CARBIS Final Report

Detecting soil carbon and its spatial variability by Imaging Spectroscopy

Antoine Stevens*,a, Bas van Wesemaela, Harm Bartholomeusb, Damien Rosillonc, Bernard Tychonc and Eyal Ben-Dord

department of Geography, Université catholique de Louvain, Place Pasteur, 3, 1348 Louvain-La-Neuve, Belgium

b Centre for Geo-Information, Wageningen University, Droevendaalsesteeg 3, NL 6708 PB, Wageningen, The Netherlands

c Department of Environmental Sciences and Management, University of Liège, Campus of Arlon, 185, avenue de Longwy, 6700 Arlon, Belgium

d Department of Geography and Human Environment, Tel Aviv University, P.O. Box 39040 Ramant Aviv, Tel Aviv, Israel 69978

*Corresponding author
Antoine Stevens
3, place Pasteur, 1348 Louvain-La-Neuve, Belgium
Phone : 32 10 47 28 61
e-mail: stevens@geog.ucl.ac.be
Index

Index .................................................................................................................................. 2
General introduction ........................................................................................................ 3
I. Calibration of hyperspectral data for soil organic carbon determination ....................... 5
I.1 Introduction ............................................................................................................. 6
I.2 Materials and Methods ............................................................................................ 7
  I.2.1 Study site, field campaign, spectral measurements ........................................ 7
  I.2.2 Statistical Analysis ........................................................................................... 9
I.3 Results and Discussion ......................................................................................... 11
  I.3.1 Comparison between field and remote spectra ............................................. 11
  I.3.2 Calibration and Validation .......................................................................... 12
I.4 Conclusions .......................................................................................................... 15
I.5 References ............................................................................................................. 15
II. The influence of fractional vegetation cover on the Quantification of Soil Organic Carbon with Imaging Spectroscopy ................................................................. 18
II.1 Introduction .......................................................................................................... 18
II.2 Influence of Vegetation Cover on SOC prediction .............................................. 18
II.3 Quantification of Fractional Vegetation Cover ................................................... 20
II.4 Results .................................................................................................................. 22
II.5 Conclusions .......................................................................................................... 25
II.6 Further research .................................................................................................... 26
II.7 References ............................................................................................................ 26
III. Effects of other disturbing factors on soil reflectance ............................................... 28
III.1 Introduction ......................................................................................................... 28
III.2 Material and method ........................................................................................... 29
  III.2.1 Field measurements ..................................................................................... 29
  III.2.2 Laboratory measurements .......................................................................... 30
  III.2.3 Data treatment .............................................................................................. 31
III.3 Results and discussion ....................................................................................... 35
  III.3.1 Impact of soil shadow on soil reflectance .................................................... 35
  III.3.2 Roughness – soil shadow relationship ......................................................... 44
III.4 Conclusions ......................................................................................................... 46
III.5 Annexes ............................................................................................................... 49
  Annex 1 : Soil models composition ........................................................................ 49
  Annex 2 : Correlation between L_{abs} and disturbing factors (G2 plot) ................. 51
III.6 References .......................................................................................................... 52
General Conclusion ........................................................................................................ 53
General 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. This has been motivated notably by the possibility to sequester CO$_2$ into soils by increasing the Soil Organic Carbon (SOC) stocks and by the role of SOC in maintaining soil quality. Within the EU soil thematic strategy, the depletion of OM is mentioned as one of the major threats to the soil resource. This increasing interest resulted in a better understanding of SOC dynamics as influenced by human activities and environmental changes.

Several difficulties in estimating SOC stocks and their temporal evolution remain challenging. One of them is linked to the spatial variability of SOC that masks its slow accumulation or depletion. When assessing the role of specific management practices on carbon stocks at the field scale, this variability masks the slow accumulation of SOC and hampers the detection of a statistical significant change. The sampling density required to detect a change in SOC content is often very high and thus very labour intensive. Basically, the area that a sample can represent depends on the degree of variation of the soil property of interest. Increasing the number of samples in a given area will decrease the confidence limit around the estimate. At the local or field scale, the relative small number of samples often reduces the predictive power of the inference model. At the regional scale, soil databases are in general well populated, but fail to reproduce the variation in biogeochemical trends observed at the local scale. This is due to the fact that the variation in soil properties observed at the local scale is likely to be small in comparison with the spatial heterogeneity generated by processes acting at broader scales (e.g. climate, geology).

Hence, there is 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. From the discussion above, it is obvious that there is a need for better analytical methods as soil science faces new challenges. Such analytical techniques would allow rapid sampling and instant determination of SOC values, at the field and regional level.
Visible and Near InfraRed (VNIR) Spectral analysis and diffuse reflectance analysis can be a suitable technique to rapidly quantify various soil characteristics simultaneously. This technique would be more rapid to assess several soils properties than classical pedotransfer functions relying on conventional measurements. Every substance possesses its own characteristic spectrum that can be detected by a sensor. This spectrum constitutes a unique “fingerprint” which is determined by the composition of the object in question. The form of the spectrum in a particular wavelength region depends on the selective absorption of radiations of given frequencies by molecular bonds. The Visible and Near Infrared (VNIR) region (350 nm to 2500 nm) is characterised by absorption bands due to overtones and combinations modes of the fundamental vibrations of bonds C-H, N-H and O-H in the Mid InfraRed (MIR) region (2500-14000nm).

There are three types of VNIR techniques operating at different spatial scales and in different environments: (1) Laboratory Spectroscopy (LS) ; (2) Portable field Spectroscopy (PS) and (3) Imaging Spectroscopy (IS). LS and PS rely on ground-based sensors (usually point spectrometry) and IS on air- or space- borne sensors (usually image spectrometry).

The objective of this study is to evaluate the potential of PS and LS in the estimation of SOC content in agricultural fields. Both can produce a large amount of samples. PS provides a rapid in-situ analysis of soils. LS has even a greater potential. As remote sensing might be a practical way to spatially evaluate soil carbon, the development of this technique can be an important step toward generating a useful vehicle to track after the SOC content change on large scales trough repeated measurements over time.

The impact of disturbing factors (soil roughness, soil moisture, fractional vegetation cover) on the accuracies retrieved by the two types of sensors will be also assessed. Taking in account these disturbing factors could be an important achievement in order to be able to use these techniques in real case studies in which soil conditions are not always optimal.

The report comprises three complimentary chapters dedicated respectively to (i) the calibration of hyperspectral data for soil organic carbon determination, (ii) the influence of fractional cover on the quantification of SOC with Imaging Spectroscopy and (iii) the effects of other disturbing factors on soil reflectance. The last section will draw the main conclusions of the CARBIS field campaign.
I. Calibration of hyperspectral data for soil organic carbon determination

Antoine Stevens*,a, Bas van Wesemaela

a Department of Geography, Université catholique de Louvain, Place Pasteur, 3, 1348 Louvain-La-Neuve, Belgium

*Corresponding author
Antoine Stevens
3, place Pasteur, 1348 Louvain-La-Neuve, Belgium
Phone : 32 10 47 28 61
e-mail: stevens@geog.ucl.ac.be
I.1 INTRODUCTION

Soil Organic Carbon (SOC) stock have been the objects of a growing literature last decades due to their role in the conservation of soil quality and in the global C cycle. However, the study of their temporal evolution remains difficult due to a strong intra- and inter-field variability masking the signal with a lot of noise. As a result, the detection of SOC stock changes in soil monitoring or modeling studies require a high sampling that is rarely achieved without resorting to expensive and time consuming routine soil analysis methods. New analytical methods are needed that would allow a rapid sampling and instant determination of SOC contents.

Visible Near Infrared (VNIR) diffuse reflectance spectroscopy has been extensively explored in the laboratory as a rapid means to quantify various soil properties and OC in particular (e.g. McCarty and Reeves III, 2001; McCarty et al., 2002; Martin et al., 2002). Indeed, the VNIR region (350 nm to 2500 nm) is characterised by absorption bands due to molecular bonds that are present in soil organic matter. For the time being, portable field spectroscopy and Imaging Spectroscopy (IS) have been less studied as SOC analytical tools. Nevertheless, their potential for soil inventories or monitoring seems greater as it requires minimal samples handling. Portable Spectroscopy may have different applications in the field. Either the sensor is hand held and can be utilized for a rapid in-situ monitoring of soils (Kooistra et al., 2003; Udelhoven et al., 2003), or the sensor is mounted on a tractor. In the latter case, these “on-the-go” systems are dedicated to the site-specific management of inputs in precision agriculture (Adamchuk et al., 2004). IS (or hyperspectral remote sensing) has been mainly used for the mapping of surface soil properties of agricultural fields with high resolution (Ben-Dor et al., 2002; Selige et al., 2006; Stevens et al., 2006). IS has the capability to cover large areas in a single campaign and produce thousands of samples.

With these instruments, SOC can be measured with a reasonable accuracy level varying from 0.1 to 1.5 % C (Stevens et al., 2006) depending on the statistical treatments of the dataset, field and laboratory methodologies, field conditions, and the instrument.

The objective of this chapter is to compare the predictive ability of field- and remote-based VNIR spectroscopy to determine SOC contents in cropland. Then, the chapter will study the stability of the calibration across time and space. To achieve these two goals, several spectral datasets have been collected during three hyperspectral field campaigns financed by the BELSPO along the period 2003-2005.
I.2 MATERIALS AND METHODS

I.2.1 Study site, field campaign, spectral measurements

The study area (±/−50 km²) selected is located in Belgian Lorraine (49°38'; 49°43' N and 5°27'; 5°31' E) (fig.1). The area is characterised by sandy to clayey soils with strong ferric components.

Fig. I-1. Location of the study area.

The flight campaign took place the 20th June 2005. The AHS-160 hyperspectral sensor recorded 7 north-south flight lines. This time window, constrained by other projects, required an adaptation since bare fields are extremely sparse in the region during spring. In order to circumvent the problem, we managed to clear 74 areas of 7.5 by 7.5 m in 8 maize fields within the study area. The size of these bare experimental plots corresponds to a 3 by 3 pixel size of the AHS-160. Furthermore, two ploughed bare fields of c. 0.25 ha each have been selected where 36 plots were set. The location of these plots has been determined by DGPS.

Samples composed of 9 sub-samples collected to a depth of 5 cm within the experimental plot have been taken for carbon. SOC content of each sample has been determined with the so-called Walkley and Black method (Walkley and Black, 1934).

First, the hyperspectral sensor AHS-160 provided a total of 80 spectral bands covering the Visible (VIS), Near InfraRed (NIR), Short Wave InfraRed (SWIR) and Thermal InfraRed (TIR)
parts of the spectrum for 2.6 m² pixels. The data cube obtained has been corrected for geographical, radiometric and atmospheric errors by the VITO. Despite these corrections, spectral data was still affected by noise in the SWIR and the TIR region. The TIR window was too much noisy to be used and was removed from the analysis. For each experimental plot, 9 pixels centred on the middle of the plot were extracted using ENVI (Research Systems Inc., Boulder, CO). The spectrum used for the statistical analysis is a composite of these 9 spectra. Some pixels, especially those located at the edge of the plot, are influenced by the surrounding vegetation. In order to work only with “pure” soil spectra, these pixels were removed according to their NDVI value. Based on the examination of the NDVI histogram of all measurements, the threshold above which the pixel is considered to be affected by the vegetation was set to 0.3.

The second instrument, a portable spectrometer FieldSpec Pro FR (Analytical Spectral Devices – ASD), measures every nanometer of the solar spectrum in the VIS-NIR-SWIR region. A total of 118 field spectra of the experimental plots have been taken with the ASD. Each field spectrum consists of 9 individual spectra taken across the experimental plot. Some spectra, differing clearly from other spectra of the same plot were eliminated before computing the mean. The plot of the reflectance at 670 nm against the reflectance at 780 nm of the mean spectra revealed that 7 mean spectra were still influenced by the presence of residual vegetation (fig.2). These spectra are not a part of the “soil line” and were removed from further analysis. Wavelengths between 1340-1430 nm, 1810-1970 nm and beyond 2400 nm were affected by a strong noise and were removed.

![Fig. 1-2. Plot of the reflectance of ASD field measurements at 780 nm vs reflectance at 670 nm with outliers (red circles).](image-url)
I.2.2 Statistical Analysis

Before the quantitative statistical analysis, a series of pre-treatments were applied to each spectrum using *MATLAB* (The Matworks Inc., Natick, MA) to remove a part of the noise. Firstly, the reflectance (R) was converted into absorbance (A) to reduce the non-linearity problem:

\[
A = \log_{10}(1/R) = -\log_{10}R
\]

In a second step, spectra were transformed with (combinations of) standard pre-treatments. These pre-treatments consisted of: (i) 1st and 2nd derivatives, (ii) 1st and 2nd gap derivatives, (iii) Savitzky-Golay smoothing and differentiation algorithms (Savitzky and Golay, 1964), (iv) mean centering and variance scaling, and (v) skipping.

First and second derivatives have the disadvantage to increase the noise, degrading the SNR. This problem is partially overcome by the calculation of 1st and 2nd gap derivatives with different window size. The gap derivative implies that the differentiation is not done between two adjacent points \(x_{i-1}\) and \(x_i\) but between points \(x_{i-n}\) and \(x_{i+n}\) with \(2n+1\), the window size. Savitzky-Golay's algorithms are widely used in signal pre-processing when the distance between spectral bands is constant. They are similar to a moving window averaging method but have the property to better preserve local peaks in the spectrum. The principle is that, for any window \([x_{i-n}; x_{i+n}]\) centered on a point in position \(i\) on the spectrum, a polynomial with a given degree is fitted through the points by least-square. The fitted value for the point \(i\) replaces the measured value. The algorithms involve the use of tabulated coefficient published by Savitzky and Golay instead of computing each time the polynomial. An example is given on Fig. 3. Mean centering and variance scaling result in the removal of the variability related to the background reflectance, which is not related to the measured variable. In the case of the ASD measurements, the skipping of every 10 nm permitted to use only 216 spectral bands or variables instead of 2151 in the regressions. This spectral resolution seems sufficient to preserve fine peaks in the soil spectrum.

After this step, each pre-treatment or combination of pre-treatments was related separately to carbon content by Partial Least Square Regressions (PLSR) using *SAS* (*SAS* Institute Inc., Cary, NC). This regression technique, close to Principal Component Regressions, is useful when the number of independent variable (i.e. the spectral bands) is high. The maximal number of Latent
Variables (LV) was set to 10 and their optimal number was determined by leave-one-out cross-validation. For the validation procedure, two approaches were used, depending on the number of samples available. For a small dataset as the one collected during the 2005 field campaign \((n = \pm 100)\), a leave-one-out cross validation was applied to the dataset. It has the advantage to give more accurate estimate of the 'true' prediction error than a true validation (Martens and Dardenne, 1998). However, when analysing previous spectral datasets joined with the current dataset \((n = \pm 200)\), a true validation approach was preferred. In this latter case, the dataset was split randomly into one set of calibration (2/3) and one set of validation (1/3).

Fig. 1-3. Examples of the effect of different pre-treatments on a laboratory spectrum. Absorbance (black, left scale) is converted to the Savitzky-Golay first derivative (red, right scale) with a window size of (a) 5, (b) 9, (c) 17, (d) 33, (e) 65 and (f) 129 points.

The performance of the model is traditionally measured by the Root Mean Square Error of Prediction (RMSEP) or the Root Mean Square Error of Cross Validation (RMSECV), given by the following expression:

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

where \(\hat{y}_i\) and \(y_i\) are respectively the predicted and the observed values of the sample \(i\) in the test set of \(n\) samples. As this statistic depends often on the standard deviation of the samples, the Ratio of Performance to Deviation (RPD), defined as the ratio between the Standard Deviation
(SD) and the RMSEP or RMSECV, was computed in order to compare the prediction ability of each pre-treatment and to select the best one. Three categories have been defined by Chang et al., 2001: 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 ability. In order to construct robust models, spectral outliers must be removed during the calibration phase. We have implemented a method in SAS, which is close to the USE (Uninformative Sample Elimination)-PLS algorithm developed by Koshoubu et al., 2001. Sample outliers are detected after each PLSR on the basis of their prediction error and subsequently removed.

I.3 RESULTS AND DISCUSSION

I.3.1 Comparison between field and remote spectra

Figure 4 shows two spectra of the same experimental plot, measured by the AHS and the ASD. It can be observed that the ASD and the AHS spectra diverge in absolute reflectance. Derivative spectra show closer measurements between 800 and 1800 nm. Some spectral bands in the NIR region of the ASD are affected by a strong noise (1400 nm, 1900 nm and 2300-2500 nm). Similarly, the AHS spectrum is very much noisy in the Short Wave InfraRed (SWIR, 1900-2500 nm) where its spectral resolution is high.

![Comparison between two spectra in absorbance units of the same sample](image)

Fig. 4-4. Comparison between two spectra in absorbance units of the same sample (a) without pre-treatment and (b) transformed by the 1st derivative, measured by the AHS (dashed line), the ASD in the field (solid line).
I.3.2 Calibration and Validation

I.3.2.1 Predictive ability of ground- and remote-sensing VNIR spectroscopy

As a first approach, we analyzed the ASD and AHS spectra. The predictive statistics of the best model of each sensor are given in table I. The percentage of X (wavelengths) and Y (carbon content) variation explained by the models is generally high, suggesting that the models are well constructed. RPD decreases from ground to remote measurements. This diminution of predictive ability is a result of differences in sensor characteristics and measuring conditions. According to Chang et al. (2001), field spectroscopy seems a reliable technique to measure SOC content. Compared to ground-based sensors, IS has a lower RPD. Its RPD is close to the one of a very bad models (RPD ≈ 1.4), probably due to a low SNR, even after numerous corrections (radiometric, atmospheric). This is confirmed by the analysis of observed vs predicted carbon content (fig. 5a) drawing a pattern roughly circular in the plot space.

RMSECV of the ASD reaches 0.11 % C. The RMSECV constitutes a kind of limit that is comparable to the limit of repeatability (RMSE of replicate samples) of a standard routine analytical method. For the Walkley-Black method, this limit has been evaluated at 0.1 % C by Colinet et al. (2005) in the laboratory network where the analyzes have been carried out. 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 from which the technique has been tested present a relatively low mean (1.34 % C) and standard deviation (0.27 % C) so that this is only true for soil types and range of SOC present in the study area (see fig. 5b).

![Fig. I-5. Plot of predicted vs observed organic carbon as obtained after cross validation of (a) AHS spectra and (b) field ASD spectra.](image)
I.3.2.2 Spatio-temporal stability of a spectral calibration

Secondly, in order to test the stability of the calibrations obtained, we merged our samples collected during the 2005 field campaign with previous samples coming from a previous hyperspectral flight campaign (CASI 2003; see Stevens et al., 2006), producing a dataset of more or less 200 samples having differing soil characteristics, SOC range and field surface conditions.

In order to test the stability of the calibrations, we used the spectral data from a previous study (Stevens et al., 2006) with 2005 spectral measurements. Only field spectroscopic spectra could be used because spectra retrieved by airborne sensors had a different spectral range and spectral resolution. Location of the field campaigns and statistical parameters of the joined datasets are given in Table II. Given the fact that a larger dataset was created, a true test set validation could be used rather than a cross-validation.

In a third experimental test, we analyzed the entire dataset using cross-validation. This gave better results (RPD = 3.03 %C, RMSECV = 0.23 %C, 2 outliers) indicating that the statistical procedure can effectively predict samples when they are within the same range as the training set (i.e. interpolation). In the score space, the three datasets can be discriminated on the basis of their spectral characteristics (X-score axis) and their concentration in SOC (Y-score axis; Fig. 6). Along the X-axis, we can distinguish two groups (A and B+C) corresponding to the campaigns in 2003 and 2005. Along the Y-axis, there are two groups again (A+B and C) differing in their carbon content range (see table II). These observations are encouraging because the model constructed by PLSR appears to be able to take into account this double variability in spectral shape and carbon content. This fact is particularly important when assessing the robustness of the calibration at the broader scale and between measurement campaigns. Indeed, the power of the technique would be greatly enhanced if the technique could show a high stability across time and space. 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 coming 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 conditions. However, Kemper et al., 2005 have reported an effective calibration with a large number of heterogeneous samples collected from different regions of Italy.
Table I-I. Predictive statistics of the best model for each sensor

<table>
<thead>
<tr>
<th>Type</th>
<th>X var^a</th>
<th>Y var^a</th>
<th>LV^b</th>
<th>N</th>
<th>Deviation</th>
<th>RMSEC^c</th>
<th>Bias</th>
<th>RMSECV^d</th>
<th>RPD^e</th>
<th>Outlier</th>
<th>Var scaling</th>
<th>Treatment</th>
<th>Window size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging spectroscopy</td>
<td>97.49</td>
<td>98.59</td>
<td>1</td>
<td>104</td>
<td>0.24</td>
<td>0.16</td>
<td>0.0007</td>
<td>0.16</td>
<td>1.46</td>
<td>6</td>
<td>no</td>
<td>Savitzky-Golay 1st derivative</td>
<td>7</td>
</tr>
<tr>
<td>Field spectroscopy</td>
<td>99.99</td>
<td>85.97</td>
<td>8</td>
<td>94</td>
<td>0.25</td>
<td>0.09</td>
<td>0.0429</td>
<td>0.11</td>
<td>2.29</td>
<td>5</td>
<td>yes</td>
<td>Savitzky-Golay smoothing</td>
<td>129</td>
</tr>
<tr>
<td>Field spectroscopy 2005</td>
<td>99.53</td>
<td>98.95</td>
<td>8</td>
<td>182</td>
<td>0.69</td>
<td>0.19</td>
<td>0.00001</td>
<td>0.23</td>
<td>3.04</td>
<td>2</td>
<td>no</td>
<td>Savitzky-Golay 1st derivative</td>
<td>65</td>
</tr>
</tbody>
</table>

Table I-II. Summary statistics of sampled SOC (%) for the different field campaigns

<table>
<thead>
<tr>
<th>Study area</th>
<th>Sampling date</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attert</td>
<td>October 2003</td>
<td>37</td>
<td>1.33</td>
<td>0.49</td>
<td>0.57</td>
<td>2.28</td>
<td>Stevens et al. 2006</td>
</tr>
<tr>
<td>Ortho</td>
<td>October 2003</td>
<td>65</td>
<td>2.69</td>
<td>0.35</td>
<td>1.99</td>
<td>3.73</td>
<td>Stevens et al. 2006</td>
</tr>
<tr>
<td>Tintigny</td>
<td>June 2005</td>
<td>99</td>
<td>1.32</td>
<td>0.27</td>
<td>0.59</td>
<td>2.21</td>
<td>Current paper</td>
</tr>
<tr>
<td>All campaigns</td>
<td>2003-2005</td>
<td>201</td>
<td>1.77</td>
<td>0.73</td>
<td>0.57</td>
<td>3.73</td>
<td>Current paper</td>
</tr>
</tbody>
</table>

Fig. I-6 (a) 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), (b) Plot of Y-Score vs X-Score of the second Latent Variable as obtained after a cross-validation on the joined field spectroscopic dataset. Samples are labeled according to their origin: “A” for Tintigny 2005, “B” for Attert 2003, and “C” for Ortho 2003.

The development of a regional calibration would significantly reduce the costly sampling and calibration phases. This next step can be taken through the establishment of spectral libraries regrouping soil spectra from the same region (e.g. Shepherd and Walsh, 2002) and that could be re-calibrated at fixed time intervals. Those spectral libraries should be constructed by setting the
precise conditions in which the measurements are recorded (soil surface status, number of scans per samples, etc...). Samples that do not meet criteria fixed by the procedure should be avoided. Once the measurements have been completed, a PCA can reveal spectral dissimilarities between samples (Brown et al., 2005; Cohen et al., 2005). Samples that are not from the same population as the one in the library have to be removed. Ideally, the library should include the natural variability of the property of interest in the region. A greater variability in a training set may increase the ability of the statistical model to characterize diverse samples but it may results in a decrease of the prediction accuracy of the PLSR (McCarty et al., 2002). This is the reason why RMSECV/RMSEP's often depend on the SD of the samples. The extent of the spectral library and thus the region that it covers is a trade-off between these two constraints, a trade-off that depends on the level of precision that a given application requires.

I.4 CONCLUSIONS

Field spectroscopy showed good performance when measuring SOC under specific surface conditions (low variability in soil moisture content, low roughness and absence of vegetation). The potential rapid \textit{in-situ} sampling of this technique, requiring very little prior sample manipulation, outweighs the small loss of precision in comparison with traditional methods. The technique can thus be used for monitoring studies where their speed is a valuable advantage. Airborne techniques, such as IS, appear, for the time being, not able to predict SOC with an acceptable accuracy due to a low SNR. Nevertheless, the greater potential lies in this technique and more effort has to be put in spectrum calibration.

I.5 REFERENCES


II. The influence of fractional vegetation cover on the Quantification of Soil Organic Carbon with Imaging Spectroscopy.

Harm Bartholomeus*a, Antoine Stevensb

*a Centre for Geo-Information, Wageningen University, Droevendaalsesteeg 3, NL 6708 PB, Wageningen, The Netherlands
*b Department of Geography, Université catholique de Louvain, Place Pasteur, 3, 1348 Louvain-La-Neuve, Belgium

* Corresponding author
Harm Bartholomeus
Droevendaalsesteeg 3, 6708 PB, Wageningen, The Netherlands
Phone: 0031-317-474578
e-mail: harm.bartholomeus@wur.nl

II.1 INTRODUCTION

As shown in the previous part of this report quantitative determination of the soil organic carbon content using soil reflectance is well possible under controlled conditions (Stevens et al., 2006). Although it appears that the AHS images do not meet the quality requirements to estimate the SOC for the entire spatial extent of the study area, an analysis towards the influence of vegetation on the accuracy of the prediction model was done. First, the effect of vegetation on the predictive capacity of the PLS-model was studied. Next, the possibility of quantitative determination of the fractional vegetation cover in the study area was investigated.

II.2 INFLUENCE OF VEGETATION COVER ON SOC PREDICTION

Due to the non-optimal planning of the image acquisition, fields were partially covered with maize. As a result, the spectral signatures of the fractionally covered pixels show a combined
spectral response, containing features of both soil and vegetation. This process is called spectral mixing (Hapke, 1981).

The influence of fractional vegetation cover on the accuracy of the SOC prediction was simulated by mixing maize spectra with bare soil spectra, with fractions of maize varying between 0 and 1. Three soil samples that vary in SOC content were used for the mixing (samples G1-18; P4-01 and P5-08), which was done linearly according to:

\[ R_{\text{mix}} = f_{\text{maize}} \times R_{\text{maize}} + f_{\text{soil}} \times R_{\text{soil}} \]

(1)

In which: \( R \) = reflectance of the mixed spectrum, maize and soil respectively; \( f \) = the fraction of maize and soil within the mixed spectrum.

Next, the SOC content of the mixed spectra was predicted with the PLS-model calibrated on soil spectra measured under laboratory conditions, which showed that the difference between the PLS-predicted SOC content and the measured value is very sensitive to the influence of vegetation. A small fractional maize cover already leads to an overestimation of a few percent SOC. Since the SOC concentrations in the study area are relatively low, this has a strong effect on the prediction accuracy of the model for all three soil samples. SOC was overestimated with 0.5% at a fractional maize cover of 4-6%, which means that the SOC of sample p4_01 will be estimated wrong by nearly 100% of the measured value (0.59%) at low fractional vegetation cover. The difference between the measured and predicted SOC content for several fractional covers is shown in Figure 1.

Fig. II-1. The error in SOC estimation plotted against the Fractional Maize Cover. SOC concentrations are estimated with PLS-regression, calibrated on soil spectra measured under laboratory conditions.
II.3 Quantification of Fractional Vegetation Cover

The strong sensitivity of the prediction model to fractional vegetation cover indicates that an accurate mapping of SOC over the entire spatial extent of the research area is only possible if the vegetation cover is absent or very low. At the time of image acquisition most fields had a fractional vegetation cover between 0 and 70%. The fields which were used for the collection of soil samples and spectral measurements had a fractional cover in the range of 20-40%, so according to the simulation results this would mean that the SOC content would be overestimated by 2-5% SOC, which is more than the SOC-range (0.59 to 2.21%) in the study area.

Bartholomeus et al. (2005) showed that quantitative prediction of soil variables is also possible in partially vegetated areas, when the amount of fractional vegetation cover can be estimated accurately. How accurately the $f_{\text{cover}}$ has to be determined depends on the robustness of the prediction model, but as shown above the PLS-model can predict the SOC content from laboratory spectra up to 1% SOC accurate, as long as the $f_{\text{cover}}$ is below 10%. This means that, theoretically, it would be possible to estimate the SOC content of a pixel with 1 percent accurately, if the fractional vegetation cover can be determined up to 10% accurately, and the quality of the PLS model for image data would reach the same accuracy as for the laboratory measurements.

The fractional cover of the two subsets of the study area was determined using several Linear Spectral Unmixing algorithms (LSU), with different settings. To validate the unmixing results, digital photos were taken from a height of 3-4 meters, with a digital photocamera (Nikon Coolpix 800). The pictures were classified with a maximum likelihood classifier into vegetation and non-vegetation, where the occurrence of both classes within the pictures was taken as a measure for the fractional vegetation cover (Figure 2). The locations of the 32 plots were determined with a Garmin Etrex Summit GPS. The spatial accuracy of this GPS is about 5 meters, which means that per pixel comparison of the results may lead to large differences, unless homogenous plots are taken into account only. For this reason 2 plots were excluded from the validation process, since they were located close to borders of fields. Furthermore, a number of plots were relocated manually, according to field descriptions and field maps.

Spatial subsets of two North-South AHS flight lines (P03 and P04), with a ground resolution of 2.6 m were used for the estimation of the fractional vegetation cover. The images overlap partly, so part of the 30 plots is present on both images and measured with different viewing geometry.
The spectral signatures of the endmembers (bare soil and maize) were measured with an ASD contact probe. Before unmixing the endmembers were resampled to the spectral properties of the AHS-sensor.

Fig. II-2. Photos acquired with a digital camera at a height of 3-4 meters were used to determine the fractional vegetation cover.

To determine the fractional vegetation cover we used Linear Spectral Unmixing (Adams et al., 1985) in IDL-ENVI 4.2. LSU was done with different weights for the constraints and different pre-processing levels of the image data. For a number of LSU tests Minimum Noise Fraction Rotation (MNF) (Van der Meer and De Jong, 2000) was applied on the image data, and same rotation factors were applied on the endmember spectra. In this way the noise in the images was reduced and uncorrelated data were used for the LSU, which is a common procedure to improve unmixing results. Furthermore, Mixture Tuned Matched Filtering (MTMF) (Boardman, 1998) was applied on a subset of MNF rotated images. Finally, the usability of the Spectral Angle Mapper algorithm (SAM) to estimate fractional vegetation cover was tested (Kruse et al., 1993). For this approach a set of 101 endmembers was created by artificially mixing maize and bare-soil spectra, in fractions ranging from 0% to 100% in steps of 1%. The spectral angle between the pixel and all simulated spectra was calculated, and the fractional cover of the most similar spectrum was assigned to the pixel.

For each plot the fractional cover determined with the digital camera photos was compared with the results of the different settings for the LSU-algorithm. The amount of correspondence
between these sets is expressed with the Standard Error of Calibration (SEC) and by the number of plots for which the fractional vegetation cover was estimated within 10% fractional cover accurate. The test of the accuracy was done for both flight strips separately, to investigate the influence of viewing geometry on the LSU results.

Besides the regular output of the unmixing results, we decided to locally tune the output results. A linear regression was done between the measured and calculated fractional cover values of four plots, which were afterwards left out of the validation process. The linear correction function was then applied on the other plots, so the estimated fractional maize cover could be corrected for structural over- and underestimation, due to deviations in the selected endmembers and the viewing geometry.

II.4 Results

Figure 3 shows the scatter plots of the results of the used algorithms, where the fractional cover measured in the field is plotted against the fractional cover derived from the image. Although the SAM results in a structural overestimation of the fractional cover it results in the highest correlation between the measured and image-derived fractional cover. The overestimation of the fractional cover with image P03 is structurally higher than the fractional cover of the same plots in image P04, which indicates that the viewing geometry has influence on the outcome. All other algorithms are unable to give a proper estimation of the fractional vegetation cover. This is expressed with a low coefficient of correlation and by a lack of structure in the scatter plots.

In table 1 a summary of the deviations between the measured and image derived fractional cover values is given. The average deviation of all techniques is rather high. All algorithms show an average deviation of at least 10%, either in positive or negative direction, which is also expressed in the number of plots which has an estimation error of +/- 10%. An unconstrained LSU on the first 10 layers of the MNF rotated image gives the best results, as is indicated by the SEC too, but there is a large dependency on the viewing geometry. All together, these results are unsatisfactory and the results can not be used for further analysis of the images.

The local tuning of the outputs increases the quality of the fractional cover estimation. Because of the linear transformation the average deviation is, in most cases, close to zero (table 1) and coefficients of correlation remain the same. SAM clearly outperforms the other unmixing techniques, which is shown by the low SEC-values and the relatively large number of plots which is estimated within 10% accurately. Figure 4 shows that most samples are nicely situated close to the 1:1 line. Plot A8, which has fractional vegetation cover of 61.9% shows the largest
deviation, which can be explained by the fact that this plot contains a lot of grass and a relative low percentage of maize. Because the reflectance of the grass and maize are different, this plot shows a structural under-estimation for the fractional vegetation cover. Since only maize spectra have been used for the composition of the SAM endmembers and LSU, this has an effect on all used algorithms. Local tuning requires additional field measurements, but shows to be worth the extra effort.

Fig. II-3. Scatter plots of unmixing results. The fractional cover estimated from the image is plotted against the fractional vegetated cover measured in the field.

Finally, the fractional vegetation cover of the study area was calculated with the locally tuned SAM (Figure 5), which has an average absolute deviation of the fractional cover of 8.4%. Although there are of course a large number of individual pixels which have a larger deviation, this would mean that the SOC content in the study area can theoretically be mapped with an accuracy of about 1% SOC. This is a relatively large deviation in this area, where SOC concentrations are low, but the large number of measurements that can be done instantaneously with an airborne imaging spectrometer makes this a valuable technique. Unfortunately the quality of the AHS sensor does not come close to the quality of spectral measurements under controlled conditions.
Fig. II-4. Scatter plots of locally tuned unmixing results. The fractional cover estimated from the image is plotted against the fractional vegetation cover measured in the field.

<table>
<thead>
<tr>
<th>Method</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
<th>P03</th>
<th>P04</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU-MNF 1-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU-MNF 2-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTMF on MNF 2-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSU-MNF 2-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Constrained</th>
<th>Unconstrained</th>
<th>Constrained</th>
<th>Unconstrained</th>
<th>Constrained</th>
<th>Unconstrained</th>
<th>Constrained</th>
<th>Unconstrained</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²-value</td>
<td>0.55</td>
<td>0.66</td>
<td>0.15</td>
<td>0.24</td>
<td>0.41</td>
<td>0.37</td>
<td>0.40</td>
<td>0.33</td>
</tr>
<tr>
<td>STDEV</td>
<td>11.00</td>
<td>11.28</td>
<td>16.83</td>
<td>17.83</td>
<td>12.06</td>
<td>15.33</td>
<td>12.10</td>
<td>15.18</td>
</tr>
<tr>
<td>FC &gt; 10% wrong</td>
<td>26</td>
<td>8</td>
<td>26</td>
<td>13</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-10% &lt; FC &lt; 10%</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>14</td>
<td>10</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>FC &lt; -10% wrong</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>16</td>
<td>5</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>SEC (SD = 15.45)</td>
<td>23.51</td>
<td>16.00</td>
<td>34.18</td>
<td>23.92</td>
<td>18.83</td>
<td>14.91</td>
<td>19.40</td>
<td>25.04</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>18</td>
<td>30</td>
<td>18</td>
<td>30</td>
<td>18</td>
<td>30</td>
<td>18</td>
</tr>
</tbody>
</table>

|          | Manual |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|          |        |     |     |     |     |     |     |     |     |     |     |     |     |     |
|          | AVERAGE | -1.95 | -1.26 | -8.32 | 1.89 | 6.01 | 2.50 | 6.56 | 5.89 | -0.92 | 8.25 | 1.82 | 5.34 | 5.36 | 12.88 |
|          | STDEV  | 10.82 | 11.64 | 20.17 | 17.21 | 12.41 | 16.01 | 12.62 | 16.81 | 38.94 | 25.05 | 13.71 | 20.44 | 12.68 | 22.41 |
|          | FC > 10% wrong | 2    | 4    | 5    | 4    | 12    | 4    | 12    | 4    | 9    | 5    | 7    | 5    | 7    | 5    | 12    |
|          | -10% < FC < 10% | 18   | 7    | 6    | 8    | 13    | 7    | 13    | 9    | 7    | 6    | 13   | 7    | 13   | 7    | 13    |
|          | FC < -10% wrong | 6    | 3    | 15   | 2    | 1    | 3    | 1    | 1    | 10   | 3    | 6    | 2    | 1    | 1    | 1    |
|          | N      | 26   | 14   | 26   | 14   | 26   | 14   | 26   | 14   | 26   | 14   | 26   | 14   | 26   | 14   | 26   | 14   |
Fig. II-5. Fractional vegetation cover in part of the study area. Units are in percent and derived by a locally tuned SAM fractional cover estimation. Coordinates are given in UTM Zone 31N WGS-84.

Table II-1. Summary of the accuracy assessment. The R2-value describes the coefficient of correlation between the measured and image-derived fractional cover. All other measures are calculated for the difference values of the image derived minus the measured fractional cover in the field.

The determination of fractional vegetation cover is not a final result in this research. Further study will be done towards the possibility of mapping SOC variability in the entire study area.

II.5 CONCLUSIONS

This case study shows that the influence of fractional vegetation cover on the estimation of Soil Organic Carbon from spectroscopic measurements is large. Low fractional vegetation cover leads to significant overestimation of the SOC concentration, which makes the use of imaging spectroscopy data for this purpose sensitive to a proper timing of the flight campaign. Estimation of this fractional vegetation cover from the AHS image data, using standard unmixing techniques does not yield acceptable results. SAM gives better results, although it is sensitive to differences in viewing geometry. Local tuning of unmixing results can improve the results largely and is especially useful to correct for differences in viewing geometry and
differences between endmembers derived under controlled conditions and the endmembers present in the mixed pixels.

**II.6 FURTHER RESEARCH**

In a next stage of this research we will investigate if it is possible to remove the effect of fractional vegetation cover from the spectral signature of mixed pixels, in such a way that quantitative determination of the SOC-content can be done over a large area.

**II.7 REFERENCES**


III. Effects of other disturbing factors on soil reflectance

Rosillon Damien*, Tychon Bernard*

* Corresponding author
Damien Rosillon
Phone: +32 63.23.09.75
e-mail: damien.rosillon@ulg.ac.be

III.1 INTRODUCTION

During field campaigns, a lot of factors may disturb the measurements and decrease the signal to noise ratio of soil surface reflectance. According to Atzberger (2000), the main disturbing factors are soil water content, vegetation residues and surface roughness. Independent from wavelength, reflectance decreases exponentially with soil water content. This decrease in soil reflectance, even outside the water absorption bands is due to the so-called “trapping effects”. Moreover small amounts of living or dead vegetation residues change the soil reflectance dramatically. During this field campaign, the hyperspectral measurements were achieved over bare and dry soils. So soil roughness is supposed to be the main disturbing factor and will therefore be analysed in more details.

Some investigations have been led to study the effect of roughness on soil reflectance. For a relatively rough soil surface, Arnfield (1975 in Matthias et al., 2000) showed that the albedo will generally be lower than the corresponding smooth surface due to self shadows. Atzberger (2000) simulated the influence of the soil roughness on the soil reflectance by using the SOILSPEC model. He came to the conclusion that, due to decreasing shadows, soil reflectance increases throughout the visible range when the soil becomes smoother.

Geometrical models have been developed to simulate bidirectional reflectance of light from rough soil surface based on the assumption that reflection is strongly correlated with the area of shadowed soil as well as on the illumination and the viewing geometry. Even if those models
are validated (Cierniewski, 1987, Cierniewski and Verbrugghe, 1997) their application in practical on real time and on the field is not easy. A lot of inputs have to be known and the assessment of those parameters is quite complex in practical cases. Moreover such models are a simplification of the reality. For example, some models assume only Lambertian reflection from directly illuminated soil fragments.

Several geometrical models predict soil reflectance based on the assumption that shadowing of soil aggregates or clods has a greater influence than the scattering properties of a soil (Cierniewski and Verbrugghe, 1997). Our study is also based on this assessment. So we will bring the study of the roughness effects to the study of the shadowing effects.

The first object of this chapter is to study the impact of disturbing factors, principally of soil self shadowing, on soil reflectance. Two methods are used. The first method focuses on soil shadow. The second method takes into account all the disturbing factors. For each method, the impact of disturbing factors at different wavelength is detailed and correction factors are calculated to correct soil reflectance measurements. The second objective is to understand the relationship between roughness and relative shadow.

**III.2 MATERIAL AND METHOD**

The study of the roughness-shadow relationship in the field depends on parameters that are uncontrollable such as lightening conditions. The sun position varies with the date and the hour and can not be considered as constant. So it makes the analysis more complex. That is why two types of measurements were led: field measurements and laboratory measurements. The field measurements took place during the day of the flight campaign (20th June 2005) and 2 days later (22nd June 2005).

**III.2.1 Field measurements**

For each square plot of 9m x 9m, 4 or 6 photographs of the soil surface (jpeg format) are taken randomly with a digital camera (Nikon Coolpix 3200) in high resolution (2048 x 1536 pixels). The number of photographs depends on the plot. The zenithal angle of the camera is supposed to be zero. On each photo, a compass indicates the north direction and a rule is used as length reference. The hour at which the photos are taken is noted. The roughness is measured with a roughness meter (50 X 50 cm size), 2 measurements per square (Fig 1). The rows of nails are parallel to the furrow. The roughness meter is compounded of 10 rows of 10 nails ordered in a 5
cm x 5 cm grid on a holed plank. Unlike a linear roughness meter, this tool permits to measure roughness in different directions.

![Roughness meter](Fig. III-1. Roughness meter)

**III.2.2 Laboratory measurements**

The laboratory measurements were led on an artificial rough soil in order to control parameters influencing soil shadow (i.e. lightening, soil roughness, illumination and viewing geometry,…) (Fig.2). As Cierniewski’s approach (1999) simulating the soil surface by equal-sized opaque spheroids, the soils are modelled using painted half spheres. There are 3 sizes of half spheres: 50, 30 and 15 mm diameter. For each size, they are arranged regularly or randomly for different surface occupations (20, 40, 60, 80 and 100%). Soil surface is also modelled with a mix of different size half spheres. Soil model compositions are included in annex 1.

<table>
<thead>
<tr>
<th>Occupied surface (%)</th>
<th>100</th>
<th>80</th>
<th>60</th>
<th>40</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regular</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
</tr>
<tr>
<td>random</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
<td>![Image](Fig. III-2: soil models with 50 mm half spheres)</td>
</tr>
</tbody>
</table>
Photos are taken vertically under three illumination conditions. Models are lighted by a 500W spotlight.

### III.2.3 Data treatment

The aim of the photographs is to assess the shadow due to soil aggregates and clods for each plot. This operation is realized with eCognition professional 4.0, a Definiens software permitting an object oriented image classification (Fig.3). The first step consists in removing the disturbing objects from the photo (the compass, the rule, the date and hour display) in order to facilitate the classification and to remove the shadow due to the objects. It is realized with Adobe® Photoshop® 6.0. The second step, realized in eCognition, is a segmentation of the image in order to convert the pixel-based image in an image compounded of small homogeneous objects. Tests have been realized to find the best segmentation parameters. During the last step, the objects are classified into 3 classes: light, shadow and cut area. Once the classification is done, 2 .csv files are exported. The first one contains the number of pixels in each classes and the second one contains the number of pixels for each objects, its area and its classification.

![Step 1: removing disturbing objects](image1)

![Step 2: image segmentation](image2)

![Step 3: image classification](image3)

**Fig. III-3:** Image treatment
III.2.3.1 Impact of shadow on soil reflectance

Two methods are used to assess the impact of the shadow on the soil reflectance. In method 1, the impact is assessed by comparing the weighed average reflectance of the soil with shadow and without shadow for each band. In method 2, the impact is assessed by comparing the ASD measurements in the field and in laboratory, assuming that laboratory observations are done on soils whose clods were withdrawn.

a) method 1

The impact is assessed by comparing the weighed average reflectance of the whole soil with light and shadow area (R) and with light area only (R_l) for each band red, green and blue (RGB). The reflectance of each object is weighed by its area. The reflectances are calculated with the equations 1a and 1b.

\[
R_j = \frac{1}{n} \sum_{i=1}^{n} a_i R_{i,j} \quad (1a)
\]

with \( R_j \), the average soil reflectance (shadow and light areas) in the band j
\( R_{i,j} \), the soil reflectance for the object i (classified in shadow or light) in the band j (DN value)
\( n \), the number of objects in the shadow and the light classes
\( a_i \), the area of the object i (pixels)

\[
R_{i,j} = \frac{1}{m} \sum_{i=1}^{m} a_i R_{i,j} \quad (1b)
\]

with \( R_{i,j} \), the average soil reflectance without shadow in the band j
\( R_{i,j} \), the soil reflectance for the object i (classified in light only) in the band j (DN value)
\( m \), the number of objects in the light class
\( a_i \), the area of the object i (pixels)

The soil reflectance correction factor for the band j (K_j) is calculated by the equation 2.

\[
K_j = \left( \frac{R_{i,j} - R_j}{R_{s,j}} \right) \times 100 \quad (2)
\]

with \( K_j \), The soil reflectance correction factor for the band j (%)
\( R_j \), the average soil reflectance of the photo in the band j
\( R_{ij}, \) the average soil reflectance without shadow in the band \( j \)

Finally, the soil reflectance estimated with the AHS sensor in the band \( j \), corrected from the shadow effects is calculated with the equation 3. It represents the reflectance of a soil without shadow.

\[
R_{c,j} = R_{0,j} \times (1 + \frac{K_j}{100})
\]

(3)

with \( R_{c,j} \), corrected soil reflectance (%)

\( R_{0,j} \), field soil reflectance measured by AHS (%)

This method is used to correct AHS data from shadow effects.

\( b) \) method 2

The impact of the disturbing factors on the soil reflectance is assessed by comparing for each plot, the ASD measurements in the field and in laboratory. The reflectance measured in laboratory is considered free from disturbing factors (the shadow and the moisture are nil). The reflectance (\( R \)) is converted into absorbance (\( A \)). Then, a correction factor is calculated for each band and for each pixel (equations 4).

\[
L_{abs,i} = \left( \frac{A_{i, field} - A_{i, lab}}{A_{i, field}} \right) \times 100
\]

with : \( L_{abs,i} \), correction factor of soil absorbance for wavelength \( i \) (%)

\( A_{i, field} \), the absorbance measured in field in the wavelength \( i \) (%)

\( A_{i, lab} \), the absorbance measured in laboratory in the wavelength \( i \) (%)

III.2.3.2 Roughness – soil shadow relationship

The second objective is to study the roughness-shadow relationship. The soil shadow is quantified by the relative shadow (\( S \)) calculated by the equation 5.

\[
S = \frac{P_s}{P_s + P_l} \times 100
\]

(5)
with $S$, the relative shadow (%)

$P_s$, the amount of pixels in shadow class

$P_l$, the amount of pixels in light class

Classically, the roughness, $\rho$, is calculated by the standard deviation of the height of the 100 nails (equation 6).

$$\rho = \sqrt[2]{\frac{1}{n-1} \sum_{i=1}^{n} (Z_i - Z_m)^2}$$

(6)

with $\rho$, the soil roughness (cm)

$n$, the number of nails

$Z_i$, the height of the nail $i$ (cm)

$Z_m$, the average height of the nails (cm)

But this method biases the results if the plot has been ploughed and that furrows are pronounced (Fig. 4).

Fig. III-4. Influence of furrow on roughness.

In such a case, the roughness measured perpendicularly to the furrow includes the variation of the surface due to the furrow in addition to the variation due to earth clods. That is why for some
squares, the roughness calculated by the standard deviation of the height of 100 nails is important even though soil surface is smooth. In our case, we are interested by the roughness due to soil aggregates only because they are responsible for the soil shadow. Furrow impact on the shadowing is negligible on our field observations. To avoid this problem, the roughness is given by the average of the standard deviations calculated for each row of nail parallel to the furrows (equation 7).

\[ \rho = \frac{1}{m} \sum_{j=1}^{m} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Z_{ij} - Z_{mj})^2} \]  

with \( \rho \), the soil roughness (cm)
\( m \), the number of nails rows parallel to the furrow
\( n \), the number of nails in each row
\( Z_{ij} \), the height of the nail i in the row j (cm)
\( Z_{mj} \), the average height of the nails in the row j (cm)

### III.3 RESULTS AND DISCUSSION

#### III.3.1 Impact of soil shadow on soil reflectance

**III.3.1.1 Method 1**

\( a) \) Correction factors assessment

The correction factors, K, have been calculated for each plot. The difference between the soil reflectance with and without shadow is quite important. K average varies from 19 to 34% in function of the plots (Table 1). These results are slightly higher than the results of Obukhov and Orlov (1964 in Cierniewski, 1987) which maintained that structureless soils reflect from 15 to 20% more light than soils having a well-developed structure. The shadow decreases strongly the soil reflectance in the visible range and has to be taken into account in the SOM assessment.

For each plot, the influence of the shadow in the visible part of the spectrum decreases slightly when the wavelength increases. Cierniewski (1986 in Cierniewski, 1987) studied the reduction of soil reflectance due to soil shadow on the basis of laboratory spectrophotometric studies of soils. He also found that the influence of shadow is slightly higher to short wavelengths in the visible range.
Table III-1: Correction factors in RGB for the different plots

<table>
<thead>
<tr>
<th>Plot</th>
<th>G1</th>
<th>G2</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>P6</th>
<th>P7</th>
<th>P8_1</th>
<th>P8_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>K red</td>
<td>33</td>
<td>31</td>
<td>22</td>
<td>27</td>
<td>19</td>
<td>23</td>
<td>25</td>
<td>19</td>
<td>25</td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td>K green</td>
<td>34</td>
<td>31</td>
<td>22</td>
<td>28</td>
<td>19</td>
<td>23</td>
<td>25</td>
<td>19</td>
<td>26</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>K blue</td>
<td>36</td>
<td>33</td>
<td>23</td>
<td>29</td>
<td>19</td>
<td>24</td>
<td>27</td>
<td>20</td>
<td>29</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>K average</td>
<td>34</td>
<td>32</td>
<td>22</td>
<td>28</td>
<td>19</td>
<td>24</td>
<td>26</td>
<td>19</td>
<td>27</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>S</td>
<td>14</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>12</td>
<td>26</td>
<td>21</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>24</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>4</td>
<td>16</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>R²</td>
<td>0.17</td>
<td>0.80</td>
<td>0.95</td>
<td>0.04</td>
<td>0.58</td>
<td>0.10</td>
<td>0.62</td>
<td>0.49</td>
<td>0.24</td>
<td>0.65</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Rem: P8_1 and P8_2 are observations from the P8 plot but they were taken at different moments and at different dates. That is why they are considered as two distinct datasets.

We studied the correlation between K average and relative shadow (S). By considering the relationship for each plot separately, some good correlations appear, depending on the plot. In all cases, K average increases with S but the slope of the regression lines depends on the plot.

There is no general relationship if all the plots are taken together in one dataset but if the intensity of soils colour is also taken into account in a multiple linear regression, we find the following relation (equation 8).

\[ K_{average} = -12.48 + 0.0003 \cdot I + 1.0391 \cdot S \]  \hspace{1cm} (8)

with \( K_{average} \), the average of the correction factors for the RGB bands

I, the intensity of soil colour (equation 9)

S, the relative shadow

\[ I = \frac{1}{3} \sum_{i=1}^{3} R_{i,j} \]  \hspace{1cm} (9)

with I, the intensity of soil colour

\( R_{i,j} \), the average soil reflectance without shadow in the band j

3, the number of measured bands
I is a characterization of the lightness of a soil. If the soil is dark, I is low. If the soil is clear, I is high.

The relationship given by the equation 8 is not so bad. In cross validation, prediction $R^2$ is 0.42 (Fig.5). Relationship is significant. Including I to the $K_{\text{average}} - S$ relationship permits to generalize the relationship to all the plots. The impact of shadow on soil reflectance depends on intensity of soil colour. Brighter is the soil, higher is the shadow impact on reflectance.

![Fig. III-5. $K_{\text{average}}$ predicted vs $K_{\text{average}}$ observed](image)

b) AHS measurements correction

Correction factors are tested to correct AHS measurements. Only the VNIR part of IS spectrum is used to assess SOC content (400 – 900 nm). For each plot, correction factors do not vary a lot with wavelength so we suppose that average correction factor may be applied to the IS spectrum. The corrections are applied to the pixels where the photographs are taken around 12:00 am (flight time). We correct the pixels where the photographs are taken between 12:00 am and 13:45 pm. During this period, solar zenithal angle did not change a lot and soil shadow is supposed to be similar (Fig. 6). 27 pixels are corrected.
SOC content is assessed for the 27 pixels with the corrected IS spectrum following the methodology detailed above (signal pre-processing and PLSR). SOC content is also assessed for the same pixels with the uncorrected signal. RPD for the 2 treatments are compared by a paired t-test. Difference is positive (1.432) and highly significant (p < 0.0001). The correction factors calculated by method 1 improve positively and significantly SOC content assessment with AHS data.

Rem. : Those corrections factors have not been used to correct ASD data because soil carbon content estimation with ASD data requires wavelength in the whole spectrum. Method 1 corrects only visible wavelengths.

III.3.1.2 Method 2

a) Correction factors assessment

Method 2 is used to study the impact of all disturbing factors on soil absorbance in VIS and SWIR. Correction factors (L_{abs,i}) are estimated for each wavelength by comparing ASD field and ASD laboratory measurements (Eq.4). For some square plots, the absorbance measured in the field (A_{field}) is lower than the absorbance measured in laboratory (A_{lab}). Normally, A_{field} should be higher than A_{lab} because disturbing elements increase absorbance.

---

1 Sources : University of Oregon, Solar Radiation Monitoring Laboratory. http://solardat.uoregon.edu/cgi-bin/SolarPositionCalculator.cgi
However, if a crust is observed in the field, the soil reflectance may increase (Kondratyev et Fedchencko, 1980 in Cierniewski, 1987). But by analysing the photographs, the crust does not look more important on these squares than on the neighbouring squares. The observations where $A_{\text{field}}$ is inferior to $A_{\text{lab}}$ are excluded from our analysis. Another hypothesis may be the difference of incident ray in the field and in laboratory.

$L_{\text{abs,i}}$ average is calculated for each plot (Fig. 7). The evolution of $L_{\text{abs,i}}$ with the wavelength is similar. In the UV part of the spectrum, $L_{\text{abs,i}}$ decreases with the wavelength until a minimum value at the beginning of the visible part (400 nm). In the visible, $L_{\text{abs,i}}$ increases strongly until 750 nm. In the NIR and SWIR part, $L_{\text{abs,i}}$ increases lightly until 1500 nm. The maximum value is reached around 2200 nm.

![Graph showing the evolution of Labs,i (%)](image)

**Fig. III-7. Average correction factors (Labs,i) vs wavelength**

The impact of the disturbing factors is particularly important in the IR. In this part of the spectrum, they may increase the soil absorbance of 40%. Their influence is minimal in the first part of the visible (from 400 to 500 nm). At those wavelengths, $L_{\text{abs,i}}$ varies from 11 to 18% in function of the plot.

**Comparison with method 1**

To permit a comparison of the impact of soil shadow on soil reflectance by the 2 methods, $L_{\text{ref,i}}$ is calculated for the reflectance with equation 10.

$$L_{\text{ref,i}} = \left( \frac{R_{i,\text{lab}} - R_{i,\text{field}}}{R_{i,\text{field}}} \right) \times 100$$  \hspace{1cm} (10)
With, \( L_{\text{ref},i} \), the reflectance correction factor for wavelength \( i \)
\( R_{i,\text{lab}} \), reflectance measured in laboratory at wavelength \( i \)
\( R_{i,\text{field}} \), reflectance measured in the field at wavelength \( i \)

In figure 8 average correction factors are compared. Only the plots where photographs are taken in the same time than ASD field measurements are taken in consideration. For method 1, \( K_{\text{average}} \) data (from table 1) are reported above. For method 2, average \( L_{\text{ref},i} \) is calculated for wavelength from 400 to 750 nm.

![Fig. III-8. Comparison of correction factors](image)

Correction factors calculated by method 2 (\( L_{\text{ref},i} \)) are systematically superior to the correction factors calculated by method 1 (\( K \)). For p3 and p4, \( L_{\text{ref},i} \) is more than 2 times superior to \( K \). With method 1, only shadow effects are corrected. With method 2, all disturbing effects are corrected because soil samples are dried and sieved before measuring soil reflectance with ASD in laboratory. With such results, we can assume that shadow is not the only disturbing factor.

\( b) \ Correlation \ between \(L_{\text{lab}}, \text{and \ the\ disturbing\ factors})\)

We study here the correlation between correction factors and parameters able to modify soil reflectance (shadow, soil moisture, carbon content) to quantify the impact of each parameter. Firstly, a multiple regression is realized with a dataset including all the plots whose photographs are taken at the same time as ASD field measurements (= 48 observations).
Secondly, a multiple regression is realized for 2 plots separately. If there is a significant correlation, we study the correlation between $L_{\text{abs}}$ and each parameter separately. As for method 1, the impact of soil colour intensity on the correlations will be tested.

When we consider all plots in the same dataset for regression, the correlation between $L_{\text{abs}}$ and the 3 independent variables is not good ($R^2_{\text{max}} = 0.18$) and not significant ($p_{\text{min}} = 0.0509$).

The impact of the soil colour intensity on the correlations is studied below. The dataset includes 14 points where soil colour intensity data are available. For this dataset, the correlation is calculated excluding (Fig. 9a) and including soil colour intensity factor from/into the multiple regression (Fig.9b). Including soil colour intensity increases correlations particularly in the visible but they are not significant. $R^2$ maximum is observed around 470 nm ($R^2 = 0.38$). $p$ statistic minimum is also observed around 470 nm and equals 0.3.
There are no significant correlations between \( L_{\text{abs}} \) and disturbing factors if the dataset includes all the plots. We are going to study the correlations for two plots where the number of observations is sufficient to realize a multiple regression and where the ASD measurements in the field have been realized at the same time as the photographs. P2 (12 observations) and G2 plots (14 observations) are retained.

For p2, there is no significant relationship between \( L_{\text{abs}} \) and the disturbing factors. For G2, \( L_{\text{abs}} \) and the disturbing factors are highly and significantly correlated (Fig. 10). Correlations are important in the visible and decrease in IR but they are significant in the whole spectrum. The highest correlations are observed in the blue bands range (\( R^2 \) is around 0.79).

![Fig. III-10: Correlation coefficients (blue) and p statistic (red) for G2](image)

A typical model of the visible part is given at equation 11 (\( \lambda = 450 \) nm).

\[
L_{\text{abs,450}} = -0.121 + 0.0164 * S + 0.2183 * m + 0.0184 * c
\]  

(11)

With,  
\( L_{\text{abs,450}} \), the correction factor for \( \lambda = 450 \) nm  
\( S \), the relative shadow (%)  
\( m \), the soil moisture (%)
Soil moisture is the best correlated factor for this plot. A linear regression between $L_{\text{abs}}$ and soil moisture reveals that $R^2$ is maximum in the visible ($R^2 \text{ max} = 0.6$) and the relationship is significant. The highest p value in the whole spectrum is 0.027, observed in the SWIR. $L_{\text{abs}}$ is also significantly correlated with the carbon content in the UV, visible and the short IR wavelengths (from 350 to 960 nm). Correlations are important in the short wavelengths. From 350 to 529, $R^2$ is higher than 0.4. $L_{\text{abs}}$ is not significantly correlated to the relative shadow. Graphs are included in annex 2.

For this plot, even if soil moisture is low (average moisture = 0.53 %) it is the most disturbing factor. In the visible, soil moisture explains nearly 60% of $L_{\text{abs}}$ variation. Observed SOC also influences $L_{\text{abs}}$ but only for the short wavelength (inferior to 900 nm). For this plot, relative shadow does not influence $L_{\text{abs}}$.

Unfortunately, those observations cannot be compared with other results. p2 is the only plot where the number of observations is sufficient and the correlations are not significant. We cannot draw conclusions about the impact of the different disturbing factors. Firstly, because only 2 plots have been studied. Secondly, the field conditions were selected to minimize the disturbing effects (dry soil surface, weak roughness, minimum zenithal angle). For G2 plot, average relative shadow was low (8%), perhaps too low to influence soil reflectance significantly. Besides, the standard deviation of the disturbing factors is low compared with $L_{\text{abs}}$ standard deviation. They cannot explain $L_{\text{abs}}$ variation significantly. Ideally, the range of disturbing factors should be more important. Those considerations will be taken into account in the methodology of MOCA project which is a continuation of the present project.

c) **ASD measurements correction**

Correction factors are applied to the whole spectrum of ASD field measurements (61 samples). SOC content is assessed for this dataset with the corrected and the uncorrected signal. RPD for the 2 treatments are compared by a paired t-test. The difference is positive (1.141) but it is not significant ($p = 0.3206$). Such results are encouraging. Even if prediction improvement is not significant, it is positive.
III.3.2 Roughness – soil shadow relationship

Soil shadow influences soil reflectance. However, assessing soil shadow is difficult and requires specialised software of image analysis. In this chapter, we are going to study the relationship between soil shadow and roughness, more easily measurable.

2 datasets are available, field measurements and laboratory measurements, including soil roughness measurements and soil surface photographs.

There is no clear relationship between roughness (calculated with equation 7) and soil shadow field measurements even by considering each plot separately (Fig 11).

![Fig. III-11. Relative shadow – roughness relationship per plot](image)

This may be due to the fact that, on a given plot, photographs are taken in a period of 30 minutes. During this time, sun position varies (Fig. 6) and illumination conditions change from one measure to another. Besides, the resolution of the roughness meter is quite low (distance between nails is 5 cm). It may lead to an undersampling of the roughness. The measured roughness may be not representative of the actual roughness.

In laboratory, the illumination conditions are under control and 4 roughness measurements are realized for each soil model. Roughness is calculated on 400 nails and not on 100 nails as during the field campaign. In this way, possible problems caused by an undersampling are avoided. While working with all laboratory data (Tab. 2), there is a linear relationship between roughness
and shadow but the correlation is quite weak. Correlation coefficient in cross validation (R²cv) equals 0.30.

Studying separately the correlations for each model increases R²cv. For the soils modelled with half spheres of the same size, R²cv is higher than 0.6 but it decreases when the model soil gets more complex (mix model). In the field, the soil surface is more complex than all these soil models. It could explain why there is no relationship between relative shadow and roughness with the field dataset.

<table>
<thead>
<tr>
<th>Soil model</th>
<th>R²cv</th>
</tr>
</thead>
<tbody>
<tr>
<td>all data</td>
<td>0.30</td>
</tr>
<tr>
<td>15</td>
<td>0.64</td>
</tr>
<tr>
<td>30</td>
<td>0.74</td>
</tr>
<tr>
<td>50</td>
<td>0.62</td>
</tr>
<tr>
<td>Mix</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Table. III-2. Coefficient correlation in cross validation between soil shadow and roughness for the soil models.

The total area occupied by soil clods (Sclods, given by equation 12) is much better correlated to the shadow than roughness.

\[ S_{\text{clods}} = \sum n_i \pi \left( \frac{d_i}{2} \right)^2 \]  \hspace{1cm} (12)

With \( S_{\text{clods}} \), the surface occupied by soil clods (cm²)
\( n_i \), the number of clods of dimension \( i \)
\( d_i \), the clods diameter (1.5, 3 or 5 cm)

In an ascending stepwise linear multiple regression, Sclods is the first selected explanatory variable. The correlation coefficient in cross validation between soil shadow and Sclods for all the laboratory data equals 0.94 (Fig.12). The correlation is highly significant (p<0.0001). It means that the soil shadow depends more on the surface occupied by the clods than on the roughness.
S\textsubscript{clods} depends on clods diameter but also on clods number. Soil surface covered by big clods, therefore with high roughness, does not infer high relative shadow. With 16 clods of 50 mm diameter, roughness equals 0.63 but relative shadow equals 12 % (Table 3). Under identical lightening conditions, numerous small clods produce the same relative shadow than few big clods, even if the roughness measured in the first case is lower than in the second case. 81 clods of 50 mm diameter produce the same relative shadow as 196 clods of 30 mm diameter but roughness is 2 times superior for big clods.

<table>
<thead>
<tr>
<th>Diameter (mm)</th>
<th>Number of clods</th>
<th>Relative shadow (%)</th>
<th>S\textsubscript{clods} (cm²)</th>
<th>Roughness</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>196</td>
<td>34</td>
<td>1385</td>
<td>0.65</td>
</tr>
<tr>
<td>50</td>
<td>16</td>
<td>12</td>
<td>314</td>
<td>0.63</td>
</tr>
<tr>
<td>50</td>
<td>81</td>
<td>37</td>
<td>1590</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Table III-3. Relative shadow, S\textsubscript{clods}, and roughness for 3 soil models.

Relative shadow is mainly correlated to S\textsubscript{clods} and not to soil roughness. Consequently, 2 soils with a different roughness can produce the same relative shadow.

III.4 CONCLUSIONS

The first objective was to assess the impact of disturbing factors on soil reflectance.
According to method 1, difference between soil reflectance with and without shadow is important. The correction factors vary from 19 to 34% depending on the plots. For each plot, the influence of the shadow in the visible part of the spectrum decreases slightly when the wavelength increases.

Method 2 is used to study the impact of all disturbing factors (carbon content, soil shadow and soil moisture) on soil absorbance in VIS and SWIR. The impact of the disturbing factors is particularly important in the IR. In this part of the spectrum, they may increase soil absorbance of 40%. Their influence is minimal in the first part of the visible (from 400 to 500 nm). At those wavelengths, $L_{abs}$ varies from 11 to 18 % in function of the plot.

Disturbing factors decrease strongly soil reflectance. They have to be taken into account in the SOC assessment. The correction of shadow effects improves significantly SOC predictions with AHS data. It influences positively SOC predictions with ASD data but RPD improvement is not significant.

Intensity of soil colour may influence disturbing factors effects. For the shadow, it appears that brighter is the soil, higher is the shadow impact on reflectance.

Correlation between disturbing parameters and correction factors are studied. For method 1, correlation between relative shadow, soil colour intensity and correction factors for all plots is significant. Correlation coefficient in prediction equals 0.42.

For method 2, there are no general correlation for all plots to estimate $L_{abs}$ in function of soil moisture, soil shadow and carbon content. For G2 plot, $L_{abs}$ and the disturbing factors are highly and significantly correlated. Correlations are important in the visible and decrease in IR but they are significant in the whole spectrum. The highest correlations are observed in blue band ($R^2$ is around 0.79).

For this plot, unlike our suppositions, soil moisture is the most disturbing factor. Observed SOC also influences $L_{abs}$ but only for the short wavelengths (inferior to 900 nm). For this plot, relative shadow does not influence $L_{abs}$. However, we can not draw conclusions about the impact of different disturbing factors because field conditions were selected to minimize the disturbing
effects. Besides, the standard deviation of the disturbing factors is low compared with $L_{\text{abs}}$ standard deviation. They can not explain $L_{\text{abs}}$ variation significantly. Ideally, the range of disturbing factors should be more important.

Those considerations will be taken into account in the methodology of the MOCA project which is a continuation of the present project. For example, we could increase moisture range by dampening soil surface. To increase the shadow range, we could take measurements at different period (early in the morning, at midday and later in the afternoon). The other limit of this study was the number of usable observations. A lot of IS data could not be used because we had no linked shadow measurements that were measured with camera but not necessarily at the same time. For MOCA project, it will be important to take photography at the same time as IS measurements.

The **second objective** was to understand the relationship between roughness and relative shadow.

Soil shadow influences soil reflectance. However, assessing soil shadow is difficult and requires specialised software of image analysis. With the field dataset, there is no clear relationship between roughness and soil shadow field measurements even by considering each plot separately. With the laboratory dataset, correlations between roughness and shadow are good for the soils modelled with half spheres of the same size. However, correlations decrease when the model soil gets more complex (mix model).

In an ascending stepwise linear multiple regression, $S_{\text{clods}}$, that represents the surface occupied by soil clods, is the first selected explicative variable. Soil shadow depends more on the surface occupied by the clods than on the roughness. The coefficient correlation in cross validation between soil shadow and $S_{\text{clods}}$ for all the laboratory data equals 0.94 (Fig.11). The correlation is highly significant ($p<0.0001$).
## Annex 1: Soil models composition

<table>
<thead>
<tr>
<th>Half spheres size (mm)</th>
<th>Pattern</th>
<th>Surface occupied (%)</th>
<th>Number of half spheres</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>Ra</td>
<td>100</td>
<td>81</td>
</tr>
<tr>
<td>50</td>
<td>Ra</td>
<td>80</td>
<td>65</td>
</tr>
<tr>
<td>50</td>
<td>Ra</td>
<td>60</td>
<td>49</td>
</tr>
<tr>
<td>50</td>
<td>Ra</td>
<td>40</td>
<td>33</td>
</tr>
<tr>
<td>50</td>
<td>Ra</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>50</td>
<td>Re</td>
<td>80</td>
<td>64</td>
</tr>
<tr>
<td>50</td>
<td>Re</td>
<td>60</td>
<td>49</td>
</tr>
<tr>
<td>50</td>
<td>Re</td>
<td>40</td>
<td>33</td>
</tr>
<tr>
<td>50</td>
<td>Re</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>30</td>
<td>Ra</td>
<td>100</td>
<td>196</td>
</tr>
<tr>
<td>30</td>
<td>Ra</td>
<td>80</td>
<td>157</td>
</tr>
<tr>
<td>30</td>
<td>Ra</td>
<td>60</td>
<td>118</td>
</tr>
<tr>
<td>30</td>
<td>Ra</td>
<td>40</td>
<td>78</td>
</tr>
<tr>
<td>30</td>
<td>Ra</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>30</td>
<td>Re</td>
<td>80</td>
<td>156</td>
</tr>
<tr>
<td>30</td>
<td>Re</td>
<td>60</td>
<td>121</td>
</tr>
<tr>
<td>30</td>
<td>Re</td>
<td>40</td>
<td>77</td>
</tr>
<tr>
<td>30</td>
<td>Re</td>
<td>20</td>
<td>39</td>
</tr>
<tr>
<td>15</td>
<td>Ra</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td>15</td>
<td>Ra</td>
<td>80</td>
<td>320</td>
</tr>
<tr>
<td>15</td>
<td>Ra</td>
<td>60</td>
<td>240</td>
</tr>
<tr>
<td>15</td>
<td>Ra</td>
<td>40</td>
<td>160</td>
</tr>
<tr>
<td>15</td>
<td>Ra</td>
<td>20</td>
<td>80</td>
</tr>
<tr>
<td>15</td>
<td>Re</td>
<td>60</td>
<td>240</td>
</tr>
<tr>
<td>15</td>
<td>Re</td>
<td>40</td>
<td>156</td>
</tr>
<tr>
<td>15</td>
<td>Re</td>
<td>20</td>
<td>81</td>
</tr>
<tr>
<td>Mix 1</td>
<td>Ra</td>
<td>45</td>
<td>15 (50mm)</td>
</tr>
<tr>
<td>Mix 2</td>
<td>Ra</td>
<td>30</td>
<td>40 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>165 (15mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10 (50mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>110 (15mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 (50mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>15 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>55 (15mm)</td>
</tr>
<tr>
<td>Mix 2</td>
<td>Ra</td>
<td>43</td>
<td>22 (50mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>121 (15mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14 (50mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>22 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>80 (15mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7 (50mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11 (30mm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40 (15mm)</td>
</tr>
</tbody>
</table>
Annex 2: Correlation between $L_{abs}$ and disturbing factors (G2 plot)
III.6 REFERENCES


General Conclusion

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 VNIR spectroscopy. The study showed the potential of field and airborne spectroscopy and presented their accuracy in estimating SOC content of various soil samples.

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.

Nevertheless, the limit of accuracy of the method is sufficient to capture changes in SOC stocks at the field and regional scale. Even for IS, having still rather low precision level, the large number of sample that can be taken could improve estimates of mean SOC stocks compared to more precise chemical analyses of an inherently smaller number of samples.

The study of disturbing factors (soil roughness, moisture, fractional vegetation cover) showed that their influence is generally large. The correction of shadow effects due to soil roughness improved significantly SOC predictions with AHS data. Low fractional vegetation cover lead to significant overestimation of the SOC concentration, which makes the use of imaging spectroscopy data for this purpose highly sensitive to a proper timing of the flight campaign. Some methodologies to correct for these disturbing factors have been discussed and further research needs have been highlighted.

Obviously, there is a need for the establishment of standard spectral measurement protocols in the field (surface conditions characteristics required, sampling needs in relation to spatial variability, robustness of the calibration across physiographic region) and soil spectral libraries covering broad zones. This is the next development foreseen in order to bring spectroscopy from a potential measuring technique to a routine soil analysis method operational in real case studies.