



Snapshot Hyperspectral Imaging for Soil Diagnostics – Results of a Case Study in the Spectral Laboratory

ANDRÁS JUNG & MICHAEL VOHLAND, Leipzig

Keywords: hyperspectral camera, proximal soil sensing, multivariate calibration, spectral variable selection, partial least squares regression

Summary: Field reflectance spectroscopy has been widely used in proximal soil sensing. Results of spectroscopic approaches depend, inter alia, from the experimental setup and the applied spectroradiometric instrumentation. Beyond the traditional instrument concepts (acquisition of ground truth data with field spectroradiometers, air- and space-borne scanners), there are currently alternative developments in the ground-based or near-ground spectroscopy: The hand-held and thus mobile non-scanning hyperspectral imaging technique might be one previously missing part in the operational spectral data chain to be used for down- and up-scaling purposes. It should effectively bridge the gap between point and image data as it enables a very rapid data acquisition.

This study describes how readings of a hyperspectral frame camera (in the nominal spectral range from 450 nm to 950 nm) could be utilised for soil detection and analysis. The proximally sensed hyperspectral images were compared to 1D spectroradiometric data, both acquired in the lab using raw, sieved and grinded soil samples. Measured spectral datasets were then used to define multivariate calibration models, i.e., the spectra were analysed to extract quantitative models between spectral data and soil constituents of interest determined by wet chemical analysis. We used partial least squares regression (PLS) as statistical calibration method to estimate soil organic carbon (OC), hot-water extractable carbon (HWE-C) and nitrogen (N). The results that we obtained from the camera data were satisfactory (with coefficients of determination (R^2) between 0.62 and 0.84 in the cross-validation), but only with crushed samples and when combining PLS with CARS (competitive adaptive reweighted sampling), an effective spectral variable selection technique. For in-field studies without any sample preparation, stratified approaches considering soil surface roughness and/or the elimination of shadow pixels from the acquired images might both be promising to improve the accuracy of obtainable estimates.

Zusammenfassung: Einblick in die hyperspektrale Abbildung zur Untersuchung von Böden – Ergebnisse einer Laboruntersuchung. Anwendungen der Feldspektroskopie zur Charakterisierung von Böden sind in zahlreichen Studien aufgezeigt worden. Erzielte Ergebnisse sind unter anderem von den eingesetzten Spektroradiometern und der gewählten Messkonfiguration abhängig. Neben klassischen Instrumentierungskonzepten (Erhebung von Referenzdaten mit Feldspektroradiometern, Scanner auf Flugzeug- und Satellitenplattformen) gibt es aktuell in der bodengestützten bzw. -nahen Spektroskopie eine Reihe von Neuentwicklungen. So könnte sich eine handgeholtene und somit mobil einsetzbare bildgebende (nicht-scannende) Hyperspektralkamera als ein bislang fehlendes Element in der operationellen Spektraldatenkette erweisen, nutzbar sowohl zum Up- als auch zum Downscaling. Diese Technik sollte die Lücke zwischen Punktmessung und Bilddaten effektiv schließen können, da sie eine sehr schnelle Datenaufnahme möglich macht.

Eine Vollformat-Hyperspektralkamera (nominaler Spektralbereich 450 nm bis 950 nm) wurde in der vorliegenden Studie zur spektralen Erfassung von Böden eingesetzt. Bodennah aufgenommene Hyperspektralbilder wurden dazu mit 1D Spektroradiometerdaten verglichen. Beide Datensätze wurden im Labor für jeweils unaufbereitete, gesiebte und fein gemahlene Bodenproben aufgenommen. Die gemessenen Spektren wurden genutzt, um multivariate Kalibrierungen aufzustellen, d.h. den Zusammenhang zwischen Spektraldaten (nach adäquater Transformation) und nasschemisch gemessen Bodenparametern zu modellieren. Als Methode zur Kalibrierung wurde die „partial least squares“-Regression (PLS) genutzt, um die Gehalte an organischem Kohlenstoff (OC), heißwasserlöslichem C (HWE-C) und Stickstoff (N) abzuschätzen. Die aus den Kameradaten abgeleiteten Schätzergebnisse waren zufriedenstellend. Die Bestimmtheitsmaße (R^2) der Schätzmodelle lagen zwischen 0.62 und 0.84 in der Kreuzvalidierung.

Diese Resultate konnten aber nur erzielt werden, wenn die Bodenproben vor der spektralen Messung zerkleinert und die PLS mit einer effektiven Spektralvariablenelektion (CARS, „competitive adaptive reweighted sampling“) kombiniert wurden. Für Feldstudien ohne mögliche Probenaufberei-

tung erscheinen Submodelle für aus den Bildern geschätzte unterschiedliche Bodenrauigkeiten und/oder die Beseitigung von Schattenpixeln als aussichtsreiche Möglichkeiten, um die Schätzergebnisse verbessern zu können.

1 Introduction

Spectroscopy in the visible and near-infrared (VNIR) has been widely used in soil sensing, either in the laboratory (e.g. BEN-DOR & BANIN 1995, SUDDUTH et al. 1989, VISCARRA ROSSEL & McBRATNEY 1998) or for in-situ soil monitoring (e.g. KOOISTRA et al. 2003, UDELHOVEN et al. 2003). Typically, field reflectance spectra are collected by portable field spectrometers (1D high-resolution spectra), which are often complemented by 2D data of air- or spaceborne imaging spectrometers with a more limited spatial resolution. Compared to portable field spectroscopy, airborne imaging spectroscopy has a greater potential to cover large areas during a single flight campaign, but accuracies of estimated soil properties are usually lower due to a lower signal-to-noise ratio and disturbing atmospheric influences, for example. Variable soil and surface properties (moisture content, roughness, crusting) induce spectral variability that is critical for large area approaches and may be accounted for by using stratified (local) calibrations (STEVENS et al. 2008, 2010, HILL et al. 2010).

With traditional instrument concepts, i.e., ground truthing with field spectrometers to link ground spectra with data of air- and spaceborne scanners, there is a gap in the “point-pixel-image”-upscaling approach as proximally sensed hyperspectral image data are missing. However, ground-based imaging line-scanners are currently less widespread in ground truthing than portable field spectrometers. One reason for this is the time factor as operating a field line scanner on a tripod set-up is very time consuming compared to the use of a 1D-field spectrometer. One of the concepts to overcome this limitation and to bridge the gap in the hyperspectral data chain is non-scanning snapshot hyperspectral imaging, which enables rapid (1 ms) data acquisition in a hand-held mode (JUNG et al. 2013).

Due to the novelty of this technique there are no available references for non-scanning hyperspectral cameras used in proximal soil sensing. However, there is a comprehensive list of studies and works conducted with line-scanners. Recently, STEFFENS & BUDDENBAUM (2013) utilised a hyperspectral scanner from 400 nm to 1000 nm to determine the concentrations of carbon, nitrogen, aluminium, iron and manganese of a stagnic Luvisol profile under laboratory conditions. For air- and spaceborne scanners e.g. STEVENS et al. (2010) provide an overview of available soil studies. In addition, numerous studies exist dealing with ground-based imaging spectroscopy using the line-by-line-scanning principle for applications in geology and vegetation analysis (KURZ et al. 2013, VIGNEAU et al. 2011, YE et al. 2008).

Irrespective of the instrumentation, hyperspectral measurements provide large sets of spectral variables which are strongly correlated and often contain noise. These data have to be processed to model the relationship between spectral values and soil constituents, which can make use of a simple spectral index, for example, or a number of factors or components extracted after data projection; these factors/components should ideally represent the underlying structure and contain the most relevant information of the data. In the calibration approach, the statistical components are then modelled against the constituents determined by wet chemical analysis. After its validation, either internal with e.g. leave-one-out cross-validation or – more appropriate – external with an additional dataset, this model may be applied to unknown samples.

The statistical method that is used in the calibration process should reflect the inherent structure of the hyperspectral data and be able to handle correlated and noisy data. In chemometrics, partial least squares regression (PLS) has firmly established as a robust multivariate

calibration tool. However, improvements of accuracy are usually achieved by selecting the most informative spectral variables instead of using the full spectrum. The selection of spectral variables also tends to reduce the complexity of the multivariate model that is finally retrieved for quantification purposes (XIAOBO et al. 2010).

This paper focuses on the use of the recently available non-scanning UHD 285 hyperspectral frame camera (Cubert GmbH, Ulm, Germany) for VNIR soil sensing in a laboratory experiment. The studied sample set consisted of 40 soil samples, which were analysed for their contents of organic carbon (OC), hot-water extractable carbon (HWE-C) and nitrogen (N); spectral readings were taken with the UHD 285 and an ASD (Analytical Spectral Devices, Boulder, Colorado, USA) full range spectroradiometer for three different

states of crushing (raw, sieved, grinded). For multivariate calibration, we applied both full spectrum-PLS and PLS combined with a key wavelengths selection procedure, the CARS method (competitive adaptive reweighted sampling; LI et al. 2009). The workflow followed in this study is illustrated in Fig. 1.

2 Data Acquisition

2.1 Study Site, Sampling and Soil Wet-Chemical Analysis

The soil sampling area was situated in the Northwest Saxon Basin (Geopark Muldenland), which is characterised by Permian bedrock geology (rhyolites and ignimbrites), Cretaceous-Tertiary weathering products (like Kaolin) and quaternary sediments (loess, Pleistocene terrace gravel).

Within the study area 40 randomly selected soil samples were taken on different agricultural fields from the very top layer (Ap, 0–10 cm depth). For further wet-chemical analysis, soil samples were air-dried, sieved ≤ 2 mm, and subsequently homogenised by grinding using an agate mortar. Soil texture was determined with the Köhn sieve-pipette method (E DIN ISO 11277:1994-06 1994) and ranged from loamy sand ($n = 2$), sandy loam (11), loam (4), silt loam (22) to silt clay loam (1) (after FAO classification; FAO 2006) (Tab. 1).

The total contents of OC and N were measured by gas chromatography after dry combustion at 1100°C with an EuroEA elemental analyser (HekaTech, Wegberg, Germany), all soil samples were free of carbonate-C. Determination of HWE-C followed the method of

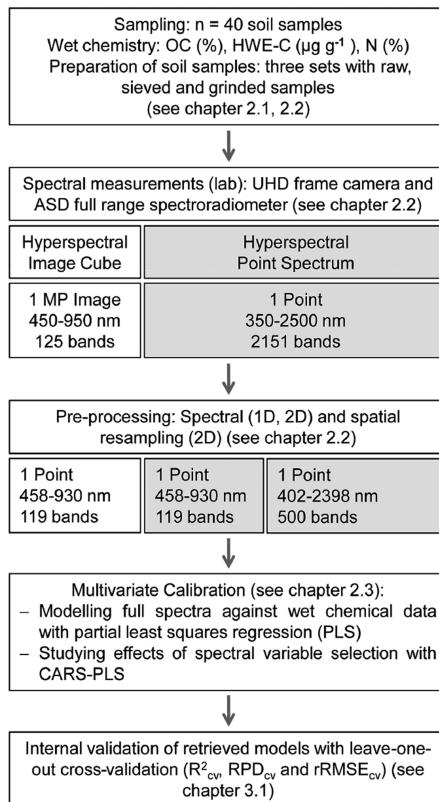


Fig. 1: Workflow to assess soil constituents from spectral data.

Tab. 1: Soil texture of three selected soil samples (from loess and sandy moraine material).

	Sand (%)	Silt (%)	Clay (%)
Soil from loess (silt loam)	5	79	16
Soil from sandy loess (sandy loam)	31	56	13
Soil over sandy moraine (loamy sand)	82	9	6

Tab. 2: Wet-chemical parameters of the studied soil samples.

	Mean	Min	Max	Std
OC (%)	1.54	0.62	4.31	0.74
HWE-C ($\mu\text{g g}^{-1}$)	652	306	1568	265
N (%)	0.145	0.048	0.377	0.068
Quotient C \times N ⁻¹	10.9	8.5	18.0	2.2

KÖRSCHENS et al. (1998) and was examined by an one hour extraction of 10 g soil with distilled water (50 ml) at 100 °C using a Gerhardt Turbotherm TT 125 (Gerhardt, Bonn, Germany). After the extraction, cooling, adding of MgSO₄ and centrifugation at 2600 rpm for 10 minutes, the dissolved organic carbon of the supernatant was analysed with a TOC-V_{CPN}-analyzer (Shimadzu, Duisburg, Germany). Tab. 2 illustrates mean, minimum, maximum, and standard deviation (std) of the analysed soil properties. In total, soil parameters given in Tab. 2 cover the values that are typical for agricultural soils.

2.2 1D and 2D Spectral Data Acquisition and Pre-Processing

For the acquisition of image data we used the UHD 285 hyperspectral frame camera. A silicon CCD chip with a sensor resolution of 970 \times 970 pixel captures the full frame images. The dynamic image resolution is 14 bit. At normal sun light illumination, the integration time of taking one hyperspectral data cube is 1 ms. The camera is able to capture more than 15 spectral data cubes per second, which facilitates hyperspectral video recording. The high-resolution imaging spectrometer coupled with the camera chip was designed and developed by ILM (Institute of Laser Technologies in Medicine and Metrology) at the University of Ulm and the Cubert GmbH. For our analysis, we used the spectral range from 450 nm to 950 nm that is covered by 125 channels. The hyperspectral data cube has a spectral resolution of 4 nm.

1D measurements were performed with an ASD FieldSpec 4 Wide-Res spectroradiometer with an available spectral range from

350 nm to 2500 nm. The spectral resolution of this instrument is 3 nm at 700 nm and 30 nm at 1400/2100 nm. The sampling interval is 1.4 nm in the VNIR range from 350 nm to 1050 nm and 2 nm in the SWIR range; spectra are provided with 1 nm increments (2151 channels).

For the data collection both instruments were mounted on a single tripod (Fig. 2). As illumination source we used an ASD Pro-Lamp model, which is also tripod-mountable for indoor laboratory diffuse reflectance measurements. The size of the calibrated reference panel (Zenith Polymer®) was 30 cm \times 30 cm. For imaging and non-imaging measurements, the same white reference panel was used to keep the referencing standardised.

The air-dried soil samples were prepared at three different degrees of fineness (raw, sieved \leq 2 mm and grinded, Fig. 3) in order to vary micro-shadowing and to possibly maximise the spectral significance of the chemical soil components. The distance between sensor and soil sample was set to 35 cm in the nadir position, the illumination zenith angle was 45°. All samples were prepared on a reflection neutral plate (spectrally tested before) and covered, prior to the spectroscopic measurement, by a black passepartout (reflectance under 5% over the entire spectral range from 400 nm to 2500 nm) with a window of 20 cm \times 20 cm. After each measurement, the soil sample was rotated by 90°, so that each sample was archived with 4 spectra. The spectra of the ASD instrument were pre-processed by ViewSpec (ASD software) and exported as mean spectra for the subsequent statistical analysis. The same measurement scheme was followed for

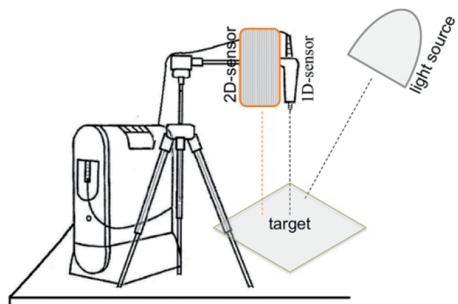


Fig. 2: Experimental set-up for 1D and 2D spectral measurements in the laboratory.

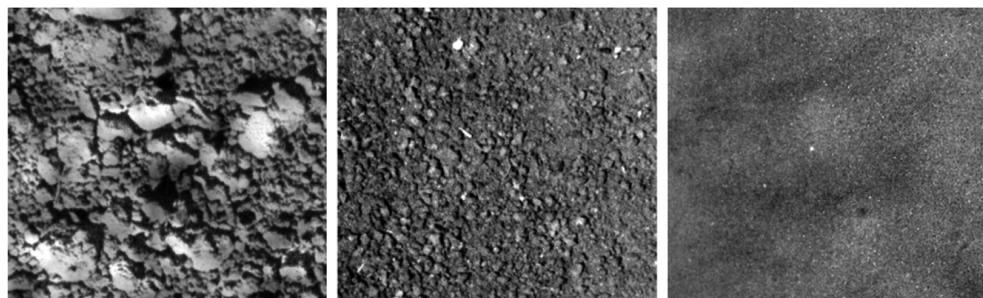


Fig. 3: Examples of the differently prepared soil samples (raw, sieved and grinded) with typical micro-structures and micro-shadowing due to surface roughness.

the 2D reflectance measurements. The native hyperspectral data cubes were converted into .bsq (band sequential) format and processed by the image analysis software ENVI (Exelis Visual Information Solutions).

The spectral resolution of both datasets was adjusted prior to the multivariate calibration procedure. In detail, both sets were reduced to 458 nm – 930 nm and spectrally resampled to the 4 nm resolution of the native image spectra. The camera's first spectral bands below 458 nm showed non-correctable artefacts and were removed. The spectral region over 930 nm suffered from a distinct Si-induced loss of sensitivity. Therefore, the last spectral band was set to 930 nm. It is a general and known phenomenon in CCD imaging technology that the light efficiency of a silicon

detector decreases from around 800 nm up to the response limit of the detector (MAGNAN 2003), which can to some extent be observed for the now reduced image spectra in Fig. 4 too. Based on the described pre-processing, 119 spectral dimensions (458 nm – 930 nm, see Fig. 4) were used for both the 1D reflectance vectors and the 2D image pixels. To enable a direct comparison to the 1D spectra, the 2D reflectance data (hyperspectral data cubes) were converted to virtual 1D measurements by averaging the entire image of each sample. Additionally, spectra were transformed by converting reflectances to absorbances with $\log(\text{reflectance}^{-1})$ and by applying the standard normal variate approach, that is assumed to partly remove the multiplicative interferences of scatter and particle size. Fig. 4 shows also the known effect of soil particle size on reflectance level, i.e., a general decrease with increasing particle size, which has often been described in soil spectroscopic studies (e.g. BOWERS & HANKS 1965).

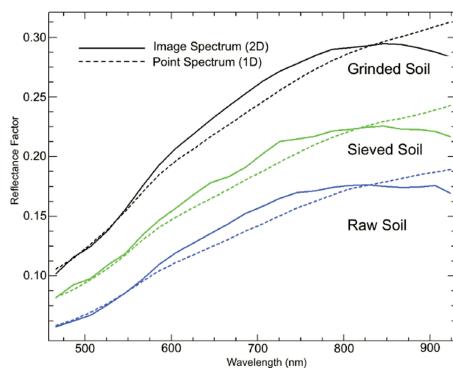


Fig. 4: 1D and 2D reflectance curves after spectral resampling for raw, sieved and grinded soil samples captured by the hyperspectral camera (2D) and the ASD field spectrometer (1D).

2.3 Statistical Methods

PLS has established as multivariate standard tool in the field of chemometrics. PLS is similar to principal component regression (PCR), as both employ statistical rotations to overcome the problems of high-dimensionality and multicollinearity. Different from PCR, PLS maximises the covariance between the spectral matrix (X) and the chemical concentration matrix (Y) to maximise the predictive power of the resulting model (WOLD et al.

2001). To calibrate a PLS model for each constituent, the optimum number of latent variables was identified by performing a leave-one-out cross-validation; the minimum root-mean-squared error (RMSE) in the cross-validation was used as decision criterion (with a predefined maximum of ten latent variables). For an overall description of the PLS method, please refer to ABDI (2003).

Many studies have shown that more accurate calibration models may be achieved by selecting the most informative spectral variables instead of using the full spectrum. For this purpose, we used the CARS approach, which was combined with PLS to CARS-PLS. For a detailed description of the CARS procedure, please refer to LI et al. (2009). Briefly, it uses two successive steps of wavelengths selection in a series of Monte Carlo sampling runs: In a first step, an exponentially decreasing function is used for an enforced removal of wavelengths with relatively small PLS regression coefficients. In a second step, an adaptive reweighted sampling of variables is employed to further eliminate wavelengths in a competitive way. In this step, random numbers are generated to pick variables; the probability of each spectral variable to be kept depends on its weight (calculated from the respective PLS regression coefficient). For our data, 50 successive sampling runs with both steps described above were performed; at the end, the optimal subset of variables with the lowest RMSE in the cross-validated PLS model is kept.

Due to the Monte Carlo strategy and the generation of random numbers in the second step, CARS does not provide a unique solution. Thus, CARS was rerun 50 times to generate 50 estimates for each sample and each constituent; these 50 solutions were averaged to obtain the final estimates.

To assess the accuracy of the multivariate calibrations, we used as measures the residual prediction deviation (RPD, defined as the ratio of standard deviation of the reference values to standard error of the cross-validated estimates), the RMSE, the relative RMSE ($rRMSE = RMSE \times \text{measured arithmetic mean}^{-1}$) and R^2 . Obtained accuracies (cross-validated (cv) values) were evaluated following the guideline of SAEYS et al. (2005): RPD_{cv} and R^2_{cv} values greater than 3.0 or 0.90, respectively, are

considered to indicate excellent predictions, whereas values from 2.5 to 3.0 (RPD_{cv}) and 0.82 to 0.90 (R^2_{cv}) denote a good prediction. Approximate quantitative predictions are indicated by RPD_{cv} values between 2.0 and 2.5 and R^2_{cv} values in the range from 0.66 to 0.81. The possibility to distinguish between high and low values is shown by values between 1.5 and 2.0 (RPD_{cv}) and 0.50 and 0.65 (R^2_{cv}). Unsuccessful predictions have RPD_{cv} or R^2_{cv} values lower than 1.5 or 0.50, respectively.

3 Results and Discussion

3.1 Estimates from Full Range and VNIR Spectra (Spectroradiometer and Image Data)

Estimates from full range spectroradiometer data (cross-validated results) obtained with both approaches, PLS and CARS-PLS, are summarised in Tab. 3. Excellent results were obtained with CARS-PLS for OC (raw samples) and N (raw and grinded samples); for HWE-C, results were slightly worse (good predictions using raw or grinded samples). As HWE-C represents a comparatively small carbon pool (as a measure of labile C), indirect correlations to the spectral data triggered by OC may be of relevance. High correlations between OC and HWE-C (Persons's $r = 0.93$, $p < 0.01$) also support the assumption of such an indirect correlation (see also VOHLAND et al. 2011). CARS-PLS outperformed PLS (without variable selection) in all cases and resulted in at least "approximate quantitative predictions". Due to the selection procedure, CARS-PLS models were – as a general rule – more parsimonious than PLS models. With the exception of sieved samples (OC and N), the number of latent variables reduced slightly; the number of spectral variables that were integrated in CARS-PLS models ranged between 13.9 (HWE-C, sieved samples – averaged from 50 runs) and 23.5 (N, sieved samples) and was thus distinctly lower than the original number of $n = 500$ spectral variables. With respect to the preparation level, worst results were obtained with the use of sieved samples.

Tab.3: Cross-validated results from field spectrometer data (1D) and image mean spectra (2D) for raw, sieved and grinded soils ($n = 40$, l.v.: number of latent variables (for CARS-PLS averaged from 50 runs); cv: leave-one-out cross-validation).

Spectral Range	Soil Preparation Level	Model	OC			N			HWE-C					
			l.v.	R ² _{cv}	RPD _{cv}	rRMSE _{cv}	l.v.	R ² _{cv}	RPD _{cv}	rRMSE _{cv}	l.v.	R ² _{cv}	RPD _{cv}	rRMSE _{cv}
ID 402 nm - 2398 nm 500 variables	raw	PLS	9	0.73	1.94	0.25	9	0.72	1.89	0.26	8	0.61	1.58	0.26
		CARS-PLS	8.2	0.93	3.85	0.13	8.0	0.91	3.30	0.15	7.2	0.86	2.73	0.15
	sieved	PLS	6	0.65	1.66	0.29	6	0.64	1.66	0.29	7	0.51	1.37	0.30
		CARS-PLS	6.0	0.84	2.49	0.19	6.0	0.84	2.49	0.19	6.0	0.72	1.91	0.21
	grinded	PLS	9	0.73	1.90	0.25	9	0.75	2.02	0.24	10	0.63	1.58	0.26
		CARS-PLS	6.7	0.88	2.93	0.16	7.3	0.92	3.53	0.14	8.2	0.88	2.86	0.14
ID 458 nm - 930 nm 119 variables	raw	PLS	7	0.41	1.28	0.37	10	0.34	1.19	0.41	7	0.43	1.28	0.32
		CARS-PLS	6.2	0.77	2.08	0.23	8.0	0.78	2.09	0.23	5.6	0.70	1.84	0.22
	sieved	PLS	9	0.63	1.64	0.29	8	0.45	1.34	0.36	9	0.64	1.64	0.25
		CARS-PLS	8.4	0.82	2.37	0.20	8.0	0.77	2.09	0.23	7.8	0.77	2.14	0.19
	grinded	PLS	10	0.66	1.73	0.28	10	0.49	1.38	0.35	9	0.56	1.49	0.27
		CARS-PLS	9.2	0.85	2.61	0.18	8.5	0.79	2.17	0.22	7.5	0.71	1.88	0.22
2D 458 nm - 930 nm 119 variables	raw	PLS	3	0.38	1.27	0.38	4	0.30	1.19	0.40	6	0.39	1.27	0.32
		CARS-PLS	3.5	0.49	1.42	0.34	3.0	0.42	1.33	0.36	4.1	0.58	1.56	0.26
	sieved	PLS	7	0.59	1.57	0.31	7	0.54	1.47	0.33	7	0.54	1.45	0.28
		CARS-PLS	6.5	0.72	1.89	0.25	6.7	0.70	1.83	0.26	5.6	0.62	1.63	0.25
	grinded	PLS	10	0.65	1.70	0.28	8	0.60	1.58	0.30	10	0.65	1.67	0.24
		CARS-PLS	8.4	0.84	2.49	0.19	7.1	0.75	2.01	0.24	8.9	0.83	2.45	0.17

excellent

good

Fig. 5 illustrates the selection frequencies of spectral variables for assessing OC, which were obtained in 50 runs of CARS-PLS with the full spectra of raw and sieved samples. Selection frequencies of grinded samples (not illustrated) showed peaks partly similar to raw and partly similar to sieved samples. Some regions in the visible were rather frequently selected for all preparation levels (e.g. 406 nm – 418 nm, 662 nm – 674 nm and – for sieved and grinded samples – 526 nm – 538 nm), whereas selection frequencies in the very near IR (a region covered by the UHD camera) were very low (Fig. 5). In the NIR range up to 1300 nm, selection peaks were found only at about 1100 nm (raw and grinded samples)

and at 1140 nm (sieved samples). Compared to the visible and the NIR domain, the shortwave IR (SWIR) region was distinctly more relevant for the CARS approach with markedly higher peaks of selection frequencies at 1374 nm – 1394 nm (raw and grinded samples), 1850 nm – 1902 nm (raw, grinded), 2078 nm – 2194 nm (all preparation levels) and some wavelengths beyond 2300 nm (e.g. 2306, 2334/2338 nm) (Fig. 5). In the wavelength range between 1406 nm and 1798 nm, selection frequencies were low in all cases.

The found selection peaks indicate some wavelength regions (VIS range, prominent water bands, hydroxyl band, C-H absorption bands) that were already identified in other

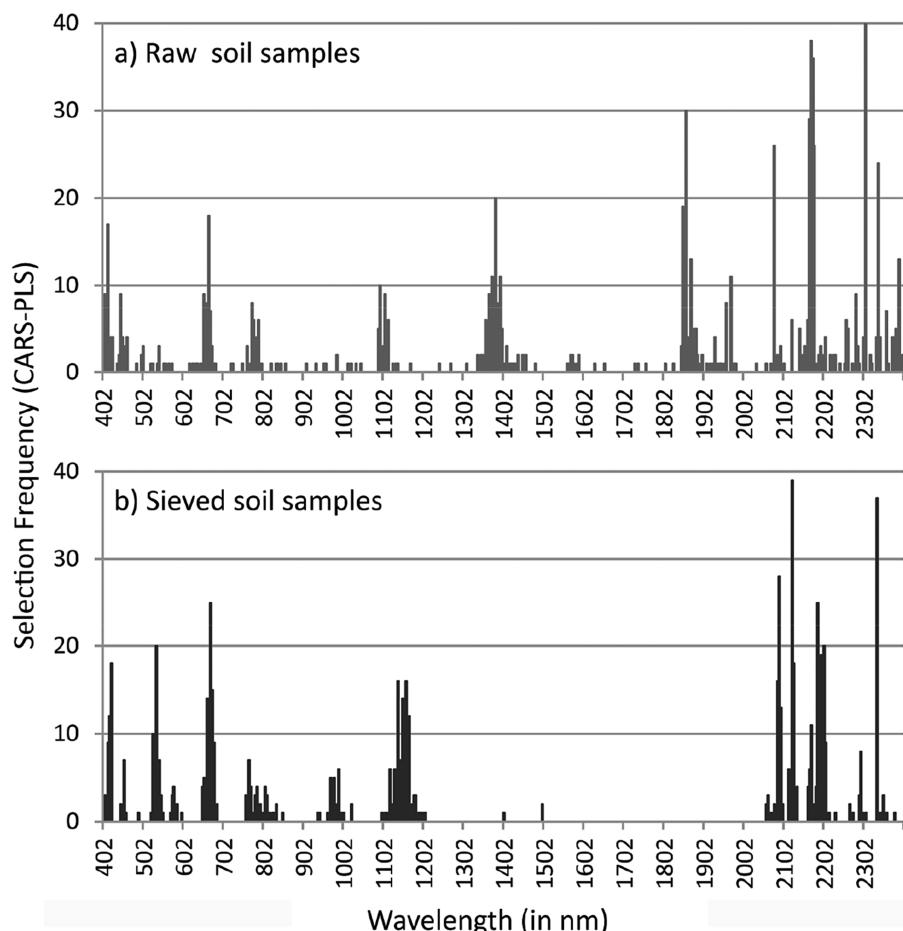


Fig. 5: Selection frequencies (raw (a) and sieved (b) samples) for spectral variables in the full range from 402 nm to 2398 nm (realised in 50 runs of CARS-PLS with OC as target soil constituent).

studies to be relevant for assessing OC with spectroscopy (e.g. MOUAZEN et al. 2007, VIS-CARRA ROSEL & BEHRENS 2010, VOHLAND et al. 2014). BELLON-MAUREL & McBRATNEY (2011) quote the 1600 nm – 2500 nm range to be the most relevant for measuring OC. CÉCILLON et al. (2009) present a compilation of important NIR wavelengths for OC and also N that are all beyond 1100 nm.

As most of these essential wavelength regions were not included in the pruned ASD spectra and the UHD data, a drop of accuracies was a priori to be expected for the multivariate calibrations with these datasets. The obtained results were, de facto, inferior to those with full range spectra (Tab. 3), but especially for OC these differences were small when using grinded samples and the CARS-PLS method; for both datasets (pruned ASD and averaged UHD Spectra) the cross-validation resulted in “good” estimates. In addition, very similar results were obtained for OC from sieved samples using full and pruned ASD spectra and also for HWE-C from grinded samples with full range ASD and UHD data (Tab. 3).

Differences of accuracy between the three instrumental settings were most pronounced when raw samples were measured and used for the calibration with CARS-PLS; in this case, the general order for all soil variables was ASD (full spectra) >> ASD (pruned spectra) >> UHD data. The CARS variable selection procedure obviously highlighted these differences, as accuracies obtained with PLS

from raw samples were on a similar level at least for pruned ASD and UHD data (Tab. 3). As the UHD spectra were obtained by averaging all pixel values over the complete image, it is not entirely clear why these data performed generally worse than the ASD readings. The differences were probably due to the irregular surface roughness of the raw samples which caused multidirectional light scattering effects and strong contrasts between illuminated and shadowed image regions (Fig. 3). Soil surface roughness is known as one main disturbing factor for e.g. SOC estimates from proximally sensed VNIR data, which may be compensated, in case of larger samples sets, by stratified models specified for different surface roughness classes (RODIONOV et al. 2014). Thus, for future in-field tests of the UHD hyperspectral frame camera its potential to assess soil roughness from the images (e.g. by analysing the extent of shadowed image portions; GARCÍA MORENO et al. 2008) should be fully exploited.

3.2 Spatial Analysis: How Much Variability is in One Image?

We selected three samples with different levels of mean OC contents to analyse the variability in the hyperspectral images (Tab. 4). Within each image, 500 randomly determined pixels were extracted and then used to obtain pixel-wise estimates for OC. For this estima-

Tab. 4: Statistics for OC estimates from image data obtained with CARS-PLS (each sample and preparation level: 500 randomly selected pixels).

	Raw			Sieved			Grinded		
	Mean ²	Std	Min Max	Mean ²	Std	Min Max	Mean ²	Std	Min Max
Sample #1 ¹	1.62	0.77	-1.21 3.77	1.72	0.61	-0.06 3.56	1.95	1.88	-4.05 6.18
Sample #2 ¹	2.28	0.66	0.51 6.40	2.21	0.85	-0.24 4.69	2.53	1.91	-4.04 8.20
Sample #3 ¹	1.10	0.41	0.17 2.37	0.98	0.82	-1.44 3.51	1.26	2.09	-5.58 6.93

¹ Wet-chemical reference values (OC): #1: 1.90, #2: 2.70, #3: 0.62

² Previous estimates from averaged images (see 3.1):

Sample #1: raw: 1.83, sieved: 1.81, grinded: 1.90; Sample #2: raw: 2.34, sieved: 2.20, grinded: 2.33; Sample #3: raw: 1.28, sieved: 1.28, grinded: 1.09

tion, the already available models calibrated before (for $n = 40$ reference values) were applied.

The results we obtained were similar for all three images. In all cases, the estimates followed the normal distribution. As an example, the Q-Q (quantile-quantile) plot is illustrated for sample #1 (grinded) in Fig. 6. Mean values calculated from the 500 image pixels provided useful estimates close to the measured reference values (Tab. 4). The range of estimates indicated, on the one hand, highly variable OC contents; on the other hand, however, we often received inconsistently low and high values, i.e. negative values and obvious overestimates (OC contents of more than 6%, for example; see boxplots in Fig. 6 and Tab. 4). Evidently, the calibration set was not sufficient to represent all situations (including illumination differences) contained in the image data; this may be due to the small number of calibration samples and the way the calibration set was compiled, as only averaged and thus “smoothed” image spectra were used.

4 Summary and Conclusions

We tested two multivariate calibration methods; in all cases, PLS combined with CARS outperformed full spectrum-PLS (without

variable selection). CARS was found to be an effective selection method that improved accuracies at least in the cross-validation; however, it should be tested with an independent validation set. The majority of CARS-selected wavelengths were physically meaningful, as they were related – in the case of organic carbon – to water absorption bands, the hydroxyl band at 2200 nm or C-H absorption bands in the region beyond 2300 nm.

In soil spectroscopy, the SWIR domain is of high relevance. Although the tested hyperspectral camera (UHD 285) does not cover this region, obtained values indicated, at least in part, good estimates. These results were restricted to crushed (grinded) samples; accuracies dropped distinctly for raw samples (which represent the normal in-field situation with rough soil surfaces). Thus, for in-field studies, the full potential of the image data should be used, which is to estimate soil roughness directly from the images to define stratified models or to eliminate shadowed pixels which have very restricted information content.

The image data that we analysed in detail showed a great variability of the contained (spectral) information. These fine-scale variations are relevant for chemometric approaches, as they require a careful definition of calibration sets that have to cover these variations adequately.

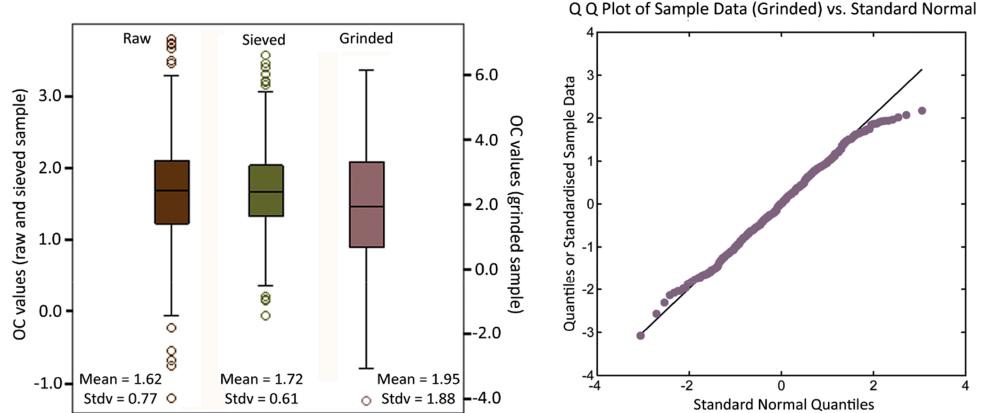


Fig. 6: Statistics for OC estimates (in %) obtained from 500 randomly selected pixels for one sample (#1) and its different preparation levels (for clarification, boxplot for grinded sample is scaled differently).

Acknowledgements

Our special thanks go to Cubert GmbH that made it possible to use and test the hyperspectral frame camera, by name to RAINER GRASER and Dr. RENÉ MICHELS. We would like to thank the Soil Science Department at University of Trier, headed by SÖREN THIELE-BRÜHN, for carrying out chemical analyses. CARS calculations were carried out using the freely available CARS package (Copyright HONGDONG Li, 2011; <https://code.google.com/p/carspls/>). The project was supported by a grant of the German Research Foundation (DFG, VO 1509/3-1).

References

- ABDI, V., 2003: Partial least squares (PLS) regression. – LEWIS-BECK, M., BRYMAN, A. & FUTING, T. (eds.): Encyclopedia for research methods for the social sciences: 772–795, Sage Publications, Thousand Oaks, CA, USA.
- BELLON-MAUREL, V. & McBRATNEY, A., 2011: Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils – Critical review and research perspectives. – *Soil Biology and Biochemistry* **43** (7): 1398–1410.
- BEN-DOR, E. & BANIN, A., 1995: Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. – *Soil Science Society of America Journal* **59** (2): 364–372.
- BOWERS, S.A. & HANKS, R.J., 1965: Reflection of radiant energy from soils. – *Soil Science* **100** (2): 130–138.
- CÉCILLON, L., BARTHÈS, B.G., GOMEZ, C., ERTLEN, D., GENOT, V., HEDDE, M., STEVENS, A. & BRUN, J.J., 2009: Assessment and monitoring of soil quality using near-infrared reflectance spectroscopy (NIRS). – *European Journal of Soil Science* **60** (5): 770–784.
- E DIN ISO 11277:1994-06, 1994: Bodenbeschaffenheit – Bestimmung der Partikelgrößenverteilung in Mineralböden – Verfahren durch Sieben und Sedimentation nach Entfernen der löslichen Salze, der organischen Substanz und der Carbonate. – Beuth-Verlag, Berlin.
- FAO (Food and Agriculture Organization of the United Nations), 2006: Guidelines for soil description. – 4th ed., 97 pp., Rome, Italy.
- GARCÍA MORENO, R., SAA REQUEJO, A., TARQUIS ALONSO, A.M., BARRINGTON, S. & DÍAZ, M.C., 2008: A shadow analysis method to measure soil surface roughness. – *Geoderma* **146** (1–2): 201–208.
- HILL, J., UDELHOVEN, T., VOHLAND, M. & STEVENS, A., 2010: The Use of Laboratory Spectroscopy and Optical Remote Sensing for Estimating Soil properties. – OERKE, E.C., GERHARDS, R., MENZ, G. & SIKORA, R.A. (eds.): Precision crop protection – the challenge and use of heterogeneity: 67–86. – 441 pp., Springer Science, Dordrecht, The Netherlands.
- JUNG, A., VOHLAND, M., HEINRICH, J., MICHELS, R. & GRASER, R., 2013: Soil analysis with hyperspectral data – an experiment with a hyperspectral frame camera and VIS-NIR spectrometers. – 8th EARSeL SIG Imaging Spectroscopy Workshop, Nantes, France, in press.
- KOOISTRA, L., WANDERS, J., EPEMA, G.F., LEUVEN, R.S.E.W., WEHRENS, R. & BUYDENS, L.M.C., 2003: The potential of field spectroscopy for the assessment of sediment properties in river floodplains. – *Analytica Chimica Acta* **484** (2): 189–200.
- KÖRSCHENS, M., WEIGEL, A. & SCHULZ, E. 1998: Turnover of Soil Organic Matter (SOM) and Long-Term Balances – Tools for Evaluating Sustainable Productivity of Soils. – *Zeitschrift für Pflanzenernährung und Bodenkunde* **161** (4): 409–424.
- KURZ, T.H., BUCKLEY, S.J. & HOWELL, J.A., 2013: Close-range hyperspectral imaging for geological field studies: workflow and methods. – *International Journal of Remote Sensing* **34** (5): 1798–1822.
- LI, H., LIANG, Y., XU, Q. & CAO, D., 2009: Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. – *Analytica Chimica Acta* **648** (1): 77–84.
- MAGNAN, P., 2003: Detection of visible photons in CCD and CMOS: A comparative view. – *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **504** (1): 199–212.
- 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 and Tillage Research* **93** (1): 13–27.
- RODIONOV, A., PATZOLD, S., WELP, G., PALLARES, R.C., DAMEROW, L. & AMELUNG, W., 2014: Sensing of soil organic carbon using VIS-NIR spectroscopy at variable moisture and surface roughness. – *Soil Science Society of America Journal* **78** (3): 949–957.
- SAEYS, W., MOUAZEN, A.M. & RAMON, H., 2005: Potential for onsite and online analysis of pig ma-

- nure using visible and near infrared spectroscopy. – *Biosystems Engineering* **91** (4): 393–402.
- STEFFENS, M. & BUDDENBAUM, H., 2013: Laboratory imaging spectroscopy of a stagnic Luvisol profile – High resolution soil characterisation, classification and mapping of elemental concentrations. – *Geoderma* **195**: 122–132.
- 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** (1–2): 395–404.
- STEVENS, A., UDELHOVEN, T., DENIS, A., TYCHON, B., LIOY, R., HOFFMANN, L. & VAN WESEMAEL, B., 2010: Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. – *Geoderma* **158** (1–2): 32–45.
- SUDDUTH, K.A., HUMMEL, J.W. & FUNK, R.C., 1989: NIR soil organic matter sensor. – *ASAE-Paper 89-1035*.
- 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** (2): 319–329.
- VIGNEAU, N., ECARNO, M., RABATEL, G. & ROUMET, P., 2011: Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in wheat. – *Field Crops Research* **122** (1): 25–31.
- VISCARRA ROSSEL, R.A. & BEHRENS, T., 2010: Using data mining to model and interpret soil diffuse reflectance spectra. – *Geoderma* **158** (1–2): 46–54.
- VISCARRA ROSSEL, R.A. & McBRATNEY, A.B., 1998: Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content. – *Geoderma* **85** (1–2): 19–39.
- VOHLAND, M., HARBICH, M., SCHMIDT, O., JARMER, T., EMMERLING, C. & THIELE-BRÜHN, S., 2011: Use of imaging spectroscopy to assess different organic carbon fractions of agricultural soils. – NEALE, C.M.U. & MALTESE, A. (eds.): *Remote Sensing for Agriculture, Ecosystems, and Hydrology XIII*. – Proceedings of SPIE 8174, Bellingham, WA, USA.
- VOHLAND, M., LUDWIG, M., THIELE-BRÜHN, S. & LUDWIG, B., 2014: Determination of soil properties with visible to near- and mid-infrared spectroscopy: effects of spectral variable selection. – *Geoderma* **223–225**: 88–96.
- WOLD, S., SJÖSTRÖM, M. & ERIKSSON, L., 2001: PLS-regression: a basic tool of chemometrics. – *Chemometrics and Intelligent Laboratory Systems* **58** (2): 109–130.
- XIAOBO, Z., JIEWEN, Z., POVEY, M.J., HOLMES, M. & HANPIN, M., 2010: Variables selection methods in near-infrared spectroscopy. – *Analytica Chimica Acta* **667** (1–2): 14–32.
- YE, X., SAKAI, K., OKAMOTO, H. & GARCIANO, L.O., 2008: A ground-based hyperspectral imaging system for characterizing vegetation spectral features. – *Computers and Electronics in Agriculture* **63** (1): 13–21.

Address of the Authors:

Dr. ANDRÁS JUNG & Prof. Dr. MICHAEL VOHLAND, Institute for Geography, University of Leipzig, Geoinformatics and Remote Sensing, Johannisallee 19a, D-04103 Leipzig, Tel.: +49-341-97-32785, Fax: +49-341-97-32799, e-mail: {andras.jung}, {michael.vohland}@uni-leipzig.de

Manuskript eingereicht: Juni 2014

Angenommen: September 2014