

Applications

Simon Schreiner*, Dubravko Culibrk, Michele Bandecchi, Wolfgang Gross and Wolfgang Middelmann

Soil monitoring for precision farming using hyperspectral remote sensing and soil sensors

Bodenüberwachung für Precision Farming durch hyperspektrale Fernerkundung und Bodensensoren

<https://doi.org/10.1515/auto-2020-0042>

Received March 30, 2020; accepted February 7, 2021

Abstract: This work describes an approach to calculate pedological parameter maps using hyperspectral remote sensing and soil sensors. These maps serve as information basis for automated and precise agricultural treatments by tractors and field robots. Soil samples are recorded by a handheld hyperspectral sensor and analyzed in the laboratory for pedological parameters. The transfer of the correlation between these two data sets to aerial hyperspectral images leads to 2D-parameter maps of the soil surface. Additionally, rod-like soil sensors provide local 3D-information of pedological parameters under the soil surface. The goal is to combine the area-covering 2D-parameter maps with the local 3D-information to extrapolate large-scale 3D-parameter maps using AI approaches.

Keywords: Hyperspectral remote sensing, soil monitoring, AI, 2D/3D pedological maps

Zusammenfassung: Diese Arbeit beschreibt einen Ansatz zur Erstellung bodenkundlicher Parameterkarten mittels hyperspektraler Fernerkundung und Bodensensorik, als Informationsgrundlage für automatisierte und präzise landwirtschaftliche Anwendungen durch Traktoren und

***Corresponding author: Simon Schreiner**, Fraunhofer IOSB, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation, Gutleuthausstrasse 1, 76275 Ettlingen, Germany, e-mail: simon.schreiner@iosb.fraunhofer.de

Dubravko Culibrk, University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, 21102 Novi Sad, Serbia, e-mail: dculibrk@uns.ac.rs

Michele Bandecchi, SmartCloudFarming GmbH, Rosenthaler Strasse 72a, 10119 Berlin, Germany, e-mail: bandecchi@smartcloudfarming.com

Wolfgang Gross, Wolfgang Middelmann, Fraunhofer IOSB, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation, Gutleuthausstrasse 1, 76275 Ettlingen, Germany, e-mails: wolfgang.gross@iosb.fraunhofer.de, wolfgang.middelmann@iosb.fraunhofer.de

Feldroboter. Dazu werden Bodenproben hyperspektral untersucht und pedologische Parameter im Labor analysiert. Die Übertragung der Korrelation zwischen diesen beiden Datensätzen auf hyperspektrale Luftbilder erzeugt 2D-Parameterkarten der Bodenoberfläche. Zusätzlich werden stabähnliche Bodensensoren im Feld versenkt, die lokal 3D-Information über pedologische Parameter liefern. Ziel ist die Verknüpfung der flächendeckenden 2D-Parameterkarten mit lokaler 3D-Information durch KI, um flächendeckende 3D-Parameterkarten zu erstellen.

Schlagwörter: Hyperspektrale Fernerkundung, Bodenüberwachung, KI, 2D/3D pedologische Karten

1 Introduction

The vast industrialization in agriculture during the last decades led to a significant increase in productivity. At the same time, this progress also associates with negative effects on the biosphere like faunal and botanical species decline, soil degradation, water consumption, and other environmental damages due to the use of harmful chemicals and their spread in the environment. Organic farming tries to counteract these effects accepting losses in productivity [1]. Precision farming as an alternative to conventional and organic farming, exploits the increasing digitalization in agriculture and aims to balance economic and ecological benefits. Autonomous and intelligent tractors, field robots, and trailers try to increase the efficiency in terms of time, the input of resources, and cost by enabling individual treatments of field segments [2]. For this kind of agriculture, the most essential basis is information about the condition of soils and plants. Without the monitoring of soils and plants with a high spatio-temporal resolution, precise and automated cultivation like weed control, fertilization, irrigation, and harvest is not possible. Since the costs for traditional methods to gather soil properties (via

grid sampling) increase exponentially as the spacing of grids decreases, an alternative way to collect this information is bio-monitoring by remote sensing [3]. The use of visible to near-infrared (VNIR) and short-wave infrared (SWIR) spectroscopy in soil science has a long history and experienced a boom in the last decades [4]. After initial studies on the relationship between soil properties and soil spectra in a more qualitative way [5], other scientists successfully introduced more quantitative approaches using tools of multivariate statistics [6]. Since then, numerous studies have been based on these quantitative approaches to mapping soil properties, such as mineral composites [7], soil salinity [8], texture [9], moisture [10] or chemical soil properties like carbon, nitrogen, pH, etc. [11].

This work describes an approach, which derives soil properties by producing pedological parameter maps of soil surfaces from hyperspectral (HS) remote sensing data. Correlations between the parameters nitrogen (N), phosphorus (P), calcium (Ca), magnesium (Mg), potassium (K), total organic carbon (TOC), humus, and spectral data are examined and applied to spectral airborne and spaceborne images. This endeavor is an ongoing process and will be further refined in the future. Bio-monitoring by remote sensing only enables the analysis of materials or soils on the surface due to the recorded electro-magnetic spectrum. However, in agriculture, various decisions require sub-surface 3D-information of the soil. After a rain, nitrate concentrations on the soil surface can be low and induce the farmer to fertilize nitrate. If the nitrate concentrations below the surface were actually high, the excess nitrate seeps into the groundwater [12].

Therefore, information about the variability of soil nutrients below the surface within a specific field is of fundamental importance for precision farming. Rod-like soil sensors provide local and frequently sampled below the surface data as deep as 60 cm. The presented approach studies the dependency between electrical conductivity (EC) of the soil, acquired by sensors in the soil, and soil parameters at different soil depths, determined in the laboratory. In addition, the proposed concept aims to combine hyperspectral remotely sensed 2D-soil-parameter maps pertinent to the field of interest with the local sub-surface data to create large-scale 3D-parameter maps using state-of-the-art machine learning approaches [13, 14, 15, 16, 17].

2 Proposed approach

Figure 1 visualizes the following description of our proposed approach as a flowchart. The structure of this pa-

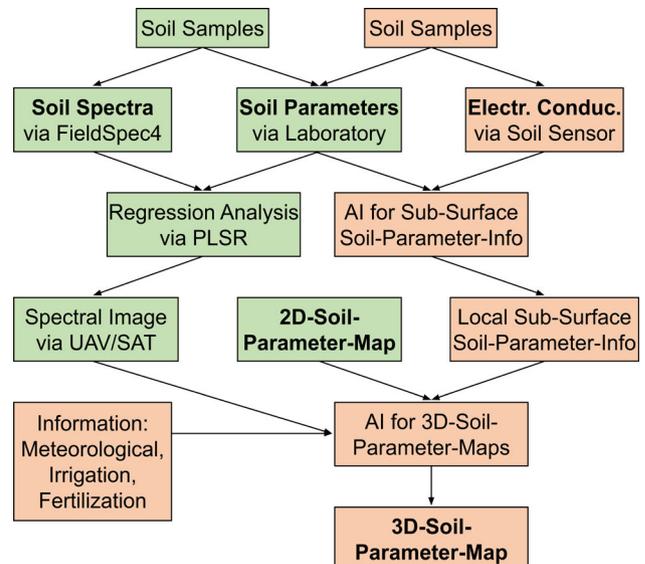


Figure 1: Flowchart of the proposed approach. Steps highlighted in green have been tested successfully, while parts in orange show current and future experiments. Main parts of the approach are marked in bold.

per and its sections are based on this workflow. Steps highlighted in green have already been tested successfully. Orange highlighted steps are part of current and future experiments. Main parts of the approach are marked in bold.

The corpus of data used in our study pertains to soil samples (see Section 3.1), which are (hyper-)spectroscopically examined and wet-chemically analyzed in a laboratory. Usually, soil samples are sieved, ground, and dried for the spectroscopic analysis to minimize noisy signals. Our approach, however, relies on a spectroscopic survey, which was performed directly in the field (in-situ) to represent a more realistic and implementable scenario. A FieldSpec 4 handheld spectroradiometer from Analytical Spectral Devices Inc. was used to record the soil spectra (see Section 3.2) in the field. Soil samples of the same area were collected and sent to a laboratory for pedological analysis. Subsequently, for each parameter and each soil sample, the FieldSpec 4 soil spectrum was processed (see Section 4.1) and labeled with the corresponding parameter value. As a multivariate statistic method, Partial Least-Square Regression (PLSR) was then used to create models able to estimate the parameter values based solely on hyperspectral data (see Section 4.2). In case of a significant correlation (see Section 5.1) between the model estimates and the data collected by sampling, these models of relationship between pedological parameters and spectral information can finally be used to estimate the

value of the parameters for each pixel of an aerial or spaceborne hyperspectral image (see Section 4.3). The results are 2D-parameter maps consisting of the parameter values for each pixel (see Section 5.2).

As part of our efforts to extend this information below the soil surface, we are planning a new field study. In order to obtain pedological information at six soil depths, SEN-TEK Drill & Drop TRISCAN soil sensors will be installed and EC signal recorded (see Section 3.3). At the same time, soil samples are taken at different depths alongside the soil sensor (see Section 3.1) and analyzed in a soil laboratory. These data will first be used to train machine learning models, such as simple neural nets and PLSR models, to derive local soil parameter values corresponding to the sub-surface readings of each soil sensor (see Section 4.4). Since the concentration of nutrients at different depths is the result of complex propagation processes in the soil and their extraction by plants over time, we need more complex approaches to create a model of these dynamics. Therefore, we will use a state-of-the-art machine learning approach for time series modeling (a class of deep recursive neural networks called Long Short-Term Memory – LSTM, see Section 4.4) to estimate the value of each sensor reading below the surface, based solely on data that can be obtained for the surface. The underlying data of the surface include hyperspectral images, our own surface 2D-soil-parameter maps, as well as meteorological and fertilization data. Using the simpler models created based on our soil sampling campaign, we intend then to convert these estimates of sensors reading to parameter values in real-time, in order to create 3D-soil-parameter maps stretching from the soil surface to the depth of our deepest sensor.

3 Data basis

3.1 Pedological data

A sufficient amount of soil samples is required to build reliable spectroscopic-chemometric models. Robust correlations have been found with 50 to 150 samples. As the laboratory analysis of soil samples is expensive, the number of soil samples is often limited. Due to these costs, there is always a trade-off between the desire to create a universally valid model and the number of analyzed pedological parameters. Based on the essential plant nutrients, nitrate (N), phosphorus (P), calcium (Ca), magnesium (Mg), and potassium (K) were chosen for this study

Table 1: Statistics of the soil samples by mean, minimum and maximum value of seven soil parameters. Data basis are 55 soil samples of a field in Brandenburg, Germany.

	Mean	SD	Min	Max
N [%]	0.18	0.03	0.11	0.26
P [mg/100 g]	1.92	0.68	1.0	4.5
Ca [cmol+/kg]	4.97	1.08	2.4	9.0
Mg [mg/100 g]	5.59	1.52	3.1	9.8
K [cmol+/kg]	0.34	0.13	0.2	0.7
TOC [%]	0.83	0.23	0.32	1.33
Humus [%]	1.45	0.4	0.6	2.3

and analyzed in the laboratory. Other pedological parameters like total organic carbon (TOC) and humus were also determined. Finally, laboratory results were checked for outliers. Table 1 shows the statistics by mean, standard deviation, minimum, and maximum value of the seven pedological parameters in order to obtain an overview of typical value ranges. While the unit of TOC, humus, and nitrogen is a percentage, phosphorus and magnesium are described in mg/100 g. Calcium and potassium in cmol+/kg provide plant available (cation exchangeable) nutrient amounts.

For a reliable soil-sensor/-parameter model, another experiment will be performed in June 2021. Soil samples will be taken by a hand sampling toolset from Arts Machine Shop Inc. at three different depths alongside the installed soil sensor. Following the parameter selection of the spectroscopic-chemometric model approach, samples are laboratory analyzed for total nitrogen, organic carbon, phosphorus, potassium, and calcium. Additionally, available nitrogen, pH, cation exchange capacity (CEC), and electrolytic conductivity (EC1:5) are determined. In order to investigate the relation between volumetric ion concentration of the soil sensor and different level of N, P and K in different level of soil depth, one soil sample is going to be collected for each of the three depths (0–10 cm, 10–20 cm, and 20–40 cm) and each of nine soil sensors installed in the experiment field. The current experimental design enables five fertilization iterations with subsequently soil sampling that will result in 135 soil samples.

3.2 Spectral data

The principle functionality of (hyperspectral) remote sensing is as follows: Sunlight in the form of electro-magnetic radiation enters the atmosphere and reacts with the earth's surface, where each material transmits, absorbs

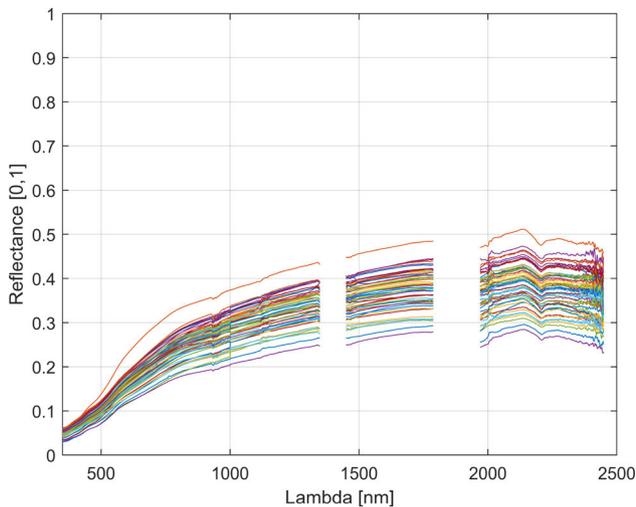


Figure 2: Soil spectra of 55 soil samples surveyed by an ASD FieldSpec 4. The wavelengths (350–2500 nm) of each band are plotted on the x-axis and the reflectivity [0,1] of the soil sample at this wavelength on the y-axis. Due to interaction with the atmosphere, noisy signals between 1349–1449 nm, 1789–1969 nm, and 2449–2500 nm were deleted.

and reflects a certain amount of radiation depending on its specific properties. While the human eye and conventional digital cameras only record the colors red, green, and blue (RGB), hyperspectral sensors can capture the whole wavelength range of the reflected light from the visible to the near and shortwave infrared [11]. After passing the converging lens, the light is collimated and split by a diffraction grating. Depending on the wavelength, the light beam is deflected in a specific angle and incidents upon a photosensitive chip. Thus, hyperspectral sensors divide and capture the reflected light in hundreds of narrow wavelengths (bands or channels) [18]. The resulting measurements are visualized by plotting the spectral signature as reflectance over the wavelengths. Band numbers or the corresponding wavelengths are plotted on the x-axis and the reflectivity of the surveyed object at this wavelength on the y-axis (see Figure 2).

Handheld hyperspectral sensors like the FieldSpec 4 of Analytical Spectral Devices Inc. have the highest spectral resolution of over 2150 bands, which generally cover the wavelength range from 350–2500 nm. Those sensors are well suited to collect spectral ground truth data of one single point using lenses of different aperture. With a sensor distance of 20 cm above the ground, each area of the soil sampling (5 by 5 cm) was captured.

In contrast to single-point hyperspectral sensors, the most common hyperspectral sensors are imaging sensors based on the push-broom principle. They record a single

spatial line with full spectral resolution. An image can then be created by moving the sensor relative to the scene. For remote sensing, the sensor is installed on a carrier platform such that it records data perpendicular to the flight direction and records subsequent frames of adjacent areas on the ground due to the movement of the platform. Choosing the right number of frames per second according to the speed of the carrier platform is essential to acquire 2D-images without gaps.

Possible carrier platforms are UAVs, planes, and satellites. For agricultural and precision farming applications, a high spatial resolution and mobile sensor system are required. The Headwall VNIR-SWIR Co-Aligned hyperspectral sensor covers the wavelength range from 400–2500 nm with 537 bands and only weights 2.83 kg, which makes this sensor well suited for drones. Mounted on a DJI Matrice 600 Pro and flying at an altitude above ground level (AGL) of 80 m, one battery-set lasts approximately 12–18 minutes. During this time and in autonomous flight mode, an area in the range of 4.2 ha is covered. An effectively continuous flight is possible with several battery-sets.

The development of several hyperspectral satellites by different countries started years ago (i. e., EnMAP-Mission). It is planned to provide hyperspectral satellite data covering large areas. Satellites and their periodic coverage guarantee a reliable data source. The disadvantage compared to sensors that are mounted on aerial platforms is the decrease in spatial resolution (for EnMAP: 30 m ground sampling distance (GSD)).

Other issues with space-borne passive sensors are the difficult atmospheric correction process and the occlusion by clouds, especially in German winters, when fields lie dormant and are well suitable for remote sensing soil analysis. Nevertheless, hyperspectral satellite images will be a valuable data source, especially for extensive farms and fields. Table 2 illustrates an overview of a selection of hyperspectral sensor system specifications.

3.3 Soil sensors

In June 2021, nine 60 cm long SENTEK Drill & Drop TRISCAN probes will be installed in an experimental field. These probes consist of six capacitance sensors (one every 10 cm, starting at 10 cm depth) for measuring the soil moisture, the electrical conductivity (salinity), and the (probe internal) temperature. More specifically, the capacitance sensors monitor changes in the dielectric properties of soils and return two outputs: the first is a signal that is converted via a normalized equation and then via default cali-

Table 2: Specifications of hyperspectral sensor systems.

	ASD FieldSpec 4 (handheld sensor)	Headwall VNIR-SWIR Co-Aligned (UAV)	EnMAP (satellite)
Wavelength range [nm]	350–2500	400–2500	420–2450
No. of bands	2151	537	242
Spatial Pixel	1	640	1024
GSD [m] (m AGL)	0.05 (0.2)	0.2 (80)	30

bration into volumetric water content (VWC). VWC is given in millimeters of water per 100 mm of soil depth. The normalized equation and the default calibration is adaptable by the user. The second output is a signal that is elaborated together with the first signal in a manufacturer data model and provides the volumetric ion content (VIC). This VIC value will be correlated with the soil EC and other specific pedological parameters of the soil sample laboratory analysis. The probe measurement of VIC ranges between 0 and 17 deciSiemens/m in sand, loamy sand, and sandy loam textures. The resolution of the VWC electronic sensors is as low as 1 microSiemens/cm (0.001 mS/cm) in dry soil condition and as high as 14 microSiemens/cm (0.014 mS/cm) in water-saturated soil conditions.

The resolution of this probe for moisture (VWC) is 1:10000, for salinity (VIC) 1:6000 and for temperature 0.3 °C. Moisture precision is ± 0.03 %vol., while the temperature accuracy is ± 2 °C at 25 °C. The operating temperature ranges between -20 °C and 60 °C.

The VWC and VIC data are downloaded and analyzed with the IRRIMAX Live software to access moisture, salinity, and temperature data and combine multiple probes to observe two-dimensional data dynamics. This software allows the user to visualize the movement dynamics of salt and water through the soil. By adding data on fertilizer input to the surface, it will be possible to get a better understanding of soil drainage characteristics of nutrients.

4 Data processing and analysis

4.1 Processing of spectral data

During data acquisitions with an aerial platform, the sensor is subjected to undesired pitch, roll, and yaw movements, which need to be corrected. Therefore, an integrated Inertial Measurement Unit (IMU) tracks these three angles continuously with a highly accurate timestamp and frequency. In a preprocessing step, called georeferencing, the tracked sensor movement is combined with the GPS location to correct the distortion and projects each recorded spectrum onto its correct location on a reference ellipsoid.

If the test area shows significant height differences, a precise digital elevation model helps to correct this inaccuracy. Since the internal GPS always tracks the exact position of the drone and sensor, geo-coordinates are assigned to every pixel of the corrected image.

When light is detected by a photosensitive chip, the sensor translates the resulting signal into a digital number with a typical range of values in 12 bit. This raw format is converted into radiance using internal and sensor-specific configuration and calibration files. Radiance depends on the intensity and direction of the illumination, the position and orientation of the target, and atmospheric conditions. These effects restrict radiance to a difficult physical quantity for the comparison of data sets. Reflectance, as the ratio between the amount of incident and reflected light, describes the properties of the materials more reliably and regardless of the effects listed. In order to convert the radiance into the reflectance, a reference panel with known reflectivity must be recorded over the entire wavelength range and set in relation to the corresponding radiance values.

Soil spectra collected by the handheld HS-sensor (FieldSpec 4) are used to carry out the correlation analysis with the pedological parameters. The goal is to transfer this correlation to airborne or space-borne hyperspectral data. Since the spectral resolution of the handheld HS-sensor is much higher, its spectral resolution has to be adapted to match the airborne or space-borne HS-sensors. Therefore, the nearest handheld HS-sensor bands are selected so that they correspond to the bandwidth of the corresponding air- and space-borne HS-sensors. This step is called spectral resampling.

For this approach, continuum removal (CR) was tested to improve the signal-to-noise ratio. CR is a tool for amplifying reflection features and especially for amplifying absorption bands in a spectrum. CR is performed by fitting a convex hull to the spectrum and dividing the reflectance values for each wavelength by the reflectance of the continuum line (convex hull) at the corresponding wavelength. This pre-processing returns a CR value of 1 to all parts of the spectrum that lie on the convex hull (i. e., wavelength regions that are not in an absorption band).

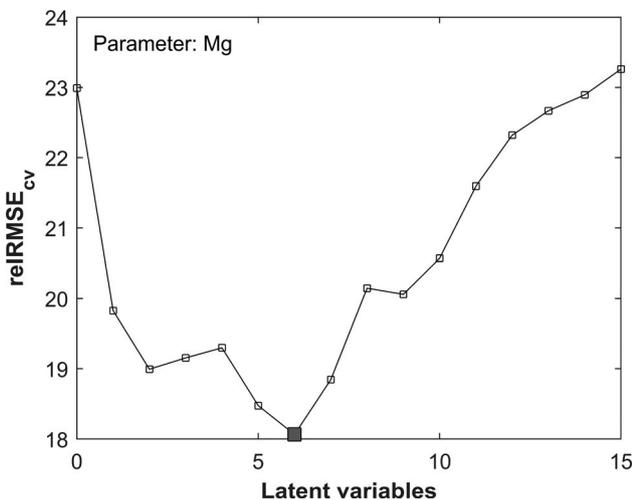


Figure 3: PLSR of magnesium. Relative cross validated RMSE (reRMSEcv) as a function of the number of latent variables, best model at six latent variables.

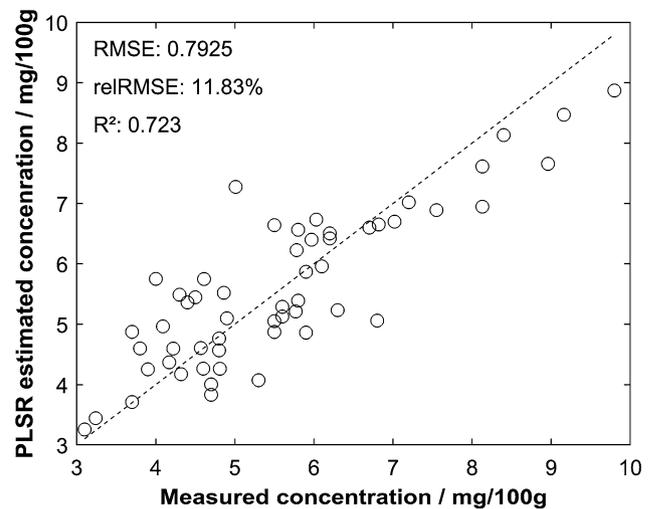


Figure 4: PLSR of magnesium. Measured versus estimated magnesium values using the model with six latent variables and the 1:1 line.

Values between 0 and 1 return to regions inside absorption bands. CR thus accentuates the absorption bands in the spectra while minimizing differences in brightness [19].

4.2 Correlation analysis – PLSR

Partial Least-Square Regression (PLSR) was chosen as an algorithm of multivariate statistics to find correlations between soil parameter values and their spectral signature values. PLSR, according to Wold et al. [20] is a widely used approach for quantitative analysis in chemometrics and hyperspectral remote sensing [21]. Closely related to principal components regression (PCR), PLSR combines features from PCR and multiple regression [22]. PLSR projects the data (chemical concentrations and reflective properties with a high number of correlated variables) into a lower-dimensional space, formed by a set of orthogonal latent variables, that maximizes the covariance between X and Y by a simultaneous decomposition of X (spectral matrix) and Y (chemical matrix) [23]. This projection reduces a large number of measured collinear spectral variables to a few non-correlated latent variables, which also implies a reduction of the data volume and the subsequent calculation time. The method is well suited for the calibration of a small number of samples with experimental noise in both chemical and spectral data, even if the number of observations is smaller than the number of wavelengths [24]. A detailed explanation of the application using PLSR on hyperspectral data can be found in [11].

For each pedological parameter analyzed, PLSR models with up to 15 latent variables were calculated in order

to determine the optimum number of latent variables. This calculation was performed separately in-/excluding continuum removal and considering the un-/resampled data. Five-fold cross-validation and a 100-fold Monte Carlo repetition were applied to each model run to avoid overfitting. The cross-validated relative root-mean-square error (reRMSEcv) was calculated to evaluate each model run. Figure 3 illustrates the relationship between the number of latent variables and reRMSEcv. The number of latent variables with the lowest reRMSEcv was selected. Figure 3 presents a PLSR for magnesium with the lowest reRMSEcv using six latent variables. As there is the reRMSE of the cross-validation plotted, the reRMSEcv can also increase with an increasing number of latent variables. With the optimal number of latent variables, a final PLSR and its errors were calculated. Figure 4 shows the resulting estimates compared to the laboratory measurements for magnesium.

4.3 Transfer to airborne/space-borne spectral images

The result of each parameter model calibration using PLSR is a correlation coefficient file. These files each contain one factor β_i ($\in \mathbb{R}$) for band i and one scalar offset α , which describes the correlation between the pedological parameter value and its reflection values. For each pixel of a hyperspectral image and each pedological parameter, these weighting factors β_i were multiplied by their corresponding sensor band values ρ_i . The resulting values were sum-

marized. After adding the offset value to this sum, the resulting amount was the final pedological parameter estimate \hat{y} of the HS-image pixel.

$$\hat{y} = \alpha + \sum_{i=1}^{n_{\text{Bands}}} \rho_i * \beta_i \quad (1)$$

A complete pedological parameter map can be calculated by applying this approach to every pixel of the airborne/space-borne HS-image.

4.4 Concept of a 3D-soil-parameter estimation using AI-approaches

While the PLSR approach proposed in the previous section should enable us to generate high-resolution spatio-temporal continuous 2D products (see also Figure 1, green boxes), we further aim to utilize the data collected by the SENTEK Drill & Drop TRISCAN sensors and deep learning to derive a detailed local 3D profile of nutrient distribution (see Figure 1, orange boxes).

State-of-the-art studies have demonstrated that estimating the nutrients' concentrations in soil-less cultivation conditions [13], based on electro-conductivity, is feasible. Whether and how this can be achieved in the soil remains an open research problem. The study of Moon et al. [13] shows that the best performance can be expected when using state-of-the-art recurrent neural networks, in particular, Long Short-Term Memory (LSTM) [14, 15, 17]. LSTMs are a favorite tool in the deep learning arsenal when it comes to modeling time series data and are able to learn both short-term dependencies, as well as long-term relationships. When it comes to remote sensing, they have recently been successfully employed to produce high-resolution spatio-temporal continuous soil moisture data from daily satellite images and more frequently sampled meteorological data at surface level [15].

We want to use LSTMs to estimate nutrient and soil moisture information at various depths based on hyperspectral images of the crops and soil surface, PLSR estimated nutrient values of the soil surface, and meteorological, irrigation, and fertilization data. Moon et al. [20] show that information about the development stage of the crop has a profound influence on the ability of their models to estimate the nutrients available in the substrate correctly. In our application scenario, we need additional information on the soil composition and soil moisture fluctuations at the surface level, as well as precipitation, in order to be able to estimate the flow of the nutrients through the soil.

Hyperspectral imagery should allow our models to infer the development stage of the crops and their influence on nutrient concentration at a certain depth due to the uptake of nutrients. In contrast, land surface meteorological data, irrigation, fertilization, and PLSR-derived surface nutrient concentration maps should enable our models to account for the propagation of the nutrients through the soil from the surface.

The downside of all deep learning approaches is the fact that they require significant amounts of training data. Since our soil-sampling campaign does not provide sufficient ground truth data concerning nutrients at a certain depth to train the LSTMs, we aim to train deep networks to predict the readings of the SENTEK sensors at different depths using the input data from the surface. These networks will output estimates of Volumetric Water Content (VWC), Volumetric Ion Content (VIC), and temperature for the spatial resolution of hyperspectral images and the temporal resolution of meteorological data from the ground surface.

The actual data from the soil sampling campaign will be used to train more conventional machine learning models, such as simple neural nets and PLSR models, to estimate the nutrient content based on real and estimated VWC, VIC and temperature values to provide detailed 3D information of nutrient distribution at each time step.

5 First results & discussion

5.1 Correlation analysis – PLSR

Table 3 shows the results of the correlations analysis between spectral and pedological data using PLSR. We compare results with the full number of bands of the Field-Spec 4 sensor (no resampling) and those with spectral resampling to emulate Headwall Co-Aligned and EnMAP data. Without resampling, TOC, humus, and potassium are reliably detected with an R^2 of 0.82–0.94. Magnesium ($R^2 = 0.72$) and nitrogen ($R^2 = 0.62$) present a less strong correlation. Phosphorus and calcium indicate the worst correlation among the parameters with R^2 of 0.29–0.39. The resampling to the Headwall spectral resolution led to similar results, only the quality of potassium decreased to an R^2 of 0.51. The resampling to the EnMAP spectral resolution increased the quality of the correlation for TOC, humus, and potassium to R^2 over 0.95. In the case of potassium, this increase is especially visible in comparison to the relRMSE, which indicates the error in percent-

age (RMSE normalized by the range of parameter). With errors of approximately 13 % for nitrogen and magnesium, a qualitative assumption can be considered possible. The spectral resampling to EnMAP improves the R^2 of calcium to 0.52, while phosphorus remains the worst correlation among the parameters with an R^2 of 0.28. The reduction in the number of bands taken into account (resampling to aerial/space-borne HS-sensors) had no negative impact on the result in most cases (except potassium resampled to Headwall resolution). In contrast, the results for TOC, humus, potassium (EnMAP), and especially calcium even improved. This fact can be explained by the improvement of the signal to noise ratio by the binning of adjacent bands.

Overall, without the need of a preprocessing, PLSR is robust even under variable and naturally occurring illumination conditions and shadow effects during the spectral data acquisition in the field. Continuum Removal (CR) only leads to slightly better results for some PLSR models, e. g., for potassium (no resampling), calcium (EnMAP resampling) as well as TOC and humus (Headwall resampling).

5.2 2D-soil-parameter maps

Due to the lack of real EnMAP hyperspectral satellite images, the German Aerospace Center simulated EnMAP images by adapting the sensor specifications of an airborne hyperspectral image to those of the planned EnMAP sensor [25]. As the spectral resolution of our ASD FieldSpec 4 was resampled to the resolution of EnMAP, corresponding PLSR models were ready to be applied to a simulated EnMAP image.

Figure 5 shows on the left a section of the image presenting an arable field and a beach section in Mecklenburg-Western-Pomerania in a true color composite. PLSR models were applied to pure soil pixels in the image (field and beach section), resulting in 2D-soil-parameter maps. The figure in the middle represents the potassium content. Humus is shown on the right. While the qualitative value distribution seems reasonable after a visual review of the exemplary maps, the underlying quantitative parameter value ranges do not match with the expected values in Table 1. The most likely reason is different parameter value ranges due to different soil types of the experimental field (Brandenburg) and the EnMAP field in Mecklenburg-Western-Pomerania. However, as soil data of the EnMAP field soil are not available, a reliable validation process is not possible.

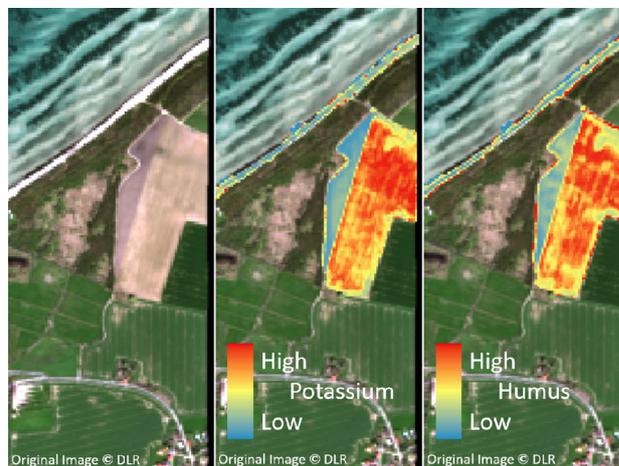


Figure 5: Simulated EnMAP image as true color image (left) and 2D-soil-parameter maps of Potassium (middle) and Humus (right). The database consists simulated EnMAP images provided by the DLR, Germany.

6 Conclusion & outlook

We proposed an approach to deriving detailed 2D-/3D-soil-parameter maps that cover the whole root zone of crops. For this purpose, we use hyperspectral remote sensing, underground soil sensors, and state-of-the-art machine learning algorithms to create computational models able to estimate the values that can currently only be determined using prohibitively expensive soil-sampling and underground sensing.

Fundamental steps of the 2D-soil-parameter map approach have been realized, and results indicate that our approach is promising. To continue our work, especially the test of the 3D-soil-parameter-maps concept using machine learning, we need to collect more data. In the future, we plan to conduct extensive field experiments involving soil sampling, underground sensing, as well as aerial sensing using a hyperspectral Headwall UAV system. In this way, we can validate our approach to generating high-resolution 2D-surface-soil-parameter maps and further complete our pipeline. If we are able to transfer the results obtained from soil samples and data collected with our handheld spectrometers to UAV hyperspectral images, we will eliminate the need for future soil sampling and handheld spectroscopic data acquisition. This makes the proposed 2D-surface-soil-parameter maps approach much more likely to become a feasible method of agricultural practice.

Several potential issues can be expected when applying our approach to real hyperspectral remotely sensed images. With UAV-acquired images, we are currently trying

Table 3: Statistical comparison of the PLSR models. The best statistical results (R^2 and reRMSE) and the corresponding number of latent variables (LV) are listed for each parameter. If the best result was achieved including CR, it is referred with “CR”.

	No resampling (2151 bands)				Headwall resampling (537 bands)				EnMAP resampling (242 bands)			
	R^2	reRMSE [%]	#LV	CR	R^2	reRMSE [%]	#LV	CR	R^2	reRMSE [%]	#LV	CR
N	0.62	12.54	5	–	0.60	12.88	6	–	0.59	13.09	6	–
P	0.29	16.1	2	–	0.30	16.06	2	–	0.28	16.25	2	–
Ca	0.39	12.69	2	–	0.34	12.76	2	–	0.52	11.23	5	CR
Mg	0.72	11.83	6	–	0.70	12.46	6	–	0.67	12.85	6	–
K	0.82	11.07	8	CR	0.51	18.38	6	–	0.95	6.15	12	–
TOC	0.94	5.67	10	–	0.92	6.35	10	CR	0.98	3.37	12	–
Humus	0.93	5.94	10	–	0.93	5.99	11	CR	0.97	3.76	12	–

to determine whether the model needs to be adapted to be effective in neighboring fields and soils. When it comes to using satellite data, the quality of the results is likely to decrease due to the noise introduced by the atmosphere, and the reduced spectral and spatial resolution. Feedback from farmers indicate, that even data of lower accuracy would be a benefit. Whether the quality of our described results is sufficient for high-precision treatments remains to be determined.

The main limitation of our 3D-soil-parameter estimation concept using machine learning is the relative data scarcity that can be obtained using soil sampling and underground sensors. This limits our ability to address the problem using state-of-the-art machine learning techniques, which require large amounts of data to train deep neural nets. Until more such data become available, a prospective workaround is to create ways to synthesize data using simulators.

Acknowledgment: The authors would like to thank the German Aerospace Center for providing EnMAP simulations based on HySpex data and Dylan Warren Raffa for the agronomic insights and expertise.

References

- Gomiero, T., D. Pimentel and M. G. Paoletti. 2011. Environmental impact of different agricultural management practices: conventional vs. organic agriculture. *Critical Reviews in Plant Sciences* 30 (1-2): 95–124. DOI: 10.1080/07352689.2011.554355.
- Finger, R., S. M. Swinton, N. El Benni and A. Walter. 2019. Precision farming at the nexus of agricultural production and the environment. *Annual Review of Resource Economics* 11 (1): 313–335.
- Ge, Y., J. A. Thomasson and R. Sui. 2011. Remote sensing of soil properties in precision agriculture: A review. *Front. Earth Sci.* 5: 229. DOI: 10.1007/s11707-011-0175-0.
- Bellon-Maurel, V. and A. McBratney. 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. DOI: 10.1016/j.soilbio.2011.02.019.
- Stoner, E. R. and M. F. Baumgardner. 1981. Characteristic variations in reflectance of surface soils 1. *Soil Science Society of America Journal* 45(6): 1161. DOI: 10.2136/sssaj1981.03615995004500060031x.
- Stevens, A., T. Udelhoven, A. Denis, B Tychon, R. Liroy, L. Hoffmann and B. van Wesemael. 2010. Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma* 158 (1-2): 32–45. DOI: 10.1016/j.geoderma.2009.11.032.
- Viscarra Rossel, R. A., S. R. Cattle, A. Ortega and Y. Fouad. 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* 150 (3-4): 253–266. DOI: 10.1016/j.geoderma.2009.01.025.
- Aldabaa, A. A. A., D. C. Weindorf, S. Chakraborty, A. Sharma and B. Li. 2015. Combination of proximal and remote sensing methods for rapid soil salinity quantification. *Geoderma* 239–240: 34–46. DOI: 10.1016/j.geoderma.2014.09.011.
- Stenberg, B. 2010. Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma* 158 (1-2): 15–22. DOI: 10.1016/j.geoderma.2010.04.008.
- Keller, S., F. M. Riese, N. Allroggen, C. Jackisch and S. Hinz. 2018. Modeling subsurface soil moisture based on hyperspectral data – first results of a multilateral field campaign. In: *Tagungsband der 38. Wissenschaftlich-Technischen Jahrestagung der DGPF e. V.*, vol. 27, pp. 34–48.
- Schreiner, S., H. Buddenbaum, C. Emmerling and M. Steffens. 2015. VNIR/SWIR laboratory imaging spectroscopy for wall-to-wall mapping of elemental concentrations in soil cores. *Photogrammetrie - Fernerkundung - Geoinformation* 2015 (6): 423–435. DOI: 10.1127/pfg/2015/0279.
- Qian, J. H., J. W. Doran, K. L. Weier, A. R. Mosier, T. A. Peterson and J. F. Power. 1997. Soil denitrification and nitrous oxide losses under corn irrigated with high-nitrate groundwater. *Journal of Environmental Quality* 26 (2).
- Moon, T., T. I. Ahn and J. E. Son. 2018. Forecasting root-zone electrical conductivity of nutrient solutions in closed-loop soilless cultures via a recurrent neural network using

environmental and cultivation information. *Frontiers in Plant Science* 9: 859.

14. Hochreiter, S. and J. Schmidhuber. 1997. Long short-term memory. *Neural Comput.* 9: 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.
15. Fang, K., M. Pan and C. Shen. 2018. The value of SMAP for long-term soil moisture estimation with the help of deep learning. *IEEE Transactions on Geoscience and Remote Sensing* 57 (4): 2221–2233.
16. Zhu, J.-H., M. M. Zaman and S. A. Anderson. 1998. Modelling of shearing behavior of a residual soil with Recurrent Neural Network. *International Journal for Numerical and Analytical Methods in Geomechanics* 22 (8).
17. Li, Q, Y. Zhao and F. Yu. 2020. A novel multichannel long short-term memory method with time series for soil temperature modeling. *IEEE Access* 8: 182026–182043. DOI: 10.1109/ACCESS.2020.3028995.
18. Lenz, A., H. Schilling and W. Middelmann. 2014. Umwelt- und Katastrophenschutz mit einem luftgetragenen hyperspektralen Multisensorsystem. *Technische Sicherheit* 4 (2014): 33–36. ISSN: 2191-0073.
19. Kokaly, R. and R. N. Clark. 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment* 67 (3): 267–287. DOI: 10.1016/S0034-4257(98)00084-4.
20. Wold, S., M. Sjöström and L. Eriksson. 2001. PLS-regression. A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58 (2), 109–130. DOI: 10.1016/S0169-7439(01)00155-1.
21. Steffens, M. and H. Buddenbaum. 2013. Laboratory imaging spectroscopy of a stagnic Luvisol profile – high resolution soil characterisation, classification and mapping of elemental concentrations. *Geoderma* 195–196: 122–132. DOI: 10.1016/j.geoderma.2012.11.011.
22. Marabel, M. and F. Alvarez-Taboada. 2013. Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. *Sensors* 13 (8): 10027–10051. DOI: 10.3390/s130810027.
23. Vohland, M. and C. Emmerling. 2011. Determination of total soil organic C and hot water-extractable C from VIS-NIR soil reflectance with partial least squares regression and spectral feature selection techniques. *European Journal of Soil Science* 62 (4): 598–606. DOI: 10.1111/j.1365-2389.2011.01369.x.
24. Ben-Dor, E., D. Heller and A. Chudnovsky. 2008. A novel method of classifying soil profiles in the field using optical means. *Soil Science Society of America Journal* 72 (4): 1113. DOI: 10.2136/sssaj2006.0059.
25. Becker, M., S. Schreiner, S. Auer, D. Cerra, P. Gege, M. Bachmann, A. Roitzsch, U. Mitschke and W. Middelmann. 2018. Reconnaissance of coastal areas using simulated EnMAP data in an ERDAS IMAGINE environment. In: *SPIE Remote Sensing, Proceedings*, vol. 10790. DOI: 10.1117/12.2325402.

Bionotes



Simon Schreiner

Fraunhofer IOSB, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation, Gutleuthausstrasse 1, 76275 Ettlingen, Germany
simon.schreiner@iosb.fraunhofer.de

Simon Schreiner received his Master of Science degree in Applied Geoinformatics at the University of Trier, Germany, in 2016. He was research assistant at this University from 2012 until 2016. In 2017, he was cartographer at the City Department of Surveying in Stuttgart, Germany. Since 2018, he has been employed at the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB in Ettlingen, Germany. His research focusses on hyperspectral remote sensing for defense technologies and precision farming applications.



Dubravko Culibrk

University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, 21102 Novi Sad, Serbia
dculibrk@uns.ac.rs

Dubravko Culibrk is a Full Professor of Information Systems Engineering at the Department of Industrial Engineering and Management of the Faculty of Technical Sciences, University of Novi Sad, Serbia. While pursuing his PhD degree (2003–2006) at the Florida Atlantic University, USA, he was affiliated with the Center for Coastline Security and the Center for Cryptology and Information Security. He spent two years (2013–2015) as a postdoc researcher at the University of Trento, Italy with the Multimedia and Human Understanding Group. His current research interests include: Neural Networks and Deep Learning, Computer Vision, Machine Learning and Data Science, Multimedia, Remote Sensing and Image/Video Processing. He is an NVIDIA Deep Learning Institute University Ambassador.



Michele Bandecchi
SmartCloudFarming GmbH, Rosenthaler
Strasse 72a, 10119 Berlin, Germany
bandecchi@smartcloudfarming.com

Michele Bandecchi is the CEO of SmartCloudFarming GmbH and he is supervising the business and technology development of the company's products. He received his two Master degrees in (i) Plant Biotechnology with specialization plant genomics and (ii) Management, Marketing and Consumer Studies with a specialization in Innovation Management at the Wageningen University Research (The Netherlands). He gained experience in marketing and sales during an internship with GRASSLANDZ TECHNOLOGY LTD. in New Zealand and for ten years in his family business in marketing and sales and business development.



Wolfgang Gross
Fraunhofer IOSB, Fraunhofer Institute of
Optronics, System Technologies and Image
Exploitation, Gutleuthausstrasse 1,
76275 Ettlingen, Germany
wolfgang.gross@iosb.fraunhofer.de

Wolfgang Gross was born in Karlsruhe, Germany, in 1985. He received the Diploma in technomathematics from the Karlsruhe Institute of Technology KIT, Germany, in 2011. In 2019, he received his Ph. D. in engineering from the University of Stuttgart, Faculty for Aerospace Engineering and Geodesy. He is currently employed by the Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, His current research includes the development of algorithms for hyperspectral analysis regarding data-driven automatic preprocessing, classification and manifold alignment for data synthesis.



Wolfgang Middelmann
Fraunhofer IOSB, Fraunhofer Institute of
Optronics, System Technologies and Image
Exploitation, Gutleuthausstrasse 1,
76275 Ettlingen, Germany
wolfgang.middelmann@iosb.fraunhofer.de

Dr. Wolfgang Middelmann is the head of the image interpretation group at Fraunhofer IOSB. He got his diploma in mathematics at University Dortmund (Germany) in 1993. In 1998 he received his PhD (Dr. rer. nat.) in mathematics at Friedrich-Schiller-University Jena (Germany). Until 2002 he worked as software engineering manager for ESA projects and as project manager of international military projects at Daimler-Chrysler Aerospace. Since 2002 he was project manager and scientist for ATR and reconnaissance systems at FGAN-FOM Research Institute for Optronics and Pattern Recognition (2010 merged into Fraunhofer IOSB), Ettlingen (Germany). In 2008 he advanced to the head of the image interpretation group. Currently he is responsible for concepts, development and assessment of airborne and spaceborne ISR systems.