OPTICAL MONITORING OF SOIL-FORMING PROCESSES IN FOREST SOILS USING RGB AND HYPERSPECTRAL IMAGING

Patrick Menz[1], Philipp Koal[2*], Paul Gescher[1], Joshua Moritz[1], Steven Ackermann[2], Ina Ehrhardt[1]*

[1] Fraunhofer Institute for Factory Operation and Automation IFF, Magdeburg, Germany
[2] Forestry Research and Competence Centre - Gotha, ThüringenForst AöR, Germany

ABSTRACT

Forest soil formation is characterized by soil organic matter (SOM) accumulation, browning, clay migration, podzolization, and hydromorphic redox processes. These processes form genetically differentiated horizons and alter soil color, mineralogy, and iron oxide composition, creating optical signatures. We compare RGB, VNIR, and SWIR imaging for estimating laboratory properties relevant to forest soil formation. In total, 312 soil samples representing Cambisols, Luvisols, Podzols, podzolized Cambisols and Stagnosols were analyzed with all three sensor systems. Soil organic carbon (SOC), clay content, dithionite-extractable iron (Fe_d), and the Fe_o/Fe_d ratio served as pedologically relevant targets. Among the tested sensors, VNIR achieved the highest performance for SOC and SWIR for clay content and the Fe_o/Fe_d ratio, whereas Fe_d remained poorly predictable with all three sensors. These findings support optical estimation of pedologically relevant properties, while highlighting substantial differences among soil-type-specific models.

Index Terms— Pedogenesis, forest soils, hyperspectral imaging, soil spectroscopy, soil color, RGB imaging, machine learning

1. INTRODUCTION

Sustainable forest management depends on a reliable assessment of forest site conditions, which in turn requires a thorough understanding of long-term soil development. In contrast to agricultural systems, where management decisions are often based on individual soil properties that can be specifically influenced by fertilization or soil improvement measures, forest site evaluation is primarily based on the expression of long-term soil-forming processes. These processes determine root penetration, water balance, nutrient availability, and ultimately the suitability of sites for various tree species [1, 2]. Cambisols, Luvisols, Podzols, podzolized Cambisols, and Stagnosols represent relevant characteristic development pathways in pedogenesis, including browning, clay migration, podzolization, and hydromorphic redox processes. These processes alter soil color, mineralogy, clay mineral composition, and the abundance and crystallinity of pedogenic iron oxides, thereby generating characteristic optical signatures [3–7].

Studies using visible, near-infrared, and hyperspectral imaging focus primarily on predicting individual laboratory parameters such as soil organic carbon, texture, or various iron fractions [5, 7–11]. These established approaches can support forest site evaluation when their predictions are interpreted together with horizon and profile observations. When assessing forest sites, it is often not a single laboratory parameter that is decisive, but rather the extent of soil-forming processes, such as the degree of podzolization, hydromorphic differentiation, or clay migration. These processes form the basis of forest site mapping and support decisions regarding e.g. species selection, liming, and the long-term assessment of site productivity [1, 2]. Pedological assessment is conducted primarily on the basis of morphological characteristics in accordance with pedological mapping guidelines (e.g., KA6 in Germany [12]). Surveying these characteristics is time-consuming and requires substantial pedological expertise. Reproducible optical measurements can complement conventional assessment methods by providing objective and standardized data.

In this study, pedogenic information refers to information on soil properties associated with soil-forming processes and horizon development. These laboratory-measured properties were modeled as separate prediction targets. The study investigates whether extending optical observations from RGB to hyperspectral VNIR and SWIR imaging improves their prediction. The aim of this study is to systematically compare the predictive performance of RGB, VNIR, and SWIR imaging for soil properties associated with soil-forming processes in forest soils, using the same soil samples, target variables, and validation framework. VNIR and SWIR data were evaluated separately to assess the predictive value of each spectral range, as they capture different absorption mechanisms and were acquired with sensors differing in both spectral and spatial resolution. To this end, soil organic carbon (SOC), clay content, and extractable iron fractions (Fe_d and Fe_o) are used as reference variables reflecting selected aspects of soil formation. The predictive performance of machine-learning models based on RGB, VNIR, and SWIR data is then compared.

* These authors contributed equally to this work.

2. MATERIAL AND METHODS

2.1. Study design and selection of soil samples

The aim of the study was to investigate the extent to which AI-based methods can be used to distinguish and analyze optical signatures of soil-forming processes in forest soils. The study is based on soil samples collected from genetically distinct horizons in soil profiles classified as Cambisols, Luvisols, Podzols, podzolized Cambisols, and Stagnosols at representative selected forest sites. These soil types represent characteristic pedogenic soil-forming processes such as browning, clay migration, podzolization, and hydromorphic redox processes, and thus constitute typical site conditions for forest site evaluation.

Two horizons were sampled from each of 156 forest soil profiles in central Germany, yielding a total of 312 soil samples. Across all profiles, the dataset comprises 14 different genetic horizon designations. The samples, which were air-dried in the laboratory and sieved to < 2 mm, were homogenized, divided into three replicates and placed in identically sized Petri dishes. To ensure optimal focus and uniform coverage, they were smoothed to a defined height using a Drigalski spatula and used for all optical measurements and laboratory analyses.

2.2. Camera measurement technology

To investigate the information content of different optical imaging modalities, standardized RGB images were captured using a Canon EOS 6D Mark II (see Fig. 1). In addition, hyperspectral line scan cameras, consisting of a NEO Hyspex VNIR 1600 with 160 channels and a NEO Hyspex SWIR-320m-e with 256 channels, were used to capture spectral signatures ranging from 400 to 2500 nm (see Fig. 1).

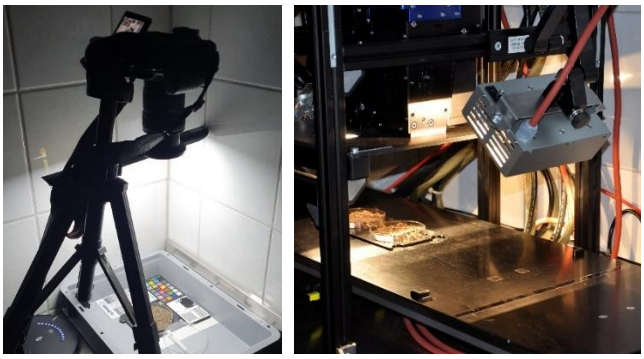


Fig. 1: Setup for RGB (left) and spectral imaging (right)

The RGB images were captured under standardized laboratory conditions in a completely darkened room using controlled photographic lighting (6000K LED studio light) and a Calibrite ColorChecker Mini in each shot to ensure reproducible color values. The hyperspectral measurements were performed under identical conditions. Line-forming halogen lamps with a stabilized DC power supply were used for illumination. A graduated Zenith Polymer® standard with certified reflectance levels of 90%, 50%, 25%, 15%, and 5% was used to calibrate the raw spectral data. This enabled the acquisition of both soil color information and spectral features associated with organic matter, iron oxides, and clay minerals. The goal was to investigate the extent to which the different spectral ranges contribute to the characterization of soil-forming processes.

2.3. Laboratory reference variables

To characterize soil properties associated with soil-forming processes, soil organic carbon (SOC), clay content, and two extractable iron fractions were determined in three analytical replicates per soil sample. SOC was measured by dry combustion using an elemental analyzer (DIN ISO 10694). Clay content (<2 µm) was determined by sieving and sedimentation using the pipette method (ISO 11277). Dithionite-extractable iron (Fe_d) was determined after citrate–bicarbonate–dithionite extraction, and oxalate-extractable iron (Fe_o) after acid ammonium oxalate extraction in the dark. The Fe_o/Fe_d ratio was calculated from these two measured fractions.

These parameters can serve as reference variables for describing key soil-forming processes. While SOC primarily reflects SOM accumulation, clay content and iron fractions are influenced by clay migration, iron transformation, podzolization, and hydromorphic soil development. In practice all variables are interpreted within a process-oriented framework that links visual observations with pedogenesis. SOC, clay, Fe_d, and Fe_o/Fe_d thus serve as complementary indicators of the dynamics of organic matter, mineral weathering, clay redistribution, iron oxide formation, and redox-driven soil development. Taken together, these variables form a pedogenic reference framework.

In addition to the distribution of the laboratory reference values in Table 1, the variability among the three replicates was considered to assess the subsequent model performance in relation to laboratory measurement variability.

Table 1: Descriptive statistics for the laboratory reference variables, including the mean, standard deviation, range, and mean replicate variance

Lab. Value	n	Mean	STD	Min	Max	Repl. Var.
SOC [%]	312	4.79	4.95	0.21	27.99	0.11
Clay [%]	312	19.36	7.65	4.53	43.47	0.60
Fe_d [g kg-1]	312	14.85	6.28	2.95	33.99	0.39
Fe_o/Fe_d [-]	312	0.21	0.11	0.03	0.53	0.01

2.4. Dataset

As described in Section 2.1, each soil sample was divided among three identically prepared Petri dishes and assessed. At the same time, each laboratory parameter was determined using three analytical replicates of the same sample. Since the laboratory values were determined using separate subsamples, a 1:1 correspondence between a single Petri dish and a single laboratory value is not possible. It is therefore assumed that the subsamples taken from the respective sample bags for optical measurements and laboratory analysis, respectively, are nearly identical due to the prior preparation. Initial evidence supporting this assumption is provided by the very low mean replicate variance of the laboratory values in Table 1, which indicates a homogeneous distribution of the target parameters within a sample.

Accordingly, both the optical measurements and the laboratory values were averaged at the sample level. From the three spectra or RGB values of each sample, an average value was calculated, and from the three analytical replicates, an average reference value was calculated. Each sample is thus assigned exactly one averaged optical signal and one averaged laboratory value.

Before averaging, the average spectrum for each Petri dish was extracted from the spectral cube and cleaned of noise using a third-order median filter. In the RGB images, color correction was performed using the ColorChecker-based linear correction including offset. The mean values of the red, green, and blue channels within the sample served as input for the training described in Section 2.5. In total, this resulted in 312 soil samples, each with an averaged optical data set per measurement system and an averaged laboratory reference value.

To illustrate the spectral variability of the soil samples, Fig. 2 and Fig. 3 show averaged reflectance spectra across three SOC and clay content ranges. This already demonstrates that a high SOC results in lower reflectance.

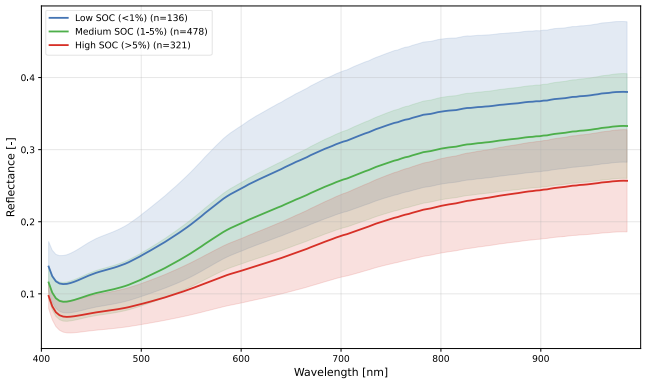


Fig. 2: Mean spectra VNIR with SOC content ranges

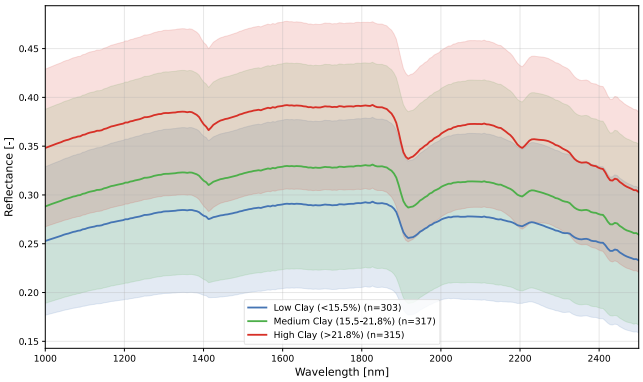


Fig. 3: Mean spectra SWIR with clay ranges

2.5. Modeling and validation

Calibration models were developed to predict the laboratory reference variables from the optical input data. The modeling was performed using the XGBoost machine learning algorithm. A separate model was trained for each target variable, each sensor system and each soil type.

Soil-type-specific calibration grouped samples whose optical signatures are shaped by comparable pedogenic processes, which reduces pedological heterogeneity within individual models. Genetic horizons were included jointly within each soil-type model. No horizon-specific calibration or performance assessment was performed.

Testing was performed using leave-one-out cross-validation. Averaging at the sample level ensures that measurements from the same sample are not included in both the training and test datasets at the same time, thereby preventing data leakage. Horizons from the same profile were not grouped during cross-validation, as sampling covered topsoil and subsoil, which differ too strongly to represent replicate measurements of a common soil condition.

Within each fold, the training set was further split, with 20% of the samples set aside as an internal validation set and the remaining 80% used to fit the model. The validation set was used exclusively for early stopping. The test sample for each fold was not affected by this division.

To minimize the impact of initialization on the model results, each model was trained as a seed ensemble. To do this, five models with different, fixed random seeds were fitted for each run, and their predictions were averaged to obtain the final estimate. The fixed seeds also ensure the reproducibility of the results.

Hyperparameter optimization was performed using Bayesian optimization based on 5-fold cross-validation. The best parameter set identified in each case was then used for modeling in the leave-one-out approach. For the hyperspectral data, the second derivative of the spectra was used as preprocessing. The RGB data were used as they are.

3. RESULTS AND DISCUSSION

The predictive performance of the optical measurement systems was evaluated using the reference variables of soil organic carbon (SOC), clay content (Clay), dithionite-extractable iron (Fe_d), and the Fe_o/Fe_d ratio. These key targets were chosen for their relevance to soil-forming processes. For each target variable and each soil type, calibration models were trained using RGB, VNIR, and SWIR data. Model performance was assessed using R^2 and RMSE with the model predictions averaged across seeds (Table 2). Since the modeling was performed on a soil-type-specific basis, the metrics refer to the predictions from the leave-one-out cross-validation that were pooled across all soil types.

Table 2: Results of the models for seed-averaged predictions for each optical sensor and target variable pooled across all soil types; the best result for each target is shown in bold

Target	R^2			RMSE		
	RGB	VNIR	SWIR	RGB	VNIR	SWIR
SOC	0.53	**0.6**	0.50	3.41	**3.13**	3.48
Clay	0.40	0.60	**0.68**	5.94	4.84	**4.32**
Fe_d	0.19	**0.37**	0.25	5.66	**4.98**	5.41
Fe_o/Fe_d	0.41	0.51	**0.64**	0.09	0.08	**0.07**

Across the five members of the seed ensemble, the standard deviation of R^2 remained below 0.02 in all cases, indicating limited sensitivity to the random seed under the applied evaluation procedure.

The best-performing spectral range varies among the laboratory targets. For SOC, VNIR data achieved the highest predictive performance. The comparatively small difference between RGB and VNIR indicates that the darkening of the soil caused by SOM accumulation already provides a substantial part of the relevant information in the visible range, while the VNIR range adds further spectral detail. For clay content, the highest performance was obtained in the SWIR. This is plausible, as clay minerals exhibit characteristic absorption features associated with O-H and Al-OH bonds in the short-wave infrared region.

For the iron-related targets, the results differ markedly between the two parameters. The Fe_o/Fe_d ratio was predicted best in the SWIR. Although iron oxides show characteristic electronic absorptions mainly in the VNIR range, the higher performance of the SWIR models suggests that the prediction is not based solely on direct iron-oxide absorptions but also reflects indirect covariation with SWIR-active constituents, in particular OH-bearing and clay minerals associated with pedogenic development. Fe_d remained poorly predictable across all sensors. One possible explanation is that Fe_d reflects both lithological supply and pedogenic transformation, whereas reflectance is more sensitive to specific Fe phases and their associations with other minerals than to bulk extractable Fe.

Fig. 4 shows the predicted versus measured values together with the soil-type-specific performance measures. Sample sizes and predictive performance differ markedly among soil types. For SOC in the VNIR, the Luvisols and Cambisols achieve comparatively high within-group R^2, Stagnosols a lower positive R^2, and Podzols and podzolized Cambisols negative R^2. A comparable ranking occurs for clay in the SWIR. For Fe_o/Fe_d in the SWIR, only the smallest group remains negative.

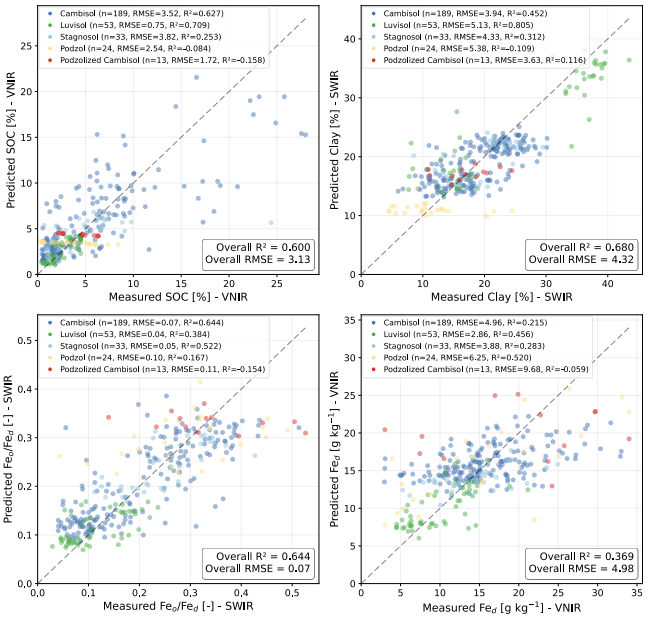


Fig. 4: Predicted versus measured values for SOC, clay, Fe_d and Fe_o/Fe_d, with soil-type-specific R^2, RMSE and numbers of samples

The groups also differ strongly in number of samples (Fig. 4). Small sample sets are unlikely to cover the full within-group range of soil composition, horizon characteristics and reference-value concentrations, which makes cross-validated estimates sensitive to individual observations and may contribute to the weak performance of the smallest groups. A second, potentially interacting influence is the lithological diversity of the parent materials, which were designated according to the German soil survey guidelines KA6 [12]. The dataset comprises 39 distinct parent materials among the Cambisols, 8 among the Stagnosols, 6 among the Podzols, 5 among the podzolized Cambisols and 2 among the Luvisols. Lithological inheritance governs the primary mineral assemblage and influences particle-size composition, weathering products and the abundance and speciation of Fe-bearing phases. It can therefore modify both the spectral response and its covariance with SOC, clay, Fe_d and Fe_o/Fe_d [3, 7]. Soils of the same pedogenic group do not need to show

a uniform relationship between spectra and properties if they developed from mineralogically dissimilar substrates. Stratification by parent material has been reported to improve Vis-NIR predictions of soil organic matter [13]. Lithological effects are moreover superimposed by pedogenesis, as clay illuviation, podzolization and redox-driven Fe redistribution generate horizon-dependent contrasts even within comparable substrate settings [1, 4]. Splitting the already small soil-type groups by horizon would have further reduced the number of samples per model. The number of recorded parent material classes is not a quantitative measure of mineralogical dissimilarity and is itself a function of sampling intensity. Because samples representation, parent material composition and horizon distribution cannot be disentangled with the present data, the observed variation in model performance is not attributed to a single cause. The uneven representation of soil types mirrors their naturally heterogeneous occurrence in the study area, so the dataset reflects the pedological diversity of the investigated forest landscape. A more rigorous assessment would require sample-size-matched comparisons, residual analysis by parent material and genetic horizon, and a larger number of samples for the weakly represented soil types.

4. CONCLUSION AND FURTHER WORK

This study shows that optical imaging techniques can provide a basis for interpreting soil-forming processes in forest soils. Based on the reference indicators SOC, clay content, and Fe_o/Fe_d, the spectral data proved effective in capturing soil properties related to pedogenic processes. Fe_d, in contrast, was poorly predicted by all three sensor systems, indicating that either the optical inputs alone or the available sample size were insufficient for robust estimation in this dataset. While SOM accumulation can already be primarily described using visible color information, hyperspectral data (particularly in the SWIR range) provide additional information for mineralogically driven processes such as clay migration.

This approach opens new perspectives for forest site assessment. The combination of RGB and hyperspectral imaging with AI-supported analysis methods can help to characterize soil-forming processes more objectively, reproducibly, and efficiently. Thus, the approach holds potential for digital site mapping as well as for future monitoring programs in forest ecosystems.

The soil-type-specific differences in predictive performance are consistent with a substrate influence, but sample-size effects and horizon-related variation cannot be separated from it in the present data. Future investigations should combine more balanced sampling across parent materials and diagnostic horizons with substrate-aware calibration strategies to assess the transferability of the models across diverse forest soils.

5. ACKOWNLEDGMENT

Funding within the BMLEH and FNR project KI-SOIL (2224NR083A) under the program Renewable Raw Materials is gratefully acknowledged.

6. REFERENCES

[1] R. J. Schaetzl and M. L. Thompson, Soils: Genesis and Geomorphology, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[2] IUSS Working Group WRB, World Reference Base for Soil Resources: International Soil Classification System for Naming Soils and Creating Legends for Soil Maps, 4th ed. Vienna, Austria: International Union of Soil Sciences, 2022.

[3] J. A. M. Demattê et al., “Spectral pedology: A new perspective on evaluation of soils along pedogenetic alterations,” Geoderma, vols. 217-218, pp. 190-200, 2014, doi: 10.1016/j.geoderma.2013.11.012.

[4] F. S. Terra, J. A. M. Demattê, and R. A. Viscarra Rossel, “Proximal spectral sensing in pedological assessments: Vis-NIR spectra for soil classification based on weathering and pedogenesis,” Geoderma, vol. 318, pp. 123-136, 2018, doi: 10.1016/j.geoderma.2017.10.053.

[5] R. A. Viscarra Rossel et al., “Diffuse reflectance spectroscopy for estimating soil properties: A technology for the 21st century,” Eur. J. Soil Sci., vol. 73, no. 4, Art. no. e13271, 2022, doi: 10.1111/ejss.13271.

[6] U. Schwertmann and R. M. Cornell, Iron Oxides in the Laboratory: Preparation and Characterization, 2nd ed. Weinheim, Germany: Wiley-VCH, 2000.

[7] B. Stenberg, R. A. Viscarra Rossel, A. M. Mouazen, and J. Wetterlind, “Visible and near infrared spectroscopy in soil science,” Adv. Agron., vol. 107, pp. 163-215, 2010, doi: 10.1016/S0065-2113(10)07005-7.

[8] R. A. Viscarra Rossel, D. J. J. Walvoort, A. B. McBratney, L. J. Janik, and J. O. Skjemstad, “Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties,” Geoderma, vol. 131, nos. 1-2, pp. 59-75, 2006, doi: 10.1016/j.geoderma.2005.03.007.

[9] D. J. Brown, K. D. Shepherd, M. G. Walsh, M. D. Mays, and T. G. Reinsch, “Global soil characterization with VNIR diffuse reflectance spectroscopy,” Geoderma, vol. 132, nos. 3-4, pp. 273-290, 2006, doi: 10.1016/j.geoderma.2005.04.025.

[10] R. A. Viscarra Rossel et al., “A global spectral library to characterize the world's soil,” Earth-Sci. Rev., vol. 155, pp. 198-230, 2016, doi: 10.1016/j.earscirev.2016.01.012.

[11] C. Piccini et al., “In-field soil spectroscopy in Vis-NIR range for fast and reliable soil analysis: A review,” Eur. J. Soil Sci., vol. 75, no. 2, Art. no. e13481, 2024, doi: 10.1111/ejss.13481.

[12] AG Boden, Bodenkundliche Kartieranleitung (KA6): 6th completely revised and expanded ed. Hannover, Germany: Bundesanstalt für Geowissenschaften und Rohstoffe (BGR), 2024.

[13] S. Xu, X. Shi, M. Wang, and Y. Zhao, “Effects of subsetting by parent materials on prediction of soil organic matter content in a hilly area using Vis–NIR spectroscopy,” PLOS ONE, vol. 11, no. 3, e0151536, 2016. doi: 10.1371/journal.pone.0151536