



## RESEARCH ARTICLE

# A LUCAS-based mid-infrared soil spectral library: Its usefulness for soil survey and precision agriculture<sup>#</sup>

Matthias Leenen<sup>1</sup> | Stefan Pätzold<sup>1</sup> | Gergely Tóth<sup>2,3</sup> | Gerhard Welp<sup>1</sup>

<sup>1</sup>Institute of Crop Science and Resource Conversation (INRES), Soil Science and Soil Ecology, University of Bonn, Bonn, Germany

<sup>2</sup>Institute of Advanced Studies, Kőszeg, Hungary

<sup>3</sup>Institute for Soil Sciences, Agricultural Research Centre, Budapest, Hungary

## Correspondence

Stefan Pätzold, Institute of Crop Science and Resource Conversation (INRES), Soil Science and Soil Ecology, University of Bonn, Nussallee 13, 53115 Bonn, Germany.  
Email: [s.paetzold@uni-bonn.de](mailto:s.paetzold@uni-bonn.de)

<sup>#</sup>This article has been edited by Horst Gerke.

## Funding information

German Federal Ministry of Education and Research

## Abstract

**Background:** Mid-infrared spectroscopy (MIRS) is commonly recognized as a rapid and high throughput measurement technology for numerous soil properties, given that appropriate prediction models are calibrated. Soil spectral libraries (SSL) may reduce effort and costs for MIRS practical application.

**Aims:** To calibrate MIR-SSL-based prediction models for soil properties and to test their applicability to independent sample sets at regional scale (e.g., for soil survey) and at field scale (i.e., for precision agriculture, PA).

**Methods:** Spectra of 1013 arable topsoil samples of the European Land Use/Land Cover Area Frame Survey 2009 (LUCAS) from Belgium, the Netherlands, Luxembourg, and Germany formed the basis for the MIR-SSL. Leave-one-out cross-validation (LOOCV) via partial least squares regression served to calibrate (1) generic prediction models including all samples, and (2) stratified models for different parent materials. Test-set validation (TSV) was conducted on samples from independent campaigns at (1) regional scale with a sample set from Schleswig-Holstein (Germany;  $n = 385$ ) and (2) field scale for four individual fields in Germany ( $n = 513$ ).

**Results:** Generic LOOCV models successfully predicted soil organic carbon, total nitrogen, sand, silt, clay, carbonate, and pH. Calibration for available nutrients failed. The TSV was successful for the regional sample set for all variables ( $2.5 \leq \text{RPIQ} \leq 5.9$ ), except for carbonate ( $\text{RPIQ} = 0$ ). At field scale, the validation was highly variable for different sites and parameters. Stratified models using soil parent material as auxiliary variable improved only occasionally the applicability at field scale, that is, on single fields and only for clay and carbonate.

**Conclusions:** Although the MIR-SSL in its present state cannot be recommended for nutrient management, it provides valuable support for soil survey and PA.

## KEYWORDS

diffuse reflectance spectroscopy, parent material, partial least squares regression, precision agriculture, soil sensing, within-field soil heterogeneity

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Journal of Plant Nutrition and Soil Science published by Wiley-VCH GmbH.

## 1 | INTRODUCTION

There is large demand for sensor techniques that provide consistent, reliable, and inexpensive information on soil properties (Lobsey et al., 2017; White et al., 2021). An application that relies on such soil data is precision agriculture (PA). In PA, information on soil properties with high spatial resolution is required to characterize the occurring within-field soil heterogeneity and to subsequently address this heterogeneity by spatially adapted management measures, such as variable rate fertilization (Mulla & Khosla, 2016). Yet, the acquisition of soil information with conventional laboratory analyses at the required spatial resolution for PA is too expensive and time consuming. In addition, spatially high-resolution information is also required for other applications such as environmental monitoring, modeling, and mapping (Viscarra Rossel et al., 2006).

With regard to rapid and precise quantification of soil properties, different soil sensing techniques made enormous progress in the last few decades (Kuang et al., 2012; Viscarra Rossel & Bouma, 2016). The optical measurement techniques, diffuse reflectance visible and near infrared (vis-NIRS) and mid-infrared spectroscopy (MIRS), have proven to predict various soil properties at a level of preciseness and accuracy that principally meets the requirements of various data demanding applications such as PA (reviewed by Viscarra Rossel et al., 2006).

However, there are substantial differences between vis-NIRS and MIRS, which cannot only be traced back to different spectral ranges (vis-NIRS from 25,000 to 4000  $\text{cm}^{-1}$  and MIRS from 4000 to 400  $\text{cm}^{-1}$ ). In the MIR range, intense fundamental molecular frequencies occur, which are related to diverse organic and inorganic soil components, while the vis-NIR region is dominated by overtones and combinations of fundamental bands. Thus, in the MIR range, stronger and more distinct features appear (Viscarra Rossel et al., 2006). A number of studies have highlighted that MIRS outperforms vis-NIRS in predicting soil properties at least in temperate regions (Bellon-Maurel & McBratney, 2011; Reeves III, 2010; Vohland et al., 2014).

In this context, major advantages of vis-NIRS are (1) that in contrast to MIRS less effort for sample preparation for spectra acquisition is required and (2) that vis-NIRS has already proven to be generally appropriate for mobile measurements in the field (Mouazen et al., 2007; Rodionov et al., 2015). These are important prerequisites for the envisaged application in PA. Portable MIRS instruments have only recently become available and only few studies with satisfactory results recording MIR spectra in the field are published up to now (e.g., Dhawale et al., 2015; Izaurrealde et al., 2013; Ji et al., 2016). However, once adapted, portable MIRS will provide an attractive opportunity to collect soil data for PA.

Many studies investigated in detail the relationship between diffuse reflectance spectra and soil properties from underlying chemical perspective (e.g., Bornemann et al., 2010; Ellerbrock & Gerke, 2021). Nevertheless, practical application in lab routine analyses or in field studies is still subjected to methodological restrictions. In order to derive soil properties from MIR spectra (irrespective of benchtop or portable instruments), modeled functions based on multivariate statistics are used, such as the commonly applied partial least squares regression

(PLSR) method. To calibrate respective prediction models, corresponding conventional soil analytical data (reference data) are needed. However, the collection of conventional soil data is time and cost intensive. Hence, calibrating individual models for single fields requires efforts that diminish the advantages of MIRS over conventional methods. In general, soil spectral libraries (SSL) reduce this effort, provided that the performance of SSL models fulfills the demands of the envisaged application.

The potential of vis-NIR- and MIR-SSL on global, continental, national, and local scales has already emphasized in several studies (vis-NIR-SSL; e.g., Araújo et al., 2014; Tóth et al., 2013; Viscarra Rossel et al., 2016; Wetterlind & Stenberg, 2010; MIR-SSL: e.g., Breure et al., 2022; Clairotte et al., 2016; Hicks et al., 2015; Terhoeven-Urselmans et al., 2010; Viscarra Rossel et al., 2008; Wijewardane et al., 2018). To our knowledge, the largest vis-NIR-SSL in Europe is based on approximately 20,000 topsoil samples from the European Land Use/Land Cover Area Frame Survey (LUCAS) which was carried out in 2009 (Orgiazzi et al., 2018; Tóth et al., 2013). The LUCAS survey is particularly characterized by consistency of sampling and analysis of basic chemical and physical soil properties (Tóth et al., 2013). The most comprehensive MIR-SSL ( $n = 4329$  MIR spectra) against a geochemical background at a European scale was generated in the framework of the Geochemical Mapping of Agricultural Soils and Grazing Land of Europe (GEMAS) project (Reimann et al., 2012).

Numerous studies were conducted on MIR-SSL with detailed consideration of spectral response of the underlying soil properties (e.g., Terhoeven-Urselmans et al., 2010; Viscarra Rossel et al., 2008). However, there is still a need for research on the applicability of MIR-SSL-based prediction models. Especially under European agro-ecological conditions, the applicability of MIR-SSL-based prediction models to independent sample sets from smaller scales has not yet been studied in detail. Against this background, the goal of this study was to test the applicability of MIRS-SSL-based prediction models using the entire (generic prediction models) or selected set of MIR-SSL samples (stratified prediction models) to characterize soil properties of independent sample sets that originated from different institutions and projects in order to serve for public soil survey (regional scale) or PA purposes (field scale).

In this study, we recorded MIR spectra of arable topsoil samples from Belgium, the Netherlands, Luxembourg, and Germany which were collected during the LUCAS survey 2009 to build the MIR-SSL. The international LUCAS survey was conducted by the European Union and therefore ensures the same standards beyond methodological decisions. To the best of our knowledge, this study represents the first study using LUCAS samples and data for MIR spectroscopy.

The hypotheses of this study are as follows: (1) generic prediction models calibrated based on the spectra and reference data of the MIR-SSL (LUCAS sample set) via PLSR are capable of predicting basic physical and chemical soil properties with similar accuracy to conventional soil analytical methods. This is independent of the scale on which these prediction models are applied. (2) Calibrating stratified models by using the soil parent material as auxiliary variable can improve the applicability of MIR-SSL at local scale for predictions of parameters that are





**TABLE 1** Parent materials and reference soil groups observed in field investigations (acc. to IUSS Working Group, 2015) at the four individual fields under study

Location	Parent material	Reference soil groups
Ascheberg	Cretaceous marls partially covered by Saalian glacial till, aeolian sand, and fluvial sediments	Cambisols, Stagnosols
Bölingen	Pleistocene periglacial slope deposits consisting of Weichselian loess with variable amounts of weathered Devonian sand-, silt-, and claystones and scattered Tertiary basalt bombs	Cambisols, Luvisols, Stagnosols
Görzig	Weichselian loess	Chernozems, Regosols
Wilmersdorf	Glacial till, partly covered with Weichselian glaciofluvial sands	Arenosols, Retisols, Luvisols, Planosols

more or less directly related to soil parent material. As a consequence, applicability of an MIR-SSL for PA could be promoted by an improved capability to characterize within-field soil heterogeneity.

## 2 | MATERIALS AND METHODS

### 2.1 | Soil samples

The Institute for Environment and Sustainability of the Joint Research Center (JRC) in Ispra (Italy) provided subsamples of 1013 arable topsoils (0–20 cm) from the LUCAS survey in 2009 which were spatially distributed over Germany ( $n = 895$ ), Belgium ( $n = 31$ ), the Netherlands ( $n = 87$ ), and Luxembourg ( $n = 1$ ). In the LUCAS survey, the number of samples for Belgium, the Netherlands, and Luxembourg was generally lower than for Germany. In addition, the proportion of arable samples for these countries was lower than for Germany (Tóth et al., 2013). All soil samples had already been analyzed by conventional methods as part of the LUCAS survey (for details see Tóth et al., 2013). In this study, 2 g of each LUCAS soil sample were submitted to MIR spectra recording.

To predict properties of independent target samples, independent surveys at regional and field scales without direct spatial coincidence with the LUCAS survey were selected. Therefore, 385 arable topsoil samples (0–30 cm) were provided by the State Agency for Agriculture, Environment and Rural Areas (LLUR) of Schleswig-Holstein, Germany. This archived sample set representatively covered the entire arable land of the federal state Schleswig-Holstein (S-H; 15,800 km<sup>2</sup>). Compared to the LUCAS sample set, the spatial sample density for S-H was higher for the LLUR sample set (LUCAS:  $n = 36$  samples from S-H). Geographic coordinates were not available for the individual LLUR samples for data protection reasons.

To test the MIR-SSL models at field scale, 513 topsoil samples (0–30 cm) from four arable fields (2.6–30 ha) in different German regions were investigated (in the following denoted as field scale validation [FSV] sample set). These fields were located in Ascheberg (51°45′32″N, 7°34′47″E;  $n = 115$ ), Bölingen (50°34′1″N, 7°5′5″E;  $n = 71$ ), Görzig (51°39′50″N, 11°59′48″E;  $n = 112$ ), and Wilmersdorf (53°6′39″N, 13°54′21″E;  $n = 217$ ) and were sampled within the “BonaRes-I4S” PA project at high spatial resolution to characterize

within-field soil heterogeneity. The parent materials and reference soil groups observed in the field are summarized in Table 1.

Figure 1 localizes the four FSV sample sets and the federal state S-H (LLUR) against the background of the European Parent Material Map (EPM; part of the Soil Regions Map of the European Union and Adjacent Countries 1:5 000 000, freely accessible from the German Federal Institute for Geosciences and Natural Resources [BGR] via web map service; Bundesanstalt für Geowissenschaften und Rohstoffe, 2005). For the LLUR samples, no geographic information was available (see above). The spatial distance of the FSV sample sets to the nearest LUCAS sample was between 8 (Ascheberg) and 23 km (Wilmersdorf).

### 2.2 | Conventional soil analyses

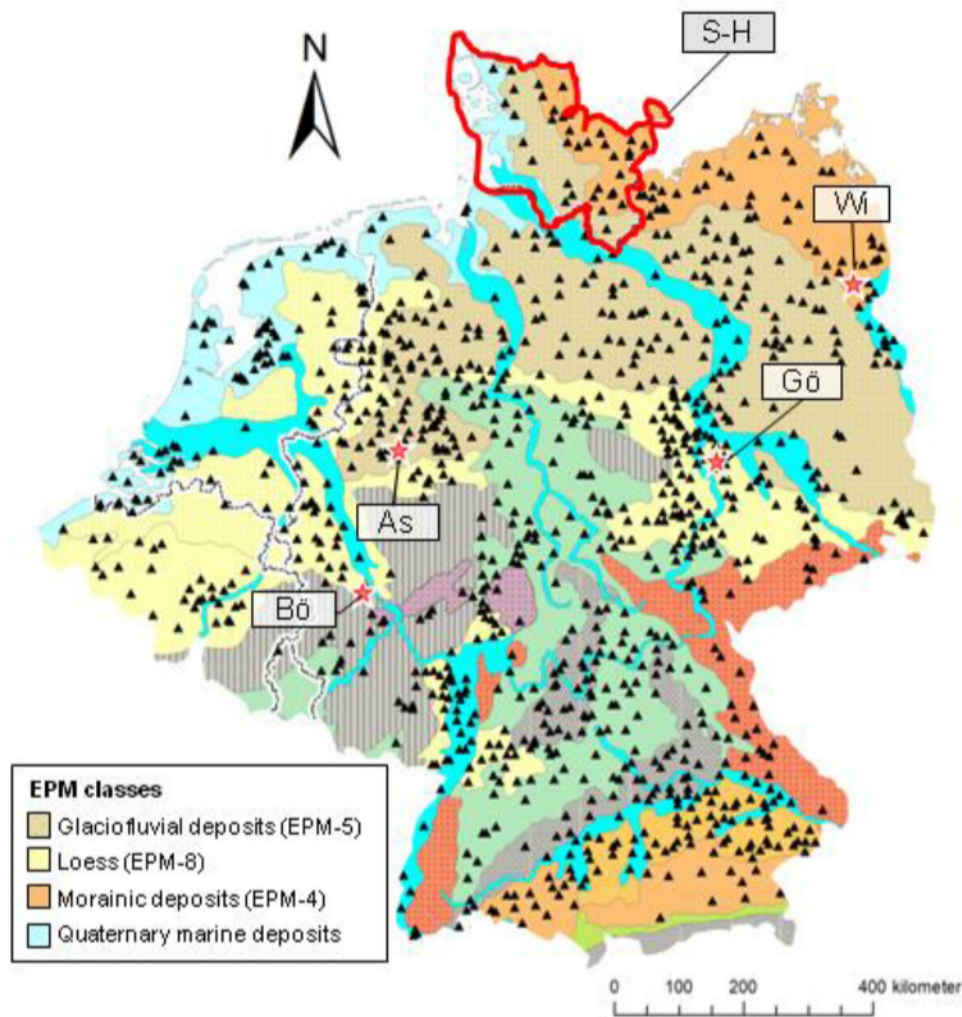
The conventional soil analyses of the LUCAS sample set were performed by a single laboratory on behalf of the JRC and were made available online on the website of the European Soil Data Centre (ESDAC, 2013). The LLUR provided analytical data along with the 385 soil samples from Schleswig-Holstein, which were in parts supplemented and extended by own analyses.

The FSV samples were analyzed in the frame of the BonaRes-I4S project. For all sample sets, the applied analytical methods and a summary of statistics for the conventional soil analyses are presented in Table 2.

### 2.3 | Spectroscopic measurements

In preparation for spectral measurements, 2 g of air dried, sieved (<2 mm) soil per sample were ground in a ball mill to a particle size <500 μm. About 20 mg of ground soil were filled in fivefold repetition into microtiter plates and compacted with a plunger to ensure a plain and dense surface. Diffuse reflectance mid-infrared Fourier transform (DRIFT) spectroscopy was performed in the laboratory with a Bruker Tensor 27 spectrometer, equipped with an HTS-XT device for automated high-throughput measurement (Bruker Optik, Ettlingen, Germany). The spectrometer operated with a liquid N<sub>2</sub>-cooled mercury-cadmium telluride (MCT) detector and a broadband KBr





**FIGURE 1** Location of data sources displayed on the section of the European Parent Material (EPM) Map. The map was used to stratify the LUCAS sample set for subsequent model calibration. Black triangles indicate investigated arable topsoil samples of the Land Use/Land Cover Area Frame Survey (LUCAS) survey ( $n = 1013$ ). The federal state Schleswig-Holstein (S-H), from where the LLUR sample set originates, is framed in red. Locations of LLUR samples ( $n = 385$ ) could not be illustrated because geographic coordinates were not available for privacy reasons. The four individual fields under study were As = Ascheberg, Bö = Bölingen, Gö = Görzig, and Wi = Wilmersdorf. The map legend contains only the EPM classes being relevant for this study. For further information, refer to Bundesanstalt für Geowissenschaften und Rohstoffe (2005)

beam splitter. For each repetition, 120 scans at a resolution of  $4 \text{ cm}^{-1}$  and a spectral range of  $7500\text{--}550 \text{ cm}^{-1}$  were carried out.

## 2.4 | Spectra treatments and calibration of prediction models

The calibration of prediction models to derive soil properties from MIR spectra was carried out with the commercial spectroscopy software OPUS Quant (Bruker Optik) after averaging the repetitions of (calculated) absorbance spectra. This study was focused on the MIR range. Therefore, the spectral range was narrowed to  $3800\text{--}550 \text{ cm}^{-1}$ . Prior to model calibration, a principal component analysis (PCA) was carried out for the first two scores (PC-1 and PC-2) with centered data and a correlation matrix using spectra of all sample sets (LUCAS, LLUR,

Ascheberg, Bölingen, Görzig, and Wilmersdorf). The PCA was carried out in order to identify similarities or differences in spectral characteristics between the investigated sample sets. Model calibration was performed using partial least squares regression (PLSR), which was based on the PLS1 algorithm (Brereton, 2018) and leave-one-out cross validation (LOOCV), a default method in OPUS Quant (Ludwig et al., 2019). PLSR was used since it is one of the most common and reliable chemometric techniques for calibration of spectroscopic prediction models and has been shown to be suitable to derive various soil properties from soil spectra (e.g., Soriano-Disla et al., 2014; Viscarra Rossel et al., 2006).

The number of latent variables with smallest root mean square error of cross-validation (RMSECV) was selected for each model. Analogous to the study of Ludwig et al. (2019), the default maximum number of 10 latent variables in OPUS Quant was used in order to



**TABLE 2** Statistical summary of the observed soil properties for all investigated sample sets

	LUCAS <i>n</i> = 1013	LLUR <i>n</i> = 385	Ascheberg <i>n</i> = 115	Bölingen <i>n</i> = 71	Görzig <i>n</i> = 110	Wilmersdorf <i>n</i> = 217 <sup>a</sup> /59 <sup>b</sup>
<b>Sand (%)<sup>c</sup></b>						
Mean (median)	41 (37)	55 (62)	62 (67)	12 (10)	17 (16)	64 (63) <sup>b</sup>
Min-max	1–97	1–96	21–80	7–23	13–33	47–82 <sup>b</sup>
SD; IQR	30; 60	27; 46	14; 16	4; 4	4; 4	8; 10 <sup>b</sup>
<b>Silt (%)<sup>c</sup></b>						
Mean (median)	42 (42)	33 (29)	14 (14)	57 (59)	61 (62)	26 (26) <sup>b</sup>
Min-max	1–88	2–78	9–22	37–70	49–66	15–45 <sup>b</sup>
SD; IQR	23; 41	19; 31	3; 4	8; 12	3; 3	7; 6 <sup>b</sup>
<b>Clay (%)<sup>c</sup></b>						
Mean (median)	17 (16)	12 (8)	22 (17)	30 (26)	20 (20)	8 (8) <sup>b</sup>
Min-max	1–56	0–48	9–55	18–57	15–23	2–18 <sup>b</sup>
SD; IQR	11; 16	9; 12	12; 13	9; 12	1; 1	3; 4 <sup>b</sup>
<b>SOC (g kg<sup>-1</sup>)<sup>d</sup></b>						
Mean (median)	16.7 (15)	15.7 (14.1)	17.8 (16.2)	12.6 (11.9)	16.7 (16.4)	8.9 (8.7)
Min-max	4.3–39.5	1.8–39.7	12.2–32.9	9.4–19.0	13.6–21.9	4.6–20.0
SD; IQR	6.8; 8.0	6.9; 7.9	4.4; 5.9	2.2; 2.9	1.7; 2.0	2.3; 2.4
<b>N<sub>total</sub> (g kg<sup>-1</sup>)</b>						
Mean (median)	1.6 (1.5) <sup>e</sup>	1.4 (1.3) <sup>f</sup>	1.7 (1.5) <sup>f</sup>	1.3 (1.3) <sup>f</sup>	1.4 (1.4) <sup>f</sup>	0.9 (0.9) <sup>f</sup>
Min-max	0.4–3.9	0.2–4.1	1.1–3.3	1.0–1.9	1.2–1.7	0.4–2.0
SD; IQR	0.5; 0.6	0.5; 0.5	0.5; 0.6	0.2; 0.2	0.1; 0.1	0.2; 0.3
<b>CaCO<sub>3</sub> (g kg<sup>-1</sup>)<sup>g</sup></b>						
Mean (median)	19 (1)	3 (0)	11 (0)	0 (0)	0.3 (0)	8 (0)
Min-max	0–715	0–46	0–191	0–0	0–9	0–81
SD; IQR	57; 6	9; 0	28; 7	n.d.; n.d.	1; 0	17; 6.5
<b>pH (CaCl<sub>2</sub>)<sup>h</sup></b>						
Mean (median)	6.24 (6.37)	5.73 (5.60)	6.29 (6.28)	6.01 (6.06)	6.46 (6.50)	6.24 (6.20)
Min-max	3.42–7.61	3.60–7.60	4.48–7.42	5.08–6.53	5.56–7.34	4.10–7.80
SD; IQR	0.86; 1.39	1.00; 1.70	0.81; 1.59	0.32; 0.27	0.39; 0.66	1.08; 2.10
<b>P<sub>avl</sub> (mg kg<sup>-1</sup>)</b>						
Mean (median)	58 (54) <sup>i</sup>					
Min-max	0–212	n.d.	n.d.	n.d.	n.d.	n.d.
SD; IQR	30; 38					
<b>K<sub>avl</sub> (mg kg<sup>-1</sup>)</b>						
Mean (median)	221 (185) <sup>j</sup>					
Min-max	0–1724	n.d.	n.d.	n.d.	n.d.	n.d.
SD; IQR	163; 177					

Abbreviations: CaCO<sub>3</sub>, calcium carbonate; IQR, interquartile range; n.d., not determined; N<sub>total</sub>, total nitrogen; K<sub>avl</sub>, plant available potassium; P<sub>avl</sub>, plant available phosphorus; SOC, soil organic carbon.

<sup>a,b</sup>Two sample subsets: only sample subset (b) was analyzed for particle size distribution.

<sup>c</sup>Combined sieve/pipette method (ISO, 1998a).

<sup>d</sup>Determined by the difference of total carbon (elemental analysis; ISO, 1995a) and inorganic carbon.

<sup>e</sup>Modified Kjeldahl (ISO, 1995b).

<sup>f</sup>Elemental analysis (ISO, 1998b).

<sup>g</sup>Gas volumetric determination (ISO, 1994b).

<sup>h</sup>Potentiometric measurement (ISO, 1994a).

<sup>i</sup>Olsen P (ISO, 1994c).

<sup>j</sup>NH<sub>4</sub>OAc (USDA-NRCS, 2004).



avoid overfitting. To eliminate the physical effect of light scattering and to improve model performance, different mathematical treatments of the spectra were tested: none, first derivative, standard normal variate (SNV), multiplicative scatter correction (MSC), first derivative + SNV, first derivative + MSC, and second derivative. The default settings in OPUS Quant were used for all of these treatments. For spectra treatments that included the calculation of derivatives, this meant 17 data points for smoothing. The treatment method with best predictive power was selected for each prediction model. The predictive power was evaluated by calculating  $R^2$ , the RMSECV or root mean square error of prediction (RMSEP), the bias, and the ratio of performance to interquartile range (RPIQ). The RPIQ relates the prediction error to the variation of observed values and was calculated from interquartile distance of the observed values divided by RMSECV or RMSEP (Bellon-Maurel et al., 2010). We used the RPIQ instead of the widely used RPD because RPIQ relies on the normality assumption of the variable distribution over the sample set, whereas these values obtained by the RPD on skewed sample values are misleading (Bellon-Maurel et al., 2010). In general, as RPIQ values increase, so does the performance of the model. The threshold of RPIQ  $<1.89$ , which was used in the study of Ludwig et al. (2019), was used to unravel unsuccessful predictions. However, whenever predictions were evaluated by these values, it should be noted that the thresholds are not based on any theory or experiment and the usefulness of a model should always be defined in its specific context (Ludwig et al., 2019). With regard to the hypotheses of the current study, the threshold of 1.89 was considered as an appropriate criterion for evaluating the predictions.

## 2.5 | Generic and stratified modeling

The spectra and reference data of the LUCAS sample set were used to calibrate generic and stratified prediction models. The generic models served to predict soil properties in all independent validation sample sets (LLUR and FSV). For validation at field scale (FSV sample sets), that is, for the envisaged future PA application, stratified models were calibrated with consideration of the parent material as auxiliary variable. Unfortunately, this was not feasible at the regional scale because coordinates were not available for the LLUR samples. Prior to model calibration, the degree of collinearity of the reference data was verified by determining the Pearson correlation coefficient ( $r$ ) after inspection for normality. The calibration of generic prediction models was carried out on all 1013 LUCAS spectra and corresponding reference data, while stratified models were calibrated separately for different soil parent materials (Bundesanstalt für Geowissenschaften und Rohstoffe, 2005; see Figure 1). For this purpose, the respective EPM classes according to the EPM map classification (Figure 1) were assigned to the LUCAS samples and FSV sample sets. In total, three different EPM classes were relevant for the four FSV sample sets. Though, there were differences in the number of LUCAS samples assigned to the EPM classes: (1) glaciofluvial deposits (EP-5,  $n_{\text{LUCAS}} = 206$ ), (2) loess (EPM-8,  $n_{\text{LUCAS}} = 180$ ), and (3) morainic deposits (EPM-4,  $n_{\text{LUCAS}} = 95$ ). Separate (stratified) prediction models for sand, silt, clay, and carbonate ( $\text{CaCO}_3$ )

were calibrated for each of these three LUCAS (sub)sample sets. These soil properties were chosen because they are closely related to parent material. The EPM classes derived from the map (Figure 1) were consistent to observed parent material by field investigations (Table 1) with only small variation in detail.

## 3 | RESULTS

### 3.1 | Characteristics of the mid-infrared spectra

The calibration samples (LUCAS) revealed large variability in spectral characteristics (e.g., intensities and distribution of peaks). This became particularly clear when the difference between minimum and maximum absorbance in the LUCAS spectra was investigated. For example, in the range of  $3800\text{--}3600\text{ cm}^{-1}$ ,  $2900\text{--}2800\text{ cm}^{-1}$ , and  $2600\text{--}2400\text{ cm}^{-1}$ , obvious differences were found in peak intensities (Figure 2A). A large diversity of spectral characteristics could also be documented on the basis of the first two PCA scores (PC-1 and PC-2; Figure 2B). No striking deviations in peak appearance and spectral domain were found between the LUCAS and the LLUR validation samples. This became obvious by comparing the averaged spectra and the PCA scores (Figure 2A,B). There also appeared no obvious differences in spectral domain between the LUCAS and the FSV sample sets. Only very few samples from Ascheberg, Bölingen, and Wilmersdorf were outside the spectral domain of the LUCAS sample set. Overall, Wilmersdorf samples were represented by the smallest number of LUCAS samples.

Within the FVS sample set, the four sub-sets from the individual fields differed in terms of spectral characteristics. For example, differences in the occurrence and intensity of peaks in the range of  $3800\text{--}3600\text{ cm}^{-1}$ ,  $2900\text{--}2800\text{ cm}^{-1}$ , and  $2600\text{--}2400\text{ cm}^{-1}$  were observed (Figure 2A). The diversity between these fields was also reflected by different spectral domains they covered in the scores plot (Figure 2B).

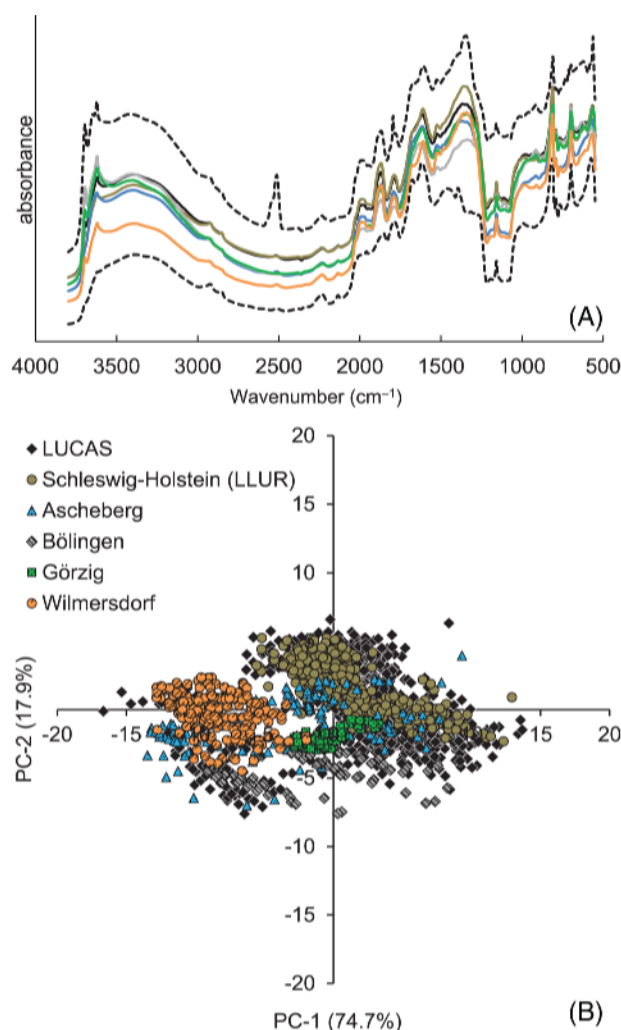
### 3.2 | Calibration of generic LUCAS-based prediction models

The indicators of performance for calibration of generic LUCAS prediction models are listed in Table 3. Large  $R^2$  and RPIQ were achieved for six out of nine soil properties (entire range of values:  $R^2 = 0.23\text{--}0.98$  and RPIQ = 0.7–9.4) with largest  $R^2$  for  $\text{CaCO}_3$  and RPIQ for sand. The RPIQ for  $\text{CaCO}_3$  was small (RPIQ = 0.7) and below the threshold for successful prediction according to Ludwig et al. (2019). For plant available phosphorus ( $P_{\text{avl}}$ ) and potassium ( $K_{\text{avl}}$ ),  $R^2$  and RPIQ were small and also below the threshold for successful prediction ( $R^2 = 0.23\text{--}0.44$  and RPIQ = 1.4–1.5). With regard to particle size, smallest RMSE was achieved for clay content.

It must be taken into account that analyzed soil properties of the LUCAS sample set were related to each other. A strong correlation was found, for example, for silt and sand ( $R^2 = -0.95$ ) as well as for total nitrogen ( $N_{\text{total}}$ ) and soil organic carbon (SOC;  $R^2 = 0.86$ ).







**FIGURE 2** (A) Averaged mid-infrared spectra of the individual investigated sample sets and minimum and maximum absorbance of the 1013 Land Use/Land Cover Area Frame Survey (LUCAS) mid-infrared spectra. (B) Scores plot of the principal component analysis (first two principal components: PC-1 vs. PC-2) for untreated (raw) spectra of all sample sets under study

### 3.3 | Validation of generic LUCAS prediction models with an independent regional sample set (LLUR samples)

The results of the validation of generic LUCAS prediction models (Table 3) with the independent regional LLUR sample set are shown in Figure 3. This validation was performed for all properties under study except for  $P_{\text{avl}}$  and  $K_{\text{avl}}$  because the calibration of generic LUCAS models had already failed.

As shown in Figure 3, large  $R^2$  and RPIQ were achieved for sand, silt, clay, SOC,  $N_{\text{total}}$ , and pH ( $\text{CaCl}_2$ ). The  $R^2$  ranged from 0.83 (silt) to 0.92 (SOC). Large RPIQ ( $>2.5$ ) was achieved, except for  $\text{CaCO}_3$ . However, in most cases,  $R^2$  and RPIQ for validation were smaller than for calibration, and predictions were more biased. Only SOC and pH ( $\text{CaCl}_2$ ) revealed larger  $R^2$  and RPIQ.

**TABLE 3** Model performance indicators of leave-one-out cross validation (LOOCV) for the Land Use/Land Cover Area Frame Survey (LUCAS) sample set under investigation ( $n = 1013$ ). All samples originate from arable topsoils from Belgium, the Netherlands, Luxembourg, and Germany of the LUCAS topsoil survey in 2009

Soil property	Preprocessing	LV	Leave-one-out cross validation			
			$R^2$	RMSECV	Bias	RPIQ
Sand (%)	2 <sup>nd</sup> der.	10	0.96	6.4	0.0	9.4
Silt (%)	2 <sup>nd</sup> der.	10	0.93	6.3	0.0	6.5
Clay (%)	1 <sup>st</sup> der. + SNV	9	0.95	2.4	0.0	6.7
SOC ( $\text{g kg}^{-1}$ )	1 <sup>st</sup> der.	10	0.89	2.3	0.0	3.5
$N_{\text{total}}$ ( $\text{g kg}^{-1}$ )	1 <sup>st</sup> der.	10	0.89	0.18	0.00	3.3
$\text{CaCO}_3$ ( $\text{g kg}^{-1}$ )	1 <sup>st</sup> der. + MSC	9	0.98	8.3	0.0	0.7
pH ( $\text{CaCl}_2$ )	1 <sup>st</sup> der.	10	0.84	0.35	0.00	4.0
$P_{\text{avl}}$ ( $\text{mg kg}^{-1}$ )	2 <sup>nd</sup> der.	8	0.23	26.5	0.0	1.4
$K_{\text{avl}}$ ( $\text{mg kg}^{-1}$ )	2 <sup>nd</sup> der.	9	0.44	122	0	1.5

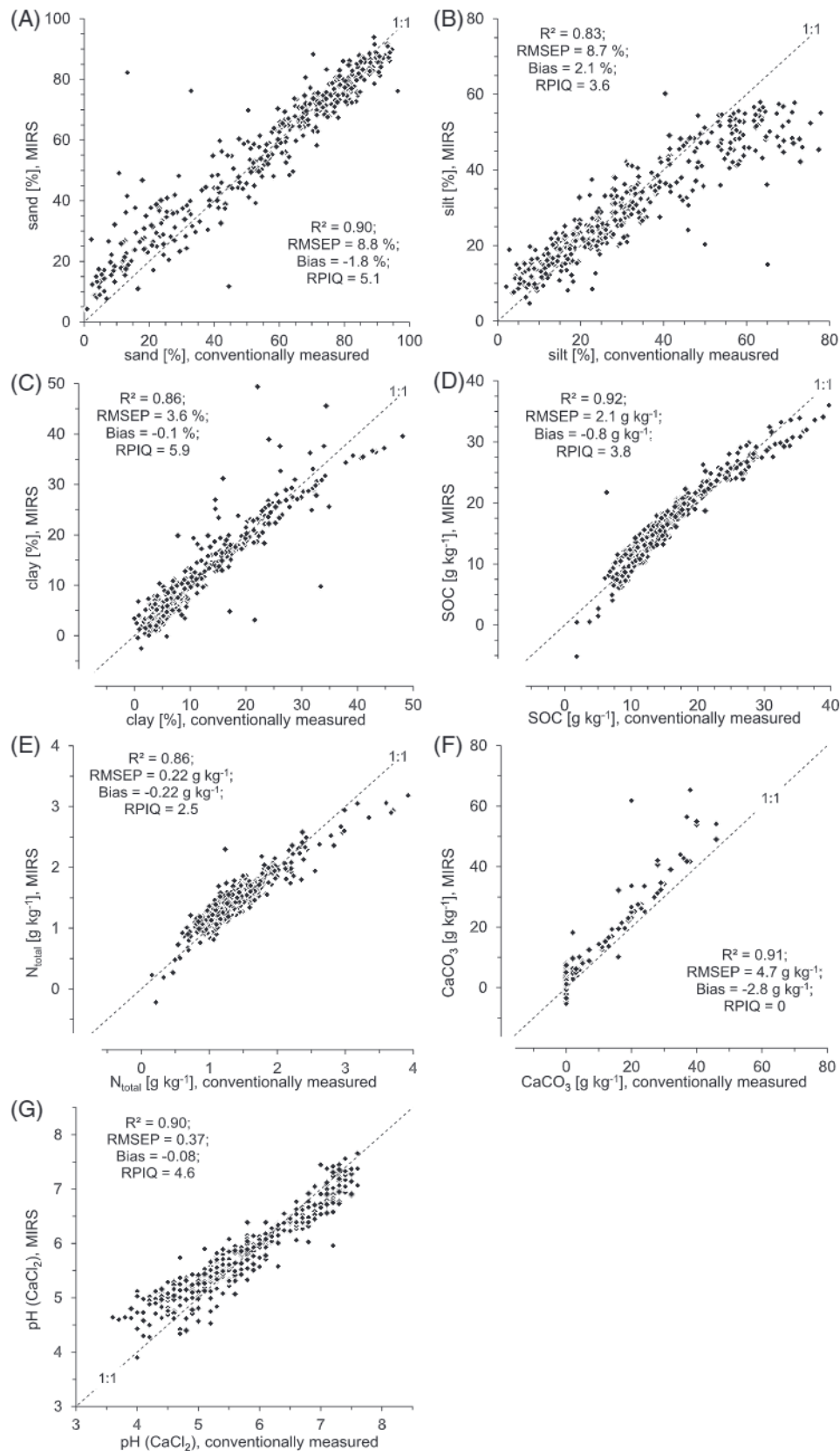
Abbreviations:  $\text{CaCO}_3$ , calcium carbonate; LV, latent variables; MSC, multiplicative scatter correction;  $N_{\text{total}}$ , total nitrogen;  $K_{\text{avl}}$ , plant available potassium;  $P_{\text{avl}}$ , plant available phosphorus;  $R^2$ , coefficient of determination; RMSECV, root mean square error of cross validation; RPIQ, ratio of performance to interquartile range; SOC, soil organic carbon; SNV, standard normal variate; 1st der., first derivative; 2nd der., second derivative.

### 3.4 | Validation of generic LUCAS prediction models with the independent local sample sets of four individual fields (FSV samples)

The generic LUCAS models served to predict contents of sand, silt, clay, SOC,  $N_{\text{total}}$ ,  $\text{CaCO}_3$  as well as pH ( $\text{CaCl}_2$ ) for the FSV sample sets (Ascheberg, Bölingen, Görzig, and Wilmersdorf) with the aim of characterizing the within-field variability of these properties. It should be noted that the FSV sample sets revealed different extents of variability (Table 2). Overall, small  $R^2$  and RPIQ were achieved for fields with low within-field variability of the respective soil property ( $R^2 = 0.00$ – $0.61$ ; RPIQ =  $0.0$ – $1.1$ ; Table 4, footnote 2). For fields and properties with more pronounced within-field heterogeneity, predictions were better ( $R^2$  from 0.55 for sand prediction at Wilmersdorf to 0.91 for clay prediction at Ascheberg). Smallest RPIQ was achieved for  $\text{CaCO}_3$  prediction at Ascheberg (RPIQ = 0.6) and largest for pH ( $\text{CaCl}_2$ ) prediction at Wilmersdorf (RPIQ = 4.5).

### 3.5 | Calibration and validation of stratified prediction models

Calibration of stratified prediction models for the three relevant EPM classes (EPM-4, EPM-5, and EPM-8) and four soil properties that are closely related to parent material (sand, silt, clay, and  $\text{CaCO}_3$ ) yielded  $R^2$  between 0.80 for silt (EPM-8) and 0.98 for  $\text{CaCO}_3$  (EPM-5 and EPM-8, results not shown). Small RPIQ was only achieved for calibration of  $\text{CaCO}_3$  prediction models (RPIQ =  $0.6$ – $1.3$ ). For all other EPM classes and soil properties, RPIQ was  $\geq 1.9$ . Overall, the stratified models for



**FIGURE 3** Independent validation of generic Land Use/Land Cover Area Frame Survey (LUCAS)-based prediction models with the regional sample set from Schleswig-Holstein (LLUR). Independent validation was performed for (A) sand, (B) silt, (C) clay, (D) soil organic carbon (SOC), (E) total nitrogen (N<sub>total</sub>), (F) calcium carbonate (CaCO<sub>3</sub>), and (G) pH (CaCl<sub>2</sub>)





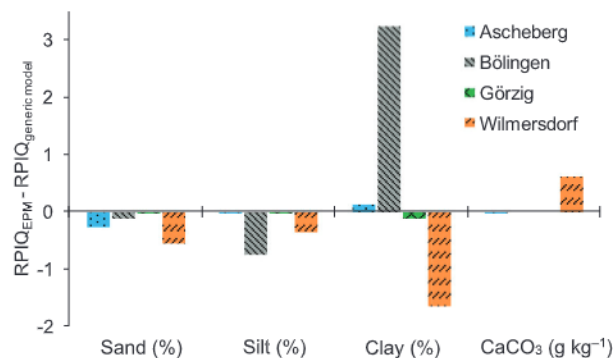
**TABLE 4** Performance of different generic Land Use/Land Cover Area Frame Survey (LUCAS)-based prediction models to predict soil properties at individual fields (field scale validation)

	Ascheberg <i>n</i> = 115	Bölingen <i>n</i> = 71	Görzig <i>n</i> = 110	Wilmersdorf <i>n</i> = 217/59 <sup>a</sup>
<b>Sand (%)</b>				
<i>R</i> <sup>2</sup>	0.79	0.00	0.04	0.55
RMSEP (%)	7.7	8.0	8.3	5.7
Bias (%)	-4.0	4.0	4.7	-1.7
RPIQ	2.1	0.5 <sup>b</sup>	0.5 <sup>b</sup>	1.8
<b>Silt (%)</b>				
<i>R</i> <sup>2</sup>	0.16	0.80	0.03	0.25
RMSEP (%)	5.8	7.7	10.0	5.9
Bias (%)	-0.8	-5.6	-8.1	0.1
RPIQ	0.7 <sup>(2)</sup>	1.6	0.3 <sup>(2)</sup>	1.0 <sup>b</sup>
<b>Clay (%)</b>				
<i>R</i> <sup>2</sup>	0.91	0.89	0.36	0.88
RMSEP (%)	4.5	3.5	1.2	1.2
Bias (%)	1.8	0.5	-0.1	-0.6
RPIQ	2.9	3.4	0.8 <sup>b</sup>	3.3 <sup>b</sup>
<b>SOC (g kg<sup>-1</sup>)</b>				
<i>R</i> <sup>2</sup>	0.86	0.78	0.61	0.72
RMSEP (g kg <sup>-1</sup> )	1.8	1.5	2.0	1.7
Bias (g kg <sup>-1</sup> )	0.0	1.1	1.6	0.0
RPIQ	3.3	1.9	1.0 <sup>b</sup>	1.4
<b>N<sub>total</sub> (g kg<sup>-1</sup>)</b>				
<i>R</i> <sup>2</sup>	0.90	0.53	0.56	0.76
RMSEP (g kg <sup>-1</sup> )	0.2	0.2	0.1	0.1
Bias (g kg <sup>-1</sup> )	-0.1	0.1	0.0	0.0
RPIQ	3.0	1.1 <sup>b</sup>	1.0 <sup>(2)</sup>	2.1
<b>CaCO<sub>3</sub> (g kg<sup>-1</sup>)</b>				
<i>R</i> <sup>2</sup>	0.96	n.d.	0.56	0.98
RMSEP (g kg <sup>-1</sup> )	11.8		1.0	6.4
Bias (g kg <sup>-1</sup> )	-5.9		-0.2	-3.5
RPIQ	0.6		0.0 <sup>b</sup>	1.1
<b>pH (CaCl<sub>2</sub>)</b>				
<i>R</i> <sup>2</sup>	0.83	0.62	0.82	0.85
RMSEP	0.49	0.28	0.20	0.47
Bias	-0.36	-0.17	0.06	-0.14
RPIQ	3.2	1.0	3.3	4.5

Abbreviations: n.d., not determined; *R*<sup>2</sup>, coefficient of determination, RMSEP, root mean square error of prediction; RPIQ, ratio of performance to interquartile range.

<sup>a</sup>Sample subset: sand, silt, and clay content were only analyzed for 59 of the 217 topsoil samples from Wilmersdorf.

<sup>b</sup>Fields with unsuccessful prediction of soil property associated with low within field heterogeneity (based on small IQR). For IQR values see Table 2.



**FIGURE 4** Effect of using stratified prediction models instead of generic prediction models on the ratio of performance to interquartile range (RPIQ). Stratified prediction models were calibrated with Land Use/Land Cover Area Frame Survey (LUCAS) samples allocated to the map units of the European Parent Material (EPM) Map (see Figure 1). Generic prediction models were calibrated with the entire LUCAS sample set (*n* = 1013). The effect of using stratified prediction models was tested for contents of sand, silt, clay, and calcium carbonate (CaCO<sub>3</sub>) at the four individual fields Ascheberg, Bölingen, Görzig, and Wilmersdorf

the respective EPM classes were less accurate than the generic models (see Table 3).

Finally, these stratified LUCAS prediction models were validated by predicting soil properties at local scale. This step yielded largely differing success for the different fields (Figure 4). In comparison to the generic models, performance of predictions for sand and silt declined at any field. The only considerable improvements in RMSEP (not shown) and RPIQ were achieved for clay prediction at Bölingen and for CaCO<sub>3</sub> prediction at Wilmersdorf.

## 4 | DISCUSSION

### 4.1 | Suitability of the LUCAS sample set for building a mid-infrared spectral library

The preselection of arable soils narrowed the overall variability of soil properties. For example, SOC content varies less than it would be expected when grassland and forest soils were included. However, due to the large number of LUCAS samples under study and the size of the sampled area, soil properties revealed large variability. Thus, the samples of the MIR-SSL are considered representative for arable soils in the study region. The pronounced spectral domain and the striking spectral differences between minimum and maximum spectrum of LUCAS samples indicate a large variability of the sample characteristics. For example, variability in the spectra between 3600 and 3800 cm<sup>-1</sup> can be attributed to variability in clay mineralogy (Clark et al., 2003). The spectral variability at 2920 and 2850 cm<sup>-1</sup> is explained by differences in methyl and methylene groups of organic matter (Wijewardane et al., 2018). Peaks around 2513 cm<sup>-1</sup> in a number of LUCAS spectra indicate the presence of carbonates

(Leenen et al., 2019). Yet, most soil properties are affected by spectral overlaps from fundamentals, combinations, and overtones of other soil properties (Viscarra Rossel et al., 2008). In consequence, for most soil properties, direct annotation of spectral features to specific soil properties and quantitative interpretation is difficult or even impossible without the use of chemometric methods (e.g., PLSR). Here, it is important to note that unknown samples can only be predicted if they fall within the property range (Terhoeven-Urselmans et al., 2010). Due to the mentioned characteristics of the LUCAS MIRS-SSL, a wide applicability for basic soil properties of arable soils by use of chemometric methods is potentially given. With regard to the validation sample sets, the achieved results indicate no evidence that fields are underrepresented by the MIR-SSL which could be caused, for example, by differences in parent materials and soil development processes, or by different management practices. However, if a MIR-SSL does not adequately address the location of the samples to be predicted, methods such as spiking can be advantageous, which allow to extend the applicability of an MIR-SSL with less effort (Breure et al., 2022).

#### 4.2 | Performance of LUCAS-based MIR models to predict soil properties at supra-regional scale

The generic LUCAS prediction models for the eight soil properties with large  $R^2$ , RPIQ, and small RMSECV confirmed the applicability of the MIR-SSL at supra-regional scale. Texture prediction performed generally well; however, the superior performance for sand and clay compared to silt can be explained by the closer association with infrared active minerals which dominate distinct grain size classes (Hicks et al., 2015). Nevertheless, the rather good silt prediction ( $R^2 = 0.93$ ; RMSECV = 6.3%) was insofar surprising as it was superior to many earlier findings as reviewed by Soriano-Disla et al. (2014). The authors reported a maximum  $R^2$  of 0.84 and a median  $R^2$  of 0.63 for silt. However, indirect effects on predicted silt could be the reason for the achieved good results. The strong negative correlation between conventionally measured silt and sand contents points to this assumption. Furthermore, this was also confirmed by a comparison of the PLSR regression coefficients for sand and silt. Similar wavenumbers contributed to the prediction of sand and silt contents (data of regression coefficients are not shown).

The good results for SOC were not surprising because MIRS is sensitive to C—H, C—O, and C—N functional groups that are prevalent in soil organic matter (SOM) (Soriano-Disla et al., 2014). Best model performance in terms of  $R^2$  was achieved for  $\text{CaCO}_3$ . The RPIQ was comparatively small for  $\text{CaCO}_3$  because the LUCAS sample set included predominantly noncalcareous samples. Carbonate peaks in the MIR range are hardly influenced by other soil components and thus  $\text{CaCO}_3$  contents can usually be precisely predicted (Leenen et al., 2019; Tatzber et al., 2010). Soil pH ( $\text{CaCl}_2$ ) models performed well in this study. In general, it should be considered that soil pH is usually only indirectly predictable due to its correlation with other soil properties such as SOM composition and clay mineralogy (Soriano-Disla et al., 2014). For a global MIR-SSL, which was even more diverse regarding, for example,

soil mineralogy, Terhoeven-Urselmans et al. (2010) achieved slightly worse predictions for pH ( $\text{H}_2\text{O}$ ) ( $R^2 = 0.80$ , RMSEP = 0.75,  $n = 291$ ). Similar results for pH ( $\text{H}_2\text{O}$ ) were obtained by Wijewardane et al. (2018) for an MIR-SSL with >20,000 soil samples from the United States ( $R^2 = 0.80$ , RMSEP = 0.57), which differed from the LUCAS sample set with regard to soil forming factors and mineralogy and thus also in the spectral domain.

Since the calibration of models for  $\text{P}_{\text{avil}}$  and  $\text{K}_{\text{avil}}$  failed, our large-scale MIR-SSL cannot be considered suitable for prediction at this scale. Therefore, a diverse MIR-SSL is inferior to conventional analytical methods for available nutrients. It is known that these operationally defined nutrient fractions have weak or no spectral features in the IR region (Kuang et al., 2012), and various IR-relevant factors that control P availability impede universal predictability (Pätzold et al., 2020). Thus, the low performance of the  $\text{P}_{\text{avil}}$  and  $\text{K}_{\text{avil}}$  prediction models for this diverse large-scale sample set was not surprising. Wijewardane et al. (2018) reported similar findings for their MIR-SSL of US soils. In their study, besides PLSR, also artificial neural network was tested for model calibration, but none of the calibration methods resulted in reliable prediction models. Nevertheless, some studies reported on approximated quantitative  $\text{P}_{\text{avil}}$  prediction models (e.g., review of Kuang et al., 2012). This is because  $\text{P}_{\text{avil}}$  is under certain conditions correlated with soil properties that can be predicted via MIRS-PLSR. Using 586 topsoil samples from different arable fields with loess soils in Germany, Pätzold et al. (2020) achieved satisfactory results for MIRS-PLSR prediction of calcium-acetate-lactate extractable P (CAL-P;  $R^2 = 0.72$ , RPD = 1.9). However, the authors also demonstrated that, for example, harvest residues from pre-crops and other factors affecting SOM quality influence the performance of  $\text{P}_{\text{avil}}$  prediction. Therefore, the authors proposed a standardized sampling strategy considering management factors to improve  $\text{P}_{\text{avil}}$  prediction via MIRS. In this study, no correlation between conventionally measured SOC content and  $\text{P}_{\text{avil}}$  or  $\text{K}_{\text{avil}}$  was observed. However, the variability of SOM composition in the LUCAS sample set was probably much larger than in the sample sets investigated by Pätzold et al. (2020). In consequence, soil properties controlling  $\text{P}_{\text{avil}}$  extractability were most probably too diverse for model calibration using the MIR-SSL.

#### 4.3 | Performance of LUCAS-based MIR models to predict soil properties of an independent regional sample set

The results of an independent validation of LUCAS prediction models at regional scale (LLUR samples) considering  $R^2$ , RMSEP, and RPIQ (Figure 3) indicated a reliable prediction for contents of sand, silt, clay, SOC, and  $\text{N}_{\text{total}}$ , as well as for pH ( $\text{CaCl}_2$ ). While for  $\text{CaCO}_3$  the  $R^2$  and RMSEP also confirmed a satisfactory prediction, the RPIQ was below the threshold proposed by Ludwig et al. (2019). However, the small RPIQ can be attributed to the large number of noncalcareous LLUR soil samples ( $n = 315$ ). In order to explain the differences between performance of LOOCV and independent validation, it is necessary to consider the history of the different sample sets. In contrast to the LUCAS





samples, which were taken in a uniform way with scope of creating the first harmonized and comparable data on soil at European level (Tóth et al., 2013), the LLUR sample set and the corresponding texture analyses of the sample set originate from different sampling campaigns at state level over many years. However, harmonized and consistent analytical procedures are important for transferability of prediction models (Minasny et al., 2009). For the LLUR sample set, methodological changes in sampling procedures and analytical methods are assumed due to the time factor in sampling, which caused a greater inconsistency in the analytical data and a larger analytical bias. This might have contributed to smaller  $R^2$  and RPIQ as well as larger RMSE and more biased predictions in the independent validation of sand and silt. In contrast, SOC,  $N_{\text{total}}$ , and pH ( $\text{CaCl}_2$ ) of the LLUR sample set were uniformly reanalyzed in the framework of this study. In consequence, the performance of predictions for these properties was more similar or even better than for LOOCV on LUCAS sample set. This underlines the crucial importance of consistent analyses when selecting sample sets for building SSL. Thus, a carefully built MIR-SSL can provide accurate predictions at the regional scale for basic soil properties to fully exploit the advantages of spectroscopic techniques.

#### 4.4 | Performance of generic LUCAS-based MIR models to characterize within-field heterogeneity

In order to evaluate the performance of generic LUCAS models for characterizing within-field heterogeneity, RMSEP and RPIQ are particularly suitable because the RMSEP provides the absolute prediction error, while the RPIQ relates the prediction error to the prevailing variation of observed values at the respective field (see Bellon-Maurel et al., 2010). Thus, in general, it can be stated from a small RPIQ that the prevailing within-field heterogeneity at a given field cannot be characterized by the prediction model. If, in addition, the RMSEP is also small, this indicates that the within-field variability of the respective soil property is only weakly expressed.

Small RPIQ values were generally achieved for soil properties and fields revealing only small within-field heterogeneity. This was found for prediction of, for example, sand at Bölingen and Görzig, silt at Ascheberg, Görzig, and Wilmersdorf, clay at Görzig, and SOC at Görzig (see Table 4, footnote 2). In these cases, the characterization of heterogeneity with generic LUCAS models was not possible. However, at least in the context of PA, the characterization of within-field heterogeneity is only required if a pronounced variability is observed. For the fields with larger within-field heterogeneity, generally more reliable predictions were achieved; this applied, for example, to sand, clay, and SOC prediction at Ascheberg and Wilmersdorf as well as pH ( $\text{CaCl}_2$ ) prediction at Ascheberg, Görzig, and Wilmersdorf. Regarding the soil texture properties, the most accurate predictions were possible for clay content, similar to the previous section.

The  $\text{CaCO}_3$  prediction must be evaluated as unsuccessful with respect to the RPIQ threshold of Ludwig et al. (2019). This was due to the dominance of noncalcareous soils in the validation sample sets. However, for agronomic applications, such as PA, the level of  $\text{CaCO}_3$

content is not important; rather it is sufficient to know whether the sample is calcareous or not. To easily decide upon this question, Leenen et al. (2019) proposed an approach for the identification of calcareous soil samples via the height of a carbonate peak in a given MIR spectrum. Thus, an SSL model is dispensable for evaluating the presence of  $\text{CaCO}_3$ .

Mostly large RPIQ values were achieved for pH ( $\text{CaCl}_2$ ; Table 4), indicating the general suitability of the generic model. However, if more precise pH predictions are required, locally calibrated models should be preferred because they are more specific to field conditions (see also Minasny et al., 2009).

It also applies to other properties that locally calibrated models are in principle more precise (e.g., Pätzold et al., 2020). Leenen et al. (2019) calibrated local MIR prediction models for SOC, clay, and pH ( $\text{CaCl}_2$ ) to determine lime requirement at the same fields as in this study. In comparison to prediction with generic LUCAS models, the local calibrations in the previous study were occasionally more accurate, but no general rule could be deduced. However, the disadvantage of local models is that their calibration is more costly and time consuming. Certainly, MIR-SSL-based predictions are also inferior to conventional analytical methods in terms of accuracy. However, it has to be decided, depending on the problem, which accuracy has to be achieved. In many management applications, a classification of soil properties replaces exact determination (e.g., liming: classification of SOC and clay; see Leenen et al., 2019). Here, the advantages of the MIR-SSL certainly outweigh local MIRS prediction models as well as conventional analytical methods.

#### 4.5 | Stratified modeling using soil parent material as auxiliary variable to improve local transferability

Prediction models for soil texture rely in particular on the mineralogical composition of the soil, such as quartz (sand) and various IR-active clay minerals (Soriano-Disla et al., 2014). This also applies in a similar way to  $\text{CaCO}_3$ . Thus, it was hypothesized that due to the expected relationship between mineralogical composition and the auxiliary variable parent material, stratification of the MIR-SSL-based on this auxiliary variable will lead to improvement in local transferability. However, considerable improvements in characterization of within-field heterogeneity with stratified compared to generic models were only achieved in single cases, for example, for predicting clay at Bölingen and  $\text{CaCO}_3$  at Wilmersdorf. The reason is probably the smaller concentration range in the stratified models. An example is clay prediction at Ascheberg with the EPM-5 model. Despite of a well-performing calibration, only a negligible improvement for clay prediction at Ascheberg was achieved compared to the generic model (Figure 4). Clay content in the EPM-5 calibration sample set ranged from 3% to 37%; thus, it did not completely cover the range at Ascheberg. However, it is a basic requirement that the concentration range of the calibration sample set covers entirely the sample set to be predicted (Terhoeven-Urselmans et al., 2010). Hence, LUCAS was not designed to depict regional extremes of soil properties but to yield consistent soil data at European level



(Tóth et al., 2013). In consequence, to fully benefit from the stratification approach, a higher sample density that includes the extremes in the respective region is necessary. Yet, also the already widespread approach of spiking can help to better adapt the calibration to the local conditions (e.g., Breure et al., 2022; Seidel et al., 2019). In principle, the geological map used here represented the best available basis for stratified models at this spatial scale. Of course, the small scale of the map and the rough classification of the EPM classes contributed to possible inaccuracies in the allocation of LUCAS samples to the EPM classes.

Wijewardane et al. (2018) showed that using auxiliary variables can provide improvements in model performance. This cannot be confirmed by our results. However, Wijewardane et al. (2018) calibrated stratified models on more than 20,000 samples that originated from different land use forms and soil horizons. In this study, by focusing on topsoil of arable fields, a preselection had already taken place.

## 5 | CONCLUSIONS

The enlargement of the LUCAS database to MIR spectra offers a further valorization of the LUCAS survey. In particular, the fundamental vibrations that occur in the MIR contribute to improve the informative value of the LUCAS samples over the already available vis-NIR spectra.

With regard to hypothesis (1), this hypothesis can be confirmed for various basic soil properties. The MIR-SSL built with LUCAS samples which were taken in various geologic-pedogenic environments and consistently analyzed in a single laboratory formed an ideal basis for successful calibration of reliable and accurate MIR prediction models for basic soil properties that were considered in the framework of LUCAS: contents of SOC,  $N_{\text{total}}$ , sand, silt, clay, and  $\text{CaCO}_3$ , and pH ( $\text{CaCl}_2$ ). However, for  $P_{\text{avl}}$  and  $K_{\text{avl}}$  the calibration was not successful. Consistency of reference analyses is crucial for calibrating reliable models as shown for the LLUR sample set.

Varying success was achieved for characterization of soil properties of unknown sample sets at regional and field scales. In particular, a precise characterization of SOC and clay content via MIRS-SSL models was possible. Thus, it can be assumed that this MIR-SSL can complement or even replace conventional soil analyses in, for example, soil mapping campaigns at regional scale. Further, composite soil samples as usually taken as basis for management recommendations to farmers can be rapidly analyzed at small effort and costs. In addition, the MIRS-SSL can be used to precisely characterize the within-field heterogeneity of soil properties. Consequently, MIR-SSL can also contribute to applications in PA such as soil zoning or variable rate liming. Yet, for precise fertilizer dosage, model performance was not satisfactory despite the standardized analysis protocol. All in all, effort and costs of model calibration can be drastically reduced by the use of an MIR-SSL. This also increases the attractiveness of MIRS for data demanding applications such as PA.

In contrast, hypothesis (2) cannot be confirmed. Using soil parent material auxiliary information for stratified model calibration did not lead to the expected distinct improvement in characterizing within-field soil heterogeneity for an already well-focused MIR-SSL. The expected advantage of adapting models to specific regional spectral

properties (e.g., with respect to mineralogical composition) in the stratified approach is compensated by the narrowed data range. In order to take advantage of stratified models, the spatial sample density of the MIR-SSL should be increased. This includes, in particular, additional soil samples that increase the value range of the relevant soil properties.

## ACKNOWLEDGMENTS

This study was funded by the German Federal Ministry of Education and Research (BMBF) within the *BonaRes Project "Integrated System for Site-Specific Soil Fertility Management (IAS), part F"* (FKZ 031A564F). We thank *Oihane Fernández-Ugalde* [European Commission, Joint Research Center (JRC), Ispra, Italy] for discussion and valuable support in providing the LUCAS samples. We are also grateful to the State Agency for Agriculture, Environment and Rural Areas (LLUR) in Flintbek (Schleswig-Holstein, Germany), namely to *Marek Filipinski* and *Eckhard Cordsen*, for providing samples and soil data. Soil sampling at Görzig and Wilmersdorf was supported by *Peter Wagner* (Martin Luther University Halle-Wittenberg, Germany) and by *Robin Gebbers* [Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB), Potsdam, Germany], respectively. *Ralf Wehrle* gave important advice in data analysis. We thank the laboratory staff of INRES-Soil Science and Soil Ecology for conducting numerous soil analyses.

Open access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

The ground truth data of LUCAS samples that support the findings of this study are available from European Soil Data Centre (ESDAC), European Commission, Joint Research Centre. Restrictions apply to the availability of these data, which were used under license for this study. Data are available at <https://esdac.jrc.ec.europa.eu/content/lucas-2009-topsoil-data> with the permission of European Soil Data Centre (ESDAC), European Commission, Joint Research Centre. All other data that support the findings will be available in the BonaRes Data Portal at <https://datenzentrum.bonares.de/research-data.php> following an embargo from the date of publication.

## ORCID

Matthias Leenen  <https://orcid.org/0000-0002-0095-4184>

## REFERENCES

- Araújo, S. R., Wetterlind, J., Demattê, J. A. M., & Stenberg, B. (2014). Improving the prediction performance of a large tropical vis-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *European Journal of Soil Science*, 65(5), 718–729.
- Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *Trends in Analytical Chemistry*, 29(9), 1073–1081.
- Bellon-Maurel, V., & McBratney, A. (2011). Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils—Critical review and research perspectives. *Soil Biology and Biochemistry*, 43(7), 1398–1410.
- Bundesanstalt für Geowissenschaften und Rohstoffe (2005): Soil regions map of the European Union and adjacent countries





- 1:5,000,000 (Version 2.0). *Special Publication*. EU catalogue number S.P.I.05.134. [https://www.bgr.bund.de/EN/Themen/Boden/Projekte/Informationsgrundlagen\\_abgeschlossen/EUSR5000/EUSR5000\\_en.html](https://www.bgr.bund.de/EN/Themen/Boden/Projekte/Informationsgrundlagen_abgeschlossen/EUSR5000/EUSR5000_en.html)
- Bornemann, L., Welp, G., & Amelung, W. (2010). Particulate organic matter at the field scale: Rapid acquisition using mid-infrared spectroscopy. *Soil Science Society of America Journal*, 74(4), 1147–1156.
- Brereton, R. G. (2018). *Chemometrics: Data driven extraction for science* (2nd ed.). John Wiley and Sons.
- Breure, T. S., Prout, J. M., Haeefele, S. M., Milne, A. E., Hannam, J. A., Moreno-Rojas, S., & 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 and Tillage Research*, 215, 105196. <https://doi.org/10.1016/j.still.2021.105196>
- Clairotte, M., Grinand, C., Kouakoua, E., Thébault, A., Saby, N. P., Bernoux, M., & Barthès, B. G. (2016). National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma*, 276, 41–52.
- Clark, R. N., Swayze, G. A., Livo, K.E., Kokaly, R. F., Sutley, S. J., Dalton, J. B., McDougal, R. R., & Gent, C. A. (2003). Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *Journal of Geophysical Research: Planets*, 108(E12), 5131. <https://doi.org/10.1029/2002JE001847>
- Dhawale, N. M., Adamchuk, V. I., Prasher, S. O., Viscarra Rossel, R. A., Ismail, A. A., & Kaur, J. (2015). Proximal soil sensing of soil texture and organic matter with a prototype portable mid-infrared spectrometer. *European Journal of Soil Science*, 66(4), 661–669.
- Ellerbrock, R. H., & Gerke, H. H. (2021). FTIR spectral band shifts explained by OM–cation interactions. *Journal of Plant Nutrition and Soil Science*, 184(3), 388–397.
- ESDAC (2013). LUCAS 2009 TOPSOIL data, European Commission. *Joint Research Centre*. <https://esdac.jrc.ec.europa.eu/content/lucas-2009-topsoil-data>
- Hicks, W., Viscarra Rossel, R., & Tuomi, S. (2015). Developing the Australian mid-infrared spectroscopic database using data from the Australian Soil Resource Information System. *Soil Research*, 53(8), 922–931.
- ISO (1994a). ISO 10390:1994. *Soil quality—Determination of pH*. International Organization for Standardization.
- ISO (1994b). ISO 10693:1994. *Soil quality—Determination of carbonate content: Volumetric method*. International Organization for Standardization.
- ISO (1994c). ISO 11263:1994. *Soil quality—Determination of phosphorus: Spectrometric determination of phosphorus soluble in sodium hydrogen carbonate solution*. International Organization for Standardization.
- ISO (1995a). ISO 10694:1995. *Soil quality—Determination of organic and total carbon after dry combustion (elementary analysis)*. International Organization for Standardization.
- ISO (1995b). ISO 11261:1995. *Soil quality—Determination of total nitrogen: Modified Kjeldahl method*. International Organization for Standardization.
- ISO (1998a). ISO 11277:1998. *Soil quality—Determination of particle size distribution in mineral soil material - Method by sieving and sedimentation*. International Organization for Standardization.
- ISO (1998b). ISO 13878:1998. *Soil quality—Determination of total nitrogen content by dry combustion (elemental analysis)*. International Organization for Standardization.
- IUSS Working Group (2015). World reference base for soil resources 2014, update 2015. International soil classification system for naming soils and creating legends for soil maps. *World Soil Resources Reports*.
- Izaurrealde, R. C., Rice, C. W., Wielopolski, L., Ebinger, M. H., Reeves, J. B., Thomson, A. M., Harris, R., Francis, B., Mitra, S., Rappaport, A. G., Etchevers, J. D., Sayre, K. D., Govaerts, B., & McCarty, G. W. (2013). Evaluation of three field-based methods for quantifying soil carbon. *PLoS One*, 8(1), e55560. <https://doi.org/10.1371/journal.pone.0055560>
- Ji, W., Adamchuk, V. I., Biswas, A., Dhawale, N. M., Sudarsan, B., Zhang, Y., Viscarra Rossel, R. A., & Shi, Z. (2016). Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering*, 152, 14–27.
- Kuang, B., Mahmood, H. S., Quraishi, M. Z., Hoogmoed, W. B., Mouazen, A. M., & van Henten, E. J. (2012). Sensing soil properties in the laboratory, in situ, and on-line: A review. *Advances in Agronomy*, 114, 155–223.
- Leenen, M., Welp, G., Gebbers, R., & Pätzold, S. (2019). Rapid determination of lime requirement by mid-infrared spectroscopy: A promising approach for precision agriculture. *Journal of Plant Nutrition and Soil Science*, 182(6), 953–963.
- Lobsey, C. R., Viscarra Rossel, R. A., Roudier, P., & Hedley, C. B. (2017). rs-local data-mines information from spectral libraries to improve local calibrations. *European Journal of Soil Science*, 68(6), 840–852.
- Ludwig, B., Murugan, R., Parama, V. R., & Vohland, M. (2019). Accuracy of estimating soil properties with mid-infrared spectroscopy: Implications of different chemometric approaches and software packages related to calibration sample size. *Soil Science Society of America Journal*, 83(5), 1542–1552.
- Minasny, B., Tranter, G., McBratney, A. B., Brough, D. M., & Murphy, B. W. (2009). Regional transferability of mid-infrared diffuse reflectance spectroscopic prediction for soil chemical properties. *Geoderma*, 153(1–2), 155–162.
- Mouazen, A. M., Maleki, M. R., De Baerdemaeker, J., & Ramon, H. (2007). On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil and Tillage Research*, 93(1), 13–27.
- Mulla, D., & Khosla, R. (2016). Historical evolution and recent advances in precision farming. In R. Lal & B. Stewart (Eds.), *Soil-specific farming* (pp. 1–35). CRC Press.
- Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., & Fernández-Ugalde, O. (2018). LUCAS Soil, the largest expandable soil dataset for Europe: A review. *European Journal of Soil Science*, 69(1), 140–153.
- Pätzold, S., Leenen, M., Frizen, P., Heggemann, T., Wagner, P., & Rodionov, A. (2020). Predicting plant available phosphorus using infrared spectroscopy with consideration for future mobile sensing applications in precision farming. *Precision Agriculture*, 21(4), 737–761.
- Reeves, J. B. III (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*, 158(1–2), 3–14.
- Reimann, C., de Caritat, P., GEMAS Project Team, & NGS Project Team (2012). New soil composition data for Europe and Australia: Demonstrating comparability, identifying continental-scale processes and learning lessons for global geochemical mapping. *Science of the Total Environment*, 416, 239–252.
- Rodionov, A., Welp, G., Damerow, L., Berg, T., Amelung, W., & Pätzold, S. (2015). Towards on-the-go field assessment of soil organic carbon using Vis–NIR diffuse reflectance spectroscopy: Developing and testing a novel tractor-driven measuring chamber. *Soil and Tillage Research*, 145, 93–102.
- Seidel, M., Hutengs, C., Ludwig, B., Thiele-Bruhn, S., & 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*, 354, 113856. <https://doi.org/10.1016/j.geoderma.2019.07.014>
- Soriano-Disla, J. M., Janik, L. J., Viscarra Rossel, R. A., Macdonald, L. M., & McLaughlin, M. J. (2014). The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Applied Spectroscopy Reviews*, 49(2), 139–186.
- Tatzber, M., Mutsch, F., Mentler, A., Leitgeb, E., Englisch, M., & Gerzabek, M. H. (2010). Determination of organic and inorganic carbon in forest soil samples by mid-infrared spectroscopy and partial least squares regression. *Applied Spectroscopy*, 64(10), 1167–1175.





- Terhoeven-Urselmans, T., Vagen, T. G., Spaargaren, O., & Shepherd, K. D. (2010). Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. *Soil Science Society of America Journal*, 74(5), 1792–1799.
- Tóth, G., Jones, A., & Montanarella, L. (2013). The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. *Environmental Monitoring and Assessment*, 185(9), 7409–7425.
- USDA-NRCS (2004). Soil survey laboratory manual. Soil Survey Investigation Report No. 42. Version 4.0. United States Department of Agriculture Natural Resources Conservation Service.
- Viscarra Rossel, R. A., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthès, B. G., Bartholomeus, H. M., Bayer, A. D., Bernoux, M., Böttcher, K., Brodský, L., Du, C. W., Chappell, A., ... Ji, W. (2016). A global spectral library to characterize the world's soil. A global spectral library to characterize the world's soil *Earth-Science Reviews*, 155, 198–230.
- Viscarra Rossel, R. V., & Bouma, J. (2016). Soil sensing: A new paradigm for agriculture. *Agricultural Systems*, 148, 71–74.
- Viscarra Rossel, R. V., Jeon, Y. S., Odeh, I. O. A., & McBratney, A. B. (2008). Using a legacy soil sample to develop a mid-IR spectral library. *Soil Research*, 46(1), 1–16.
- Viscarra Rossel, R. V., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1–2), 59–75.
- Vohland, M., Ludwig, M., Thiele-Bruhn, S., & Ludwig, B. (2014). Determination of soil properties with visible to near-and mid-infrared spectroscopy: Effects of spectral variable selection. *Geoderma*, 223, 88–96.
- Wetterlind, J., & Stenberg, B. (2010). Near-infrared spectroscopy for within-field soil characterization: Small local calibrations compared with national libraries spiked with local samples. *European Journal of Soil Science*, 61(6), 823–843.
- White, E. L., Thomasson, J. A., Auvermann, B., Kitchen, N. R., Pierson, L. S., Porter, D., Bailloe, C., Hamann, H., Hoogenboom, G., Janzen, T., Khosla, R., Lowenberg-DeBoer, J., McIntosh, M., Murray, S., Osborn, D., Shetty, A., Stevenson, J. T., & Werner, F. (2021). Report from the conference, 'identifying obstacles to applying big data in agriculture'. *Precision Agriculture*, 22(1), 306–315.
- Wijewardane, N. K., Ge, Y., Wills, S., & Libohova, Z. (2018). Predicting physical and chemical properties of US soils with a mid-infrared reflectance spectral library. *Soil Science Society of America Journal*, 82(3), 722–731.

**How to cite this article:** Leenen, M., Pätzold, S., Tóth, G., & Welp, G. (2022). A LUCAS-based mid-infrared soil spectral library: Its usefulness for soil survey and precision agriculture. *Journal of Plant Nutrition and Soil Science*, 185, 370–383.  
<https://doi.org/10.1002/jpln.202100031>

