~~Dataset variability and carbonate concentration influence the performance~~Best performances of ~~local~~ visible-near infrared ~~spectral~~ models in soils with little carbonate - a field study in Switzerland

Simon Oberholzer[1], Laura Summerauer[2], Markus Steffens[1,3], Chinwe Ifejika Speranza[1]

[1]Institute of Geography, University Bern, 3012 Bern, Switzerland
[2]Department of Environmental System Science, ETH Zürich, 8092 Zürich, Switzerland
[3]Department of Soil Science, Research Institute of Organic Agriculture, 5070 Frick, Switzerland

Correspondence to: Laura Summerauer (laura.summerauer@usys.ethz.ch)

Abstract. ~~The application~~Conventional laboratory analysis of ~~visual~~soil properties is often expensive and requires much time if various soil properties are to be measured. Visual and near infrared ~~soil~~(vis–NIR) spectroscopy ~~(vis–NIR) is an easy~~offers a complementary and cost-efficient way to gain a wide variety of soil information ~~to cover~~in high spatial and temporal resolution ~~in large-scale soil surveys and~~. Yet, applying vis-NIR spectroscopy requires confidence in ~~local field-scale studies. However, unlike for conventional methods,~~ the prediction accuracy of ~~vis–NIR spectral~~ the infrared models ~~cannot yet be estimated before the data collection, which hampers its application at the local scale where often a high precision is required (e.g., field experiments).~~. In this study we used soil data from six agricultural fields in Eastern Switzerland and calibrated i) field-specific (local) models and ii) general models (combining all fields) for soil organic carbon~~, total carbon, total nitrogen,~~ (SOC), permanganate oxidizable carbon (POXC), total nitrogen (N), total carbon (C) and pH using partial least squares regression. ~~24 out of~~The 30 local models showed ~~an accurate or even excellent performance (~~a ratio of performance to deviation (RPD) ~~≥ 2)~~between 1.14 and 5.27 and the root mean square errors (RMSE) ~~of~~ were between 1.07 and 2.43 g kg^{-1} for SOC, between 0.03 and 0.07 g kg^{-1} for POXC, between 0.09 and 0.14 g kg^{-1} for total N, between 1.29 and 2.63 g kg^{-1} for total C and between 0.04 and 0.19 for pH. Two fields with high carbonate content and poor correlation between the target properties were responsible for six local models with a low performance (RPD < 2). Analysis of variable importance in projection as well as correlations between spectral variables and target soil properties confirmed that high carbonate content masked absorption features for SOC. Field sites with low carbonate content can be combined to general models with only limited loss in prediction ~~were, except for pH, maximum five times higher than the lab measurement error. The variability of a~~ accuracy compared to the field-specific ~~soil property and~~models. On the ~~mean~~other hand, for fields with high carbonate ~~concentration in~~contents the ~~dataset were the two factors influencing the performance of the local models. We found a significant relationship between the coefficient of variation in the dataset and the metrics for model performance (R^2, percental RMSE and RPD). Starting from a tolerable~~ prediction ~~error for the spectral measurements, the regressions can be used to develop a sampling design that matches the corresponding target variability. The five inaccurately performing local models with RPD < 2 were on the two fields with highest carbonate content raising the question if local vis–NIR models are suitable for~~ accuracy substantially decreased in

Formatvorlagendefinition: Standard
Formatvorlagendefinition: Bullets: Schriftart: Verdana, 9.5 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Kopfzeile: Schriftart: Verdana, 9.5 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Copernicus_Word_template: Schriftart: 12 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: MS title: Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Affiliation: Schriftart: 11 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Sprechblasentext: Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Equation: Schriftart: 11 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Fußzeile: Schriftart: 11 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Correspondence: Schriftart: 11 Pt., Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Authors: Deutsch (Schweiz), Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: EndNote Bibliography Title: Schriftart: 11 Pt., Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: EndNote Bibliography: Schriftart: 11 Pt., Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Kommentartext: Schriftart: 11 Pt., Ligaturen: Standard + Kontextbezogen
Formatvorlagendefinition: Überarbeitung
Formatiert: Breite: 21 cm

general models. Whether the combination of soils with high carbonate ~~concentration. General models combining the datasets from all six fields showed an accurate overall performance but the RMSE on the field level were higher compared to the local models.~~ contents in one prediction model leads to satisfying prediction accuracies needs further investigation.

1 Introduction

The application of spectroscopy in the visible and near infrared (vis–NIR) range is increasing in soil science and related disciplines with the main objective to gain information on soil properties ~~for~~of more samples at lower costs than with conventional laboratory methods. With a larger sample size, the spatial or temporal resolution can be increased which allows conclusions about the within-field or within-farm variability but might potentially also increase the statistical power in agricultural experiments (Greenberg et al., 2022). Despite its ~~often lower performance~~tendency to be less accurate compared to mid infrared (MIR) spectroscopy, vis–NIR spectroscopy is widely applied because of less sample preparation, lower costs~~,~~ and generally easier portability (Soriano-Disla et al., 2014).

On-site vis–NIR measurements are therefore feasible but laboratory measurements with dried and sieved soil samples have so far shown higher accuracy (Allory et al., 2019; Hutengs et al., 2019). In particular, soil properties related to soil organic matter can be estimated appropriately by laboratory vis–NIR spectroscopy (Angelopoulou et al., 2020). In most cases, the focus is to provide soil information over large areas (e. g. soil maps) where a high sample number is present and only a moderate prediction accuracy is needed. ~~Large~~Hence, large-scale spectral libraries have been developed, to further reduce the need for wet chemistry data. Due to the high complexity within spectral libraries, the application of a general model to ~~the~~a local context leads to high prediction errors. Recent research ~~has shown~~shows that the localization of these infrared models substantially improves the predictive performance in a local context, for example by spiking (Brown, 2007; Li et al., 2020; Ng et al., 2022; Seidel et al., 2019; Wetterlind and Stenberg, 2010; Zhao et al., 2021), memory-based learning (Ramirez-Lopez et al., 2013), resampling algorithms (Lobsey et al., 2017) or deep learning (Shen et al., 2022). However, for analyzing small-scale variability (field or farm level), a local model is often still the best choice because ~~it achieves the lowest~~of its low prediction errors. ~~On a theoretical level, the development of~~Theoretically, developing local models is supported by the finding that in the vis–NIR range, spectral features that influence specific soil properties vary strongly between different datasets, which makes highly heterogenous large datasets prone to insufficient model performance (Angelopoulou et al., 2020; Grunwald et al., 2018).

~~The main motivation for the application of vis–NIR spectroscopy in local projects are lower costs for analytics that permits to increase the number of samples at only little additional costs. With a larger sample size, the spatial or temporal resolution can be increased which allows conclusions about the within-field or within-farm variability but might potentially also increase the statistical power in agricultural experiments (Greenberg et al., 2022).~~ The development of local spectral models has the main purpose to cope with a ~~high~~large sample size at the local scale, but ~~these~~such local models ~~do not~~ have ~~any use~~no utility beyond the analysis of the specific local dataset.

Spectral vis–NIR models developed from local datasets showed a very high variability in model performance ranging from excellent models (Breure et al., 2022; Seidel et al., 2019) to those with relatively poor model performance (Camargo et al., 2022; Kuang and Mouazen, 2011). The reasons for these different performances of local models are understudied and remain ~~often~~ unclear. Among many different possible modelling approaches including support vector machine regression, artificial neural networks, cubist and random forest, partial least square regression (PLSR) is the most frequently used model type to build spectral models with small datasets (Alomar et al., 2021; Zhao et al., 2021).

~~Comparing the performance of several models is demanding because spectral models (as most other models) need several parameters for their evaluation and one parameter alone might be misleading. The most known statistical parameter is the goodness of fit (R^2) between the predicted versus the measured data. However, R^2 does not allow any conclusion about prediction accuracy. The root mean square error (RMSE) is the estimated prediction error of the spectral model and has the unit of the measured property. However, it is difficult to compare the RMSE among models and datasets since the data range varies. Therefore, the ratio of performance to deviation (RPD), which relates the RMSE with the range of the data expressed as standard deviation, is often used. A less frequently reported but important performance parameter is the bias that expresses the average over- or underestimation of the model predictions. The bias is particularly important if the model is applied on samples that are very different from the samples used in the calibration model (Bellon-Maurel and McBratney, 2011).~~

~~For local models at field or farm scale, the variability of the target soil property is very important (Alomar et al., 2021; Perez-Bejarano and Guerrero, 2018; Stenberg et al., 2010). If the variability is very small, spectral measurements might not be sensitive enough to distinguish the different samples, which may result in a low RMSE but also in a low R^2 and probably also a low RPD. On the other hand, if variability of a specific soil property is high the local spectral models tend to have a higher R^2 and eventually higher RPD, but also a higher RMSE as observed by Kuang and Mouazen (2011). However, it remains unclear what an optimal variability of a soil property in a project of local extent would be to achieve a high measurement accuracy in absolute values (low RMSE) but also relative to the range of the data (RPD).~~

~~Besides dataset variability, the~~The number of samples is crucial for local models because~~,~~ often, only a limited number of samples with reference laboratory data are available. ~~It has been shown~~Kuang and Mouazen (2012) showed that local models improve with increasing numbers of calibration samples and that a sample size of at least 50 provides accurate prediction models~~ (Kuang and Mouazen, 2012).~~. Some studies thus combined multiple target sites and develop a general model by combining all the local datasets to reach a larger sample size and potentially better model performance (Kuang and Mouazen, 2011; Singh et al., 2022). In these studies, the general model showed an intermediate performance, and the general prediction error was between the best and the poorest performing local model. However, these studies only calculated the overall prediction error of the general model and therefore it is not clear if the prediction on target sites with poorly performing local models could be improved by applying a general model.

For vis–NIR spectroscopy application at local scales, it is therefore very difficult to estimate the measurement accuracy for the predicted samples beforehand. This uncertainty is probably the main reason that hampers the application of vis–NIR spectroscopy because researchers prefer to rely on conventional lab measurements with a smaller sample size (and smaller

spatial resolution) where the measurement accuracy is known before sampling and measurements are conducted. Applying spectroscopy at field or farm scale thus bears the risk that the measurement accuracy (RMSE) may be beyond the tolerable threshold, which might then question ~~the~~a whole project. Thus, in this paper, we analyze the performance of field-specific (local) spectral models of a field experiment conducted in six fields in Eastern Switzerland and that of a general model combining the data from all six fields to ascertain their influencing factors. We ask:

1. To what extent do the prediction errors of local spectral models differ from the lab measurement error?
2. Does a general model that includes several target sites improve the prediction on a target site with a poor local model performance?
3. ~~What is the optimal variability in the dataset of a local model to achieve a good model performance?~~
4. ~~Which~~Do field and soil characteristics (e. g. field size, soil texture, carbonate ~~concentration~~content, correlations of soil properties) of the target site influence the performance of spectral models?

~~Answering~~By answering these questions, we want to provide insights for estimations of prediction accuracies for vis–NIR studies at the local scale with the objective to support decision making during the development of a sampling design~~.~~

and planning of laboratory reference measurements for subsequent calibration modeling.

2 Methods

2.1 Datasets from a cover cropping experiment on six field sites

We used datasets from six fields (A, B, C, D, E, F) of a cover cropping experiment in the Canton of Thurgau, Eastern Switzerland (paper in preparation). The six fields were ~~maximally~~up to 13 km apart from one another and soil type was for all of them Eutric Cambisol that had developed on base moraine (Table 1). The aim of the study was to compare the influence of two different cover cropping regimes on short-term soil organic matter cycling. Each field had 39 differential GPS (dGPS) referenced sampling points in an unaligned sampling design. At each dGPS referenced point, soil was sampled three to four times in three depths (0-5, 5-10 and 10-20 cm) during one long cover cropping period (August 2019 to May 2020). Fields A, B, C and D had four sampling times resulting in 468 samples per field. Fields E and F had three sampling times resulting in 351 samples per field. All samples were dried at 40° C to constant weight (around 72 h) and then gently crushed and sieved to 2 mm. For the total sample size of 2574 samples, soil properties were estimated using vis–NIR soil spectroscopy, whereas 386 samples were analyzed conventionally by wet chemistry for subsequent calibration modeling. These 386 samples for laboratory analysis were selected for each field separately using the Kennard-Stones algorithm (Kennard and Stone, 1969) to ensure a cover of the whole spectral variability. Thereby, the Kennard-Stones algorithm was run with ~~2~~two to ~~7~~seven principal components and the number of principal components was chosen that covered at least 99 % of the spectral variance and provided a reference sample selection that well represented the different sampling times, soil depths and spatial distribution. The laboratory analysis comprised soil organic C (SOC), total C, total N, permanganate oxidizable C (POXC) also ~~referred as~~called active C~~,~~ and pH.

2.2 Chemical soil analyses and its accuracy

Total C and N concentrations were measured on a ground aliquot by dry combustion (vario MICRO tube, Elementar, Germany). Inorganic C was analyzed for each sample in triplicates through the dissolution of carbonate in a Scheibler-apparatus with 10 % HCl solution and the measurement of the evolved CO2 volume. SOC was then calculated as the difference between total C and the mean of the three measurements for inorganic C. POXC was measured according to the Protocol of Weil et al. (2003) with the adaption of Lucas and Weil (2012). In brief, 2.0 mL of 0.2 M $KMnO_4$ were added to 2.5 g of soil and after a reaction time of ~~totally~~ 10 minutes, the absorption of the liquid was measured at 550 nm with a Spectrophotometer (UV-1800, Shimadzu ~~corporation~~Corporation, Japan). Measurement of pH was done in a 0.01 M $CaCl_2$ solution. ~~Mean soil texture was measured for each field on a representative composite sample with the improved integral suspension pressure method (ISP+; Durner and Iden (2021)) on a PARIO Plus Soil Particle Analyzer (METER Group, Germany/USA).~~

To estimate the lab measurement error, we took three samples per field (in total 18) where we conducted the measurements for total C, total N, POXC~~,~~ and pH in triplicates to calculate a standard deviation. ~~To estimate~~We estimated the lab measurement error ~~of~~for SOC ~~we took the sum of~~(σ_{SOC}) according to Equation 1:

$$\sigma_{SOC} = \sqrt{\sigma_{Total\ C}{}^2 + \sigma_{Inorganic\ C}{}^2} \quad (1)$$

Where $\sigma_{Total\ C}$ is the standard deviation of the total C measurement ~~with~~and $\sigma_{Inorganic\ C}$ the standard error of the inorganic C measurement because inorganic C measurement were for all samples done in triplicates. The measurement error of all 18 triplicates were then averaged to obtain the overall lab measurement error for a soil property.

To characterize the spatial variability of soil texture in the field, we measured grain size on 20 samples per field (every second sampling point in 10-20 cm soil depth). Organic matter in the samples was oxidized with hydrogen peroxide (H2O2) and then grain size was measured with laser-diffraction analysis (LDA) after dispersion of the sample (22 mM socium carbonate and 18 mM sodium hexaphosphate) using a Mastersizer 2000 (Malvern Panalytical, UK). Since the LDA is underestimating the clay content compared to the standard grain size methods (Taubner et al., 2009), we measured one composite sample per field with the improved integral suspension pressure method (ISP+; Durner and Iden (2021)) on a PARIO Plus Soil Particle Analyzer (METER Group, Germany/USA). We rescaled the mean sand, silt and clay content of the LDA-data to the mean of the IPS+ method, while keeping the coefficient of variation constant (See Table S3 in the supplementary material).

2.3 Spectral measurement and pre-processing of spectra

All samples were measured with a vis–NIR spectrometer (ASD FieldSpec 4 Hi-Res, Malvern Panalytical, USA) with a sampling interval of 1.4 nm from 350-1000 nm and 1.1 nm from 1000-2500 nm. The device then provides a reflectance spectrum with a resolution of 1 nm and 2151 wavelengths. Measurements were done with a contact probe, containing an internal halogen bulb, which was in a fixed position and soil samples, placed in a petri dish of 1.5 cm height and 3 cm diameter, were lifted with a laboratory scissor jack until close contact with the probe to ensure a stable measurement position. For each sample, five petri dishes were filled to provide five replicate spectra per sample. Each of these five replicates consists of 30 internal repetitive scans that were automatically averaged by the device internal RS3 software. Between samples, the contact probe was carefully cleaned with water and ethanol. After the 5 replicates of a sample, the calibration of the spectrometer was checked with a 100 % reflectance white reference panel (Spectralon, 12×12 cm, Labsphere, USA). The infrared data of each sample was kept in two versions, once as reflectance spectra, as provided by the spectrometer, and once as absorbance spectra using the log(1/reflectance) transformation. Several pre-processing options and their combination were tested on both, the reflectance and the absorbance spectra: a) resampling of the spectra in an interval from 1 to 6 nm, b) cutting of the beginning (350-400 nm) or the end (2450-2500 nm) of the spectra c) first or second order derivative d) Savitzky-Golay (SG) smoothing in a third order polynomial ~~order of 3~~ with window sizes ranging from 5 to 51, e) gap segment derivative (GSD) with window width between 5 and 51 and segment size between 1 and 21, f) standard normal variate (SNV) combined with GSD, g) SG smoothing combined with multiplicative scatter correction (MSC). All applied pre-processing techniques are frequently used in soil spectroscopy and well described in Ellinger et al. (2019). The pre-processing techniques from a) to g) led to around 100

meaningful combinations that were tested in model building and the final pre-processing option was selected based on the ~~lowest~~smallest RMSE.

2.4 Development and evaluation of field-specific local models

We used for all 30 local models (6 fields x 5 properties) a PLSR modelling approach (Wold et al., 1983). ~~Since all soil properties showed a limited skewness (see Table S1 in the supplementary material) that was always in the range of $-2 \leq$ skew ≤ 2 which was proposed as acceptable to assume a normal univariate distribution (George and Mallery, 2010), we consider the application of PLSR appropriate, especially since it is robust to minor deviations from a normal distribution (Goodhue et al., 2012).~~

Model performance was assessed using the statistics of the hold-out folds of each five times repeated five-fold cross-validation because it was evaluated as a robust method for smaller datasets (Kuhn and Johnson, 2013; Molinaro et al., 2005). To avoid model overfitting, we set the maximum of latent variables in the PLSR model to 12. For each number of latent variables (1, 2, …., 12) the dataset was five times randomly split into five folds of which four were used for model training and the remaining fold was held out and used for model validation. The RMSE (Eq. ~~1~~2) of the hold-out samples was averaged among the five repeats resulting in a cross-validated RMSE per number of latent variables. The final number of latent variables was then chosen according to the "one standard error rule" which means that instead of directly choosing the number of latent variables with ~~lowest~~smallest mean RMSE, the most parsimonious (less latent variables) model within one standard error of the mean RMSE of the optimal model was selected (Hastie et al., 2017). The "one standard error rule" was also applied during optimization of pre-processing to avoid model overfitting. The final model was trained using all training data with optimized number of latent variables.

A proper validation of a spectral model is very crucial and particularly important in this study where soil was repeatedly sampled in different depths at the same GPS point. To analyze the correlation among the samples and define a grouping factor for the cross-validation, we calculated the mean Euclidean distances between all samples and compared it with the mean distance 1) between samples at the same GPS point but different depths, 2) between samples at the same point and depth but different sampling times and 3) between samples at the same point but different depth and sampling times (Fig. S1 in the supplementary material). Thereby, we ~~have seen~~observed that the soil samples from the three different soil depths sampled at the same GPS-point at the same sampling time had a substantially lower mean Euclidean distance compared to the overall mean. Consequently, we grouped the samples from the same GPS point at the same sampling time and kept them in the same fold to avoid a too ~~optimist~~optimistic model evaluation during cross-validation.

Since we used a cross-validation approach on the field scale, all models showed a very small bias~~.~~ (see Table 2). We therefore do not discuss the bias in this paper and focus on R^2, RMSE and RPD (Eq. ~~2~~3) for the evaluation and comparison of different models. RMSE was calculated according to Equation ~~1~~2 where $\hat{y}_i$ is the prediction of the spectral model for sample *i* and y_i the actual measured value for the same sample in the laboratory.

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

RPD compares the RMSE with the standard deviation (SD, Eq. 3) of the data:

$$RPD = \frac{SD}{RMSE}\,, \quad (2)$$

For all model performance parameters (R^2, RMSE and RPD) of the cross-validation, we calculated the uncertainty with the standard deviation of the prediction of the hold-out folds across the five repetitions.

To classify the model performance, we combined the RPD based classification of Chang et al. (2001) and Zhang et al. (2018). We considered spectral models with RPD < 1.4 as poor, models with RPD between 1.4 and 2 as approximate, models with RPD between 2 and 3 as accurate and models with RPD > 3 as excellent. Even though in spectroscopy project of a local extent the RMSE is probably the most important model performance parameter, RPD is the best parameter to compare models of different scales. Model metrics (R^2, RMSE and RPD) mentioned in the text are based on the cross-validation and metrics for the model calibration in Table 2 are specifically labelled as R^2_{cal}, $RMSE_{cal}$, RPD_{cal}.

2.5 Development and evaluation of general models

In addition to the field-specific local models, we built general models for the five soil properties that included all reference samples (n = 386) of the six fields. Even though for this sample size an independent test-set would be more suitable than a cross-validation approach, we evaluated the model performance using the hold-out samples in the five times repeated 10-fold cross-validation, keeping, as for the local models, samples from same GPS point and same sampling time in the same fold. The first reason for not using an independent validation set is that the modelling approach of the general model should be similar to the one of the local models to make them comparable. The second reason is that a representative split of the dataset into a calibration and a validation set according to the spectral variability would not result in equal number of samples per field in the validation set. Conversely, if we selected an equal sample size per field for the validation set, we would not have been able to cover the entire spectral variability. Evaluating the general models with hold-out samples of the cross-validation allowed us to calculate not only the RMSE over all samples but also the RMSE for the samples of each field individually and compare it. These field specific RMSE of the general model could then be compared with the RMSE achieved byof the local models. Since the only purpose of the general models was to increase modelling efficiency for a specific combined dataset, we did not group the samples according to fields during cross-validation because the same share of samples from the same field would also be in the prediction dataset. For the general models we cannot indicate uncertainties on a field-specific level since the folds did not always contain the same number of samples per field.

2.6 Model interpretation

To interpret spectral models, it is crucial to find relevant spectral features that are consistently important for a certain soil property. To identify the most important wavelength ranges in the final chosen models, we used the variable importance in

projection (VIP) method first published by Wold et al. (1993) and evaluated by Chong and Jun (2005). The VIP method can deal with multicollinearity and is therefore suitable for the interpretation of spectral models ~~(~~as it was for example applied by Baumann et al~~.~~, (2021). Wavelengths that have an above-average impact on the model have a VIP score above 1. We classified spectral ranges in groups of VIP scores between 1 and 1.5, 1.5 and 2 as well as VIP scores above 2.

2.7 Assessment of site characteristics influencing model performance

~~The 35 developed spectral models differed in their performance. Therefore we tested various factors that might explain under what field and soil characteristics the application of vis–NIR spectroscopy is most beneficial at the local scale. We focused on the relationship between variability of a soil property in a dataset and model performance. We looked at the coefficient of variation (CV), also known as the relative standard deviation, of the modeled soil property and examined its relationship with R^2, RMSE and RPD. Thereby, we expressed RMSE in percentages of the mean (PRMSE) to make all models comparable. We used linear regressions to fit the PRMSE and the RPD with the CV. Since R^2 cannot be higher than one, we fitted a two-parameter Weibull function (Seber and Wild, 2004) to the CV. For each regression we calculated the 95 % confidence and the 95 % prediction interval. In the discussion part, we compare the significant regressions between CV and model metrics found in our datasets with model performance in literature. Since the 95 % prediction interval for RPD independence of CV was quite high, we only discuss the found regressions for R^2 and PRMSE with the performance of local models in the literature. Furthermore, we fitted a linear regression for RPD in the range $0 < CV < 44$ % but we expect that with higher CV, RPD might also show saturation and therefore an extrapolation of our fitted linear regression would not make sense. In addition to the variability in the dataset, we analyzed the influence of mean carbonate concentrations, soil texture and field size on the model metrics but only the mean carbonate concentration showed effects and is therefore presented in the results.~~

To understand the reasons for the varying performance of the 35 developed spectral models, we studied the influence of various site characteristics on the models. To do so, we correlated the model performance parameters (R^2, RPD and RMSE) with field size, soil texture, carbonate content and with the correlation coefficients between SOC and total N in the dataset. With six local datasets as independent variables it is hardly possible to apply statistical tests and therefore, we relied on visible inspection. For the identified site characteristics that had a clear visual influence on model performance (carbonate content, correlation coefficient between SOC and N and variability in clay content) we looked for possible explanations in the spectral features. Thereby, we relied on the VIP-analysis of the trained models, on the correlation coefficients between soil properties with spectral variables and on the correlation matrices between target variables.

2.8 Data organization

All analysis were performed in R version 4.0.3 (R Core Team, 2020). The spectral datasets were analyzed using the R-package *simplerspec* version 0.2.0 (Baumann, 2019) in combination with the packages *prospectr* version 0.2.1 (Stevens and Ramirez-Lopez, 2020) and *caret* version 6.0-86 (Kuhn, 2020).

hat formatiert: Schriftart: Kursiv

hat formatiert: Schriftart: Kursiv

hat formatiert: Schriftart: Kursiv

3 Results

3.1 Description of the datasets

A comparison of the data distribution between the six different fields can be seen in Fig. 1 and the corresponding statistics in Table S1 in the supplementary material. The means for SOC, ~~Total~~total N and POXC differed between the six fields, but the distribution was relatively similar for these three soil properties. The density functions for total C and pH were highly influenced by the spatial distribution of carbonate in the soil: Fields B, D and E contain samples with and without carbonate resulting in a broad distribution for both, total C~~,~~ and pH. All soil samples of fields A and C contained carbonate in varying concentrations resulting in a broad distribution for total C but narrow distribution for pH. Field F showed high and only slightly varying carbonate ~~concentrations~~content and therefore a very narrow distribution for total C and pH. Field C had highest mean clay content and field A highest mean sand content whereas field F showed highest variability in soil texture.

3.2 Performance of spectral models

Based on RPD, 13 out of 30 local models showed an excellent performance (RPD > 3), ~~12~~11 models an accurate performance (RPD > 2), ~~four~~five models an approximate (RPD > 1.4) and one model a poor performance (RPD > 1.4; Table 2). The ~~five~~six models without accurate performance were ~~from field A (~~SOC, POXC and pH~~)~~ on field A and ~~field~~ F ~~(SOC and pH).~~.
However, the RMSE of the local models for pH of fields A (0.~~07~~08 ± 0.02) and F (0.~~03~~04 ± 0.01) were ~~lower~~similar or smaller than the RMSE of the other three local models (between 0.~~13~~08 ± 0.02 and 0.~~18~~19 ± 0.03) whose performance was classified as accurate. Differently, the local models for SOC on fields A and F with only approximate performance showed a higher RMSE (~~1.88~~2.43 ± 0.55 and 2.~~42~~00 ± 0.38 g kg^{-1}) than the other accurately performing local models for SOC (between 1.~~05~~07 ± 0.19 and 1.~~58~~59 ± 0.28 g kg^{-1}). The five general models showed all an accurate to excellent performance with RPD ranging from 2.~~42~~60 ± 0.43 to ~~3.97~~4.16 ± 0.47.

3.3 Influence of pre-processing on spectral variability

~~There was no pre-processing combination that proved to be suitable for all models of the same field (Fig. S2 in the supplementary material) or of the same soil property (data not shown). Nevertheless, for all 35 models, pre-processing was necessary and models could be improved compared to the raw spectra.~~For all 35 models, pre-processing improved the models compared to the raw spectra (see an example of pre-processing optimization for total C in Table S2 in the supplementary material). Although pre-processing was necessary for all models, we highlight that several pre-processing options performed similarly well within one standard deviation, and the differences in RMSE were often relatively small (see Table S2 in the supplementary material). Figure S2 in the supplementary material gives an overview of the best performing pre-processing techniques. Most times, the first or second order derivatives improved the models substantially. Most models performed best when the spectra were reduced to every third wavelength and models based on absorbance were a bit more frequently used than models based on reflectance. The combined application of SG filter and MSC was the most successful pre-processing

while a single SG filter, GSD and SNV in combination with GSD were of minor importance. Cutting of the beginning (350-400) or end of the spectra (2450-2500) sometimes improved the model performance but since most pre-processing steps reduce the beginning and end of the spectra, it was not possible to evaluate the cutting. Similarly, it was not possible to evaluate the window width chosen in the SG-filter because there is an interference with the resampling interval. A detailed list of the selected pre-processing options of the final models and the corresponding metrics for model performance can be found in Table 2.

The sensitivity of model performance to pre-processing can be visualized with the biplots of principal component analysis (PCA). Figure 2 shows the first three biplots of the raw spectra and the spectra that were pre-processed according to the general models of the five soil properties. The raw spectra had a very high share of the explained variance (96.8%) on the first principal component but hardly any groups according to fields could be observed with the first two principal components. All pre-processing options used for the general models decreased the explained variance on the first principal component (32.5 to 39. 6 %) and a grouping according to fields could already be seen in the biplot of the first two principal components. Thereby, especially field F with highest carbonate ~~concentration~~content and field C with highest clay content often showed clear groups. Nevertheless, in the pre-processing for pH, field E with the highest pH variability shows a clear group in the first biplot and the pH variability is well represented with the first PC.

3.4 Comparison of general models with local models and lab measurement error

The overall cross-validated model metrics of the general model (black filled circle in Fig. 3) indicated over all fields for all soil properties a good performance, but the field-specific model evaluation showed distinct differences among fields. The field-~~specifically calculated~~specific R^2 of the general models of fields B, C, D and E was similar to the R^2 of the local model for SOC, total C, total N and POXC (only a slight slope in Fig. 3). For pH, only field C, D and ~~D~~E showed similar R^2 in the local and general model while ~~all other~~ fields A, B and F showed clearly higher R^2 in the local model. On the other hand, field F had clearly lower R^2 in the general model than in the local model ~~(strong negative slope)~~ for all ~~five~~ soil properties except POXC. For field A, R^2 was similar between the local and the general model for SOC, total C and POXC but ~~for the other four properties~~ clearly lower for total N and pH in the general model.

The field-~~specifically calculated~~specific RPD of the general model was across all soil properties on average ~~27~~31 % lower compared to the local models (Fig. 3). All property-field combinations of fields B, C, D and F showed at least an approximate (RPD > 1.4) performance in the general models, whereas the ~~eight~~seven poorly (RPD < 1.4) performing property field combinations were all from fields A and F. It can therefore ~~clearly~~ be concluded that the general models could not improve the low performing local models.

Field-~~specifically calculated~~specific RMSE of the general models was on average ~~49~~47 % higher compared to the local models. However, there were substantial differences between the different fields. For Field ~~A and~~ F ~~showed~~, the ~~highest increases of~~ field-specific RMSE in the general ~~model~~ models for SOC, total C, total N and pH (2.58 g kg^{-1}, 0.17 g kg^{-1} and 0.09,) were much higher compared to the local model ~~for SOC, total C, and total N (strong slopes in~~(2.00 ± 0.38 g kg^{-1}, 0.09 ± 0.02 g kg^{-1}

and 0.04 ± 0.01, respectively; Fig. 3). Similarly, for total N and pH, field A had much higher RMSE in the general model 0.22 g kg^{-1} and 0.14) than in the local model (0.14 ± 0.03 and 0.08 ± 0.02). On the other hand, fieldfields C and E showed quite similar RMSE in the local and in the general model for all soil properties except total C.

The RMSE of the best local models waswere close to the overall lab measurement errorerrors for SOC, total C and total N, a bit higher for POXCpH and substantially higher for pHPOXC (Fig. 3).3). The RMSE of SOC on field B (1.26 ± 0.36 g kg^{-1}) and D (1.07 ± 0.19 g kg^{-1}) were within the standard deviation of the lab measurement error for SOC (1.01 ± 0.40 g kg^{-1}). The overall lab measurement error for SOC (1.3 g kg^{-1}) was calculated as the sum offrom the measurement error for total C and inorganic C and therefore the RMSE for SOC on fields B and D with only little inorganic C (fields B and D) was lower than the overall lab measurement error. However, the RMSE of SOC on field B (1.20 g kg^{-1}) and D (1.05 g kg^{-1}) with little inorganic C was clearly above, the lab measurement error offor total C which would for these two fields(0.83 ± 0.25 g kg^{-1}) might be the better lab reference error (0.85 g kg^{-1}). The cross-validated. However, the RMSE of the local spectral models of all fields exceeded the overall lab measurement errors between factor 0.81.1 and 1.92.4 for SOC, 1.46 and 3.02 for total C, 1.3 and 2.0 for total N, 2.13 and 4.03 for POXC and between 2.83.4 and 16.517.8 for pH. The field-specifically calculatedspecific RMSE of the general model exceeded the overall lab measurement error between factor 1.03 and 2.13 for SOC, 2.42 and 4.95.2 for total C, 1.5 and 3.32 for total N, 2.78 and 4.96 for POXC and between 7.98.3 and 22.419.9 for pH.

The VIP scores (Fig. 4) show that the most important wavelengths were very dataset- specific. It can clearly be seen that on field B and differed forto a lower extent on field F, the same property substantially from one anotherwavelengths were important in all soil properties related to soil organic matter (SOC, total C, total N and POXC), whereas on the other fields the VIP pattern of the different properties were more distinct from each other. However, for all the analyzed soil properties the wavelength ranges between 400 and 750 nm (visible) as well as 1800 and 2450 nm were most important while the range in between was of lower importance. Nevertheless, some models had VIP scores above 2 in the range between 750 and 1800 nm. Prediction performance in terms of RMSE and RPD of total C of fields E and F was particularly lower in the general model than in the local model (Fig. 3). This finding can be explained with the VIP analysis (Fig. 4) that showed for the general model the most important wavelength range between 2150 and 2450 nm but for the local models of fields E and F in the range of 500 to 1020 nm. The local model for total N of field F showed very high VIP scores (>2) in a small specific range between 2345 and 2369 nm but these wavelengths were not important in the general model for total N (Fig. 4), which resulted in a much lower prediction accuracy of total N for field F in the general model compared to the local model.

3.5 Influence of dataset variability on model performance

The three model performance parameters, R^2, RPD and PRMSE increased with increasing CV in the dataset (Fig. 5). For the relationship of R^2 and CV both parameters of the Weibull function were significant ($p < 0.05$), and the function shows that above a CV of 10 % the R^2 of the spectral models only show minor increases with increasing CV. The PRMSE showed a relatively strong linear relationship with CV ($R^2 = 0.61$, $p < 0.001$) and the regression slope was higher than the slope of the regression for RPD which showed a weaker relationship with CV ($R^2 = 0.40$, $p < 0.01$). The reason is that RPD accounts for

~~the variability in the dataset because by its definition the increasing RMSE is divided by an increasing standard deviation as the CV increases, which makes RPD less influenced by the CV in the dataset. Nevertheless, we found a significant increase of RPD with increasing CV indicating a better performance of spectral models with higher variability in the dataset at least up to a CV of 44 %. The optimal variability in a spectroscopy project of local extent might mainly be limited by the PRMSE that increases relatively fast and consistent with increasing CV in the dataset. The 95 % prediction intervals in Fig. 5 show that a dataset with CV of 15 % is sufficient to reach an R^2 in the spectral models above 0.65 and keeping the PRMSE at the same time between 1.8 and 8.8 %. For RPD the 95 % prediction interval for a new dataset with 15 % CV is very wide ranging between 1.2 and 4.4. An increase of CV in the dataset provides slight increases of R^2 and RPD but PRMSE increases faster and at a CV of 40 % the PRMSE would range between 6.9 and 14.5 % with 95 % prediction probability (Fig. 5). For many projects this might still be tolerable for others it might already be too high.~~

~~3.6~~3.5 Site characteristics influencing model performance

~~Besides the variability for a specific property in the dataset, the performance of the local field-specific models for SOC were influenced by the mean inorganic C concentration. The local models of fields A and F with highest carbonate concentrations showed, aggregated over all five soil properties, lower R^2 and RPD than the other fields (Fig. 6).~~

We found an order of model performance with respect to R^2 and RPD in dependence of mean carbonate content, correlation coefficient between SOC and total N as well as coefficient of variation in clay content (Fig. 5). Fields A and F that showed lower model performance in terms of RPD had higher carbonate content, lower correlation coefficient between SOC and total N and higher variability in soil texture (compare also with density plots in Fig. 1). However, in absolute prediction performance (RMSE) we only found a clear effect for SOC (Fig. 6) and not for the other properties (Fig. S3 in the supplementary material). We did not observe an influence of field size absolute contents of sand, silt and clay or variability of carbonate content on model performance (see Fig. S4 in the supplementary material).

The influence of carbonate content on model performance of SOC is illustrated by plotting at each wavelength the correlation coefficients between pre-processed spectral variables and inorganic C as well as SOC content (Fig. 7). The correlation between SOC and spectral variables was higher on fields B, D and E than on fields A, C and F, which also explains the better model performance. On field A, SOC and carbonate content show a very similar correlation with spectral variables across the whole vis–NIR range, which makes it difficult to distinguish organic and inorganic C on field A resulting in an excellent performance of total C but much lower performance for SOC (see Table 2). Even though the correlation between spectral variables and SOC content on field C was lower than on other fields (B, D and E), the very different correlation pattern of carbonate content still resulted in good model performance for SOC. Especially the ranges between 600 and 1200 nm and the peaks at 1680 nm and 2240 nm showed different spectral features for SOC and carbonate which corresponds to the high VIP scores at that wavelengths for the SOC model on field C. On field F, correlations for both carbonate content and SOC were relatively weak, whereby carbonate content showed stronger correlations with spectral variables which probably masked the spectral features of SOC resulting, as for field A, in a better model for total C than SOC.

The better model performance on field B, D and E compared to fields A, C and F also coincided with higher correlation between SOC and total N (Fig. 5). In general, correlation coefficient between target variables tended to be higher on field B, D and E compared to fields A, C and F (see Fig. 8 as example and all correlation matrices in Fig. S5 in the supplementary material).

4 Discussion

4.1 Performance of local spectral models

Most of the developed local models showed an accurate performance and confirm the suitability of vis–NIR spectroscopy in projects of local or single plot extent. The performance (based on RPD) of the two models for pH on field A and F, which were classified as only approximate or even poor, respectively, can be explained by the low variability of pH in these datasets (see Fig. 1) and is supported by the fact that these two models ~~have~~had the ~~lowest~~smallest RMSE values for pH (Fig. 3). This explanation does not hold for the other three local models that were also classified as only approximate because SOC and POXC on field A as well as SOC on field F showed a similar variability as on the other fields (Fig. ~~1). We assume that the lower performance of these three models on field A and F is caused by the parent material containing a high carbonate concentration (see section 4.6). Nevertheless, in all local models the RMSE was mostly comparable to the lab measurement error which was also the case in the field scale models with MIR spectroscopy of Greenberg et al. (2022).~~ 1), but higher RMSE. However, considering the mean SOC concentration on fields A (22.4 ± 3.7 g kg^{-1}) and F (28.6 ± 2.7 g kg^{-1}) as well as the lab measurement error (1.00 ± 0.04 g kg^{-1}), we argue that the RMSE on fields A (2.43 ± 0.55 g kg^{-1}) and F (2.00 ± 0.38 g kg^{-1}) are probably for many research projects still acceptable, especially when taking into account that a higher sample size can be analyzed for the same costs.

In agreement with literature (Soriano-Disla et al., 2014), primary properties with a direct impact in the vis–NIR range like SOC, total C, total N~~,~~ and POXC showed a RMSE that was closer to the lab measurement error ~~(maximum four times higher).~~. On the other hand, pH has only an indirect impact on the spectra and thus showed a much higher RMSE compared to the lab measurement error ~~(maximum 16.4 times higher).~~but the RMSE for pH in the local models (between 0.04 ± 0.01 and 0.19 ± 0.03) is probably small enough for most research purposes.

4.2 Comparison of general models with local models

The general models could not improve the prediction of low performing local models. This finding is especially interesting because in this study the general model was built with datasets of six fields that were spatially close to one another (maximal distance of 13 km), had the same soil type and the same parent material. However, the base moraine as a parent material can be variable which we mainly observed in different soil texture and carbonate ~~concentration~~content but also in the high spectral variability (see PCA biplots in Fig. 2). In this sense, we confirm the conclusions of Seidel et al. (2019) and Ng et al. (2022) who suggested that the best solution is always to develop a local model if enough samples (> 30) are available. This conclusion

is supported in this study by the quite distinctive pattern of VIP scores between the different models (Fig. 4). The overall picture shows that the wavelengths between 2000 and 2450 nm followed by the visible range between 400 and 700 nm were most important for prediction of the investigated properties, which is in agreement with the literature (Munnaf and Mouazen, 2022; Soriano-Disla et al., 2014). Nevertheless, each local model has distinct and site-specific features that could not be attributed to specific soil characteristics ~~but obviously were~~while being important for the model development. The development of general models where different locations are aggregated in one dataset can save costs because the number of lab analysis per location can be reduced and less work is required for model building. Depending on the research purpose and the required measurement accuracy, the development of general models can be a very suitable and cost-effective approach. Nevertheless, this study showed that some fields (A and F) can show a poor performance in general models, hence it is crucial to consider what locations or datasets are being combined.

4.3 Pre-processing

The selection of the optimal pre-processing scheme was crucial for model performance but strongly dependent on the dataset. Often MSC was the best performing pre-processing option, which was confirmed in some studies (Cambule et al., 2012; Liu et al., 2019) but disproved in others (Knox et al., 2015; Riefolo et al., 2020). We therefore highly recommend considering MSC as a pre-processing option in spectral modelling but at the same time agree with Barra et al. (2021) that there is no general pre-processing solution that works for all datasets. The principal component analysis with the combined dataset of all fields (Fig. 2) illustrates this finding by the different grouping of individual field datasets due to different pre-processing. This leads to the conclusion that studies that did not optimize the pre-processing scheme for every soil property separately did eventually not make full use of the spectroscopy ~~potential (see also examples in Table 3),~~ which has been shown by other studies as well (Alomar et al., 2021; Rodriguez-Febereiro et al., 2022; Singh et al., 2022). Nevertheless, the property-specific optimization of spectral pre-processing is a tedious process and constrains the fast and cost-effective application of vis–NIR spectroscopy, but some progress has recently been made ~~simultaneously with our study~~ by Mishra et al. (2022).

~~4.4 Influence of dataset variability on model performance~~

~~With 35 spectral models we found significant relationships between CV in the dataset and model performance. Hereby it must be stated that the built regressions only have a validity for local datasets with relatively similar mineralogy. A dataset could potentially also have a very low CV for a certain soil property but could contain soil samples from very distinct regions with very different mineralogical background. Nevertheless, we also included the general models in that regression knowing that the different fields differ in some respects but are not fundamentally different from one another since they are in the same region. The combined dataset used for the general model contains a higher variation of soil samples but does not necessarily have a higher CV for a certain soil property. Nevertheless, it can be seen in Fig. 5 that the general models (black symbols) fit quite well into the regression for R^2, PRMSE and RPD. Dataset variability has always been an issue in vis–NIR spectroscopy application and many authors stated that "enough" variability is necessary (Alomar et al., 2021; Perez-Bejarano and Guerrero,~~

2018) without directly quantifying it. Based on the 95 % prediction interval of the found regression we suggest a minimum CV of about 15 % for a successful spectral model. The question of an optimal maximum variability in a local dataset for spectral modelling has so far not really been addressed. In many studies, a better model performance is ascribed to a higher variability in a dataset (Kawamura et al., 2021; Nawar and Mouazen, 2019) without considering the increasing prediction error. The observation that a higher dataset variability leads to a higher RMSE in local spectral model was also made by Kuang and Mouazen (2011) but so far, no quantitative relationship was provided. Our results show clearly that a higher variability in a dataset linearly increases the prediction error (PRMSE). Therefore a dataset for a local vis–NIR spectral model has an optimal maximum variability that depends on the tolerable prediction error in a research project.

4.5 Comparison of found regressions with performance of local models in literature

The comparison of our fitted regressions for R^2 and PRMSE with performance of local models in literature can be found in Fig. 7 where each number of a data point corresponds to a publication listed in Table 3. Thereby, we only plotted local models with a CV < 105 %. The R^2 found in literature lie often within the prediction interval for CV > 25 % but for lower CV our nonlinear regression seems to be too optimistic (Fig. 7). For the PRMSE, our linear regression result tends to be too optimistic over the whole range of CV but especially for datasets with CV > 60 %. The main reason for this difference in performance may be that we developed field-specific models covering an area of about 1 ha while most local models found in literature cover larger areas, which might not necessarily result in a higher CV for a certain soil property but in tendency increases the variability in the mineralogical background. However, the models that were performing the furthest outside of the 95 % prediction interval for R^2 belong all to study 19, followed by models from study 3, 2 and 20. For PRMSE, the models showing a far higher PRMSE than the upper 95 % prediction interval belong to studies 4, 19, 1, 7, 3 and 13. It can be seen that in these studies showing a relatively poor model performance, only study 1 and 7 optimized pre-processing for each soil property individually. Additionally, for properties with a low mean like in study 4 (mean SOC = 0.79 %), the high PRMSE is misleading because the absolute error was relatively low. The suggested regression for PRMSE can serve as a rough estimate for model performance but with a higher CV the PRMSE shows higher variability. A vivid example is study 7 (Singh et al., 2022), with four field-specific models and one general model for SOC, that shows R^2 values that lie all within the 95 % prediction interval, but the PRMSE values vary around the suggested 95 % prediction interval. In this respect, we consider the 95 % prediction interval for R^2 as a good predictor for the expected R^2 of local spectral models and recommend a minimum CV of 15 % in a dataset for local models. The regression for PRMSE gives a rough estimate about the error that must be expected in a spectroscopy project of a local extent, but the variability is relatively high because it is strongly influenced by the overall mean in the dataset and other factors that remain unknown and are probably related to the mineralogical background.

4.64.4 Site characteristics influencing model performance

We found anhigher model performance on fields with low carbonate content, high correlations between soil properties and low variability in clay content. We want to discuss how these identified important field characteristics influence or mask spectral features.

4.4.1 Mean carbonate content

We found an influence of carbonate concentrationcontent with lowest performance of local spectral models on fields A and F. Similar observations were made by Amare et al.,. (2013) and McCartyMccarty et al. (2002) who argued that the absorbance bands of carbonate mask those of SOC. Looking at the correlation between spectral variables and inorganic C respectively SOC (Fig. 7) we can confirm this finding but have to add that on the local scale the absorption bands for carbonate and SOC varied substantially between different datasets. In this context, Reeves (2010), who showed that the spectrum of a soil sample varied greatly with its carbonate concentrationcontent, considered the prediction of SOC in soils with high carbonate concentrationcontent as one of the open questions in vis–NIR spectroscopy research. Even though the local models of fields A and F had the lowest performance among all local models, they still showed, with exception of pH, approximate results which, depending on the research question might still provide useful information. An important point missing in this discussion is the measurement accuracy of SOC in the laboratory, which is strongly influenced by the presence of carbonate and the method used (Goidts et al., 2009). If the soil samples contain carbonate, often two measurements must be conducted, and SOC is calculated as the difference between total C and inorganic C. Especially with a high carbonate concentrationcontent, the measurement error for the inorganic C concentrationcontent can be a substantial share of the SOC concentration.content. The higher lab measurement error with higher carbonate concentrationcontent can explain the lower model performance on soils with high carbonate concentrationcontent for SOC but not for the other four soil properties where model performance (in terms of RPD) still tended to be lower than on fields with little carbonate concentrationcontent (Fig. 65). This confirms the above-mentioned observation of spectral interference between inorganic C and organic matter and is additionally substantiated by the result that most properties of fields A and F showed a poor performance in the general models (Fig. 3). It is known that carbonate has many more defined peaks and less interferences with organic matter in the MIR than in the vis–NIR (Reeves, 2010), therefore). Therefore, datasets that combine soil samples with high and low carbonate concentrationscontent might better be predicted with MIR spectroscopy.

4.4.2 Correlations between target variables

Reflectance measured with vis–NIR spectroscopy is a combined effect of all constituents present in the soil sample (Stenberg et al., 2010) and through processing and modeling one tries to distinguish the absorption feature of one specific soil property from the other constituents of the sample. Apart from pH, all our target variables were closely related to soil organic matter which was therefore for this study the most important soil constituent influencing the absorption features. In case of high

correlations between target variables, that form part of soil organic matter, the modelling is easier because the same absorption features can be used for modelling the different properties which was the case for field B (see VIP analysis in Fig. 4). On the other hand, a low correlation between target variables makes it more difficult to attribute absorption features of organic matter to specific soil properties, which probably contributed to the lower model performance of fields A, C and F compared to fields B, D and E. The literature shows that different soil properties related to soil organic matter (e.g. SOC and total N) can show different absorption features in the vis–NIR range (Chang and Laird, 2002; Kusumo et al., 2019), which is also supported in our study (see VIP analysis in Fig. 4). However, we argue that prediction accuracy improves substantially if target variables related to soil organic matter are well correlated with each other, which was also hypothesized by Martin et al. (2002) in a one location field study.

4.4.3 Variability of clay content

Unlike Stenberg et al. (2010) and Heinze et al. (2013), we did not find a better model performance with increasing mean clay content in the dataset which might also be explained by the relatively small range in mean clay contents between 18 % (Field F) and 38 % (Field C). However, we observed that the fields A and F with lower model performance also showed a higher variability in soil texture (see density plots in Fig. 1). We hypothesize that this observation is mainly an effect of our sampling design and the specific agricultural management and therefore not generalizable. Clay as well as soil organic matter are claimed to be modelled with high success rate with vis–NIR spectroscopy since they have strong absorption features (Da Silva-Sangoi et al., 2022). Unfortunately, soil texture was measured on different samples than the reference dataset for the spectral modelling, so we cannot check for the correlation between soil texture and target variables. However, in this study the correlation may be relatively low for the following reason: We took samples in different depths (0-5, 5-10 and 10-20 cm) within the past tillage layer and therefore expect that the soil texture is homogenized across the sampling depth. Since all fields are now under organic reduced tillage management, the three soil layers show quite distinct soil organic matter content (see Fig. S6 in the supplementary material) but very probably similar soil texture. Therefore, a high (horizontal) variability in soil texture on a field (e. g. clay content) without strong correlation to organic matter could have added “noise” to the spectrum which worsened the prediction accuracy in our specific sampling design. Nevertheless, in untilled soils or more distinct depth segments a high variability in soil texture may not be a disadvantage in vis–NIR modeling because it might also be correlated with organic matter.

hat formatiert: Englisch (Vereinigtes Königreich)

5 Conclusion

This study investigated the impact of ~~various factors~~site characteristics on vis–NIR modeling performances and compared a local and a general modeling approach. Among the 35 ~~built,~~ models ~~30,~~ 29 performed accurate or even excellent whereby the RMSE was ~~maximally four times higher than~~ close to the lab measurement error ~~for all analyzed soil properties except pH.~~

~~We identified the variability and the~~ and achieved prediction accuracies are probably for many research purposes acceptable. The local models with lowest performance, were all from field A and F and we found three field characteristics in their datasets that interfered with model performance. Fields A and F had higher mean carbonate content ~~in a dataset as crucial factors that~~ , lower correlation between target soil properties and higher variability in soil texture compared to the other fields. The influence ~~the performance~~ of ~~local models. We found significant relationships between the CV in a dataset and the model performance~~soil texture variability was mainly an issue in ~~terms of R^2 and PRMSE. The lower 95 % prediction interval indicates that for a good R^2 (> 0.65) a CV of at minimum 15 % is required in a dataset. Assuming a tolerable prediction error of 12 %, the upper 95 % prediction interval for PRMSE would suggest a maximum CV of 30 % in the dataset. Since from our experience the main concern to use vis–NIR spectroscopy in a project of local extent is the unknown measurement accuracy, we recommend determining first a tolerable percental measurement error and then developing a~~this specific sampling design ~~according to the targeted variability in the dataset. Such a procedure requires an attentive stratification strategy for soil sampling according to land use, management, soil depth or parent material. Since the variability of~~whereas the influence of carbonate content and correlation between soil properties can probably be generalized due to observed spectral features and VIP analysis. Before starting a local vis–NIR project, testing for inorganic C content can be done relatively easily but it is almost impossible to know beforehand the correlations between different soil properties ~~can be very different, as e.g., on cropland SOC normally shows much higher variability than pH~~. One can only be aware of the correlation issue and consider potential gradients of soil properties while designing the sampling design which is probably more important and feasible in disturbed or agricultural soils than in natural undisturbed soils. In terms of efficiency in data collection, we ~~recommend to clearly prioritize the soil properties and select the one of most interest and optimize the sampling design according to that target property. Fields with high carbonate content (A and F) showed lower (but often still approximate) performance in the local models but a strongly decreased performance in the general models compared to fields with low carbonate concentration. We therefore~~ conclude that in ~~case of low carbonate concentration~~a region, several target sites ~~of a region~~(or agricultural fields) with low carbonate content can be combined ~~to~~in a general model with only a minor reduction in model performance. A general model for multiple target sites then also allows to reduce the number of wet chemistry analyses. Whether or not several target sites with high carbonate ~~concentration~~content can be combined in one general model using vis–NIR spectroscopy is a question that ~~is not answered yet and~~ requires further research. However, since carbonates show less interferences with organic matter in the MIR than in the vis–NIR spectral range, soil samples from sites with high carbonate ~~concentration~~content might be ~~the~~ better predicted with MIR spectroscopy. ~~Considering the mentioned recommendation~~Yet, the application of laboratory vis–NIR spectroscopy in projects of local extent provides the opportunity to increase~~, at low costs, the sample size drastically and increase, as desired,~~ the spatial or temporal resolution in ~~the~~a sampling design cost-effectively with only minor decreases in measurement accuracy.

References

Allory, V., Cambou, A., Moulin, P., Schwartz, C., Cannavo, P., Vidal-Beaudet, L., and Barthes, B. G., 2019, : Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions, Sci. Total Environ., v. 686, p. 764-773, https://doi.org/10.1016/j.scitotenv.2019.05.192, https://doi.org/10.1016/j.scitotenv.2019.05.192, 2019.

Alomar, S., Mireei, S. A., Hemmat, A., Masoumi, A. A., and Khademi, H., 2021, : Comparison of Vis/SWNIR and NIR spectrometers combined with different multivariate techniques for estimating soil fertility parameters of calcareous topsoil in an arid climate, Biosys. Eng., v. 201, p. 50-66, https://doi.org/10.1016/j.biosystemseng.2020.11.007, https://doi.org/10.1016/j.biosystemseng.2020.11.007, 2021.

Amare, T., Hergarten, C., Hurni, H., Wolfgramm, B., Yitaferu, B., and Selassie, Y. G., 2013, : Prediction of Soil Organic Carbon for Ethiopian Highlands Using Soil Spectroscopy, ISRN Soil Sci., v. 2013, p. 720589, https://doi.org/10.1155/2013/720589, https://doi.org/10.1155/2013/720589, 2013.

Angelopoulou, T., Balafoutis, A., Zalidis, G., and Bochtis, D., 2020, : From Laboratory to Proximal Sensing Spectroscopy for Soil Organic Carbon Estimation - A Review, Sustainability-Basel, v. 12, no. 2, https://doi.org/10.3390/su12020443, https://doi.org/10.3390/su12020443, 2020.

Barra, I., Haefele, S. M., Sakrabani, R., and Kebede, F., 2021, : Soil spectroscopy with the use of chemometrics, machine learning and pre-processing techniques in soil diagnosis: Recent advances - A review, Trac-Trend. Anal. Chem., v. 135, https://doi.org/10.1016/j.trac.2020.116166, https://doi.org/10.1016/j.trac.2020.116166, 2021.

Baumann, P.: philipp-baumann/simplerspec: Beta release simplerspec 0.1.0 for zenodo, Zenodo [code], https://doi.org/10.5281/zenodo.3303637, 2019.

Baumann, P., Lee, J., Frossard, E., Schönholzer, L. P., Diby, L., Hgaza, V. K., Kiba, D. I., Sila, A., Sheperd, K., and Six, J., 2021, : Estimation of soil properties with mid-infrared soil spectroscopy across yam production landscapes in West Africa: Soil, v. 7, no. 2, p. 717-731. https://doi.org/10.5194/soil-7-717-2021, Soil, 7, 717-731, https://doi.org/10.5194/soil-7-717-2021, 2021.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., and McBratney, A., 2010, Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy: Trac-Trend. Anal. Chem., v. 29, no. 9, p. 1073-1081. https://doi.org/10.1016/j.trac.2010.05.006.

Bellon-Maurel, V., and 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 Biol. Biochem., v. 43, no. 7, p. 1398-1410. https://doi.org/10.1016/j.soilbio.2011.02.019.

hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.

Breure, T. S., Prout, J. M., Haefele, S. M., Milne, A. E., Hannam, J. A., Moreno-Rojas, S., and Corstanje, R~~., 2022,~~.: Comparing the effect of different sample conditions and spectral libraries on the prediction accuracy of soil properties from near- and mid-infrared spectra at the field-scale~~:~~, Soil Till. Res., ~~v.~~ 215~~.~~, ~~https://doi.org/10.1016/j.still.2021.105196.~~https://doi.org/10.1016/j.still.2021.105196, 2022.

~~Brodsky, L., Vasat, R., Klement, A., Zadorova, T., and Jaksik, O., 2013, Uncertainty propagation in VNIR reflectance spectroscopy soil organic carbon mapping: Geoderma, v. 199, p. 54-63. https://doi.org/10.1016/j.geoderma.2012.11.006.~~

Brown, D. J~~., 2007,~~.: Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed~~:~~, Geoderma, ~~v.~~ 140, ~~no. 4, p.~~ 444-453~~.~~, ~~https://doi.org/10.1016/j.geoderma.2007.04.021.~~https://doi.org/10.1016/j.geoderma.2007.04.021, 2007.

Camargo, L. A., do Amaral, L. R., dos Reis, A. A., Brasco, T. L., and Magalhaes, P. S. G~~., 2022,~~.: Improving soil organic carbon mapping with a field-specific calibration approach through diffuse reflectance spectroscopy and machine learning algorithms~~:~~, Soil Use Manage., ~~v.~~ 38, ~~no. 1, p.~~ 292-303~~.~~, ~~https://doi.org/10.1111/sum.12775.~~https://doi.org/10.1111/sum.12775, 2022.

Cambule, A. H., Rossiter, D. G., Stoorvogel, J. J., and Smaling, E. M. A~~., 2012,~~.: Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique~~:~~, Geoderma, ~~v.~~ 183, ~~p.~~ 41-48~~.~~, ~~https://doi.org/10.1016/j.geoderma.2012.03.011.~~https://doi.org/10.1016/j.geoderma.2012.03.011, 2012.

Chang, C. W. and Laird, D. A.: Near-infrared reflectance spectroscopic analysis of soil C and N, Soil Sci., 167, 110-116, https://doi.org/10.1097/00010694-200202000-00003, 2002.

Chang, C. W., Laird, D. A., Mausbach, M. J., and Hurburgh, C. R~~., 2001,~~.: Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties~~:~~, Soil Sci. Soc. Am. J., ~~v.~~ 65, ~~no. 2, p.~~ 480-490~~.~~, ~~https://doi.org/10.2136/sssaj2001.652480x.~~https://doi.org/10.2136/sssaj2001.652480x, 2001.

Chong, I. G~~.,~~, and Jun, C. H~~., 2005,~~.: Performance of some variable selection methods when multicollinearity is present~~:~~, Chemom. Intell. Lab. Syst., ~~v.~~ 78, ~~no. 1-2, p.~~ 103-112~~.~~, ~~https://doi.org/10.1016/j.chemolab.2004.12.011.~~https://doi.org/10.1016/j.chemolab.2004.12.011, 2005.

da Silva-Sangoi, D. V., Horst, T. Z., Moura-Bueno, J. M., ~~Debaene, G., Niedzwiecki, J., Pecio, A., and Zurek, A., 2014, Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale: Geoderma, v. 214, p. 114-125. https://doi.org/10.1016/j.geoderma.2013.09.022.~~

Dalmolin, R. S. D., Sebem, E., Gebler, L., and Santos, M. D.: Soil organic matter and clay predictions by laboratory spectroscopy: Data spatial correlation, Geoderma Reg., 28, https://doi.org/10.1016/j.geodrs.2022.e00486, 2022.

Durner, W~~.,~~, and Iden, S. C~~., 2021,~~.: The improved integral suspension pressure method (ISP plus) for precise particle size analysis of soil and sedimentary materials~~:~~. Soil Till. Res., ~~v.~~ 213~~.~~, 105086, ~~https://doi.org/10.1016/j.still.2021.105086.~~https://doi.org/10.1016/j.still.2021.105086, 2021.

Ellinger, M., Merbach, I., Werban, U., and Liess, M~~., 2019,~~.: Error propagation in spectrometric functions of soil organic carbon~~:~~. Soil, ~~v.~~ 5, ~~no. 2, p.~~ 275-288~~.~~, ~~https://doi.org/10.5194/soil-5-275-2019.~~https://doi.org/10.5194/soil-5-275-2019, 2019.

~~George, D., and Mallery, P., 2010. SPSS for Windows step by step: a simple guide and reference, 18.0 update, Boston, Allyn & Bacon.~~

Goidts, E., Van Wesemael, B., and Crucifix, M~~., 2009,~~.: Magnitude and sources of uncertainties in soil organic carbon (SOC) stock assessments at various scales~~:~~. Eur. J. Soil Sci., ~~v.~~ 60, ~~no. 5, p.~~ 723-739~~.~~, ~~https://doi.org/doi:10.1111/j.1365-2389.2009.01157.x.~~https://doi.org/doi:10.1111/j.1365-2389.2009.01157.x, 2009.

~~Gomez, C., and Coulouma, G., 2018, Importance of the spatial extent for using soil properties estimated by laboratory VNIR/SWIR spectroscopy: Examples of the clay and calcium carbonate content: Geoderma, v. 330, p. 244-253. https://doi.org/10.1016/j.geoderma.2018.06.006.~~

~~Goodhue, D. L., Lewis, W., and Thompson, R., 2012. Does PLS have advantages for small sample size or non-normal data? Mis Quarterly v. 36, p. 981-1001. https://doi.org/10.2307/41703490.~~

Greenberg, I., Seidel, M., Vohland, M., Koch, H. J., and Ludwig, B~~., 2022,~~.: Performance of in situ vs laboratory mid-infrared soil spectroscopy using local and regional calibration strategies~~: Geoderma, v. 409. https://doi.org/10.1016/j.geoderma.2021.115614.~~, Geoderma, 409, https://doi.org/10.1016/j.geoderma.2021.115614, 2022.

Grunwald, S., Yu, C. R., and Xiong, X~~., 2018,~~.: Transferability and Scalability of Soil Total Carbon Prediction Models in Florida, USA~~:~~. Pedosphere, ~~v.~~ 28, ~~no. 6, p.~~ 856-872~~.~~, ~~https://doi.org/10.1016/s1002-0160(18)60048-7.~~https://doi.org/10.1016/s1002-0160(18)60048-7, 2018.

Hastie, T., Tibshirani, R., and Friedman, J. H~~., 2017,~~.: The elements of statistical learning : data mining, inference, and prediction, ~~New York, NY, Springer~~Second edition, corrected at 12th printing 2017, Springer series in statistics, Springer, New York, NY, https://doi.org/10.1007/978-0-387-84858-7, 2017.

Heinze, S., Vohland, M., Joergensen, R. G., and Ludwig, B~~., 2013,~~.: Usefulness of near-infrared spectroscopy for the prediction of chemical and biological soil properties in different long-term experiments~~:~~. J. Plant Nutr. Soil Sci., ~~v.~~ 176, ~~no. 4, p.~~ 520-528~~.~~, ~~https://doi.org/10.1002/jpln.201200483.~~https://doi.org/10.1002/jpln.201200483, 2013.

Hutengs, C., Seidel, M., Oertel, F., Ludwig, B., and Vohland, M~~., 2019,~~.: In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils~~: Geoderma, v. 355.~~

~~https://doi.org/10.1016/j.geoderma.2019.113900.~~, Geoderma, 355, https://doi.org/10.1016/j.geoderma.2019.113900, 2019.

~~Kawamura, K., Nishigaki, T., Tsujimoto, Y., Andriamananjara, A., Rabenaribo, M., Asai, H., Rakotoson, T., and Razafimbelo, T., 2021, Exploring relevant wavelength regions for estimating soil total carbon contents of rice fields in Madagascar from Vis-NIR spectra with sequential application of backward interval PLS: Plant Prod. Sci., v. 24, no. 1, p. 1-14. https://doi.org/10.1080/1343943x.2020.1785898.~~

Kennard, R. W~~.,~~., and Stone, L. A~~., 1969.~~.: Computer aided design of experiments~~.~~, Technometrics~~ v.~~, 11, ~~p. ~~137-148~~. https://doi.org/10.1080/00401706.1969.10490666. ~~, https://doi.org/10.2307/1266770, 1969.

Knox, N. M., Grunwald, S., McDowell, M. L., Bruland, G. L., Myers, D. B., and Harris, W. G~~., 2015,~~.: Modelling soil carbon fractions with visible near-infrared (VNIR) and mid-infrared (MIR) spectroscopy~~:~~, Geoderma, ~~v. ~~239, ~~p. ~~229-239~~.~~, ~~https://doi.org/10.1016/j.geoderma.2014.10.019.~~https://doi.org/10.1016/j.geoderma.2014.10.019, 2015.

Kuang, B~~.,~~., and Mouazen, A. M~~., 2011,~~.: Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms~~:~~, Eur. J. Soil Sci., ~~v. ~~62, ~~no. 4, p. ~~629-636~~.~~, ~~https://doi.org/https://doi.org/10.1111/j.1365-2389.2011.01358.x.~~https://doi.org/10.1111/j.1365-2389.2011.01358.x, 2011.

~~-, 2012,~~Kuang, B. and Mouazen, A. M.: Influence of the number of samples on prediction error of visible and near infrared spectroscopy of selected soil properties at the farm scale~~:~~, Eur. J. Soil Sci., ~~v. ~~63, ~~no. 3, p. ~~421-429~~.~~, ~~https://doi.org/10.1111/j.1365-2389.2012.01456.x.~~https://doi.org/10.1111/j.1365-2389.2012.01456.x, 2012.

Kuhn, M~~.,~~.: caret: Classification and Regression Training, R package [code], https://doi.org/10.18637/jss.v028.i05, 2020.

Kuhn, M. and Johnson, K~~., 2013,~~.: Applied predictive modeling, Springer, New York, ~~Springer~~https://doi.org/10.1007/978-1-4614-6849-3, 2013.

Kusumo, B. H., Sukartono, S., Bustan, B., and Purwanto, Y. A.: Total nitrogen in rice paddy field independently predicted from soil carbon using Near Infrared Reflectance Spectroscopy (NIRS), 4th Annual Applied Science and Engineering Conference (AASEC), Univ Pendidikan Indonesia, Sch Postgraduate Studies, Tech & Vocat Educ St, Bali, INDONESIA, Apr 24, https://doi.org/10.1088/1742-6596/1402/2/022096, 2019.

Li, H. Y., Jia, S. Y., and Le, Z. C~~., 2020,~~.: Prediction of Soil Organic Carbon in a New Target Area by Near-Infrared Spectroscopy: Comparison of the Effects of Spiking in Different Scale Soil Spectral Libraries~~: Sensors, v. 20, no. 16. https://doi.org/10.3390/s20164357.~~, Sensors, 20, https://doi.org/10.3390/s20164357, 2020.

~~Li, S., Rossel, R. A. V., and Webster, R., 2022, The cost-effectiveness of reflectance spectroscopy for estimating soil organic carbon: Eur~~Liu, S., Shen, H., Chen, S., Zhao, X., Biswas, A., Xiaolin, J., Shi, Z., and Fang, J.: Estimating forest soil organic carbon content using vis-NIR spectroscopy: Implications for large-scale soil carbon spectroscopic assessment, Geoderma, 348, 37-44, https://doi.org/10.1016/j.geoderma.2019.04.003, 2019.

hat formatiert: Englisch (Vereinigtes Königreich)

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

hat formatiert: Schriftart: 10 Pt.

Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen

, J. Soil Sci., v. 73, no. 1. https://doi.org/10.1111/ejss.13202.

Liu, Y., Liu, Y. L., Chen, Y. Y., Zhang, Y., Shi, T., Z., Wang, J. J., Hong, Y. S., Fei, T., and Zhang, Y., 2019, The Influence of Spectral Pretreatment on the Selection of Representative Calibration Samples for Soil Organic Matter Estimation Using Vis-NIR Reflectance Spectroscopy: Remote Sens., v. 11, no. 4. https://doi.org/10.3390/rs11040450.

Lobsey, C. R., Viscarra Rossel, R. A., Roudier, P., and Hedley, C. B., 2017, : rs-local data-mines information from spectral libraries to improve local calibrations, Eur. J. Soil Sci., v. 68, no. 6, p. 840-852. https://doi.org/10.1111/ejss.12490., https://doi.org/10.1111/ejss.12490, 2017.

Luca, F., Conforti, M., Castrignano, A., Matteucci, G., and Buttafuoco, G., 2017, Effect of calibration set size on prediction at local scale of soil carbon by Vis-NIR spectroscopy: Geoderma, v. 288, p. 175-183. https://doi.org/10.1016/j.geoderma.2016.11.015.

Lucas, S. T., and Weil, R. R., 2012, : Can a Labile Carbon Test be Used to Predict Crop Responses to Improve Soil Organic Matter Management?: Agron. J., v. 104, no. 4, p. 1160-1170, https://doi.org/10.2134/agronj2011.0415.https://doi.org/10.2134/agronj2011.0415, 2012.

Martin, P. D., Malley, D. F., Manning, G., and Fuller, L.: Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy, Can. J. Soil Sci., 82, 413-422, https://doi.org/10.4141/s01-054, 2002.

McCarty, G., Reeves, J., Reeves, V., Follett, R., and Kimble, J., 2002, : Mid-Infrared and Near-Infrared Diffuse Reflectance Spectroscopy for Soil Carbon Measurement, Soil Sci. Soc. Am. J., v. 66, p. 640-646, https://doi.org/10.2136/sssaj2002.6400., https://doi.org/10.2136/sssaj2002.6400a, 2002.

McCarty, G. W., and Reeves, J. B., 2006, Comparison of NFAR infrared and mid infrared diffuse reflectance spectroscopy for field-scale measurement of soil fertility parameters: Soil Sci., v. 171, no. 2, p. 94-102. https://doi.org/10.1097/01.ss.0000187377.84391.54.

Mishra, P., Roger, J. M., Marini, F., Biancolillo, A., and Rutledge, D. N., 2022, : Pre-processing ensembles with response oriented sequential alternation calibration (PROSAC): A step towards ending the pre-processing search and optimization quest for near-infrared spectral modelling, Chemom. Intell. Lab. Syst., v. 222, https://doi.org/10.1016/j.chemolab.2022.104497.https://doi.org/10.1016/j.chemolab.2022.104497, 2022.

Molinaro, A. M., Simon, R., and Pfeiffer, R. M., 2005, : Prediction error estimation: a comparison of resampling methods, Bioinformatics, v. 21, no. 15, p. 3301-3307, https://doi.org/10.1093/bioinformatics/bti499.https://doi.org/10.1093/bioinformatics/bti499, 2005.

Munnaf, M. A., and Mouazen, A. M., 2022, : Removal of external influences from on-line vis-NIR spectra for predicting soil organic carbon using machine learning: Catena, v. 211. https://doi.org/10.1016/j.catena.2022.106015., Catena, 211, https://doi.org/10.1016/j.catena.2022.106015, 2022.

hat formatiert: Englisch (Vereinigtes Königreich)
hat formatiert: Englisch (Vereinigtes Königreich)
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
Formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
Formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
Formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...
hat formatiert ...

Nawar, S., and Mouazen, A. M., 2019, On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning: Soil Till. Res., v. 190, p. 120-127. https://doi.org/10.1016/j.still.2019.03.006.

Ng, W., Minasny, B., Jones, E., and McBratney, A., 2022, : To spike or to localize? Strategies to improve the prediction of local soil properties using regional spectral library: Geoderma, v. 406. https://doi.org/10.1016/j.geoderma.2021.115501., Geoderma, 406, https://doi.org/10.1016/j.geoderma.2021.115501, 2022.

Perez-Bejarano, A., and Guerrero, C., 2018, Near infrared spectroscopy to quantify the temperature reached in burned soils: Importance of calibration set variability: Geoderma, v. 326, p. 133-143. https://doi.org/10.1016/j.geoderma.2018.03.038.

R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing [code], https://www.R-project.org , 2020.

Ramirez-Lopez, L., Behrens, T., Schmidt, K., Stevens, A., Dematte, J. A. M., and Scholten, T., 2013, : The spectrum-based learner: A new local approach for modeling soil vis-NIR spectra of complex datasets: Geoderma, v. 195, p. 268-279. https://doi.org/10.1016/j.geoderma.2012.12.014, https://doi.org/10.1016/j.geoderma.2012.12.014, 2013.

Reeves, J. B., 2010, : Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done?: Geoderma, v. 158, no. 1-2, p. 3-14. https://doi.org/10.1016/j.geoderma.2009.04.005.?, Geoderma, 158, 3-14, https://doi.org/10.1016/j.geoderma.2009.04.005, 2010.

Riefolo, C., Castrignano, A., Colombo, C., Conforti, M., Ruggieri, S., Vitti, C., and Buttafuoco, G., 2020, : Investigation of soil surface organic and inorganic carbon contents in a low-intensity farming system using laboratory visible and near-infrared spectroscopy: Arch. Agron. Soil Sci., v. 66, no. 10, p. 1436-1448. https://doi.org/10.1080/03650340.2019.1674446.https://doi.org/10.1080/03650340.2019.1674446, 2020.

Rodriguez-Febereiro, M., Dafonte, J., Fandino, M., Cancela, J. J., and Rodriguez-Perez, J. R., 2022, : Evaluation of Spectroscopy and Methodological Pre-Treatments to Estimate Soil Nutrients in the Vineyard: Remote Sens., v. 14, no. 6. https://doi.org/10.3390/rs14061326.https://doi.org/10.3390/rs14061326, 2022.

Seber, G. A. F., and Wild, C. J., 2004, Nonlinear regression, New York, Wiley, Wiley series in probability and mathematical statistics.

Seidel, M., Hutengs, C., Ludwig, B., Thiele-Bruhn, S., and Vohland, M., 2019, : Strategies for the efficient estimation of soil organic carbon at the field scale with vis-NIR spectroscopy: Spectral libraries and spiking vs. local calibrations: Geoderma, v. 354. https://doi.org/10.1016/j.geoderma.2019.07.014., Geoderma, 354, https://doi.org/10.1016/j.geoderma.2019.07.014, 2019.

hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen

Shen, Z. F., Ramirez-Lopez, L., Behrens, T., Cui, L., Zhang, M. X., Walden, L., Wetterlind, J., Shi, Z., Sudduth, K. A., Song, Y. Z., Catambay, K., and Rossel, R. A. V., 2022, : Deep transfer learning of global spectra for local soil carbon monitoring : Isprs Journal of Photogrammetry and Remote Sensing, v. 188, p. 190-200, https://doi.org/10.1016/j.isprsjprs.2022.04.009, https://doi.org/10.1016/j.isprsjprs.2022.04.009, 2022.

Singh, K., Aitkenhead, M., Fidelis, C., Yinil, D., Sanderson, T., Snoeck, D., and Field, D. J., 2022, : Optimization of spectral pre-processing for estimating soil condition on small farms : Soil Use Manage., v. 38, no. 1, p. 150-163, https://doi.org/10.1111/sum.12684, https://doi.org/10.1111/sum.12684, 2022.

Sleep, B., Mason, S., Janik, L., and Mosley, L., 2022, Application of visible near-infrared absorbance spectroscopy for the determination of Soil pH and liming requirements for broad-acre agriculture: Prec. Agric., v. 23, no. 1, p. 194-218, https://doi.org/10.1007/s11119-021-09834-7.

Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., and McLaughlin, M. J., 2014, : The Performance of Visible, Near-, and Mid-Infrared Reflectance Spectroscopy for Prediction of Soil Physical, Chemical, and Biological Properties : Appl. spectrosc. Rev., v. 49, no. 2, p. 139-186, https://doi.org/10.1080/05704928.2013.811081, https://doi.org/10.1080/05704928.2013.811081, 2014.

Stafford, A. D., Kusumo, B. H., Jeyakumar, P., Hedley, M. J., and Anderson, C. W. N., 2018, Cadmium in soils under pasture predicted by soil spectral reflectance on two dairy farms in New Zealand: Geoderma Regional, v. 13, p. 26-34, https://doi.org/10.1016/j.geodrs.2018.03.001.

Stenberg, B., Rossel, R. A. V., Mouazen, A. M., and Wetterlind, J., 2010, : Visible and near infrared spectroscopy in soil science, in: Adv. Agron., edited by: Sparks, D. L., ed., Adv. Advances in Agronomy, Agron., Volume 107, p. 163-215, https://doi.org/10.1016/s0065-2113(10)07005-7, https://doi.org/10.1016/s0065-2113(10)07005-7, 2010.

Stevens, A. S. and Ramirez-Lopez, L.: An introduction to the prospectr package, R package [code], https://cran.r-project.org/web/packages/prospectr/vignettes/prospectr.html, 2020.

Taubner, H., Roth, B., and Tippkotter, R.: Determination of soil texture: Comparison of the sedimentation method and the laser-diffraction analysis, J. Plant Nutr. Soil Sci., Vaudour, E., Cerovic, Z. G., Ebengo, D. M., and Latouche, G., 2018, Predicting Key Agronomic Soil Properties with UV-Vis Fluorescence Measurements Combined with Vis-NIR-SWIR Reflectance Spectroscopy: A Farm-Scale Study in a Mediterranean Viticultural Agroecosystem: Sensors, v. 18, no. 4, https://doi.org/10.3390/s18041157.

172, 161-171, https://doi.org/10.1002/jpln.200800085, 2009.

Weil, R. R., Islam, K. R., Stine, M. A., Gruver, J. B., and Samson-Liebig, S. E., 2003, : Estimating active carbon for soil quality assessment: A simplified method for laboratory and field use : Am. J. Altern. Agric., v. 18, no. 1, p. 3-17, https://doi.org/10.1079/AJAA200228, https://www.jstor.org/stable/pdf/44503242.pdf, 2003.

hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Englisch (Vereinigtes Königreich)
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.

Wetterlind, J., and Stenberg, B., 2010, : Near-infrared spectroscopy for within-field soil characterization: small local calibrations compared with national libraries spiked with local samples: Eur. J. Soil Sci., v. 61, no. 6, p. 823-843, https://doi.org/10.1111/j.1365-2389.2010.01283.x, https://doi.org/10.1111/j.1365-2389.2010.01283.x, 2010.

Wetterlind, J., Stenberg, B., and Soderstrom, M., 2008, The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale: Prec. Agric., v. 9, no. 1-2, p. 57-69. https://doi.org/10.1007/s11119-007-9051-z.

Wold, S., Johansson, E., and Cocchi, M., 1993, : PLS-partial least squares projections to latent structures, in: 3D QSAR in drug design, edited by: Kubinyi, H., Folkers, G., and Martin, Y., eds., 3D QSAR in drug design: Escom, Leiden, Escom, p. 523– 550, https://doi.org/10.1007/0-306-46858-1, https://doi.org/10.1007/0-306-46858-1, 1993.

Wold, S., Martens, H., and Wold, H., : The multivariate calibration problem in chemistry solved by the PLS method, in Proceedings Matrix Pencils, Berlin, Heidelberg, 1983, Springer, p. 286-293, https://doi.org/10.1007/BFb0062108, https://doi.org/10.1007/BFb0062108, 1983.

Yang, H. Q., 2011, Spectroscopic Calibration for Soil N and C Measurement at a Farm Scale, in Proceedings 3rd International Conference on Environmental Science and Information Application Technology (ESIAT), Xian, PR China, Aug 20-21, 2011, v. 10, p. 672-677. https://doi.org/10.1016/j.proenv.2011.09.108.

Zhang, L., Yang, X. M., Drury, C., Chantigny, M., Gregorich, E., Miller, J., Bittman, S., Reynolds, W. D., and Yang, J. Y., 2018, : Infrared spectroscopy estimation methods for water-dissolved carbon and amino sugars in diverse Canadian agricultural soils, Can. J. Soil Sci., v. 98, p. 484-499, https://doi.org/10.1139/cjss-2018-0027, https://doi.org/10.1139/cjss-2018-0027, 2018.

Zhao, D. X., Arshad, M., Wang, J., and Triantafilis, J., 2021, : Soil exchangeable cations estimation using Vis-NIR spectroscopy in different depths: Effects of multiple calibration models and spiking: Comput. Electron. Agric., v. 182, https://doi.org/10.1016/j.compag.2021.105990, https://doi.org/10.1016/j.compag.2021.105990, 2021.

Table 1: Description of the datasets of the six different fields A to F. All fields were classified as Eutric Cambisol developed on base moraine. Soil texture was measured with the improved integral suspension pressure method (ISP+)

Field	Coordinates	Elevation [m above sea level]	Area [ha]	Mean soil texture (Sand/Silt/Clay) [%]	Number of samples	
					Spectroscopy	Wet chemistry
A	47° 40' 58" N / 08° 45' 54" E	420	0.84	Sandy loam (50/29/21)	468	70
B	47° 40' 54" N / 08° 46' 05" E	420	0.67	Sandy loam (44/35/20)	468	70
C	47° 38' 01" N / 08° 57' 02" E	600	0.44	Sandy loam (27/35/38)	468	70
D	47° 38' 43" N / 08° 42' 58" E	460	0.64	Clay loam (28/44/28)	468	70
E	47° 38' 49" N / 08° 43' 06" E	460	1.05	Sandy loam (30/48/23)	351	53
F	47° 34' 22" N / 08° 48' 52" E	380	0.3	Sandy loam (39/43/18)	351	53

hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: EndNote Bibliography, Einzug: Links: 0 cm, Hängend: 0.75 cm, Zeilenabstand: 1.5 Zeilen
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
hat formatiert: Schriftart: 10 Pt.
Formatiert: Einzug: Links: 0 cm, Erste Zeile: 0 cm
hat formatiert: Schriftart: 9 Pt., Fett
Formatiert: Abstand Nach: 10 Pt.

Table 2: Description of applied pre-processing and model performance of the final chosen models using a Partial Least Square regression. The local models (fields A to F) were evaluated with 5 times repeated 5-fold cross-validation and the general models (All) with 5 times repeated 10-fold cross-validation. Model metrics of Cross-validation are indicated as mean with the standard deviation across the repeats in brackets. RMSE = Root mean square error, RPD = ratio of performance to deviation, Refl. = Reflectance, Abs. = Absorbance, SG = Savitzky-Golay filter (m = order of derivative, w = window width), SNV = standard normal variate, GSD = gap segment derivative (m = derivative, w = window width, s = segment size), MSC = multiplicative scatter correction

Field	Property	Range of wavelengths / interval [nm]	Pre-processing	Latent variable	n	Calibration				Cross-validation				Model performance
						R^2_{cal}	$Bias_{cal}$	$RMSE_{cal}$	RPD_{cal}	R^2	Bias	RMSE	RPD	
A	SOC [g kg^{-1}]	410-2500 / 6	Abs., SG (m = 1, w=35)	2	70	0.59	0.00	2.3335	1.58	0.55	0.5701 (0.69)	2.4243 (0.55)	1.5350 (0.30)	Approximate
	POXC [g kg^{-1}]	350-2500 / 2	Refl., SNV, GSD (m=2, w=31, s=1)	7	70	0.81	0.00	0.05	2.28	0.6764 (0.17)	0.0600 (0.02)	0.07 (0.01)	1.7465 (0.41)	Approximate
	Total N [g kg^{-1}]	370-2480 / 6	Refl., SG (m=1, w=21), MSC	7	70	0.87	0.00	0.11	2.77	0.79 (0.11)	0.01 (0.05)	0.14 (0.03)	2.1618 (0.61)	Accurate
	Total C [g kg^{-1}]	390-2500 / 4	Abs., SNV, GSD (m=2, w=21, s=1)	6	70	0.94	0.00	2.14	4.21	0.9288 (0.09)	0.00 (0.84)	2.5863 (0.66)	3.4948 (1.41)	Excellent
	pH	410-2500 / 4	Abs., SG (m = 1, w=35)	5	70	0.74	0.00	0.06	1.97	0.6563	0.0700 (0.03)	0.08 (0.02)	1.6970 (0.60)	Approximate
B	SOC [g kg^{-1}]	360-2500 / 5	Abs., SG (m=2, w=21), MSC	7	70	0.98	0.00	0.66	6.47	0.9291 (0.05)	0.04 (0.37)	1.2026 (0.36)	3.5546 (1.12)	Excellent
	POXC [g kg^{-1}]	360-2480 / 3	Abs., SG (m=2, w=21), MSC	4	70	0.94	0.00	0.03	4.08	0.8584 (0.12)	0.0400 (0.02)	0.05 (0.01)	2.5760 (0.74)	Accurate
	Total N [g kg^{-1}]	360-2480 / 5	Abs., SG (m=2, w=21), MSC	4	70	0.93	0.00	0.10	3.87	0.8987 (0.08)	0.1200 (0.04)	0.13 (0.03)	2.85 (0.92) Excellent	Accurate
	Total C [g kg^{-1}]	370-2500 / 5	Abs., SG (m=2, w=21), MSC	10	70	0.99	0.4900	0.51	9.64	0.93 (0.03)	10 0.05 (0.40)	1.29 (0.9225)	1.35 3.6165 (0.84)	Excellent
	pH	350-2500 / 3	Abs., SG (m=1, w=21), MSC	7	70	0.9298	0.1200	3.810.07	0.856.64	0.1883 (0.07)	0.00 (0.06)	0.19 (0.03)	2.6246 (0.51)	Accurate
C	SOC [g kg^{-1}]	370-2480 / 1	Abs., SNV, GSD (m=1, w=11, s=1)	7	70	0.90	0.00	1.02	3.11	0.7577 (0.09)	1.580.03 (0.46)	1.59 (0.28)	2.0105 (0.46)	Accurate
	POXC [g kg^{-1}]	370-2440 / 3	Refl., SG (m=2, w =21)	7	70	0.93	0.00	0.03	3.80	0.7977 (0.15)	0.00 (0.01)	0.05 (0.01)	2.2230 (0.81)	Accurate
	Total N [g kg^{-1}]	350-2460 / 4	Abs., SG (m=2,	7	70	0.97	0.00	0.05	5.87	0.8890 (0.06)	0.00 (0.03)	0.09 (0.02)	2.933.22 (0.97)	Excellent Accurate

Eingefügte Zellen
Eingefügte Zellen
hat formatiert
Formatiert
hat formatiert
Eingefügte Zellen
Formatiert
Formatiert
hat formatiert
Eingefügte Zellen
Formatiert
Formatiert
Formatiert
Eingefügte Zellen
Formatiert
Formatiert
Formatiert
Formatiert
Formatiert
Formatiert
Eingefügte Zellen
hat formatiert
Formatiert
Formatiert
Formatiert
Eingefügte Zellen
Formatiert
Formatiert
Formatiert
Formatiert
hat formatiert
Formatiert
Formatiert
Eingefügte Zellen
Formatiert
hat formatiert
Eingefügte Zellen
Formatiert
Formatiert
Gelöschte Zellen
hat formatiert
Formatiert
Eingefügte Zellen
hat formatiert
Eingefügte Zellen
Eingefügte Zellen
Formatiert
Formatiert
Eingefügte Zellen
hat formatiert
Eingefügte Zellen
Formatiert
Formatiert
Eingefügte Zellen

			w=21), MSC											
	Total C [g kg^{-1}]	390350-2500 / 53	Refl., SG (m=1, w=21), MSC	910	70	0.9697	0.9800	0.92	5.3669	0.9193 (0.03)	1.54-0.07 (0.34)	1.44 (0.29)	3.4074 (0.98)	Excellent
	pH	390-2500 / 5	Abs., SG (m=2, w=21), MSC	6	70	0.89	0.00	0.05	3.09	0.7577 (0.12)	0.00 (0.03)	0.08 (0.02)	2.0200 (0.59)	Accurate
D	SOC [g kg^{-1}]	390-2500 / 3	Abs., SNV, GSD (m=2, w=21, s=1)	6	70	0.97	0.00	0.81	6.01	0.95 (0.02)	1.05-0.01 (0.35)	1.07 (0.19)	4.6474 (1.23)	Excellent
	POXC [g kg^{-1}]	390-2460 / 6	Refl., SG (m=2, w=21)	7	69	0.95	0.00	0.03	4.72	0.9192 (0.03)	0.00 (0.01)	0.05 (0.01)	3.3447 (0.65)	Excellent
	Total N [g kg^{-1}]	370-2500 / 4	Abs., SG (m=2, w=21), MSC	6	70	0.98	0.00	0.06	7.30	0.95 (0.04)	0.1001 (0.03)	0.11 (0.02)	4.5366 (1.16)	Excellent
	Total C [g kg^{-1}]	350-2500 / 2	Refl., SNV, GSD (m=2, w=21, s=1)	6	70	0.97	0.00	1.15	5.44	0.93 (0.03)	1.6202 (0.45)	3.861.61 (0.39)	4.07 (1.04)	Excellent
	pH	350-2500 / 6	Refl., SG (m=2, w=21), MSC	9	70	0.99	0.00	0.06	9.79	0.95 (0.02)	0.00 (0.04)	0.13 (0.03)	4.4883 (1.23)	Excellent
E	SOC [g kg^{-1}]	350-2500 / 3	Abs., SG (m=1, w=25)	3	53	0.82	0.00	1.25	2.35	0.7879 (0.11)	1.35-0.05 (0.53)	2.171.40 (0.42)	2.20 (0.70)	Accurate
	POXC [g kg^{-1}]	350-2500 / 4	Abs., GSD (m=2, w=21, s=21)	4	53	0.82	0.00	0.05	2.41	0.8182 (0.11)	0.00 (0.02)	0.05 (0.02)	2.3133 (0.69)	Accurate
	Total N [g kg^{-1}]	350-2500 / 3	Abs., SG (m=2, w=21)	4	53	0.94	0.00	0.07	4.12	0.8990 (0.04)	0.00 (0.03)	0.10 (0.02)	3.0310 (0.57)	Excellent
	Total C [g kg^{-1}]	360-2500 / 3	Refl., SG (m=1, w=21), MSC	6	53	0.98	0.00	1.20	7.83	0.96 (0.03)	1.9104 (0.56)	4.911.72 (0.51)	5.27 (1.85)	Excellent
	pH	350-2500 / 4	Refl., SG (m=1, w=21), MSC	7	53	0.98	0.00	0.10	7.15	0.9495 (0.03)	0.1701 (0.08)	0.16 (0.05)	4.1057 (1.91)	Excellent
F	SOC [g kg^{-1}]	350-2500 / 3	Abs., SG (m=1, w=21), MSC	4	53	0.5866	1.7700	1.5659	0.531.73	1.880.51 (0.18)	1.4701 (0.72)	2.00 (0.38)	1.43 (0.39)	Approximate
	POXC [g kg^{-1}]	380-2500 / 2	Refl., GSD (m =2, w=21, s=21)	5	53	0.86	0.00	0.03	2.72	0.8176 (0.16)	0.00 (0.01)	0.03 (0.00)	2.291.96 (0.60)	AccurateApproximate
	Total N [g kg^{-1}]	350-2500 / 3	Abs., SG (m=1, w=21), MSC	5	53	0.92	0.00	0.06	3.47	0.83 (0.10)	0.00 (0.04)	0.09 (0.02)	2.4151 (0.84)	Accurate
	Total C [g kg^{-1}]	370-2500 / 6	Abs., SG (m=1, w=21), MSC	5	53	0.84	0.9500	2.540.96	0.752.49	1.200.72 (0.18)	2.0000.01 (0.56)	1.29 (0.24)	1.82 (0.54)	Accurate
	pH	370-2500 / 4	Refl., SG (m=1, w=51)	1	53	0.23	0.00	0.04	1.15	0.2830 (0.20)	0.0300 (0.02)	1.170.04 (0.01)	1.14 (0.19)	Poor

hat formatiert ...
Eingefügte Zellen ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
Formatiert ...
Formatiert ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...
hat formatiert ...
Eingefügte Zellen ...
Formatiert ...
Formatiert ...

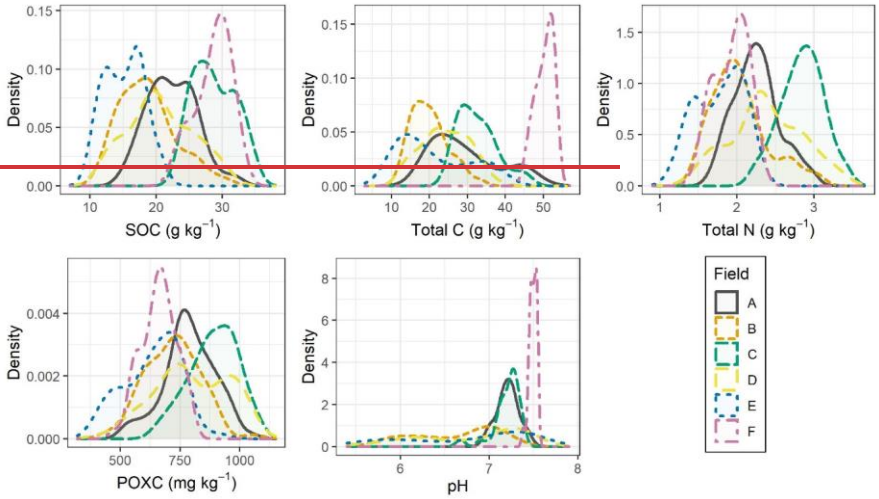
All	SOC [g kg^{-1}]	350-2500 / 3	Refl., SNV, GSD (m=2, w=5, s=1)	7	386	0.92	0.00	1.66	3.65	0.8990 (0.04)	1.99 0.02 (0.39)	3.061.93 (0.32)	3.21 (0.57)	Excellent
	POXC [g kg^{-1}]	350-2500 / 1	Refl., SG (m=1, w=21), MSC	8	385	0.88	0.00	0.05	2.90	0.8485 (0.05)	0.00 (0.01)	0.06 (0.01)	2.4260 (0.43)	Accurate
	Total N [g kg^{-1}]	350-2500 / 4	Abs., SG (m=2, w=11)	7	386	0.92	0.00	0.14	3.53	0.8889 (0.04)	0.1700 (0.03)	2.910.16 (0.02)	Accurate3.06 (0.46)	Excellent
	Total C [g kg^{-1}]	350-2500 / 2	Abs., SNV, GSD (m=2, w=5, s=1)	6	386	0.96	0.00	2.31	5.04	0.94 (0.01)	2.930.00 (0.53)	3.972.79 (0.30)	4.16 (0.47)	Excellent
	pH	350-2500 / 3	Refl., SG (m=2, w=21), MSC	9	386	0.95	0.00	0.12	4.40	0.8990 (0.03)	0.1700 (0.02)	2.970.15 (0.02)	Accurate3.34 (0.59)	Excellent

Table 3: Performance of local spectral models in the literature. All cited studies worked with lab measurements in the vis–NIR range on dried and sieved (< 2 mm) soil samples. The studies are numbered like in Fig. 7. The column pre-processing indicates if pre-processing was optimized for every soil property or not. Model types include partial least square regression (PLSR), cubist, neural network (NN) and support vector machine regression (SVM).

Field	Property	Area [ha]	N	Mean	CV [%]	R^2	RMSE	RMSE [%]	RPD	RPIQ	Model type	Pre-processing	Evaluation	Depth [cm]	Carbonate
1 Rodriguez-Febereiro et al. (2022)															
	Total N [g kg^{-1}]	29.4	96	1.9	36	0.69	0.3	16	1.89		PLSR	Optimized	Test-set	0-40	None
	SOM [g kg^{-1}]		96	43.9	47	0.73	10.6	24	1.94		PLSR				
2 Camargo et al. (2022)															
	SOC [g kg^{-1}]	200	400	7.1	19	0.54	0.9	13	1.44		PLSR	Limited	Test-set	0-40	Unknown
3 Sleep et al. (2022)															
	pH (H_2O)	124	85	6.05	13	0.49	0.78	13			PLSR	Limited	Test-set	0-10	Present
	pH ($CaCl_2$)		85	5.46	20	0.48	0.85	16			PLSR				
4 Li et al. (2022)															
	SOC [g kg^{-1}]	600	562	7.9	65	0.81	3.8	48		1.66	Cubist	Limited	Test-set	0-5.5; 13-17; 28-32; 58-62	None
5 Breure et al. (2022): 4 different fields with maximal distance of 10 km															
	SOC [g kg^{-1}]	37.7	97	13.0	26		0.8	6		7.4	PLSR	Limited	Test-set	0-25	None
	pH		97	7.31	6		0.12	2		2.6	PLSR				
6 Alomar et al., 2021															
	SOM [g kg^{-1}]	32	143	20.8	27	0.87	2.1	10	2.60		PLSR	Optimized	Test-set	0-20	High
	Total N [g kg^{-1}]		150	1.6	26	0.84	0.2	12	2.11		PLSR				
7 Singh et al. (2022)															
BOKA-PAN	SOC [g kg^{-1}]	1	128	9.9	99	0.97	1.6	16		8.55	NN	Optimized	Cross-validation	0-10; 10-30; 30-60; 60-90	Present
LAU-PAN	SOC [g kg^{-1}]	1	123	13.8	73	0.96	2.1	15		6.63	NN				
TSA-CCI	SOC [g kg^{-1}]	1	128	12.2	88	0.82	3.9	32		3.2	NN				
WAL-WIN	SOC [g kg^{-1}]	1	128	19.4	52	0.96	0.3	2		1.49	PLSR				
General model	SOC [g kg^{-1}]	4	507	13.8	78	0.96	4.2	30		1.49	NN				
8 Riefolo et al. (2020)															
	SOC [g kg^{-1}]	ca. 18	98	20.3	24	0.65	2.6	13		2.37	PLSR	Optimized	Cross-validation	0-20	High
9 Kawamura et al. (2021)															
	Total C [g kg^{-1}]	40'000	162	30.5	56	0.94	4.9	16		3.841	PLSR	Limited	Test-set	0-10	Unknown
10 Stafford et al. (2018)															
Waikato	Total C [g kg^{-1}]	81.3	172	34.1	81	0.95	6.4	19	4.33		PLSR	None	Cross-validation	0-10; 10-20; 20-30; 30-40	Very low
	Total N [g kg^{-1}]	81.3	172	3.7	71	0.91	0.8	22	3.43		PLSR				
Canterbury	Total C [g kg^{-1}]	151.9	168	20.5	63	0.91	4.0	20	3.41		PLSR				
	Total N [g kg^{-1}]	151.9	168	2.2	59	0.92	0.4	18	3.57		PLSR				

hat formatiert: Schriftart: Times New Roman, Schriftfarbe: Schwarz
Formatiert: Zentriert
Formatiert: Zentriert
Eingefügte Zellen
Formatiert: Zentriert
Formatiert: Zentriert
Formatiert: Zentriert
Eingefügte Zellen
hat formatiert: Schriftart: Times New Roman, Schriftfarbe: Schwarz
Formatiert: Zentriert
Formatiert: Zentriert
hat formatiert: Schriftfarbe: Automatisch
hat formatiert: Schriftart: Times New Roman, Schriftfarbe: Schwarz
Eingefügte Zellen
Formatiert: Zentriert
Formatiert: Zentriert
hat formatiert: Schriftart: Times New Roman, Schriftfarbe: Schwarz
Eingefügte Zellen
Formatiert: Zentriert
Formatiert: Zentriert
hat formatiert: Schriftfarbe: Automatisch

11 Seidel et al. (2019)																
TS1	SOC [g kg^{-1}]	Field-level	120	10.1	47	0.97	0.9	9	5.11		PLSR	Optimized	Test-set	0-5; 5-10; 10-20; 20-30; 30-60	None
TS2	SOC [g kg^{-1}]	Field-level	90	9.1	54	0.92	1.4	15	3.21		PLSR				
12 Nawar and Mouazen (2019)																
Hessleskew	SOC [g kg^{-1}]	12	122	20.7	18	0.6	2.0	10	1.72		PLSR	limited	Test-set	0-20	Present
Hagg	SOC [g kg^{-1}]	21	139	18.7	18	0.75	1.7	9	2.05		PLSR				
13 Vaudour et al. (2018)																
	SOC [g kg^{-1}]	6	146	10.7	48	0.81	2.2	21	2.31		PLSR	None	Cross-validation	Different depths	Present
	$CaCO_3$ [g kg^{-1}]		146	398.6	37	0.88	50.0	13	2.92		PLSR				
	Total N [g kg^{-1}]		146	0.8	44	0.92	0.1	13	3.57		PLSR				
	pH		146	8.62	2	0.9	0.06	1	3.11		PLSR				
14 Gomez and Coulouma (2018): 8 fields with each 20 to 36 topsoil samples																
General model	$CaCO_3$ [g kg^{-1}]	9.3	240	209.4	87	0.97	29.8	14	6.06		PLSR	Limited	Cross-validation	topsoil	High
15 Luca et al. (2017)																
	Total C [g kg^{-1}]	33	216	61.5	28	0.82	6.8	11		3.03	SVM	Limited	Test-set	0-20	None
16 Debaene et al. (2014)																
	pH	56	398	6.25	8	0.52	0.34	5	1.40		PLSR	Limited	Test-set	0-25	Present
	SOC [g kg^{-1}]		398	11.5	20	0.72	1.2	10	2.00						
17 Brodsky et al. (2013)																
	SOC [g kg^{-1}]	100	152	11.1	31	0.85	1.4	13	2.40		PLSR	Limited	Test-set	0-20	Unknown
18 Yang (2011)																
	Total N [g kg^{-1}]	Farm-level	122	2.0	29	0.92	0.2	8	3.63		PLSR	Optimized	Test-set	Topsoil	None
	SOC [g kg^{-1}]		122	19.7	28	0.9	1.6	8	3.21		PLSR				
19 Kuang and Mouazen (2011)																
Bramstrup Estate	Total C [g kg^{-1}]	Farm-level	70	18.4	163	0.89	11.0	60	3.15		PLSR	Limited	Test-set	0-30	Present
	SOC [g kg^{-1}]		70	17.4	144	0.96	6.2	36	4.95						
	Total N [g kg^{-1}]		70	1.7	118	0.93	0.6	35	3.88						
	pH		70	7.19	7	0.08	0.33	5	1.10						
Wimex	Total C [g kg^{-1}]	Farm-level	128	16.5	73	0.74	7.0	42	2.00						
	SOC [g kg^{-1}]		128	13.8	51	0.75	3.0	22	2.00						
	Total N [g kg^{-1}]		128	1.3	77	0.75	0.3	23	2.10						
	pH		128	7.02	27	0.16	0.63	9	1.13						
Mepol Medlov	Total C [g kg^{-1}]	Farm-level	205	15.3	13	0.14	2.2	14	0.95						
	SOC [g kg^{-1}]		205	14.8	14	0.12	1.9	13	1.07						
	Total N [g kg^{-1}]		205	1.7	14	0.09	0.3	18	0.81						
	pH		205	6.78	6	0.3	0.30	4	1.31						
General model	Total C [g kg^{-1}]	Three Farms	408	16.3	92	0.83	6.6	40	2.45						
	SOC [g kg^{-1}]		408	15.0	80	0.83	5.4	36	2.49						
	Total N [g kg^{-1}]		408	1.6	63	0.79	0.5	31	2.21						
	pH		408	6.94	7	0.03	0.70	10	0.90						
20 Wetterlind and Stenberg (2010)																
Brännеberg	SOC [g kg^{-1}]	69	86	24.0	13	0.7	1.2	5	1.90		PLSR	Limited	Test-set	0-20	Present
	pH			6.80	1	0.49	0.10	1	1.33						
Hacksta	SOC [g kg^{-1}]	97	137	24.0	25	0.85	2.2	9	2.60						
	pH			6.60	2	0.33	0.19	3	1.10						
Kärrtorp	SOC [g kg^{-1}]	62	89	49.0	35	0.71	5.3	11	1.50						
	pH			6.20	5	0.5	0.22	4	1.40						
Sjörstorp	SOC [g kg^{-1}]	78	106	19.0	26	0.57	2.7	14	1.90						
	pH			6.90	6	0.48	0.31	4	1.40						
21 Wetterlind et al. (2008)																
	SOM [g kg^{-1}]	97	50	41.0	27	0.89	3.2	8	3.00		PLSR	Limited	Test-set	0-20	None
22 McCarty and Reeves (2006)																
	SOC [g kg^{-1}]	20	544	11.0	45	0.88	1.6	15	3.06		PLSR	Limited	Cross-validation	0-10; 10-20	None
	Total N [g kg^{-1}]		544	0.9	44	0.85	0.2	18	2.50						
	pH		544	6.00	8	0.53	0.31	5	1.45						


Density
SOC (g kg^{-1})
Total C (g kg^{-1})
Total N (g kg^{-1})
POXC (mg kg^{-1})
pH
Field
A
B
C
D
E
F

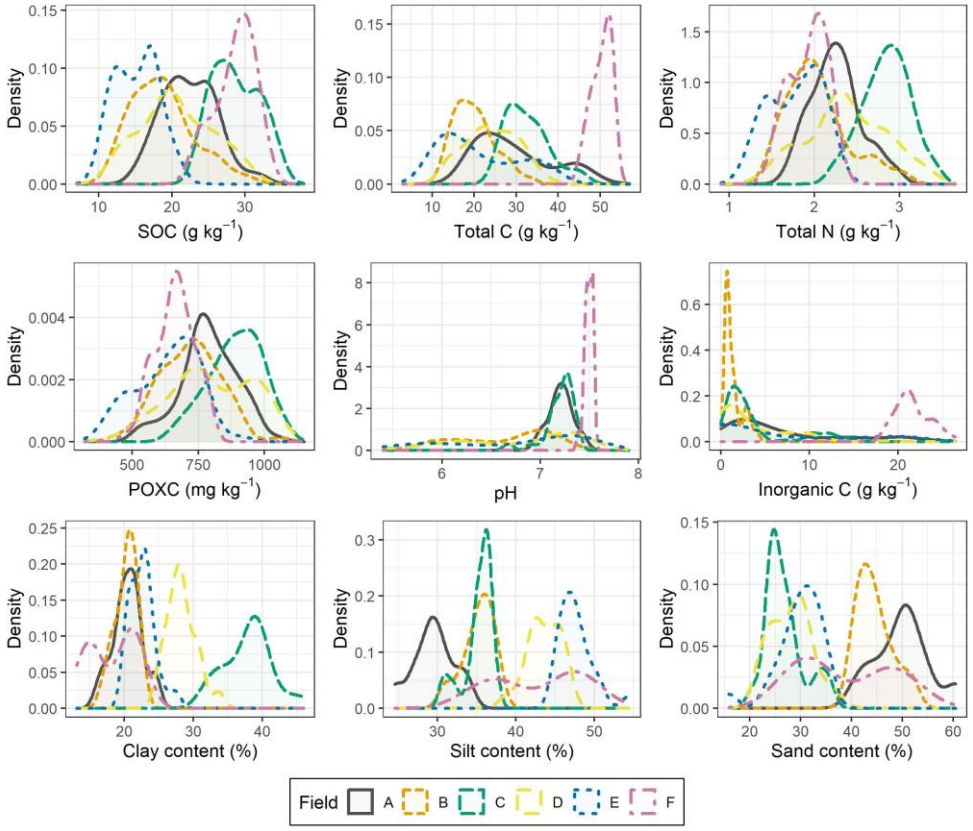
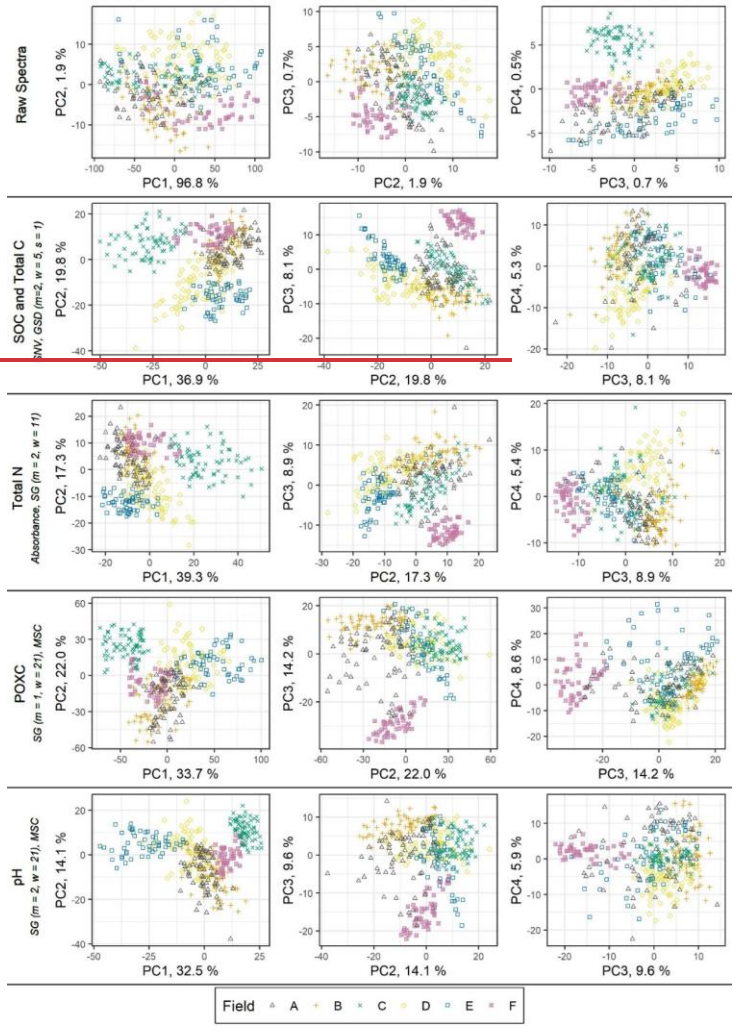


Figure 1: Density plots of the reference samples for the five ~~analyzed~~target properties (SOC, total C, total N, POXC and pH~~).~~) and inorganic C. Field A to D contained each 70 samples and field E and F each 53 samples. Soil texture was analyzed on 20 samples per field.


Raw Spectra
PC2, 1.9 %
PC3, 0.7%
PC4, 0.5%
PC1, 96.8 %
PC2, 1.9 %
PC3, 0.7 %
SOC and Total C
NV, GSD (m=2, w = 5, s = 1)
PC2, 19.8 %
PC3, 8.1 %
PC4, 5.3 %
PC1, 36.9 %
PC2, 19.8 %
PC3, 8.1 %
Total N
Absorbance, SG (m = 2, w = 11)
PC2, 17.3 %
PC3, 8.9 %
PC4, 5.4 %
PC1, 39.3 %
PC2, 17.3 %
PC3, 8.9 %
POXC
SG (m = 1, w = 21), MSC
PC2, 22.0 %
PC3, 14.2 %
PC4, 8.6 %
PC1, 33.7 %
PC2, 22.0 %
PC3, 14.2 %
pH
SG (m = 2, w = 21), MSC
PC2, 14.1 %
PC3, 9.6 %
PC4, 5.9 %
PC1, 32.5 %
PC2, 14.1 %
PC3, 9.6 %
Field A B C D E F

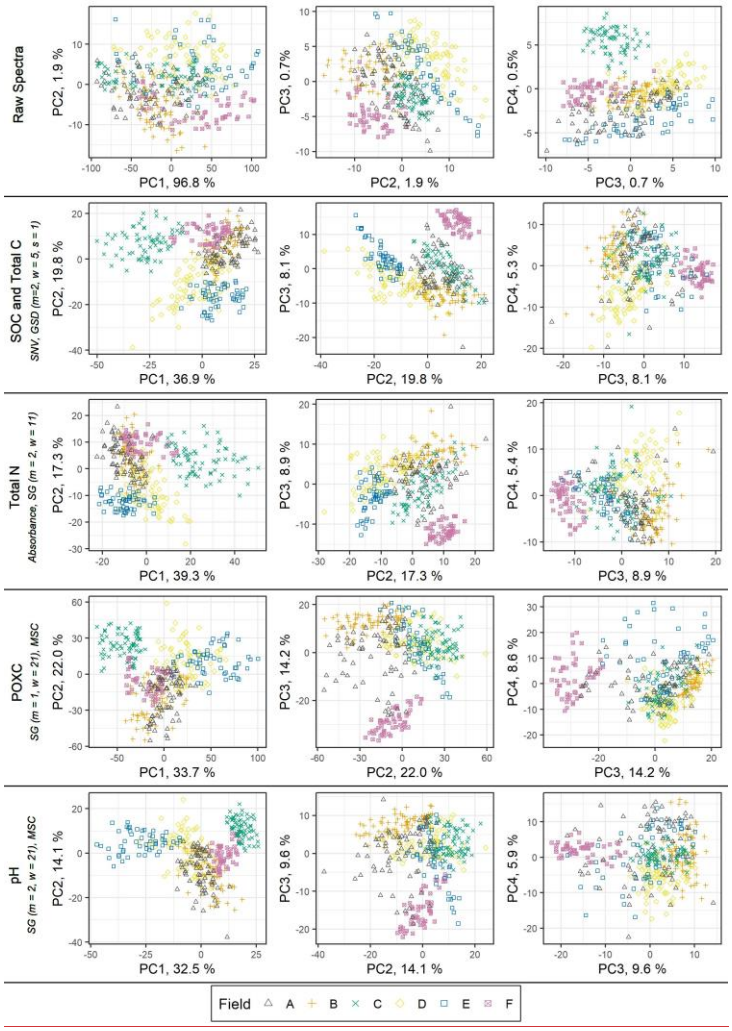
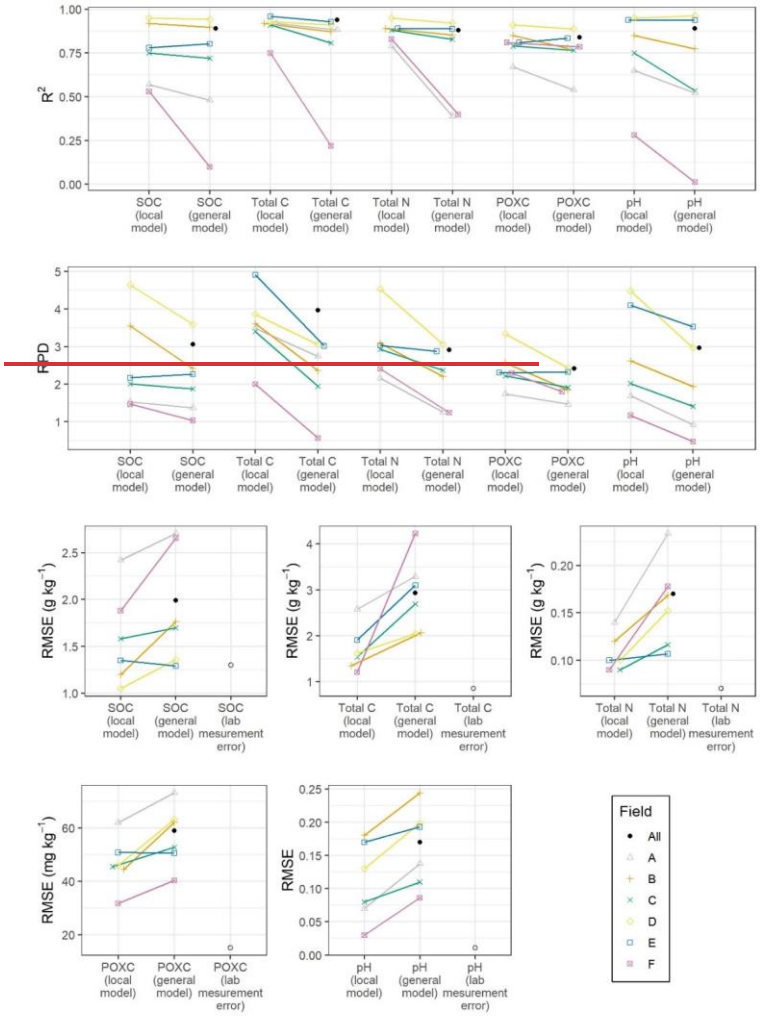


Figure 2: Biplots of principle component analysis with the first four principal components for the raw spectra and the pre-processed spectra according to the properties SOC, ~~Total~~total C, ~~Total~~total N, POXC and pH. The pre-processing is indicated in the figure and except for total N it was conducted on reflectance spectra (SG = Savitzky-Golay filter (m = order of derivative, w = window width), SNV = standard normal variate, GSD = gap segment derivative (m = derivative, w = window width, s = segment size), MSC = multiplicative scatter correction)


R²
RPD
RMSE (g kg⁻¹)
RMSE (mg kg⁻¹)
RMSE
SOC (local model)
SOC (general model)
Total C (local model)
Total C (general model)
Total N (local model)
Total N (general model)
POXC (local model)
POXC (general model)
pH (local model)
pH (general model)
SOC (lab mesurement error)
Total C (lab mesurement error)
Total N (lab mesurement error)
POXC (lab mesurement error)
pH (lab mesurement error)
Field
All
A
B
C
D
E
F

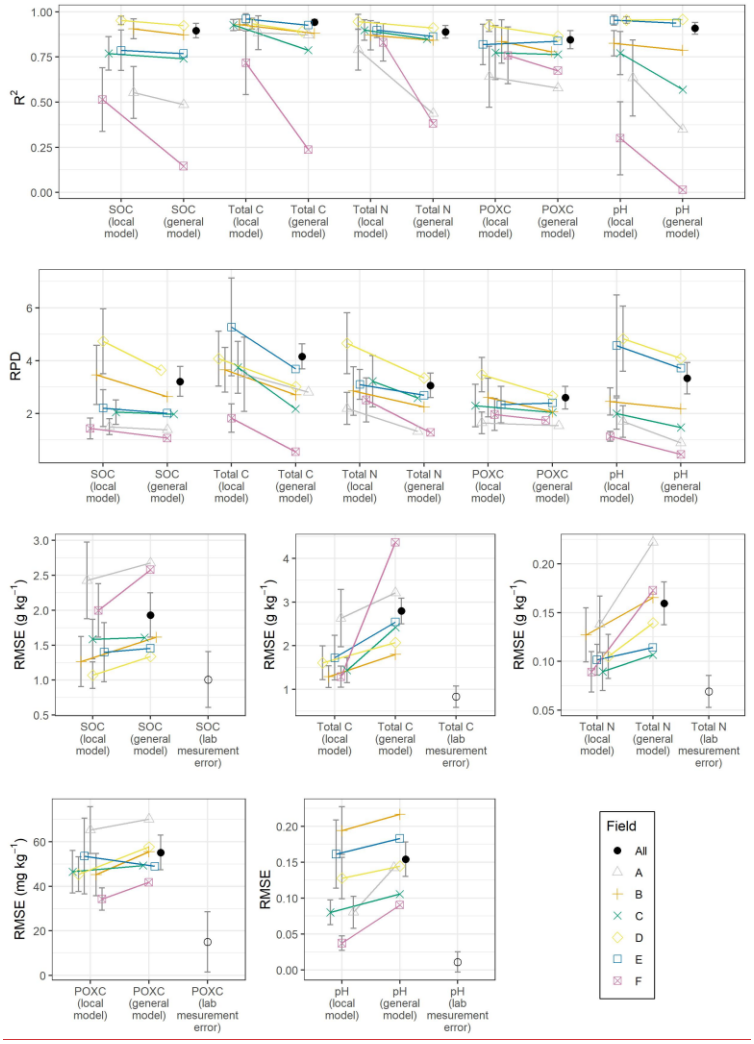
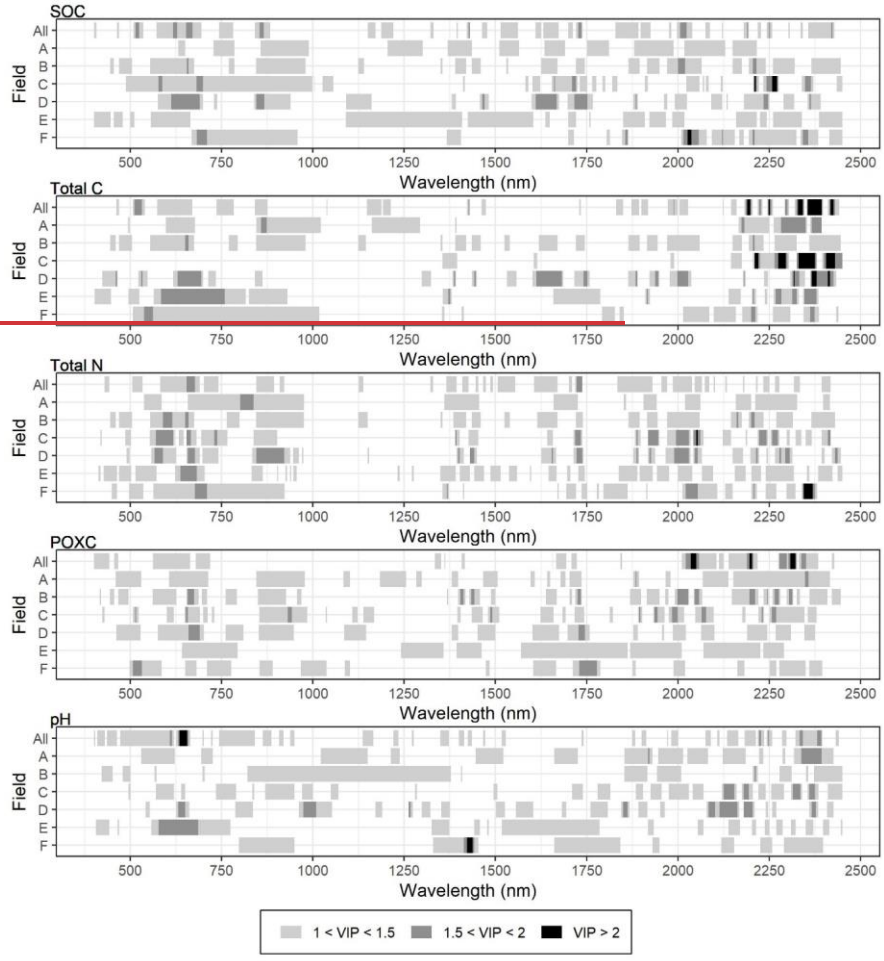


Figure 3: R^2, ratio of performance to deviation (RPD) and root mean square error (RMSE) calculated from the local models and field-specifically calculated from the general model for the six fields (A – F) and the five soil properties (SOC, total C, total N, POXC and pH). The overall RMSE of the general model is indicated with a black filled circle and the label "All". The ~~calculated~~ RMSE

are compared with the error of the lab measurements (mean standard error of 18 triplicates~~) of the lab measurements.~~ indicated with standard deviation)



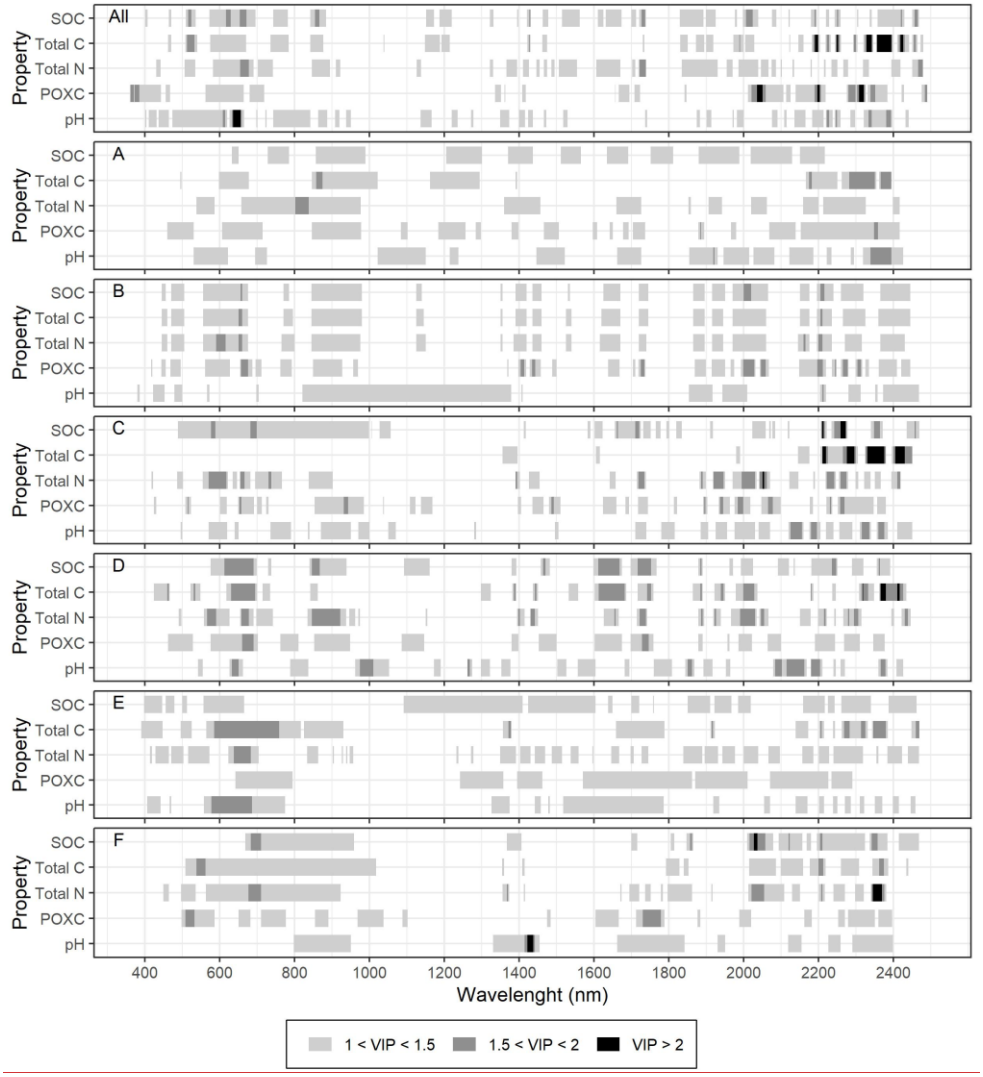
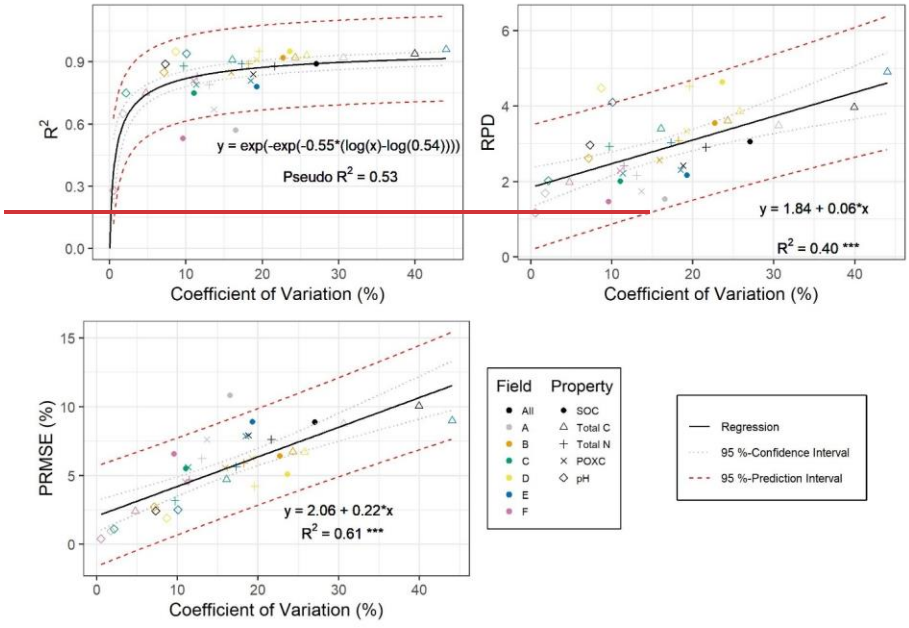
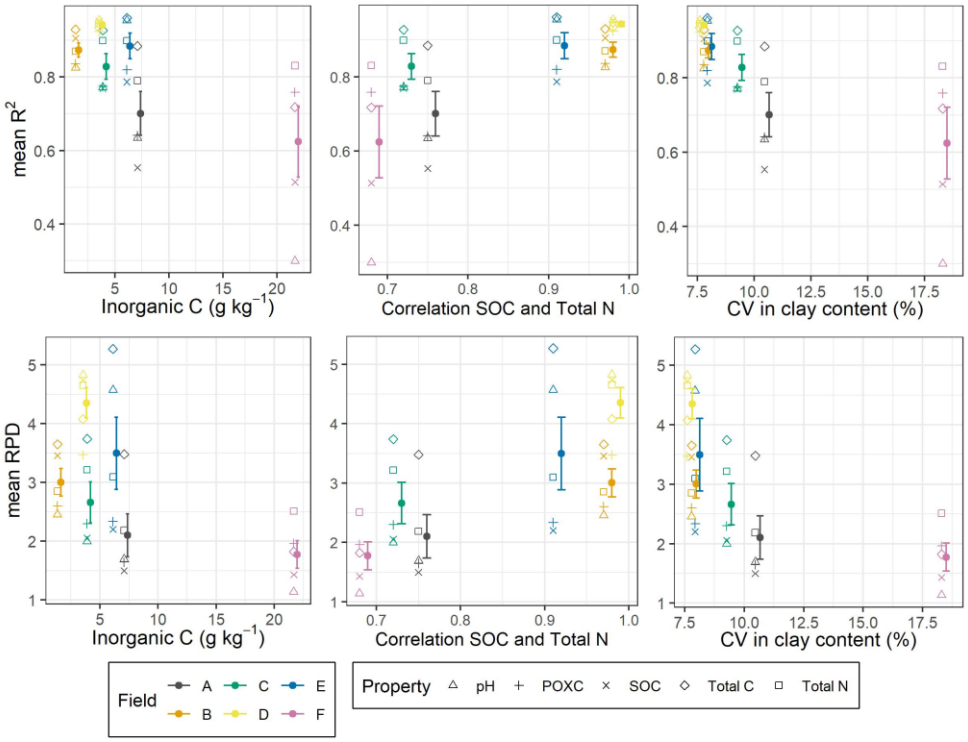
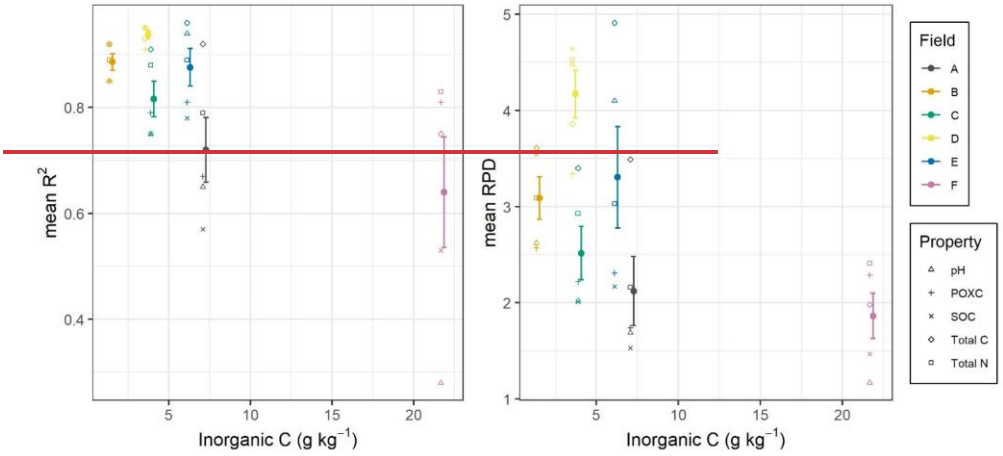


Figure 4: Variable importance in projection (VIP) for the local models of fields A to F and the general model that combined the datasets of all fields (All).


R²
y = exp(-exp(-0.55*(log(x)-log(0.54))))
Pseudo R² = 0.53
Coefficient of Variation (%)
RPD
y = 1.84 + 0.06*x
R² = 0.40 ***
Coefficient of Variation (%)
PRMSE (%)
y = 2.06 + 0.22*x
R² = 0.61 ***
Coefficient of Variation (%)
Field
All
A
B
C
D
E
F
Property
SOC
Total C
Total N
POXC
pH
Regression
95 %-Confidence Interval
95 %-Prediction Interval



865 **Figure 5: R^2, ratio of performance to deviation (RPD) and percental root mean square error (PRMSE) in dependence of coefficient of variation for all spectral models (Fields A to F plus general model (All)). Significant codes for the linear regressions: *** < 0.001, ** < 0.01, *< 0.05. For R^2 a two-parameter Weibull function was fitted, and both fitted parameters were significant (p < 0.05). 95 % confidence and 95 % prediction intervals are indicated for all regressions**



870 **~~Figure 6~~Figure 5: R^2 and ratio of performance to deviation (RPD) from the local models for SOC, total C, total N, POXC~~,~~ and pH aggregated (mean and standard error) per field (A-F) in dependence of mean inorganic C ~~concentration~~content, Pearson's correlation coefficient between SOC and total N and coefficient of ~~the field~~variation (CV) in clay content.**

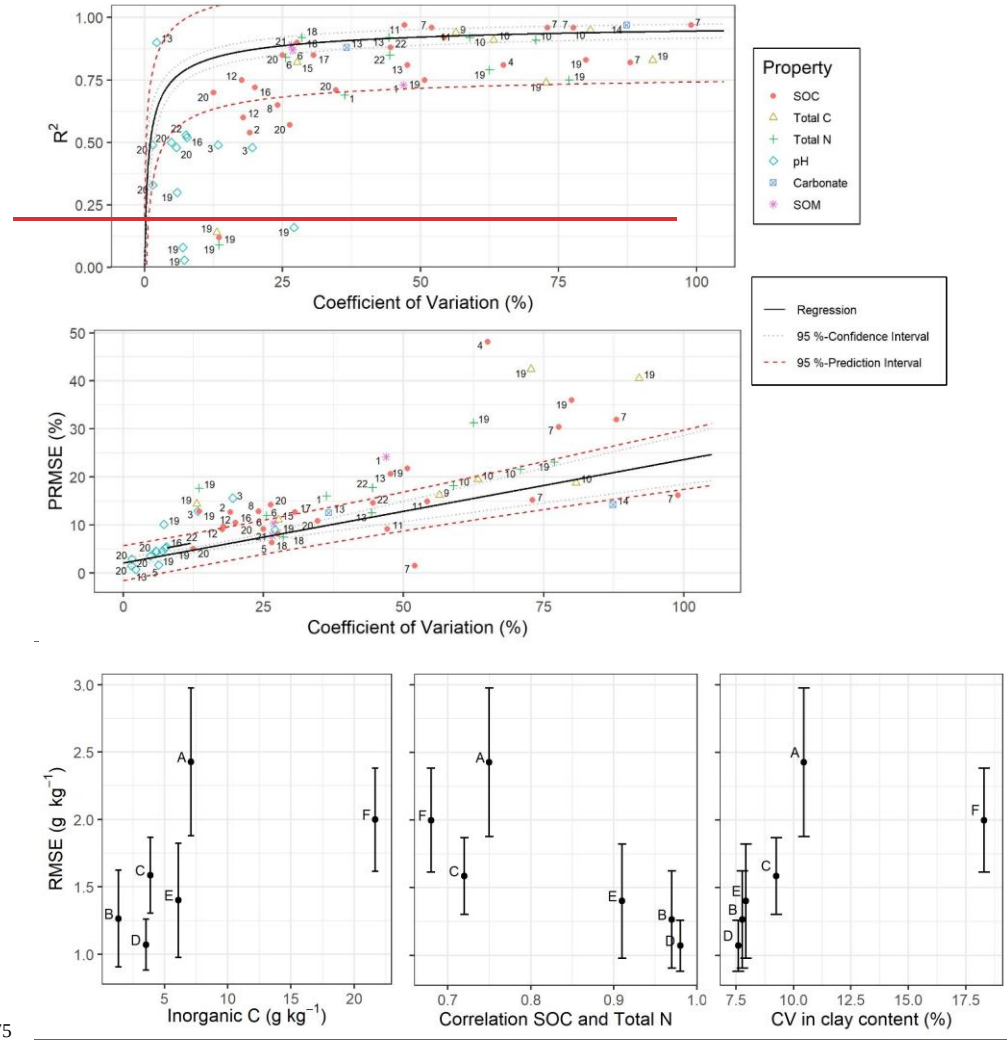


Figure ~~7: R^2 and percental root~~6: Root mean square error (~~PRMSE~~RMSE) for models with SOC for each field (A-F) in dependence of mean inorganic C content, Pearson's correlation coefficient between SOC and Total N and coefficient of variation ~~of 64 local models from 22 peer reviewed publications compared with~~ (CV) in clay content. Error bars show the ~~regressions found in this study.~~

~~Only local models with a coefficient of variation < 105 % were plotted. The numbers~~standard deviation of the ~~points correspond~~repeats in the cross-validation.

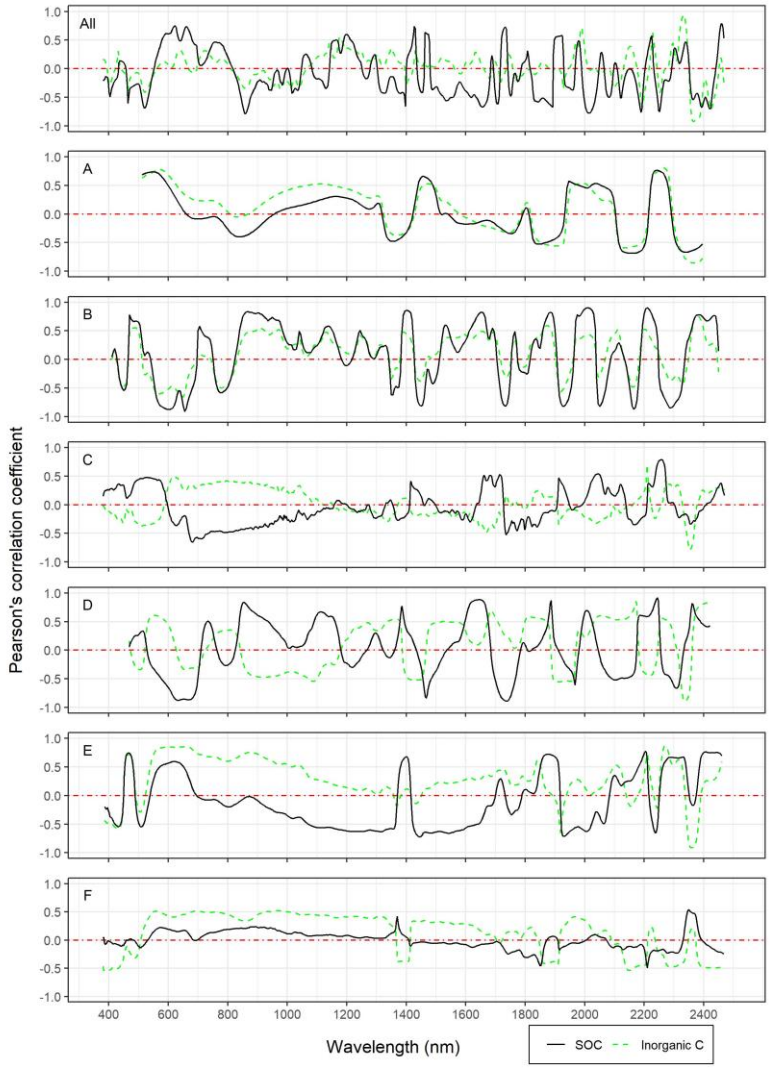


Figure 7: Correlation graphs between spectral variables at each wavelength and SOC as well as inorganic C for the combined dataset (All) and the individual fields (A-F). The spectra were pre-processed according to the ~~numbered publications in Table 3~~chosen models for SOC.

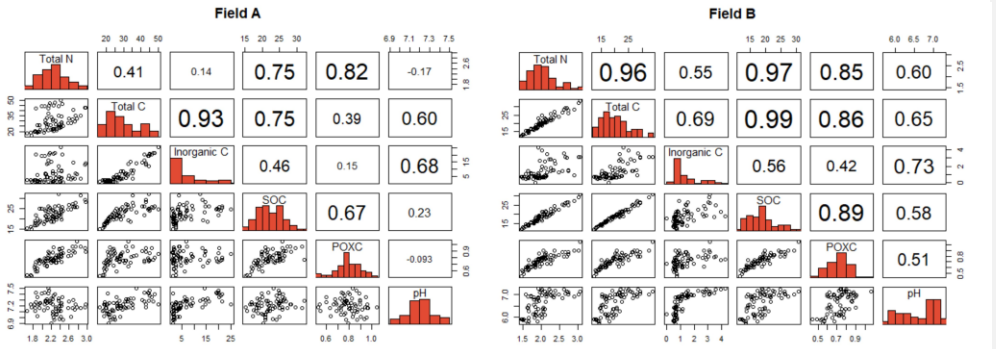


Figure 8: Pairwise correlation matrices between target soil properties and inorganic C for a field with weak (Field A) and field with strong correlations (Field B) between the target variables. The correlation matrices for all fields can be found in the supplementary material (Fig. S5).

Code and data availability:

Data and R-codes ~~are~~will be made available upon publication on ~~request~~a Zenodo repository.

Author contributions:

All co-authors conceptualized the paper. SO collected the data and conducted the analysis with the help of LS. SO wrote the original draft and all co-authors participated in the writing and editing of the manuscript.

Competing interests:

The contact author has declared that none of the authors has any competing interests.

Acknowledgements:

The authors gratefully acknowledge Daniela Fischer and Patrick Neuhaus for their engaged support during the lab work. We also warmly thank Philipp Baumann for sharing his knowledge about the handling and modeling of spectral data. Lastly, we would like to thank the two anonymous reviewers for their detailed and very constructive comments and suggestions.

Financial support:

This study was funded by the Land Systems and Sustainable Land Management Group from the Institute of Geography at the University Bern.

hat formatiert: Schriftart: 10 Pt.