS, ranging from 0.45- 2.12 µm. Reflectance spectra were transformed to Log(1/R) data and smoothed with a Savitzky-Golay filter (filter size 2, 3ʳᵈ order polynomial function). The PLSR model is validated according to the "leave one out cross validation" procedure, and yields an optimal model fitting using 8 latent variables with a R^2 of 0.619 and a RMSECV of 1.70 g/kg. With a ratio of prediction to deviation (RPD) of 1.60, the accuracy of the model falls in an intermediate category according to the definition of Chang et al. (2001). The scatter plot of observed against predicted SOC content is shown in Figure 6.4.
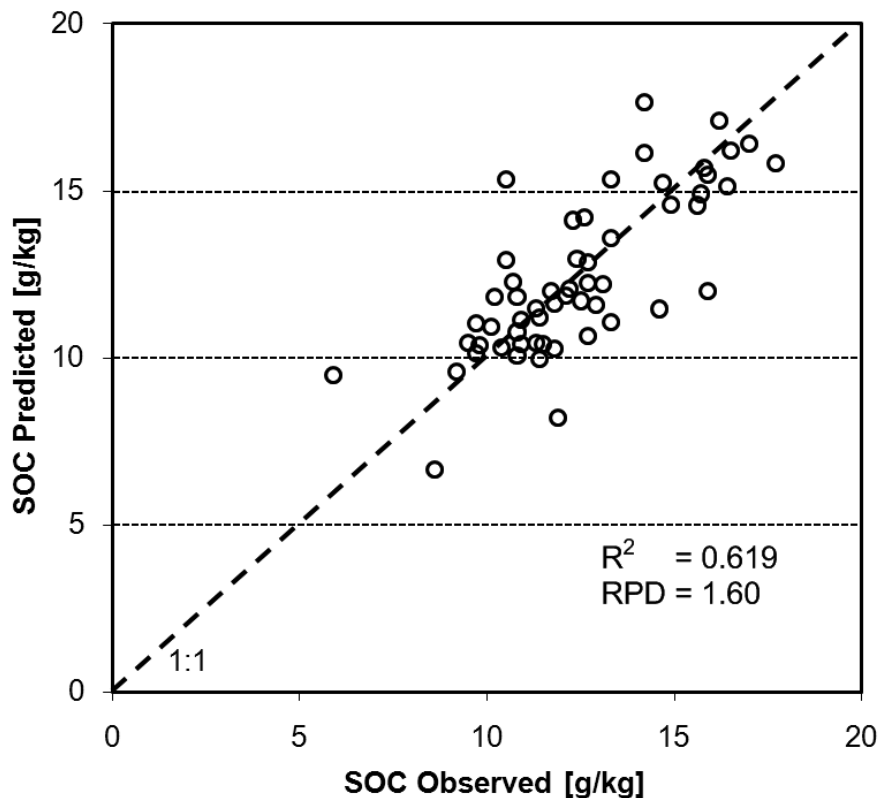

Figure 6.4: Observed versus predicted SOC content of the calibration dataset, based on PLSR of the RSU corrected AHS data (n=57).
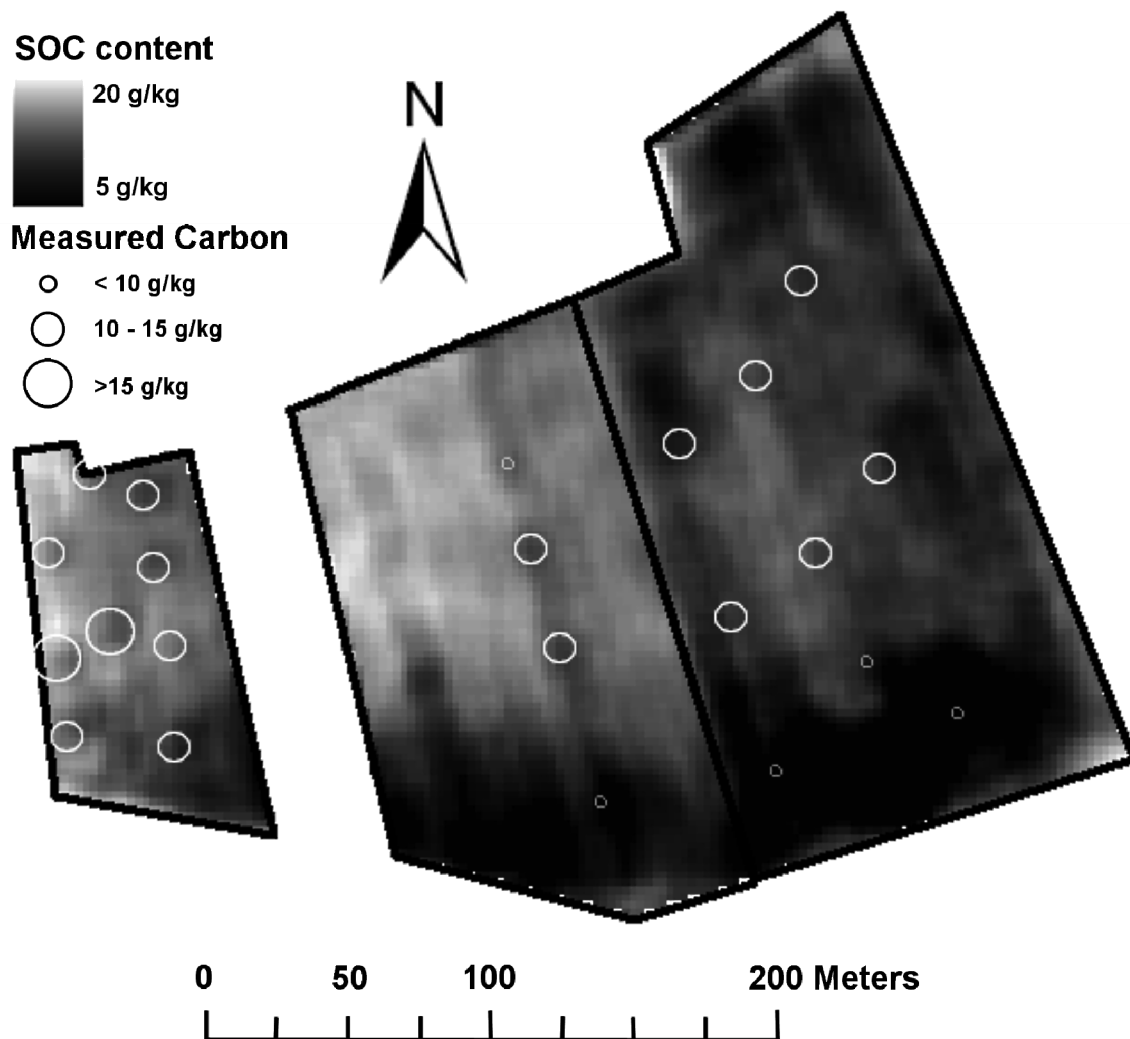
Figure 6.5: Map of soil organic carbon (SOC) concentrations for the three sampled fields. The spatial distribution of SOC is determined with PLSR, based on residual soil spectra. The circles indicate the locations where soil samples were taken for calibration and validation.

Finally, the PLSR model was used to predict the SOC content of the maize fields using the residual soil spectra obtained from the AHS image. Because of the high variation on short distances in the raw results, which are caused by the large variation on short distance in the unmixing results and neglecting the unmixing RMSE in the residual unmixing steps, we used a 5x5 low pass filter to smooth the outcomes. This resulted in the final SOC distribution map shown in Figure 6.5, which clearly shows the within field variation of SOC. The South-Eastern part of the fields shows a low carbon content, while it increases towards the North and West. The predicted SOC values show no correlation ($R^2 = 0.06$) with the fractional vegetation cover.

The geographical locations of soil samples taken within the 3 fields are also displayed in Figure 6.5. and the relative size of the circle indicates the SOC content. As mentioned before, twelve plots located within the fields were left out of the PLSR model calibration, in order to be used for validation of the final carbon distribution map. The predicted SOC content of these validation plots is plotted against the observed SOC (Figure 6.6). The 1:1 line is nicely followed and no large deviations are shown. This can also be seen from the R^2-values (R^2cal = 0.60; R^2val = 0.57) , the root mean standard error of calibration (RMSEC) of 1.6 g/kg and root mean standard error of prediction (RMSEP) of 1.7 g/kg. These values are comparable to the accuracies obtained by Uno et al. (2005) and Chen et al. (2008), who determined SOC contents using VIS-NIR spectroscopic techniques over bare soils. RMSE values reported by Uno et al. (2005) were higher than the ones calculated in our study, but the range of SOC values was larger as well. Chen et al. (2008) grouped fields according to their image properties, and found for the different groups RMSE values ranging between 0.89 g/kg and 2.16 g/kg. In our study, the two samples showing the largest deviation are located in the area with high vegetation cover, which corresponds with the large inaccuracy of the unmixing for this area.
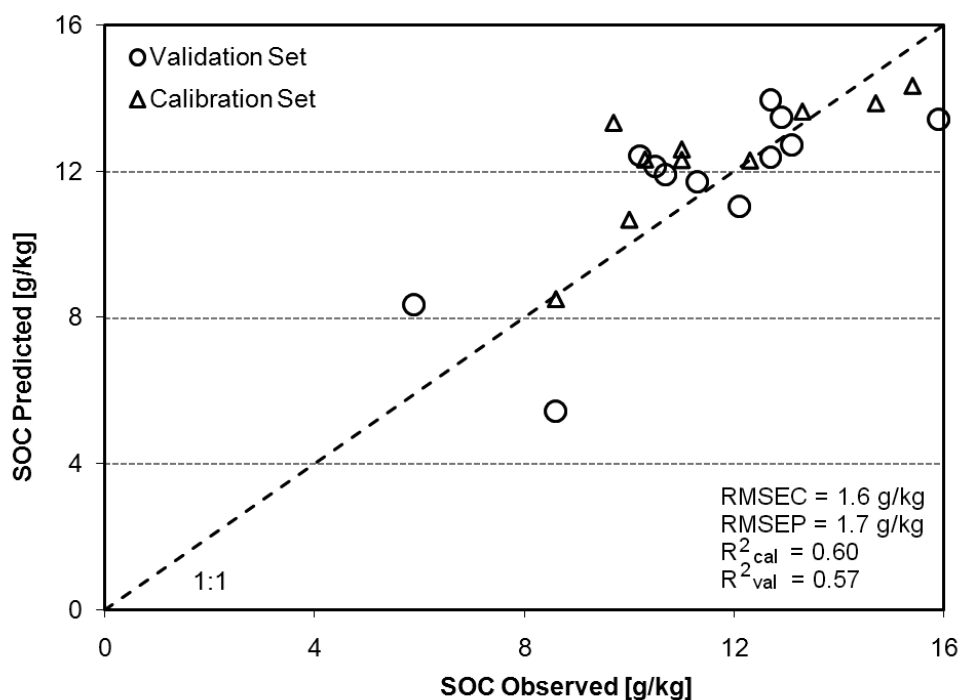

Figure 6.6: Observed vs predicted soil organic carbon (SOC) of selected soil samples, located within the three studied fields.

The results of this study show that RSU works in our study area, and allows the quantification of SOC when fields are partially covered with vegetation. The quality of the

analysis can be increased by improving the spectral unmixing results, which is the most critical step in the RSU. This can be achieved by the development of sensors with better signal-to-noise ratios and advanced unmixing techniques. In this agro-ecosystem the situation was relatively easy to model, since only two main endmembers are present, which is often the case in agricultural areas. The potential of the methodology in more complex areas (e.g. more natural ecosystems), where multiple endmembers can occur, has to be investigated. We believe that RSU can be helpful for the mapping of other soil properties as well, especially for those that have distinct absorption features compared to vegetation.

6.4. Conclusions

The results of this study show that it is possible to map SOC concentration in the plough layer of croplands with fractional vegetation cover using RSU. With the RSU approach it is possible to filter out the influence of vegetation from the mixed spectra, and the residual soil spectra contain enough information to map the SOC distribution within agricultural fields. This improves the applicability of airborne imaging spectroscopy for soil studies in temperate climates, since the use of this method can extend the flight-window.

Acknowledgements

References

Asner, G.P., Borghi, C.E., & Ojeda, R.A. (2003). Desertification in Central Argentina: Changes in Ecosystem Carbon and Nitrogen from Imaging Spectroscopy. *Ecological Applications, 13*, 629-648

Asner, G.P., & Heidebrecht, K.B. (2002). Spectral unmixing of vegetation, soil and dry carbon in arid regions: Comparing multi-spectral and hyperspectral observations. *Int. J. Remote Sens., 23*, 400-410

Bach, H., Begiebing, S., Waldmann, D., & Rowotzki, B. (2005). Analyses of hyperspectral and directional data for agricultural monitoring using the canopy reflectance model SLC progress in the upper Rhine valley and Baasdorf test-sites. *Proc. Of the 3rd ESA CHRIS/Proba Workshop, Frascati, Italy*

Barnes, E.M., Barnes, E.M., Sudduth, K.A., Hummel, J.W., Lesch, S.M., Corwin, D.L., Yang, C., Daughtry, C.S.T., & Bausch, W.C. (2003). Remote- and ground-based sensor techniques to map soil properties. *Photogrammetric Engineering and Remote Sensing, 69*, 619-630

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., & Spaargaren, O.S.P. (2008). Spectral reflectance based indices for soil organic carbon quantification. *Geoderma, 145*, 28-36

Ben-Dor, E., Irons, J.R., & Epema, G.F. (1999). *Soil Reflectance*. New-York: Wiley & Sons Inc.

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

Ben-Dor, E., Taylor, R.G., Hill, J., Demattê, J.A.M., Whiting, M.L., Chabrillat, S., & Sommer, S. (2008). Imaging Spectrometry for Soil Applications. *Advances in Agronomy, 97*, 321-392

Berk, A., Anderson, G.P., Acharya, P.K., Chetwind, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W., & Alder-Golden, S.M. (1999). *Modtran4 User's Manual.* . Hanscom, USA: Air Force Research Laboratory

Bierwirth, P.N. (1990). Mineral mapping and vegetation removal via data-calibrated pixel unmixing, using multispectral images. *International Journal of Remote Sensing, 11*, 1999-2017

Boardman, J.W., Kruse, F.A., & Green, R.O. (1995). Mapping target signatures via partial unmixing of AVIRIS data: in summaries. *Fifth JPL Airborn Earth Science Workshop, JPL Publication, 95*, 23-26

Chang, C.W., Laird, D.A., Mausbach, M.J., & Hurburgh Jr., C.R. (2001). Near-infrared reflectance spectroscopy - Principal components regression analysis of soil properties. *Soil Science Society America Journal, 65*, 480-490

Chen, F., Kissel, D.E., West, L.T., Adkins, W., Rickman, D., & Luvall, J.C. (2008). Mapping soil organic carbon concentration for multiple fields with image similarity analysis. *Soil Science Society of America Journal, 72*, 186-193

FAO-ISRIC-ISSS (1998). World reference base for soil resources. In, *World Soil Resources Report 84*. Rome, Italy: FAO

Fidencio, P.H., Poppi, R.J., De Andrade, J.C., & Cantarella, H. (2002). Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. *Commun. Soil Sci. Plant Anal., 33*, 1607-1615

Garciá-Haro, F.J., Gilabert, M.A., & Melia, J. (1996). Linear spectral mixture modelling to estimate vegetation amount from optical spectral data. *International Journal of Remote Sensing, 17*, 3373-3400

Garcia, M., & Ustin, S.L. (2001). Detection of interannual vegetation responses to climatic variability using AVIRIS data in a Coastal Savanna in California. *IEEE Transactions on Geoscience and Remote Sensing, 39*, 1480-1490

Gomez, C., Viscarra Rossel, R.A., & McBratney, A.B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma, 146*, 403-411

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions in Geoscience and Remote Sensing, 26*, 65-74

Hapke, B. (1981). Bidirectional Reflectance Spectroscopy: 1. Theory. *Journal of Geophysical research, 86*, 3039-3054

Jacquemod, S., & Baret, F. (1990). PROSPECT: A model of Leaf Optical Properties Spectra. *Remote Sensing of Environment, 34*, 75-91

Jacquemoud, S., Bacour, C., Poilvé, H., & Frangi, J.-P. (2000). Comparison of Four Radiative Transfer Models to Simulate Plant Canopies Reflectance: Direct and Inverse Mode. *Remote Sens. Environment, 74*, 471-481

Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M., & Jaggard, K. (1995). Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sens. Environment, 52*, 163-172

Kooistra, L., Leuven, R.S.E.W., Wehrens, R., Nienhuis, P.H., & Buydens, L.M.C. (2003). A comparison of methods to relate grass reflectance to soil metal contamination. *International Journal of Remote Sensing, 24*, 4995-5010

LaCapra, V.C., Melack, J.M., Gastil, M., & Valeriano, D. (1996). Remote sensing of foliar chemistry of inundated rice with imaging spectrometry. *Remote Sensing of Environment, 55*, 50-58

Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma, 123*, 1-22

Luo, H., Ye, H., Ke, Y., Pan, J., Gong, J., & Chen, X. (2005). Removing vegetation using unsupervised fully constrained least squares linear spectral mixture analysis method in soils surveying by remote sensing. In, *Proceedings of SPIE - The International Society for Optical Engineering* (pp. 90-98)

Madari, B.E., Reeves III, J.B., Machado, P.L.O.A., Guimaraes, C.M., Torres, E., & McCarty, G.W. (2006). Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma, 136*, 245-259

Murphy, R.J., & Wadge, G. (1994). The effects of vegetation on the ability to map soils using imaging spectrometer data. *International Journal of Remote Sensing, 15*, 63-86

Palacios-Orueta, A., Pinzo?n, J.E., Ustin, S.L., & Roberts, D.A. (1999). Remote sensing of soils in the Santa Monica Mountains: II. Hierarchical foreground and background analysis. *Remote Sensing of Environment, 68*, 138-151

Reeves III, J.B., McCarty, G.W., & Meisinger, J.J. (2000). Near infrared reflectance spectroscopy for the determination of biological activity in agricultural soils. *J. Near Infrared Spectroscopy, 8*, 161-170

Roberts, D.A., Smith, M.O., & Adams, J.B. (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment, 44*, 255-269

Ryan, M.G., & Law, B.E. (2005). Interpreting, measuring, and modeling soil respiration. *Biogeochemistry, 73*, 3-27

Schmidtlein, S. (2004). Coarse-scale substrate mapping using plant functional response types. *Erdkunde, 58*, 137-151

Selige, T., Boehner, J., & Schmidhalter, U. (2006). High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma, 136*, 235-244

Siegal, B.S., & Goetz, A.F.H. (1977). Effect of vegetation on rock and soil type discrimination. *Photogrammetric Engineering and Remote Sensing, 43*, 191-196

Singer, R.B., & McCord, T.B. (1979). Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In, *Proceedings of the 10th lunar and planetary science conference* (pp. 1835-1848). Houston, USA

Sohn, Y., & McCoy, R.M. (1997). Mapping desert shrub rangeland using spectral unmixing and modeling spectral mixtures with TM data. *Photogrammetric Engineering and Remote Sensing, 63*, 707-716

Sørensen, L.K., & Dalsgaard, S. (2005). Determination of clay and other soil properties by Near Infrared Spectroscopy. *Soil Sci. Soc. Am. J., 69*, 159-167

Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., & Ben-Dor, E. (2008). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma, 144*, 395-404

Stevens, A., Van Wesemael, B., Vandenschrick, G., Touré, S., & Tychon, B. (2006). Detection of Carbon Stock Change in Agricultural Soils Using Spectroscopic Techniques. *Soil Science Society of America Journal, 70*, 844-850

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 and Soil, 251*, 319-329

Uno, Y., Prasher, S.O., Patel, R.M., Strachan, I.B., Pattey, E., & Karimi, Y. (2005). Development of field-scale soil organic matter content estimation models in Eastern Canada using airborne hyperspectral imagery. *Canadian Biosystems Engineering / Le Genie des biosystems au Canada, 47*, 1.09-01.14

Urso, G.D., Dini, L., Vuolo, F., Alonso, L., & Guanter, L. (2004). Retrieval of Leaf Area Index by Inverting Hyper-Spectral, Multi-Angular CHRIS/Proba Data from SPARC 2003. *2nd CHRIS/Proba Workshop, Frascati, Italy*

Verhoef, W., & Bach, H. (1984). Light Scattering by Leaf Layers with Application to canopy Reflectance Modeling: The SAIL Model. *Remote Sensing of Environment, 16*, 125-141

Viscarra Rossel, R.A. (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems, 90*, 72-83

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., & Skjemstad, J.O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma, 131*, 59-75

Walkley, A., & Black, I.A. (1934). An estimation of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science, 37*, 29-37

Weiss, M., Baret, F., Leroy, M., Hautecoeur, O., Prevot, L., & Bruguier, N. (2002). Validation of neural net techniques to estimate canopy biophysical variables from remote sensing data. *Agronomie, 22*, 547-553

Wester, K., Lundén, B., & Bax, G. (1990). Analytically processed Landsat TM images for visual geological interpretation in the northern Scandinavian Caledonides. *ISPRS Journal of Photogrammetry and Remote Sensing, 45*, 442-460

Chapter 7

Synthesis

7. Synthesis

7.1 Discussion and conclusions

The topic of this thesis was how to determine soil properties in partially vegetated areas, based on the measured spectral reflectance. This was achieved by the development of empirical models, which can be divided in two classes: multivariate models (PLSR) and univariate spectral indices based models. The objectives were to investigate the robustness of VNIR spectroscopy based soil property prediction models and the influence of vegetation on these models. Furthermore, methods were developed for quantitative mapping of soil properties in fractionally vegetation covered agricultural fields. To achieve these objectives the questions proposed in section 1.4 were addressed, which are answered and discussed below.

A. Which scaling issues have to be considered when using VNIR spectroscopy for estimation of soil properties?
In chapter 2, several scaling issues that occur in VNIR spectroscopy are described, and most of them play a role in this research. Spatial upscaling needs to be dealt with, when the step from point spectroscopy to imaging spectroscopy is made. Mixed pixels, in the agro-ecosystem mainly because of fractional vegetation cover, limit the direct transition of techniques and models from point spectroscopy to imaging spectroscopy. Several techniques to solve the spatial scaling issue are discussed in Malenovsky et al. (2007). In this thesis, spectral unmixing was chosen to obtain sub-pixel soil information (chapter 2).
The use of different sensors to acquire spectral information (chapter 3, 5 and 6), introduces spectral scaling problems. This complicates model transfer from one sensor to another. In this thesis, this point was not investigated, but PLSR based models developed for one sensor can not directly be transferred to another sensor with deviating spectral dimensions, due to requirements to the matrix dimensions. A lower spectral resolution (less spectral bands) does not necessarily lead to less accurate models (Udelhoven et al. 2003), but PLSR models have to be calibrated for each spectral set-up. For this reason, the possibility to use spectral indices for the estimation of soil properties is investigated (chapter 4 and 5). Although transfer of indices from one sensor to another is not always directly possible, the relation between absorption features and physical properties remains. Therefore, model behavior can better be predicted and adapted to the new spectral configurations, as long as minimum requirements for specific spectral bands are met.

Finally, directional scaling issues should be considered. First, directionality should be taken into account when measurements of different setups are combined. The poor transferability

of PLSR models from the one campaign to another (chapter 3, discussed below) shows that a fixed measurement protocol is essential. This should also include the directional setup, both from the sensor and the illumination source, which becomes particularly important when models calibrated under laboratory conditions are applied in the field or on image data. Because of the differences in illumination and observation angles, this implies a conversion from biconical reflectance factor (BRF) to hemispherical conical reflectance (HCRF) (Schaepman-Strub et al. 2006). Furthermore, combining measurements of point spectrometers with varying field of views or viewing angles, can be a cause of lower model stability. It was shown, that the accuracy of iron indices depends on the illumination and viewing geometry (Bartholomeus and Mulder 2008). This may be the case for PLSR based models and other soil properties as well. Second, directionality has to be considered in the spatial domain. The use of imaging spectrometers with a wide viewing angle introduces variation in angular effects. This will have influence on the measured soil reflectance (Cierniewski et al. 2004) and affect the estimated bare soil properties, but will be of major importance when fractional vegetation cover is present.

B. What is the potential of VNIR spectroscopy for the prediction of soil properties and what is the stability of the calibrations?

Chapter 1 and 3 describe the different forms (laboratory, field or imaging spectroscopy) in which VNIR spectroscopy can be used to determine soil properties. Both approaches using point spectrometers yield good results for SOC and iron oxides. In table 7.1 the RPD and RMSE values of chapter 3-6 are summarized and compared. To provide a full overview some additional results are shown.

For estimation of SOC, PLSR in combination with both field spectroscopy and laboratory spectroscopy reaches RPD values above 2, indicating that these PLSR models accurately predict SOC. The RMSEP is comparable to the values found in previous studies, as summarized in Viscarra Rossel et al. (2006). PLSR in combination with imaging spectroscopy has a RPD of 1.47 and corrected RMSECV of 1.4 g/kg. These values are comparable to the accuracies obtained by Uno et al. (2005) and Chen et al. (2008), who determined SOC contents using VIS-NIR spectroscopic techniques over bare soils. RMSE values reported by Uno et al. (2005) were higher than the ones calculated in our study, but the range of SOC values was larger as well. Chen et al. (2008) grouped fields according to their image properties, and found for the different groups RMSE values ranging between 0.89 g/kg and 2.16 g/kg.

The accuracy of field spectroscopy is equivalent to laboratory spectroscopy, when measuring SOC under specific surface conditions (low variation in moisture content of the soil surface, low roughness, absence of vegetation) and when appropriate pre-treatments to extract information from noisy spectra are applied. The results in chapter 3 show that PLSR is able

to reach accuracies comparable to the standard analytical method (1-2 g/kg SOC, Walkley-Black), when the set is limited to a single soil type. In general, the potential of VNIR spectroscopy is high, since it requires little sample preparation, especially in the field setup. The technique can thus be potentially used for monitoring studies, where speed is a valuable advantage.

The results in chapter 4 show that biochemistry based spectral indices can be used to estimate SOC for a dataset composed of nine soil types. Prediction accuracies based on spectral indices are comparable to accuracies of PLSR based predictions for this dataset, which has a large variation in SOC content. The performance indicators for the evaluated relations between SOC and spectral reflectance are summarized in table 4.2. PLSR shows the overall best results and meets the requirements of a good model, based on the RPD value. For the spectral indices, the biochemical constituent absorption dip based index '1/Area 2050-2200 nm' and the visible wavelengths based index '1/Slope 400-600 nm' show the overall highest R^2-values. The RPD of the index '1/Slope 400-600' ($RPD_{set\ 1}$=3.18 ; $RPD_{set\ 2}$=2.72) is comparable to the prediction accuracy achieved with PLSR. These results show that indices based on the VIS or SWIR spectral region can be used for the quantification of SOC. The predictive power of VIS based indices is comparable to the results obtained with PLSR. SWIR based indices have a somewhat lower predictive power.

When more soil types are included, as is the case in chapter 4, the RPD values of the PLSR model are higher than when only a single soil type is used (chapter 3). However, the RMSEC and RMSEP values are much higher as well, because the variation in SOC within the dataset is much larger (table 4.2). Therefore, with a higher range in input values and more variability in soil types accurate calibrations can be achieved in relation to the variation within the dataset, but it is not possible to reach prediction accuracies comparable to standard analytical methods. This corresponds with the findings of McCarty et al. (2002). They stated that a greater variability in the training phase of a statistical model may lead to an improved robustness of the model and an increased ability to characterize a diverse range of samples, but it may also result in a decrease in prediction accuracy.

For the PLSR analysis in chapter 4, the wavelengths contributing most to the prediction model are found in the VIS (maximum at 650 nm), around the water absorption feature (1400 and 1900 nm), and in the SWIR region (around 2212 nm). The high loadings for the VIS are consistent with other research findings (e.g. Kooistra et al. 2001, Viscarra Rossel et al. 2006). High PLSR loadings for the water absorption features are also reported by Udelhoven et al. (2003) and Viscarra Rossel et al. (2006). Although air-dried samples were used for analysis, there is still remaining water absorbed on the surface of organic matter and OH can be present in organic compounds (Zornoza et al. 2008). The high loading for 2212 nm is less straightforward to clarify. On the one hand, a minor absorption feature due to OH is known

around 2.2 μm (Ben-Dor 2002), but cellulose, starch and lignin have known absorption features in this spectral region as well (Curran 1989). Having multiple absorbers active at the same wavelengths, makes it difficult to determine which absorber causes the high loading when predicting SOC. Wavelengths in the range of 1700-1800 nm, which correlated well with SOC in the studies of Dalal and Henry (1986), Sudduth and Hummel (1991), and Shepherd and Walsh (2002), did not show high loadings with PLSR, nor could indices active in this wavelength region be constructed. Concluding, the loadings show that PLSR selects the spectral ranges that correspond with the spectral regions used by the indices that give good results. However, the importance of the water absorption features in the PLSR model will cause difficulties when the model has to be used under field conditions or with imaging spectroscopy. Water vapor in the atmosphere limits the radiance to reach the soil surface and soil moisture content often shows variation in time and place.

Table 7.1: Performance of PLSR models to estimate soil properties

	Method	Chapter	Laboratory spectroscopy		Field spectroscopy		Imaging spectroscopy	
			RPD	RMSE	RPD	RMSE	RPD	RMSE
SOC_{local}	PLSR	3	2.03	0.7 g/kg[a]	2.11	0.7 g/kg[a]	1.47	1.7 g/kg[a]
SOC_{global}	PLSR	4	3.43	31.8 g/kg[a]	-	-	-	-
SOC_{global}	Indices	4	3.18	34.3 g/kg[a]	-	-	-	-
$SOC_{vegetated}$	RSU+PLSR	6	-	-	-	-	1.49	1.7g/kg[a]
Fe_d	Indices	5	1.38	25.9 g/kg[a]	1.07	35.0 g/kg[a]	0.65	57.9 g/kg[a]
Fe_d[c]	PLSR	-	2.32	19.6 g/kg[b]	-	-	-	-

RPD = Ratio of Performance to Deviation, RMSE = Root Mean Square Error, PLSR = Partial Least Squares Regression, RSU = Residual Spectral Unmixing, [a]root mean square error of prediction, [b]root mean square error of cross validation, [c]data not presented in previous chapters.

As shown in chapter 5, the spectral range from UV to NIR gives an indication about the dithionite extractable iron content. Four different indices were constructed The standard deviation or average of a continuum-removed spectrum in the VIS and NIR part of the spectrum, the area of the absorption dip around 550 nm and the Redness Index give comparable results. In table 5.3, it is shown that the use of laboratory measurements yields the best result, although the models are not good enough to give accurate quantitative estimates of the iron content. Field spectroscopy and imaging spectroscopy give less accurate estimates. These overall results, with relatively low RPD scores for all methods, are a result of the limited spectral range used in this study. PLSR yields a RPD of 1.48 when laboratory spectra are resampled to the ROSIS spectral range and resolution, which is comparable to the results achieved with indices. If the full spectrum (350-2500 nm) is used for PLSR on laboratory spectra, a more accurate model can be fitted (RPD = 2.32), as shown in table 7.1.

Imaging spectroscopy lacks behind in prediction accuracy when the derived soil properties are studied at the pixel level (table 7.1). The lower prediction capacities are the result of 1) instrumentation specifications (spectral resolution, low instrument signal-to-noise ratios), 2) disturbing external factors (atmospheric attenuation, geometric and optical distortions,

mixed pixels) and 3) internal factors (soil moisture, structure). Sensor development with better signal-to-noise ratios is an ongoing process. The internal factors can not be controlled when measuring with imaging spectroscopy (nor with field spectroscopy), but the correction for external factors is continuously improved (Gao et al. 2009).

The quality of products derived from imaging spectroscopy, is often determined at the pixel level (e.g. Lagacherie et al. 2008). This is understandable, since bridging the gap between the spatial scale on which ground truth is gathered and the spatial scale at which imaging spectrometers operate, is most straightforward in this way. However, for several applications, a per pixel knowledge of soil properties is not required. In precision agriculture, variable rate application of fertilizer or herbicides is usually done by dividing individual fields in management zones (Khosla et al. 2002). Imaging spectroscopy is well capable to map the within field variation in soil properties, and can be used to support this zoning. To monitor changes in soil organic matter (SOM) at the regional scale, Hanegraaf et al. (2009) state that monitoring systems require an approach in which time-series of soil organic matter content are made on a field basis. In chapter 5 and 6, it was shown that the derived soil property maps reflect the observed variation in the fields. In chapter 3, it was shown that, because of the high sampling density, imaging spectroscopy is able to detect small temporal changes in carbon stock on the parcel level. By just looking at the per pixel accuracy, the usability of imaging spectroscopy derived soil property maps is underestimated.

In chapter 3, the transferability of PLSR based SOC models from one year to another was investigated. Within the same area, with similar soil types and ranges in SOC, it was shown that when field measurements of different years are combined, accurate estimates could not be made (table 3.4). This can have several causes like stability of instrumentation or slight deviations in experimental setup. Therefore, the main message is that practical implementation of spectroscopy requires stable measurement conditions, good calibrations and standard sampling protocols.

VNIR spectroscopy under controlled conditions has become well accepted by the soil science community, because the technique is fast, easy to use, and yields good results. Multivariate statistical models, like PLSR, have become the standard in soil spectroscopy (Fidencio et al., 2002; Udelhoven et al., 2003; Madari et al., 2006; Stevens et al., 2006; Viscarra Rossel et al., 2006). Nevertheless, the results in this thesis show that it will be difficult to develop this technique beyond the level of case-specific calibrations. This limits a larger operational implementation of VNIR spectroscopy. If we want to get rid of local calibrations, we will have to accept lower accuracies, but we gain in the ease of use. Users, willing to choose one of the techniques for a particular application, have to weigh the accuracy against the area to be covered and the costs. While laboratory spectroscopy has the advantage to allow stable

calibration through time, it still requires sample preparation (drying, sieving, grinding). Field spectroscopy can reach an accuracy comparable to laboratory spectroscopy (table 3.2), but it requires a calibration before each campaign. Imaging spectroscopy is a practical way to spatially evaluate soil properties and can be an useful technique to determine soil properties and track changes at large scales. For some applications, the large number of samples (1600 pixels/ha at a 2.5 m spatial resolution) compensates for the lower accuracy. Imaging spectroscopy also requires a calibration before each campaign, due to difficulties in spectral calibration. An important question that remains is how many calibration samples are needed on what extent of the study area, to achieve a model calibration with acceptable accuracy.

C. How sensitive are models to variation in soil type and can models be used to predict soil properties for soil types that are not included in the model calibration?

To test the model sensitivity of SOC prediction models to variation in soil types, a dataset with nine soil types and a large variation in SOC content was constructed (chapter 4). When the model is calibrated for several soil types, spectral indices fail to predict the SOC variance within a single soil type in most of the cases (table 4.3). Only accidentally RPD values above two are found, with corresponding low standard error of prediction. This indicates that other soil properties (e.g. texture, parent material, mineral composition) cause variance in reflectance. Not including this in the calibration phase leads to large inaccuracies. Therefore, if these models are calibrated for several soil types they predict the variation in SOC within a single soil type poorly. Compared to indices, PLSR based models perform much better. For most soil types the RPD is above 1.4, with maximum values up to 3.16 (table 4.3). This shows that the PLSR model is reasonably accurate when applied to unknown soil types. So, in terms of model robustness towards variation in soil type, PLSR outperforms the indices presented in chapter 4. PLSR includes the spectral information in all available spectral bands, and combines the information from several wavelength ranges related to SOC. This results in a higher acceptance to the inclusion of soil types that were not used for model calibration.

PLSR also gives the best prediction for both extrapolated and non-extrapolated data-ranges When the SOC content is estimated for samples with a SOC content larger than the range of the calibration set, most models become inaccurate (table 4.2). Only for the index "1/slope 400-700", SEP values hardly change when the model is extrapolated, but in general this model performs less good than PLSR. PLSR predictions are influenced by extrapolation outside the calibration range, which was also found in section 3.4.2.2. For the dataset used in chapter 4, the SEP increases from 24.1 g/kg SOC to 40.7 g/kg SOC when samples outside the calibration range are predicted.

D. What is the influence of fractional vegetation cover on the estimation of soil properties?

Chapter 5 and 6 show that the influence of vegetation on the spectral signature of bare soil is to such extent that it limits accurate estimation of soil properties. This happens already at

low levels (<10 %) of fractional cover (figure 5.8 and 6.1). How much the predicted soil property deviates from the measured value depends on the amount of fractional vegetation cover, the index used (figure 5.8) and the amount of the soil property (figure 6.1).

When determining the iron content, it is shown that the index which calculates the area of the absorption dip around 550 nm is much more sensitive to the influence of vegetation on the spectral response than the redness index or the S.D. of the continuum removed spectrum (figure 5.8). Also the type of error depends on the index used. The redness index underestimates the amount of iron when the vegetation influence increases. This can be explained by the fact that an increase in (green) vegetation will increase the chlorophyll absorption, which is most active in the red wavelengths. As a result, the redness index will give a lower value. For the indices based on the continuum removed spectrum an opposite effect is shown, because these indices are based on the analysis of depth and width of the absorption feature. Iron and vegetation are both active absorbers in the VIS, so both the iron absorption feature and the chlorophyll absorption feature are normalized with continuum removal. Presence of vegetation causes additional absorption, which leads to overestimation of the iron content when continuum removed indices are used. The choice for a specific iron index determines the robustness of the analysis towards vegetation influence. The general behavior of the presented iron indices to fractional vegetation cover will be as described above, but the size of the error introduced will vary, depending on vegetation and soil type.

In chapter 6, the influence of vegetation on the estimation of SOC with PLSR is investigated. In this case, an increase in vegetation leads to an increase in the estimated amount of SOC (figure 6.1). Due to the statistical nature of the PLSR analysis, it is difficult to determine why the presence of vegetation leads to an overestimation of SOC. To predict the behavior of PLSR models to vegetation influence, an analysis of the factor loadings is needed. It is possible that other spectral pretreatments (e.g. smoothing, wavelets, derivatives), result in a different sensitivity to vegetation influence.

The inaccuracy due to vegetation limits the practical implementation of imaging spectroscopy for soil applications. On agricultural fields, the time that fields are completely bare is limited to a few weeks, which are timed in the period that other preconditions for a successful imaging spectroscopy campaign (stable and good weather conditions, high amounts of solar illumination) are usually poor. For regional studies this timing becomes more complex when different crops are grown with variation in phenology. In field spectroscopy, the instantaneous field of view (spatial footprint) can mostly be controlled, allowing measurements without vegetation interference, as long as there are no requirements to the minimum spot size covered with a single measurement. This gives some flexibility to avoid vegetation influence in the spectral measurements. The larger pixel size

with imaging spectroscopy makes it impossible to separate the reflectance of soil and vegetation during the measurement stage.

E. Can we determine soil properties from a mixed soil and vegetation signal?

Despite the strong influence of fractional vegetation cover on soil reflectance, it is possible to determine soil properties in partially vegetated areas. This can be done by processing the output of multiple indices (chapter 5), or by removing the vegetation influence from the spectral measurements in an early stage of the processing chain (chapter 6).

The fact that iron indices behave differently under the influence of vegetation is used to minimize the vegetation influence in the study described in chapter 5. As described, the standard deviation after continuum removal overestimates the iron content, while the redness index underestimates it. Averaging the results of both methods gives a more accurate iron prediction than using one of these indices individually. How large the improvements are depends on the vegetation type, (local) calibration of the model and soil type. If the method is applied in areas with other soil and vegetation types, or for other soil properties, the strength of the vegetation influence has to be investigated and a site-specific correction model has to be developed. Since this method is considered to be accurate up to a certain level of fractional cover, a hybrid approach is used to map the entire area. The areas with a relatively low cover are mapped with the suggested approach, while the values for the areas with higher fractional vegetation cover are interpolated.

A method that can be applied as a more general approach is presented in chapter 6. With Residual Spectral Unmixing (RSU) it is possible to filter the influence of vegetation from the mixed spectra. The residual soil spectra contain enough information to map the SOC distribution within agricultural fields with a good accuracy (RMSEC = 1.6 g/kg and RMSEP = 1.7 g/kg). The map in figure 6.5 shows that it is possible to map SOC despite fractional vegetation cover. RSU was also applied for the estimation of iron in the Álora study area. The technique was able to remove the vegetation influence from the spectral signature (Bartholomeus et al. 2005). This shows that RSU works for other soil properties as well, and that the method is more general applicable. Because RSU deals with the vegetation influence in an early stage of the processing chain, it can easily be implemented for other areas. The vegetation influence is removed before the soil property prediction model is calibrated, which gives more flexibility in the choice for the prediction model to be used.

An important step in the RSU algorithm is the estimation of the fractional vegetation cover with spectral unmixing. When evaluated at the pixel level there are occasionally large differences in fractional cover estimates from spectral unmixing compared to field estimates. If there are no pure pixels present in the fields, which is the case when the plants are still in an early phenological stage, endmembers should be obtained from other sources

like field measurements, spectral databases or modeling. Use of such endmembers should be done with care, since differences in illumination conditions or biophysical properties leads to inaccurate fractional cover estimates. When soil properties in more complex areas have to be estimated (e.g. several fields with different crop types, natural ecosystems) more advanced spectral unmixing methods have to be used (e.g. MESMA, non-linear unmixing).

7.2 General conclusions

The main contribution of this work is (i) the improved knowledge of empirical VNIR spectroscopy based soil model robustness, (ii) the quantification of uncertainties in estimation of soil properties with VNIR spectroscopy due to vegetation influence and (iii) the development of two methods to reduce this influence of vegetation which increases the applicability of imaging spectroscopy for mapping soil properties at local and regional scales.

Based on the studies of the chapters in this thesis, it can be concluded that:

- VNIR spectroscopy can be used to estimate soil properties on laboratory, field and image scale. The achieved accuracy for the laboratory and field techniques are good enough for monitoring of small temporal changes in carbon stock. The high sampling density of imaging spectroscopy offers great potential in comparison to traditional soil sampling techniques.

- High prediction accuracies can only be achieved when models are locally calibrated. In the case of field spectroscopy and imaging spectroscopy calibration for every campaign will be necessary.

- Robustness of presented SOC indices towards variation in soil types is low. PLSR is more robust to variation in soil types.

- Imaging spectroscopy based soil property models yield low prediction accuracies when the derived products are evaluated at the pixel level. However, it can be used to support the division of fields in management zones for precision agriculture applications.

- The influence of vegetation on soil reflectance is high and cannot be ignored when soil properties are estimated from mixed reflectance spectra.

- With a combination of indices or by using advanced spectral unmixing algorithms, it is possible to derive spatially continuous information about soil properties from imaging spectroscopy data in partially vegetated areas.

7.3 Outlook

It has been suggested, that the development of a global calibration of VNIR spectroscopy based soil property models is one of the first research priorities (e.g. Bartholomeus et al. 2008, Brown et al. 2006). This is based on the idea that including as many soil types as possible in the calibration phase is necessary to develop soil property prediction models that can be used worldwide. Initiatives like the construction of a global spectral library by the Soil Spectroscopy Group (Viscarra Rossel 2009) are the first combined attempts to realize this idea. As shown in chapter 4, models calibrated on highly variable datasets can be stable, but do not meet the accuracy of traditional soil analysis. It is the question if a global model can ever be achieved, but the construction of such a large spectral database is an important step for future studies. The development of an universal prediction model should not be the main purpose, but the study of such a model may lead to a better understanding of the complexity and relations within the soil-property matrix and the capability of VNIR spectroscopy to deal with this. This can help us to improve predictions by including a-priori knowledge (e.g. soil type, mineral composition, texture), or stratification according to statistical measures, as suggested in chapter 3. This may solve the problem of low mapping accuracy in areas with multiple soil types and can increase the robustness of prediction models and its applicability for extended geographical areas. Part of this a-priori knowledge can be derived from existing soil maps or from limited laboratory analysis. VNIR spectroscopy may not be directly able to meet the traditional requirements in soil scientific or precision agriculture applications and it is not obvious how the measurements at the surface should be related to sub-surface properties and processes. In combination with pedotransfer functions (Sharma et al. 2006), remote sensing can be a valuable tool to bridge the gap from point observations to regional information, on various soil properties.

Residual Spectral Unmixing performs well in the case study presented in this thesis (chapter 6), because only a limited number of endmembers are involved in the agro-ecosystem. This allows a rather simple and straightforward inversion of to the spectral unmixing algorithm. Natural ecosystems are usually more complex, which require other approaches. With the increase of the number of endmembers, and especially endmembers that are spectrally resembling, the complexity of the unmixing model increases. This automatically leads to lower accuracies of the unmixing result, because unmixing algorithms perform bad when endmembers are linearly scaled version of each other (Van der Meer and De Jong 2000).

Therefore, complex natural environments might require inverse radiative transfer modeling to extract soil reflectance information from the mixed spectral signal. A subject not dealt with in this thesis is the influence of different types of vegetation. In general, green vegetation will have a larger effect on the mixed spectrum than senescent vegetation or plant litter. The higher similarity of soil and litter implies that the negative effect of litter on the estimation of soil properties is probably less, but it will be more difficult to apply RSU to remove this influence.

One of the major concerns of the remote sensing community is how to scale and generalize information collected on the local level up through regional to the global level (Wu and Li, 2009). Remote sensing data have a significant potential to generate inputs for global as well as regional and local models of ecosystem processes (e.g. estimation of carbon fluxes DeFries et al. 2002). To optimally exploit the possibilities, remote sensing sources at different scales have to be combined, with preservation of the accuracy and further applicability of the results. Coupled radiative transfer models allow scaling from small to large scales, but the soil is usually treated as background, with little biochemical process value (Verhoef and Bach, 2003). This makes these models of limited values for soil studies. The accuracy and uncertainties of coarse resolution satellite products should properly be validated against ground-based measurements. In this study, airborne imaging spectrometer data were used, but the first results to determine soil properties with the use of satellite imaging spectrometers have been published (Gomez et al. 2008). Successful use of future space borne imaging spectrometers (e.g. ENMAP) will require a combination of several scaling techniques (statistical and empirical methods, physical models, etc.) including specific data corrections and calibration mechanisms, as well as successive validation feedbacks.

A final question that remains, is whether VNIR spectroscopy alone is able to solve the need for spatially continuous mapping of soil properties. Other techniques, like gamma ray spectroscopy (Kemmers et al. 2008), ground penetrating radar (Lambot et al. 2006) or electro conductivity (Mertens et al. 2008), are available. Gamma ray spectroscopy or electro conductivity measurements can provide the textural information that can be used as a-priori knowledge to chose the proper VNIR spectroscopy model for prediction (Egmond and Loonstra, 2008; Sudduth et al. 2005). Furthermore, these techniques provide subsurface information, where imaging spectroscopy only provides information about the surface. Because soil properties are usually sampled to a certain depth for shallow sub-surface measurements ready made systems, with the fibre optic and light source built in a shank, are available (Veristech 2009). Ground penetrating radar can provide information about soil moisture (Lambot et al. 2006), which is one of the disturbing factors for soil spectroscopy. Integration of different measurement techniques that give complementary information, may improve the estimates of soil properties and bring accurate monitoring of soil properties or

the development of on-the-go sensors closer within reach (Brown et al. 2006). EU FP7 projects like iSOIL and Digisoil are focusing at these sensor-integration approaches. However, locations where no soil data can be obtained with remote or proximal sensing will remain, for example with large amounts of vegetation cover. Therefore, also spatial interpolation techniques (Schloeder et al. 2001) will be necessary to fill in the gaps. Combined approaches, as presented in chapter 5, may offer an optimal solution. At locations without vegetation, the soil properties of the bare soil pixels are determined directly, RSU is applied on partially vegetated pixels to allow prediction of soil properties, and spatial interpolation techniques are used to fill in the gaps at the locations with full vegetation cover. The resulting soil property map can be accompanied by a quality map, indicating the accuracies of the various methods (Bartholomeus et al. 2005).

References

Bartholomeus, H.M., Epema, G.F., & Schaepman, M.E. (2005). Using imaging spectroscopy for the quantitative determination of soil iron content in partially vegetated areas. In Z. B. & S. M. (Eds.), *Imaging Spectroscopy - New Quality in Environmental Studies* (pp. 249-258). Warsaw: EARSeL

Bartholomeus, H.M., & Mulder, V.L. (2008). The Influence of Slope on the Spectroscopic Quantification of Soil Iron Content. In R. Viscarra Rossel (Ed.), *High Resolution Digital Soil Sensing and Mapping* Sydney, Australia

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., & Spaargaren, O.S.P. (2008). Spectral reflectance based indices for soil organic carbon quantification. *Geoderma, 145 (1-2)*, 28-36

Ben-Dor, E. (2002). Quantitative remote sensing of soil properties. In, *Advances in Agronomy,* 173- 243

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., & Reinsch, T.G. (2006). Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma, 132*, 273-290

Chen, F., Kissel, D.E., West, L.T., Adkins, W., Rickman, D., & Luvall, J.C. (2008). Mapping soil organic carbon concentration for multiple fields with image similarity analysis. *Soil Science Society of America Journal, 72*, 186-193

Cierniewski, J., Gdala, T., & Karnieli, A. (2004). A hemispherical-directional reflectance model as a tool for understanding image distinctions between cultivated and uncultivated bare surfaces. *Remote Sensing of Environment, 90*, 505-523

Curran, P.J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment, 30*, 271-278

DeFries, R.S., Houghton, R.A., Hansen, M.C., Field, C.B., Skole, D., & Townshend, J. (2002). Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 14256-14261

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

Egmond, F.M.v., & Loonstra, E.H. (2008). Gamma-ray sensor for topsoil mapping; the Mole. *1st Global workshop on high resolution digital soil sensing and mapping*. Sydney, Australia

Fidencio, P.H., Poppi, R.J., De Andrade, J.C. and Cantarella, H., 2002b. Determination of organic matter in soil using near-infrared spectroscopy and partial least squares regression. *Commun. Soil Sci. Plant Anal.*, 33: 1607-1615.

Gao, B.C., Montes, M.J., Davis, C.O., & Goetz, A.F.H. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment, 113 (Suppl 1), S17-S24*

Gomez, C., Viscarra Rossel, R.A., & McBratney, A.B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma, 146*, 403-411

Hanegraaf, M.C., Hoffland, E., Kuikman, P.J., & Brussaard, L. (2009). Trends in soil organic matter contents in Dutch grasslands and maize fields on sandy soils. *European Journal of Soil Science, 60*, 213-222

Kemmers, R.H., Egmond, F.M.v., & Loonstra, E.H. (2008). Kartering van fosfaatbeschikbaarheid in de bodem met behulp van natuurlijke radioactiviteit. *Alterra-report 1728*. Wageningen, The Netherlands.

Khosla, R., Fleming, K., Delgado, J.A., Shaver, T.M., & Westfall, D.G. (2002). Use of site-specific management zones to improve nitrogen management for precision agriculture. *Journal of Soil and Water Conservation, 57*, 513-518

Kooistra, L., Wehrens, R., Leuven, R.S.E.W., & Buydens, L.M.C. (2001). Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains. *Analytica Chimica Acta, 446*, 97-105

Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J., & Robbez-Masson, J.M. (2008). Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sensing of Environment, 112*, 825-835

Lambot, S., Weihermüller, L., Huisman, J.A., Vereecken, H., Vanclooster, M., & Slob, E.C. (2006). Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content. *Water Resources Research, 42*

Malenovsky, Z., Bartholomeus, H.M., Acerbi-Junior, F.W., Schopfer, J.T., Painter, T.H., Epema, G.F., & Bregt, A.K. (2007). Scaling dimensions in spectroscopy of soil and vegetation. *International Journal of Applied Earth Observation and Geoinformation, 9*, 137-164

Madari, B.E., Reeves III, J.B., Machado, P.L.O.A., Guimaraes, C.M., Torres, E. and McCarty, G.W., 2006. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma*, 136: 245-259.

Martens, H., & Naes, T. (1984). Multivariate calibration. 1. Concepts and distinctions. *Trends in analytical chemistry, 3*, 204-210

McCarty, G.W., Reeves III, J.B., Reeves, V.B., Follett, R.F. and Kimble, J.M., 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Science Society America Journal, 66: 640-646

Mertens, F.M., Pätzold, S., & Welp, G. (2008). Spatial heterogeneity of soil properties and its mapping with apparent electrical conductivity. *Journal of Plant Nutrition and Soil Science, 171*, 146-154

Schaepman-Strub, G., Schaepman, M.E., Painter, T.H., Dangel, S., & Martonchik, J.V. (2006). Reflectance quantities in optical remote sensing-definitions and case studies. *Remote Sensing of Environment, 103*, 27-42

Schloeder, C.A., Zimmerman, N.E., & Jacobs, M.J. (2001). Division S-8 - Nutrient management & soil & plant analysis: Comparison of methods for interpolating soil properties using limited data. *Soil Science Society of America Journal, 65*, 470-479

Sharma, S.K., Mohanty, B.P., & Zhu, J. (2006). Including topography and vegetation attributes for developing pedotransfer functions. *Soil Science Society of America Journal, 70*, 1430-1440

Shepherd, K.D., & Walsh, M.G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal, 66*, 988-998

Stevens, A., Van Wesemael, B., Vandenschrick, G., Touré, S. and Tychon, B., 2006. Detection of Carbon Stock Change in Agricultural Soils Using Spectroscopic Techniques. *Soil Science Society of America Journal*, 70: 844-850

Sudduth, K.A., & Hummel, J.W. (1991). Evaluation of reflectance methods for soil organic matter sensing. *Transactions of the American Society of Agricultural Engineers, 34*, 1900-1909

Sudduth, K.A., Kitchen, N.R., Wiebold, W.J., Batchelor, W.D., Bollero, G.A., Bullock, D.G., Clay, D.E., Palm, H.L., Pierce, F.J., Schuler, R.T., & Thelen, K.D. (2005). Relating apparent electrical conductivity to soil properties across the north-central USA. *Computers and Electronics in Agriculture, 46*, 263-283

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 and Soil, 251*, 319-329

Uno, Y., Prasher, S.O., Patel, R.M., Strachan, I.B., Pattey, E., & Karimi, Y. (2005). Development of field-scale soil organic matter content estimation models in Eastern Canada using airborne hyperspectral imagery. *Canadian Biosystems Engineering / Le Genie des biosystems au Canada, 47*, 1.09-01.14

Van Der Meer, F., & De Jong, S.M. (2000). Improving the results of spectral unmixing of Landsat Thematic Mapper imagery by enhancing the orthogonality of end-members. *International Journal of Remote Sensing, 21*, 2781-2797

Verhoef, W., & Bach, H. (2003). Simulation of hyperspectral and directional radiance images using coupled biophysical and atmospheric radiative transfer models. *Remote Sensing of Environment, 87*, 23-41

Veristech (2009). http://www.veristech.com/pdf_files/Optical_8thintlconf.pdf.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., & Skjemstad, J.O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma, 131*, 59-75

Viscarra Rossel, R., & The Soil Spectroscopy Group (2009). The Soil Spectroscopy Group and the development of a global soil spectral library. In, *EGU General Assembly 2009*. Vienna: Geophysical Research Abstracts

Wu, H., & Li, Z.L. (2009). Scale issues in remote sensing: A review on analysis, processing and modeling. *Sensors, 9*, 1768-1793

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., & Mataix-Beneyto, J. (2008). Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. *Soil Biology and Biochemistry, 40*, 1923-1930

Summary

Soils and its resources are of major importance for the production of fiber and food. Recently, the soil functioning as carbon pool has been added to the list of soil functions. To determine soil quality as a resource there is a need for regular monitoring of its chemical and physical properties, both in time and space.

Quantitative estimation of the exact amount, spatial distribution and temporal change of soil properties is still challenging. Conventionally, soil samples are analyzed by means of soil extraction procedures. Spatially, soil samples are collected according to a specific sampling scheme. Spatial interpolation techniques are used to prepare continuous maps, but for accurate interpolation intensive sampling is required. To achieve the high sampling density required to map the high spatial variability and slow temporal changes in soil properties new techniques are required.

Visible and Near Infrared (VNIR) spectroscopy is a promising technique for the quantitative estimation of soil properties and is used frequently for laboratory and field studies. However, the developed models are usually location specific and bridging the gap to imaging spectroscopy introduces spectral mixing problems.

The topic of this thesis is how to link spectral reflectance information to soil properties. This is done by the development of statistical models, which can be divided in two classes: multivariate models (PLSR) and univariate models (spectral indices). The objectives were to investigate the robustness of VNIR spectroscopy based soil property prediction models and the influence of vegetation on these models. Furthermore, methods were developed for quantitative mapping of soil properties in fractionally vegetation covered agricultural fields.

First, the scaling issues that play a role in VNIR spectroscopy are described. Spatial up-scaling needs to be dealt with when the step from point spectroscopy to imaging spectroscopy is made. Mixed pixels, in our case because of fractional vegetation cover, limit the direct transition of techniques and models from point spectroscopy to imaging spectroscopy. The use of different sensors to acquire spectral information introduces spectral scaling problems, which complicates model transfer from one sensor to another. Finally, directional scaling issues should be considered, including fixed protocols for the measurement setup.

Next, the accuracy and robustness of Partial Least Squares Regression (PLSR) and indices based models to estimate soil properties are investigated. Both with PLSR and indices it is possible to get good model calibrations and predictions for SOC. For the soil iron content,

indices in the VIS and beginning of the NIR can be used to get qualitative information, but the full spectrum should be used to get quantitative information about iron content. Spectral indices become inaccurate when they are used to estimate SOC beyond the range of values that were used in the calibration phase, but PLSR is less sensitive to extrapolation.

In general, model calibrations are location-specific. Including soil-types that are not used for model calibration leads to large errors in prediction. Transfer of the calibration of one year to another causes a decrease in accuracy, which indicates that practical implementation of spectroscopy requires stable measurement conditions, good calibrations and standard sampling protocols. If we want to get rid of the local calibrations, we will have to accept lower accuracies, but we gain in the ease of use. Users willing to choose one of the techniques for a particular application have to weigh the accuracy against the area to be covered. While laboratory spectroscopy has the advantage to allow stable calibration trough time, it still requires sample preparation (drying, sieving, grinding). Field spectroscopy can reach an accuracy comparable to laboratory spectroscopy, but requires calibration before each campaign. Imaging spectroscopy might be a practical way to spatially evaluate soil properties on large scales. The large number of samples compensates for the lower accuracy. Also for imaging spectroscopy, a calibration before each campaign will be necessary.

Imaging spectroscopy based models lack behind in prediction accuracy when the results are studied at the pixel level. The lower prediction accuracies are the result of 1) instrumentation specifications (spectral resolution, low instrument signal-to-noise ratios), 2) disturbing external factors (atmospheric attenuation, geometric and optical distortions, mixed pixels) and 3) internal factors (soil moisture, structure). Remote sensing studies often use the per pixel accuracies as a measure of the quality of imaging spectroscopy derived products. However, for most applications (e.g. precision agriculture, regional carbon stock estimates) a per pixel knowledge of the soil properties is not required. The power of imaging spectroscopy lies in the ease with which a high sampling density can be reached and the spatial distribution of soil properties can be determined, which offers great potential in comparison to traditional soil sampling techniques.

The influence of vegetation on the predicted soil properties is large. Already, with low amounts of fractional vegetation cover the predictions becomes inaccurate. How much the predicted soil property deviates from the measured value depends on the amount of fractional vegetation cover, the index used and the amount of the soil property.

Despite the strong influence of fractional vegetation cover on soil reflectance, it is possible to determine soil properties in partially vegetated areas. This can be done by processing of the output of multiple indices, or by removing the vegetation influence from the spectral measurements in an early stage of the processing chain.

The fact that two iron indices react differently on the influence of vegetation can be used to minimize the vegetation influence. Averaging the results of two indices gives a more accurate iron prediction than using one of these indices individually. How large the improvements are depends on the vegetation type, (local) calibration of the model and soil type.

A method that can be applied more general is Residual Spectral Unmixing (RSU). With this approach it is possible to filter the influence of vegetation from the mixed spectra, and the residual soil spectra contain enough information to map the SOC distribution within agricultural fields with a good accuracy. Because RSU deals with the vegetation influence in an early stage of the processing chain, it can easily be implemented for other areas. The vegetation influence is removed before the soil property prediction model is calibrated, which gives more flexibility in choice for the prediction models to be used.

Finally, it was concluded that VNIR spectroscopy can be used to estimate soil properties on laboratory, field and image scale. The achieved accuracy for the laboratory and field techniques are good enough for monitoring of the small temporal changes in carbon stock. The high sampling density of imaging spectroscopy offers great potential in comparison to traditional soil sampling techniques. However, high prediction accuracies can only be achieved when models are locally calibrated. In the case of field spectroscopy and imaging spectroscopy calibration for every campaign will be necessary. To get models that are robust to variation in soil types, PLSR should be used. Imaging spectroscopy derived products have low prediction accuracies when evaluated at the pixel level, but can be used to evaluate changes in soil properties over time, or to support the division of fields in management zones for precision agriculture applications. Vegetation has a large influence on soil reflectance and cannot be ignored when soil properties are estimated from mixed reflectance spectra. With a combination of indices or by using advanced pre-processing of the spectral information it is possible to derive spatially continuous information about soil properties from imaging spectroscopy data in partially vegetated areas.

Samenvatting

De bodem en zijn rijkdommen zijn van groot belang voor de productie van grondstoffen en voedsel. Onlangs is de rol van de bodem als koolstofreservoir toegevoegd aan de lijst van belangrijke bodemfuncties. Voor het bepalen van de kwaliteit van de bodem als hulpbron is er behoefte aan een regelmatige bepaling van de chemische en fysische eigenschappen, zowel in tijd als ruimte.

Kwantitatieve schatting van de exacte hoeveelheid, ruimtelijke verdeling en temporele verandering van bodemeigenschappen is nog steeds een uitdaging. Traditioneel worden bodemmonsters geanalyseerd door middel van bodemextractie procedures. Ruimtelijke analyse wordt gedaan door bodemmonsters te nemen volgens een specifiek bemonsterings-protocol. Vervolgens worden ruimtelijke interpolatietechnieken gebruikt voor het maken van continue kaarten, maar voor nauwkeurige interpolatie is intensieve bemonstering nodig. Om de hoge bemonsteringdichtheid te behalen die nodig is om de hoge ruimtelijke variabiliteit en langzame temporele veranderingen in de bodem te beschrijven, zijn nieuwe technieken nodig.

Zichtbaar en nabij-infrarood (Visible and Near Infrared - VNIR) spectroscopie is een veel-belovende techniek voor de kwantitatieve schatting van bodemeigenschappen en wordt vaak gebruikt voor laboratorium- en veldstudies. De ontwikkelde modellen zijn echter locatiespecifiek en het overbrengen van de modellen naar beeldvormende spectroscopie is lastig door problemen met spectraal gemengde pixels.

Het onderwerp van dit proefschrift is hoe spectrale reflectie informatie gelinkt kan worden aan bodemeigenschappen. Dit gebeurt door middel van statistische modellen, die kunnen worden onderverdeeld in twee klassen: multivariate modellen (Partial Least Squares Regression, PLSR) en univariate modellen (spectrale indices). De doelstellingen waren om de robuustheid van op VNIR spectroscopie gebaseerde bodemmodellen te bepalen en te onderzoeken wat de invloed van vegetatie op deze modellen is. Bovendien zijn er methoden ontwikkeld voor de kwantitatieve kartering van bodemeigenschappen in gedeeltelijk begroeide agrarische gebieden.

Om te beginnen zijn de schalen die een rol spelen bij de toepassing van VNIR spectroscopie voor de bepaling van bodemeigenschappen beschreven. Ruimtelijke opschaling speelt een rol als de stap van puntspectroscopie naar beeldvormende spectroscopie wordt gemaakt. Gemengde pixels, in dit geval vanwege gedeeltelijke vegetatiebedekking, beperken een directe toepassing van de technieken en modellen uit puntspectroscopie op beeldvormende spectroscopie data. Het gebruik van verschillende sensoren om spectrale informatie in te

winnen introduceert spectrale schalingsproblemen, waardoor modeloverdracht van de ene sensor naar de andere bemoeilijkt wordt. Ten slotte speelt directionele schaling een rol, waardoor vaste protocollen voor de meetopstelling noodzakelijk zijn.

Vervolgens worden de nauwkeurigheid en robuustheid van PLSR en indices gebaseerde modellen voor het bepalen van bodemeigenschappen onderzocht. Zowel met PLSR als indices is het mogelijk om goede modelkalibraties en voorspellingen voor organisch koolstof (SOC) te krijgen. Voor kwalitatieve schattingen van het ijzergehalte van de bodem, kunnen indices in het VIS en het begin van het NIR worden gebruikt, maar het volledige spectrum moet worden gebruikt voor kwantitatieve bepalingen van het ijzergehalte. Spectrale indices worden onnauwkeurig wanneer zij worden gebruikt voor de schatting van SOC buiten het bereik van de waarden die werden gebruikt voor de modelkalibratie, maar PLSR is minder gevoelig voor extrapolatie.

In het algemeen zijn modelkalibraties locatiespecifiek. Het gebruiken van een model voor bodemsoorten die niet zijn gebruikt tijdens de modelkalibratie leidt tot grote fouten in de voorspelling. Toepassing van de kalibratie van het ene jaar op metingen van een ander jaar veroorzaakt een afname van de nauwkeurigheid, waaruit blijkt dat de praktische uitvoering van spectroscopie stabiele metingen, goede kalibraties en standaard bemonsterings-protocollen vereist. Als we af willen van de lokale kalibraties, zullen we lagere nauwkeurigheden moeten accepteren, maar we krijgen er een groter gebruiksgemak voor terug. Gebruikers die één van de technieken willen gebruiken voor een bepaalde toepassing, moeten de nauwkeurigheid afwegen tegen de grootte van het gebied dat gemeten moet worden. Laboratoriumspectroscopie heeft het voordeel van de stabiele kalibraties in de tijd, maar het vereist monstervoorbereiding (drogen, zeven, malen). Veldspectroscopie kan een nauwkeurigheid bereiken die vergelijkbaar is met laboratoriumspectroscopie, maar er is een nieuwe kalibratie vereist voor elke campagne. Beeldvormende spectroscopie kan een praktische manier zijn om bodemeigenschappen ruimtelijk en op grote schaal te bepalen. Het grote aantal metingen dat gedaan kan worden compenseert voor de lagere nauwkeurigheid. Ook voor beeldvormende spectroscopie zal een kalibratie voor elke campagne nodig zijn.

Modellen gebaseerd op beeldvormende spectroscopie blijven achter in voorspellingsnauwkeurigheid wanneer de resultaten worden onderzocht op pixel niveau. De lagere voorspellingsnauwkeurigheden zijn het resultaat van 1) instrument eigenschappen (spectrale resolutie, lage instrument signaal/ruisverhoudingen), 2) storende externe factoren (atmosferische verstoringen, geometrische en optische verstoringen, gemengde pixels) en 3) interne factoren (bodemvocht, structuur). Remote sensing studies maken vaak gebruik van de nauwkeurigheid per pixel om de kwaliteit van afgeleide producten te evalueren. Echter, voor de meeste toepassingen (bijvoorbeeld precisie landbouw, schattingen van koolstof-

opslag) is een per pixel kennis van de bodemeigenschappen niet vereist. De kracht van beeldvormende spectroscopie schuilt in het gemak waarmee een hoge bemonsterings-dichtheid kan worden bereikt en het feit dat de ruimtelijke verdeling van de bodemeigen-schappen kan worden bepaald. Dit biedt grote mogelijkheden ten opzichte van traditionele bodembemonsteringstechnieken.

De invloed van vegetatie op de nauwkeurigheid van de gebruikte modellen is groot. Al bij een lage gedeeltelijke vegetatiebedekking worden de voorspellingen onnauwkeurig. Hoeveel de voorspelde waarde afwijkt van de gemeten waarde hangt af van het hoeveelheid vegetatiebedekking, de gebruikte index en de concentratie van de betreffende bodemeigenschap.

Ondanks de sterke invloed van gedeeltelijke vegetatiebedekking op de bodemreflectie, is het mogelijk om bodemeigenschappen in gedeeltelijk begroeide gebieden te bepalen. Dit kan gedaan worden door de uitkomst van meerdere indices te combineren, of door de vegetatieinvloed in een vroeg stadium van de procesketen uit de spectrale metingen te filteren.

Het feit dat twee ijzerindices verschillend reageren op de invloed van vegetatie kan worden gebruikt om de vegetatieinvloed te minimaliseren. Het middelen van de resultaten van twee indices geeft een nauwkeurigere voorspelling van ijzer dan het gebruik van elk van deze indices afzonderlijk. Hoe groot de verbeteringen zijn, hangt af van het vegetatietype, (lokale) kalibratie van het model en het bodemtype.

Een methode die algemeen kan worden toegepast is Residual Spectral Unmixing (RSU). Met deze methode is het mogelijk om de invloed van de vegetatie uit het gemengde spectrale signaal te filteren. De resterende bodemspectra bevatten voldoende informatie om de SOC distributie in agrarische gebieden te bepalen, met een goede nauwkeurigheid. Omdat RSU de invloed van de vegetatie in een vroeg stadium van de procesketen filtert, kan deze methode gemakkelijk worden toegepast in andere gebieden. De invloed van de vegetatie wordt verwijderd voordat het bodemmodel is gekalibreerd, wat meer flexibiliteit geeft bij de keuze voor het te gebruiken model.

Tot slot kan er worden geconcludeerd dat VNIR spectroscopie gebruikt kan worden voor de schatting van bodemeigenschappen op laboratorium-, veld- en regionale schaal. De bereikte nauwkeurigheid met laboratorium- en veldtechnieken zijn goed genoeg voor het monitoren van de kleine temporele veranderingen in de koolstofvoorraad. De hoge bemonsterings-dichtheid van beeldvormende spectroscopie biedt goede mogelijkheden vergeleken met traditionele bodembemonsteringstechnieken. Echter, een nauwkeurige voorspelling kan alleen bereikt worden als de modellen lokaal gekalibreerd zijn. In het geval van

veldspectroscopie en beeldvormende spectroscopie is kalibratie voor iedere campagne noodzakelijk. Om modellen te verkrijgen die robuust zijn voor variatie in bodemtype, is PLSR de beste keuze. Producten die afgeleid zijn van beeldvormende spectroscopie hebben een lage voorspellingsnauwkeurigheid op pixel niveau, maar kunnen worden gebruikt om de veranderingen in de bodemeigenschappen in de tijd te bepalen, of ter ondersteuning van de verdeling van de velden in beheerzones voor precisielandbouw toepassingen. Vegetatie heeft een grote invloed op de bodemreflectie en kan niet worden genegeerd wanneer de bodemeigenschappen zijn geschat op basis van gemengde spectra. Door een combinatie van indices of gebruik te maken van geavanceerde pre-processing technieken is het mogelijk om, met beeldvormende spectroscopie, ruimtelijk continue informatie over bodemeigen-schappen in gedeeltelijk begroeide gebieden te verkrijgen.

PE&RC PhD Education Statement Form

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)
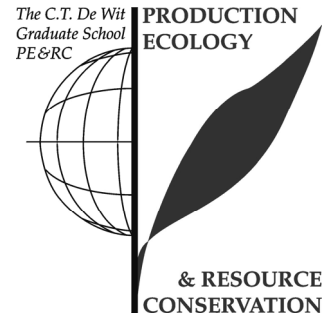
Review of Literature (5.6 ECTS)
- Book chapter / review article: Scaling directions in spectroscopy of soil and vegetation (2005)

Writing of Project Proposal (7 ECTS)
- Spatially continuous determination of organic carbon in agro-ecosystems with Imaging Spectroscopy; PE&RC (2006)

Laboratory Training and Working Visits (0.9 ECTS)
- Field spectroscopy measurements; ASD (2003)

Post-Graduate Courses (4.2 ECTS)
- ArcView 8 (ArcGIS); GISCOVER (2003)
- IDL-Programming; Creaso (2006)
- 2nd Hyperinet summer school; Hyper I net (2008)

Competence Strengthening / Skills Courses (3.2 ECTS)
- Lecturing in English; Centa (2004)
- Scientific writing; Centa (2006)

Discussion Groups / Local Seminars and Other Scientific Meetings (4.2 ECTS)
- CGI PhD Discussion group (2004-2006)
- SPAM Discussion group (2007-2008)

PE&RC Annual Meetings, Seminars and the PE&RC Weekend (1.2 ECTS)
- PE&RC Introduction weekend (2006)
- PE&RC Day (2007)

International Symposia, Workshops and Conferences (7 ECTS)
- 1st International Conference on Remote Sensing and Geomatics in the Assessment and Monitoring of Land Degradation and Desertification (RGLDD) (2005)
- EGU (2007)
- 1st Global workshop on high resolution digital soil sensing and mapping (2008)

Curriculum Vitae

Harm Bartholomeus werd op 16 februari 1977 geboren in Westerhoven. Na het behalen van het VWO-diploma aan het Anton van Duinkerken College begon hij in 1995 aan de studie Fysische Geografie aan de Universiteit Utrecht. Zijn afstudeeronderzoek richtte zich op de procesactiviteit van puinhellingen in relatie met klimaatschommelingen. De interesse in de klimaatveranderingen kwam terug in de afsluitende stage bij het IMAU, waarbij gewerkt werd aan een ruimtelijk energie balans model van de Groenlandse ijskap. Na het afronden van deze studie in 2000 volgde hij de post-doctorale opleiding tot docent Aardrijkskunde aan het IVLOS – Universiteit Utrecht. In 2001 jaar behaalde hij de eerste graads bevoegdheid Aardrijkskunde en was tevens werkzaam als vakdocent aan het Dr F.H. De Bruijne Lyceum in Utrecht. Sinds oktober 2001 is hij in dienst bij het Centrum voor Geo-Informatie aan de Wageningen Universiteit en Research Centrum. Naast de onderwijstaken begon hij hier na een aantal jaren aan een PhD-studie getiteld "The influence of vegetation on the spectroscopic estimation of soil properties", welke in september 2009 is afgerond.

Harm Bartholomeus was born on February 16, 1977 Westerhoven – The Netherlands. After obtaining his high school diploma at the Anton van Duinkerken College in 1995, he began to study Physical Geography at Utrecht University. His thesis research focused on the process activity of debris slopes in relation to climate fluctuations. The interest in climate change came back in the final internship at IMAU, where he worked on a spatial energy balance model of the Greenland ice sheet. After completing this study in 2000 he took the post-graduate training at IVLOS - University of Utrecht, to become Geography teacher. In 2001 he got his degree and was employed as a teacher at the Dr. F.H. De Bruijne Lyceum in Utrecht. Since October 2001 he works at the Center for Geo-Information at Wageningen University and Research Center. Besides teaching, he started a PhD study entitled "The influence of vegetation on the spectroscopic estimation of soil properties", which was finished in 2009.

List of publications

Peer reviewed journals

Bartholomeus, H., Epema, G., & Schaepman, M. (2007). Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy. *International Journal of Applied Earth Observation and Geoinformation, 9*, 194-203

Bartholomeus, H., Kooistra, L., Stevens, A., Wesemael, B.v., Ben-Dor, E., Tychon, B & Van Leeuwen, M. (2009). Soil Organic Carbon mapping of partially vegetated agricultural fields with imaging spectroscopy. *Remote Sensing of Environment*, Submitted

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., & Spaargaren, O.S.P. (2008a). Spectral reflectance based indices for soil organic carbon quantification. *Geoderma, 145 (1-2)*, 28-36

Bartholomeus, H., & Schaepman-Strub, G. (2009). Spectroscopic characterisation of decomposition stage and organic carbon content of arctic soils. *In preparation*

Clevers, J.G.P.W., Schaepman, M.E., Mucher, C.A., de Wit, A.J.W., Zurita-Milla, R., & **Bartholomeus, H.M**. (2007). Using MERIS on Envisat for land cover mapping in the Netherlands. *International Journal of Remote Sensing, 28*, 637-652

Clevers, J.G.P.W., Zurita Milla, R., Schaepman, M.E., & **Bartholomeus, H.M.** (2005). Using MERIS on ENVISAT for land cover mapping. In, *European Space Agency, (Special Publication) ESA SP* (pp. 771-780)

Malenovsky, Z., **Bartholomeus, H.M.**, Acerbi-Junior, F.W., Schopfer, J.T., Painter, T.H., Epema, G.F., & Bregt, A.K. (2007). Scaling dimensions in spectroscopy of soil and vegetation. *International Journal of Applied Earth Observation and Geoinformation, 9*, 137-164

Schaepman-Strub, G., Limpens, J., Menken, M., **Bartholomeus, H.M.**, & Schaepman, M.E. (2008). Towards the spatial assessment of carbon sequestration in peatlands: Spectroscopy based estimation of the fractional cover of three plant functional types. *Biogeosciences Discuss., 5 (2)*, 1293-1317

Schaepman-Strub, G., Limpens, J., Menken, M., **Bartholomeus, H.M.**, & Schaepman, M.E. (2009). Towards the spatial assessment of carbon sequestration in peatlands: Spectroscopy based estimation of the fractional cover of three plant functional types. *Biogeosciences, 6 (2)*, 275-284

Stevens, A., Wesemael, B.v., **Bartholomeus, H.**, Rosillon, D., Tychon, B., & Ben-Dor, E. (2008b). Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils *Geoderma, 144 (1-2)*, 395-404

Vrieling, A., Rodrigues, S.C., **Bartholomeus, H.**, & Sterk, G. (2007). Automatic identification of erosion gullies with ASTER imagery in the Brazilian Cerrados. *International Journal of Remote Sensing, 28*, 2723-2738

Conference Contributions

Abbasi, M., Schaepman, M.E., Darvishsefata, A., **Bartholomeus, H.M.**, Mohajer, M.R.M., & Sobhani, H. (2008). Spectroradiometric measurements of tree species in the caspian forests of Iran. In, *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* (pp. 291-296). Beijing

Bai, Z.G., Dent, D., **Bartholomeus, H.M.**, & Schaepman, M.E. (2005). Land Degradation and Improvement Assessment using NASA GIMMS in the Shaanxi Province (China) *RGLDD conference proceedings*

Bartholomeus, H.M., & Mulder, V.L. (2008). The Influence of Slope on the Spectroscopic Quantification of Soil Iron Content. In R. Viscarra-Rossel (Ed.), *High Resolution Digital Soil Sensing and Mapping* (p. 9). Sydney, Australia

Bartholomeus, H., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., & Spaargaren, O. (2007).Quantitative retrieval of Soil Organic Carbon using laboratory spectroscopy and spectral indices. In, *European Geophysical Union*. Vienna

Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B., & Spaargaren, O. (2006). Quantitative retrieval of Soil Organic Carbon using laboratory spectroscopy and spectral indices. In, *ISPRS 2006: ISPRS mid-term symposium 2006 remote sensing : from pixels to processes* (p. 5). 8-11 May 2006, Enschede, the Netherlands. - Enschede

Bartholomeus, H.M., Epema, G.F., & Schaepman, M.E. (2005). Using imaging spectroscopy for the quantitative determination of soil iron content in partially vegetated areas. In Z. B. & S. M. (Eds.), *Imaging Spectroscopy - New Quality in Environmental Studies* (pp. 249-258). Warsaw: EARSeL

Clevers, J.G.P.W., **Bartholomeus, H.M.**, Mucher, C.A., & De Wit, A.J.W. (2004a). Land cover classification with the medium resolution spectrometer (MERIS). *Earsel e-Proceedings / Warsaw 4th IS workshop proceedings, 3*, 354-362

Clevers, J.G.P.W., **Bartholomeus, H.M.**, Mucher, C.A., & De Wit, A.J.W. (2004b). Use of MERIS data for land cover mapping in the Netherlands. *Proceedings of MERIS User Workshop, 10-13 November 2003, Report SP-549*

Kooistra, L., Sanchez-Prieto, L., **Bartholomeus, H.M.**, & Schaepman, M.E. (2007a). Regional mapping of plant functional types in river floodplain ecosystems using airborne imaging spectroscopy data. In, *Proceedings of the 10th International Symposium on Physical Measurements and Spectral Signatures in Remote Sensing (ISPMSRS'07)* (pp. 291 - 296). Davos (CH)

Saito, M., Jansen, R., **Bartholomeus, H.**, Usui, Y., & Nakano, K. (2005). Non destructive detection system of potato tubers using NIRS. In, *Joint Meeting on Environmental Engineering in Agriculture 2005* Kanazawa, Japan

Saito, M., Jansen, R., **Bartholomeus, H.**, Usui, Y., & Nakano, K. (2006). Development of non-destructive detection system for Phytophthora infestans infected potato tubers. In, *EFITA, The 5th Conference of the European Federation for Information Technology in Agriculture, Food and Environment* (p. 6). Vila Real, Portugal

Stevens, A., van Wesemael, B., **Bartholomeus, H.**, Tychon, B., & Ben-Dor, E. (2008a). Performance of different VNIR spectroscopic techniques for determining soil organic carbon contents. In, *European Geophysical Union*. Vienna

Zeng, Y., **Bartholomeus, H.M.**, De Bruin, S., Epema, G.F., & Clevers, J.G.P.W. (2003). Using hyperspectral remote sensing data for identifying geological and soil units in the Alora Region, Southern Spain. *Proceedings of 3rd EARSeL workshop on imaging spectroscopy, Herrsching, Germany. 13-16 May*, 297-301

Other publications

Bartholomeus, H.M. (2007). Spatially continuous determination of organic carbon in agro-ecosystems with Imaging Spectroscopy. *PE&RC Research Proposal*

Bartholomeus, H.M., & Schaepman, M.E. (2004). Veldspectrometer voor Cenrum voor Geo-Informatie (WUR). *Agro-Informatica*, 21-24

Kooistra, L., Schaepman, M.E., Crompvoets, J.W.H.C., & **Bartholomeus, H.M**. (2007b). ESA User Services Vision 2010-2015: Results Questionnaire. In (p. 79). Wageningen Wageningen University

Schaepman, M.E., & **Bartholomeus, H.M**. (2006). Spectroradiometric measurements - From physics to applications. In, *Wageningen : Centre for Geo-Information, 2006 (CGI-Report 2006-01)* (p. 52)

Stevens, A., Wesemael, B.v., **Bartholomeus, H**., Rossilon, D., Tychon, B., & Ben-Dor, E. (2007). CARBIS Final Report - Detecting soil carbon and its spatial variability by Imaging Spectroscopy. (p. 53). Louvain-la-Neuve: Belspo