



HAL
open science

Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils

Jean-Philippe Gras, Bernard G. Barthès, Brigitte Mahaut, Séverine Trupin

► To cite this version:

Jean-Philippe Gras, Bernard G. Barthès, Brigitte Mahaut, Séverine Trupin. Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma*, 2014, 214-215, pp.126-134. 10.1016/j.geoderma.2013.09.021 . ird-04109198

HAL Id: ird-04109198

<https://hal.ird.fr/ird-04109198>

Submitted on 29 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils

Jean-Philippe Gras ^{1,2}, Bernard G. Barthès ^{1*}, Brigitte Mahaut ², Séverine Trupin ²

¹ IRD, UMR Eco&Sols, Montpellier SupAgro, 2 place Viala, 34060 Montpellier Cedex 2, France.

² ARVALIS - Institut du végétal, Route de Malesherbes, 91720 Boigneville, France.

* corresponding author: <bernard.barthes@ird.fr>; phone: +33 467 99 61 21 36

Highlights

- Field-acquired spectra allowed good VNIRS calibrations of usual soil properties.
- Spectra acquired on auger cores yielded the best results in the field.
- Prediction accuracy was similar to that achieved using 2-mm sieved dried samples.

Abstract

Diffuse infrared reflectance spectroscopy is considered a promising approach for addressing soil quality, and its use directly in the field might be an achievable challenge. The present work aimed at optimizing the acquisition procedure of visible and near infrared reflectance (VNIR) spectra of topsoils (0-20 cm) in the field, in order to predict usual soil properties. The studied set included 201 samples originating from six fields in different regions of large-scale crop cultivation in France.

Spectra were acquired using a portable spectrophotometer. Spectrum acquisition procedures included scanning on the soil surface, on raw or smoothed (cut) cores collected using an auger, and on clods resulting from core crumbling. In addition, spectra were also acquired on air-dried clods, either 2-mm sieved or not (laboratory conditions). Furthermore, 42 mathematic pretreatments were compared (including derivatives, standard normal variate SNV, multiplicative scatter correction MSC, etc.). Identifying the most appropriate scanning and pretreatment procedures was done through four-group cross-validation.

Using the most appropriate pretreatment, calcium carbonate content was very well predicted whatever the scanning procedure (RPD = 6.9-9.1; RPD is the ratio of standard deviation to standard error of cross-validation; for soil properties $RPD > 2$ denotes accurate predictions); good predictions were achieved for total nitrogen (RPD = 2.5-3.0), organic matter (RPD = 2.1-2.8) and exchangeable potassium (RPD = 2.9-3.2); but available phosphorus was poorly predicted (RPD = 1.6-1.8). Except for available phosphorus, accurate predictions of these properties could therefore be achieved whatever the scanning procedure, thus in the

field. Best predictions were often obtained using spectra acquired on 2-mm sieved air-dried samples (i.e. in laboratory conditions), otherwise using spectra acquired on raw cores. Acquiring spectra on cores, on raw cores especially, was the most appropriate field procedure; it led to predictions comparably accurate to those achieved in the laboratory with 2-mm sieved air-dried samples. Similar prediction accuracy for field and laboratory VNIRS is counterintuitive due to variable field conditions (moisture, temperature, stoniness, etc.). It might result from higher number of replicates in the field than in the laboratory (often inherent to field vs. lab conditions) and/or higher sample density and cohesion, which would improve reflectance signal. For spectra acquired on cores, best calibrations were achieved with MSC and first derivatives for calcium carbonate, total nitrogen and organic matter, but without pretreatment for exchangeable potassium and available phosphorus. Second derivatives always yielded poor results.

Keywords

Soil organic matter; nutrients; calcium carbonate; visible and near infrared reflectance spectroscopy (VNIRS); proximal soil sensing.

1. Introduction

There is increasing concern about soil quality, its degradation and its protection, and this regards agricultural soils particularly (European Commission, 2006). As a consequence, there is increasing demand for the monitoring of agricultural soil attributes (Huber et al., 2008). Decision-makers and land use planners are interested in soil quality, and of course farmers too. However, practical assessment of soil quality requires the monitoring of several attributes, which may vary meaningfully over short distances thus have to be characterized at high spatial density. In such conditions, appropriate monitoring of agricultural soil attributes using conventional soil analyses is expensive and time-consuming.

Diffuse infrared reflectance spectroscopy, especially near infrared reflectance spectroscopy (NIRS) and visible and NIRS (VNIRS), has been reported to represent a relevant alternative for assessing and monitoring soil quality (Cécillon et al., 2009). Carried out in the laboratory, this approach allows time- and cost-effective determination of numerous soil properties such as carbon and nitrogen contents, texture, pH, but also relating to organic matter quality and microbial processes, and even to the abundance of soil organisms (Malley et al., 2004; Barthès et al., 2006, 2011).

In addition, VNIRS can also be performed directly in the field, which improves time- and cost-effectiveness (no sample packaging and preparation required), and is particularly useful when high spatial resolution is needed (Stenberg et al., 2010). However, in situ measurement of soil reflectance is influenced by conditions such as soil moisture, structure, stoniness, coarse organic residues, contamination of the probe sensor by dust, possible changes in the distance between sensor and soil, and as far as spectroradiometers are used, changes in solar illumination (Baumgardner et al., 1985; Sudduth & Hummel, 1993; Lagacherie et al., 2008; Stenberg et al., 2010; Gubler, 2011). Though these conditions are not favorable to VNIRS characterization of soil properties, literature has mostly reported promising results using field VNIRS, especially for determining clay, organic carbon and total nitrogen contents (Ben-Dor et al., 2008; Stevens et al., 2008; Morgan et al., 2009; Viscarra Rossel et al., 2009; Kusumo et al., 2010); some authors nevertheless failed to achieve accurate predictions of moisture, organic and inorganic carbon contents, total nitrogen content or cation exchange capacity (Sudduth & Hummel, 1993; Udelhoven et al., 2003). In addition, most work on field VNIRS has been carried out on sample sets originating from small areas (Gubler, 2011).

The objective of the present study was to optimize spectrum acquisition for field VNIRS prediction of usual topsoil properties (including soil contents in calcium carbonate, organic matter, total nitrogen, available phosphorus, and exchangeable potassium) over a sample set representing the main areas of large-scale crop cultivation in France. Several scanning procedures were compared (on the soil surface; on raw or cut auger cores; on clods crumbled from cores, possibly mixed over several elementary sampling points, scanned after a delay, and dried; and finally on 2-mm sieved dried samples, which represent usual laboratory conditions); several usual preprocessing techniques were also tested (standard normal variate, multiplicative scatter correction, etc.).

2. Materials and methods

2.1 Study sites

Topsoil samples from 201 1-m² sites originating from six farmer's fields, each in a different French "département" (county), were studied:

- 43 sites in a 8.3-ha field in Aube, center-east of France (48°23'28"N, 04°08'20"E); according to the World reference base (IUSS-ISRIC-FAO, 2006), the soil was a clayey Rendzic Leptosol (thus rich in calcium carbonate);
- 22 sites in a 4.8-ha field in Bas-Rhin, east of France (48°18'54"N, 07°37'22"E); the soil was a sandy loam Leptic Fluvisol;

- 33 sites in a 7.3-ha field in Essonne, north-east of France (48°18'47"N, 02°23'56"E); the soil was a silty Cambic Calcisol;
- 29 sites in a 7.0-ha field in Haute-Garonne, south of France (43°27'09"N, 01°38'31"E); the soil was a loamy Cambic Fluvisol, rich in calcium carbonate;
- 51 sites in a 14.1-ha field in the Landes, south-west of France (44°24'20"N, 00°34'13"O); the soil was a Podzol (thus sandy with topsoil rich in organic matter);
- 23 sites in a 5.7-ha field with in Loire-Atlantique, west of France (47°27'20"N, 00°58'00"O); the soil was a clayey sandy loam Gleyic Cambisol.

Sites were rather evenly distributed over the fields, and were sampled in winter (cold and rainy or even snowy season) between November 2011 and January 2012. Roughly, air temperature ranged from 0 to 15°C; soil water content was not determined.

2.2 Spectrum acquisition

Soil sample reflectance was measured from 350 to 2500 nm at 1 nm interval using a portable spectrophotometer ASD LabSpec 5000 (Analytical Spectral Devices, Boulder, CO, USA). Each spectrum provided by the spectrometer actually resulted from the averaging of 32 co-added scans. In this device, light is delivered to the sample by a contact probe, which then collects the reflected signal and transmits it to the spectrometer. After every measurement the window of the contact probe was cleaned with a silk brush then lens paper and ethanol. A disc made of Spectralon (polytetrafluoroethylene) was used as white reference standard and its reflectance was measured every hour. At each 1-m² site, spectrum acquisition was carried out according to seven different procedures:

- SURFACE: one soil spectrum was acquired at each of the four corners of the site directly on the soil surface, which had been previously planned down using a knife, after possible residues had been removed; this rapid procedure characterized the very surface of the soil only, which might nevertheless be representative of underlying layers in tilled soils;
- COREraw: using an auger, one 20-cm deep core was collected at each of the four corners of the site, and spectra were collected on the raw faces of the cores at depths of 2-3, 10-12 and 15-17 cm; this represented a total of 12 COREraw spectra per site;
- COREcut: each core was cut so as to present a smooth surface on which spectra were acquired at depths of 2-3 cm, 10-12 and 15-17 cm, representing a total of 12 COREcut spectra per site; cutting aimed at avoiding possible noise due to core surface roughness;

- CLODraw: clods resulting from the crumbling of one core were placed in a plastic tray and scanned in triplicate, representing a total of 12 CLODraw spectra per site;
- CLODav: clods originating from the four cores of the site were grouped, mixed thoroughly, and four scans were carried out on this average sample;
- CLODav+0.5: four scans of the average sample of the site were also carried out in the field after a half-day; before scan, samples had been kept in closed plastic bags within a bucket in the field (where they had presumably undergone limited change in moisture and temperature); the objective of this variant was to test whether spectrum acquisition could wait a few hours after sample collection (thus whether the spectrometer could be left, e.g., at the field edge, in order not to carry it in the field);
- CLODav+1: four scans of the average sample of the site were also carried out in the field the next morning; before scan, samples had been kept in closed plastic bags within a car possibly in a garage overnight, then had been brought back in the field (they had presumably undergone limited change in moisture and temperature too); the objective of this variant was to test whether spectrum acquisition could wait until the next morning (thus whether the spectrometer could be left in a shed close to studied areas, for instance).

In situ measurements were also compared to laboratory measurements according to two procedures:

- CLODav_dry: the average sample of the site was air-dried at room temperature for two weeks then oven-dried at 40°C for one night, and then four scans were carried out;
- CLODav_siev: the average sample of the site was gently crushed to pass a 2-mm sieve, then oven-dried at 40°C for one night, and spectra were acquired in duplicate.

Air temperature was 24°C in the laboratory. For each combination of site and procedure, all spectra were averaged for spectral analysis.

2.3. *Spectrum analysis*

Reflectance spectra were recorded as absorbance, which is the logarithm of the inverse of reflectance ($\log[1/R]$). Data analysis was conducted using the WinISI III–V1.50e software (Foss NIRSystems/Tecator Infrasoft International, LLC, Silver Spring, MD, USA). The data sets were reduced by keeping the first of every eight consecutive spectral points (Sørensen & Dalgaard, 2005; Shenk & Westerhaus, 1991, kept every fifth point). In addition, the 350-

450 nm range was removed, due to noisy features, as well as transition domains between sensors (996-1004 and 1824-1834 nm).

Then preprocessing was used to eliminate or minimize additive and multiplicative effects caused for example by the heterogeneity of sample texture and physical properties, in order to improve the predictions of chemical properties (Osborne et al., 1993). These effects include baseline shift, slope changes, variations in spectrum curvilinearity. Classic preprocessing techniques include calculating first and second derivatives, de-trending (D), standard normal variate transformation (SNV) and multiplicative scatter correction (MSC), which involves three options namely standard MSC, weighted MSC (WMSC), and inverse MSC (IMSC; Geladi et al., 1985; Barnes et al., 1989, Dhanoa et al., 1994; Candolfi et al., 1999). In the present study, the preprocessing techniques used were None (no preprocessing), D, SNV, SNVD (i.e. both SNV and D), MSC, IMSC and WMSC, with different levels of derivatives, gap (interval used to calculate derivatives) and smoothing; this led to six combinations of these three factors: no derivatives with one- or five-point smoothing (denoted 0 0 1 and 0 0 5, respectively); first derivatives with five-, 15- or 20-point gap and five-point smoothing (1 5 5, 1 15 5 and 1 20 5), and second derivatives with five-point gap and smoothing (2 5 5). On the whole, 42 preprocessing methods were tested.

Modified partial least square (MPLS; Shenk & Westerhaus, 1991) regression coupled with cross-validation was used to relate the spectra with the reference data. The MPLS regression combines principal component analysis and multiple linear regression in order to reduce a complex spectral matrix into a few orthogonal components, and has often been considered an appropriate statistical method for studying soil organic properties (Fystro, 2002; Coûteaux et al., 2003, Brunet et al, 2007). Cross-validation is a technique which allows testing if a statistical model could be generalized to an independent data set. In the present case, this technique consisted of partitioning the dataset into four subsets, performing the calibration on three subsets and validating it on the fourth, the operation being repeated three times by rotating the subsets used for validation and calibration. The results were then averaged to produce a single estimation. The parameters used to assess and compare the performances of the prediction models were the coefficient of determination (R^2) and the slope of the linear regression between measured and predicted values, the standard error of cross-validation (SECV), the ratio of the standard deviation (SD) to SECV (denoted RPD) and the bias, which is the mean residue. The number of MPLS components above which SECV did not decrease meaningfully anymore was retained for calibration (Bjørsvik & Martens, 2001). Poorly calibrated samples (i.e. with residual > 2.5 SECV) were considered outliers. However, to

allow valuable comparison between scanning procedures and between preprocessing methods, the number of calibration outliers removed was maintained fixed for a given soil property. For calcium carbonate and exchangeable potassium, respectively, the number of calibration outliers was three and seven in general, and all were removed. For the other properties studied, the number of calibration outliers was much higher, up to 20 sometimes; the number of calibration outliers removed was limited to 5% of the sample population, which represented ten samples for total nitrogen, organic matter and available phosphorus.

External validation was not carried out because the objective of the study was not to build robust calibrations but to identify the most appropriate scanning procedure. In addition, the six studied fields represented different mineralogical and textural contexts, thus using one as external validation set was not relevant. Once the scanning procedure identified, the calibration set will be completed by other samples from new sites, and testing robustness will be more appropriate then.

2.4. Chemical analysis

Determination of soil contents in calcium carbonate, organic carbon, total nitrogen, available phosphorus and exchangeable potassium were carried out on aliquots of COREav samples dried at 38°C then gently crushed and sieved at 2 mm. Calcium carbonate was measured by a volumetric method with a Bernard calcimeter according to the French procedure NF ISO 10693, which involves the addition of hydrochloric acid and the measurement of the volume of carbon dioxide released (AFNOR, 1995a). Total nitrogen was determined on finely ground samples by dry combustion using an elemental analyzer, according to the French procedure NF ISO 13878; all forms of nitrogen are volatilized by combustion at 900°C to form dinitrogen (N₂), which is then measured by gas chromatography using a thermal conductivity detector (AFNOR, 1998a). Organic carbon content was determined by sulfochromic oxidation (oxidation with potassium dichromate and sulfuric acid) then spectrophotometry, according to the French procedure NF ISO 14235; it was then converted into soil organic matter content after multiplication by 1.72 (AFNOR, 1998b). Available phosphorous was determined by the Olsen method, which involves extraction using a solution of sodium bicarbonate then measurement by spectrophotometry, according to the French procedure NF ISO 11263 (AFNOR, 1995b). Exchangeable potassium was determined after extraction in a solution of ammonium acetate and measurement by inductively coupled plasma (ICP) analysis, according to the French procedure NF X31-108 (AFNOR, 2002).

3. Results and discussion

3.1 Conventional data

The studied topsoil population displayed wide ranges of soil contents in calcium carbonate (from 0 to 85 g 100 g⁻¹ dry soil), organic matter (1 to 9 g 100 g⁻¹), total nitrogen (0.04 to 0.46 g kg⁻¹), available phosphorus (18 to 178 mg P₂O₅ kg⁻¹) and exchangeable potassium (31 to 684 mg K₂O kg⁻¹; Table 1). This was consistent with the diversity of the soils sampled, which originated from six very different regions of France.

3.2 Overall prediction accuracy

For each soil property and each sampling procedure, the 42 above-mentioned preprocessing methods were compared. For a given soil property, a similar number of outliers was set for all sampling and preprocessing procedures, and never exceeded 5% of the total number of samples. Considering the preprocessing method that yielded the best predictions for each scanning procedure, the range of RPD values was 6.9-9.1 for calcium carbonate, 2.5-3.0 for total nitrogen, 2.1-2.8 for organic matter, 1.6-1.8 for available phosphorous, and 2.9-3.2 for exchangeable potassium (Table 2). According to Chang et al. (2001) and Dunn et al. (2002), RPD = 2 could be considered a threshold for accurate prediction of soil properties. Whatever the scanning procedure, accurate predictions could thus be achieved for all studied properties except for available phosphorus (Figure 1). The prediction of soil content in calcium carbonate was excellent whatever the scanning and preprocessing procedures, suggesting that conventional laboratory determination of this property is not useful anymore, except for maintaining calibration databases.

Prediction accuracy for calcium carbonate was much higher than that reported in field conditions by Lagacherie et al. (2008) for a set of vineyard topsoil samples in southern France (RPD estimation between 1.3 to 2.0 depending on scanning procedure; cross-validation); their less accurate results might be attributed to statistical procedures (using a narrow range of wavelengths only), to variable soil stoniness, and to possibly variable field illumination conditions (the study was carried out with a spectroradiometer). Comparable accuracy was achieved for inorganic carbon by Morgan et al. (2009; RPD = 2.0; external validation) over a sample set collected at different depth layers in six fields from two counties in Central Texas. This result was obtained for external validation, which is more demanding in general than cross-validation; but it is likely that variable sample depth increased set heterogeneity, which has been shown to affect prediction accuracy negatively (Brunet et al., 2007). The same reasons could explain why Morgan et al. (2009) achieved poor predictions of soil organic

matter content (RPD = 1.5). Studying Israeli soils, Ben-Dor et al. (2008) obtained accurate external validation of organic matter content (RPD = 4.1) from field VNIR spectra using calibration built with a mix of spectra acquired in field conditions and on sieved dried samples; but it seems that many outliers were removed, which could have contributed to this good result. Stevens et al. (2008) also reported good field VNIRS predictions of soil organic matter content for a regional sample set (50 km²) in southern Belgium (cross-validation; RPD = 2.8). In the Manawatu region of New Zealand, Kusumo et al. (2010) found much more accurate field VNIRS predictions of soil organic matter and nitrogen (cross-validation; RPD = 7.3 and 6.3, respectively). Two hypotheses might explain these excellent results: all studied samples originated from recent fluvial deposits under pastures, which probably resulted in set homogeneity thus improved VNIRS predictions (cf. Brunet et al., 2007); in addition, Kusumo et al. (2010) studied much smaller cores than in the present study (15 mm high × 46 mm wide vs. 200 mm high × 80 mm wide, respectively), and even though they possibly made fewer spectrum replicates (10 vs. 4 to 12), their scans were probably more representative of samples analyzed conventionally (reference chemical analyses). Over a topsoil sample set collected along a 130-km transect in South Africa, Nocita et al. (2011) achieved field VNIRS prediction of organic matter with similar accuracy than in the present study (RPD = 2.4, with scan from the soil surface apparently; cross-validation). Over a fresh topsoil set from northern Belgium, Mouazen et al. (2006) obtained less accurate VNIRS predictions of total nitrogen, available phosphorus and exchangeable potassium than in the present study, though set homogeneity was higher (RPD = 1.8, 1.4, and 1.2, respectively; cross-validation); scanning procedure was not specified. Few works have addressed field VNIRS predictions of available phosphorus or exchangeable potassium. However, as regarded laboratory VNIRS predictions over heterogeneous soil sample sets, several reviews reported good results for exchangeable potassium but poor ones for available phosphorus (Malley et al., 2004; Cécillon et al., 2009; Stenberg et al., 2010). Regarding organic matter content, the under-prediction that was noticed in the present study at high values (> 5 g 100 g⁻¹) was also observed by Ben-Dor & Banin (1995) over an air-dried and 2-mm sieved sample set from Israel representing a wide range of organic matter content. The authors considered that the sample set could be divided into two subsets, one with samples poor in organic matter (0-4 g 100g⁻¹), including highly decomposed organic matter, and the other with samples richer in organic matter (4-14 g 100g⁻¹), including less decomposed organic matter.

3.3. *Effect of the scanning procedure*

Spectrum acquisition procedure had an effect on prediction accuracy. Indeed, best RPD achieved using a given scanning procedure could increase noticeably when using another scanning procedure: by 2.2 for calcium carbonate, 0.5 for total nitrogen, 0.7 for organic matter, 0.3 for available phosphorus, and 0.4 for exchangeable potassium (Table 2). This possible increase was crucial for organic matter, as best SECV ranged from 0.42 to 0.55 g 100 g⁻¹ depending on the way spectrum were acquired. This possible increase was useful but less critical for calcium carbonate, total nitrogen and exchangeable potassium, which could be accurately predicted whatever the scanning procedure (RPD \geq 2.5); this was also less critical for available phosphorus, which prediction was always poor.

CLODav_siev yielded best predictions for three out of the five soil properties considered, which is consistent with the facts that this procedure reduced noise (through drying) and homogenized the samples (through crushing and 2-mm sieving). COREraw yielded best predictions for the two other properties, and was in the top four scanning procedures for all properties (Table 2). When removing laboratory procedures (CLODav-dry and CLODav-siev), COREraw was even in the top two scanning procedures for all properties. At least for the spectrometer, the sample population and the soil properties considered, COREraw could thus be seen as the most appropriate procedure for acquiring soil spectra in the field, though COREcut and CLODraw yielded slightly better predictions for total nitrogen and exchangeable potassium, respectively.

A possible reason for poorer performance of SURFACE, CLODav, CLODav+0.5 and CLODav+1 could be the smaller number of replicates that were carried out for these procedures (four) than for other field procedures (12). Samples scanned for CLODav, CLODav+0.5 and CLODav+1 were averaged over a site and thus more homogeneous a priori, but apparently this did not offset the smaller number of replicates. In addition, rather poor predictions for SURFACE resulted likely from the fact that the scan addressed only the very surface of the sample that was analyzed conventionally (0-20 cm); thus scanned and analyzed samples were not similar. Smaller number of replicates was also the probable reason why laboratory conditions, though they involved drying (CLODav-dry) and possible 2-mm sieving (CLODav-siev), did not always yield the best predictions. COREraw generally outperformed CLODraw, possibly due to its denser thus more homogeneous structure.

The results of the present study confirmed that the performance of field and laboratory VNIRS may be similar, as reported for organic carbon by Stevens et al. (2008; RPD = 2.1 vs. 2.0, respectively). However, these authors used a spectroradiometer and scanned a spot size of

0.45 m in the field (measurement height of 1 m), but a contact probe and a small soil sample in the laboratory (with artificial illumination). Similar numbers of replicates were made for both field and laboratory measurements, but the former incorporated much more soil variability than the latter. Though not underlined by Stevens et al. (2008), this represented an advantage that might be compared in the present study to the larger number of replicates made in the field than in the laboratory. Thus the hypothesis that similar prediction accuracy for field and laboratory VNIRS resulted from larger number of replicates in the field than in the laboratory could not be dismissed. However, as field work allows replication more easily than lab work, this point should be considered a difference between approaches rather than an artifact.

Stevens et al. (2008) considered that achieving similar VNIRS performances in the field than in the laboratory was counterintuitive. Indeed, spectrum acquisition in the field generally involves variable soil moisture and temperature, the presence of coarse mineral and organic particles (which are scanned in the field but removed by sieving, thus not taken into account by chemical analyses and laboratory VNIRS), and possibly vegetation cover (Stenberg et al., 2010). It might however be assumed that in some respects field conditions are beneficial, due to higher sample density and homogeneity, especially when cores are scanned, and/or due to higher cohesion in moist conditions, which both could improve the relationship between sample composition and absorbance. In total, these positive and negative factors could offset each other, and would result in similar prediction accuracy in field and laboratory conditions. Morgan et al. (2009), who obtained more accurate VNIRS predictions of organic carbon content on six fields from two Texas counties using dried cores than 2-mm sieved air-dried samples (RPD = 1.9 vs. 1.7, respectively ; external validation), indeed argued that higher bulk density in the former might result in stronger reflectance signal than in the latter. Viscarra Rossel et al. (2009) also invoked higher density, along with smoother surface, which would produce higher reflectance, to explain better VNIRS prediction of clay content using spectra collected on pit walls than using 2-mm sieved dried samples. However, Morgan et al. (2009) reported less accurate VNIRS predictions of inorganic carbon on dried cores than on 2-mm sieved air-dried samples (RPD = 2.2 vs. 2.4), possibly due to secondary carbonate accumulation in some samples. Moreover, these authors obtained less accurate predictions using field-moist cores than 2-mm sieved air-dried samples for organic carbon (RPD = 1.5 vs. 1.7) and inorganic carbon (RPD = 2.0 vs. 2.4). Morgan et al. (2009) concluded that soil heterogeneity, at the core scale, did not reduce VNIRS effectiveness for predicting organic carbon; nevertheless it reduced effectiveness for predicting inorganic carbon, which was more

correctly addressed on homogenized 2-mm sieved samples. They also underlined that air-drying improved prediction accuracy for organic and inorganic carbon. Over a topsoil sample set from northern Belgium, Mouazen et al. (2006) also achieved less accurate VNIRS predictions of total nitrogen, available phosphorus and exchangeable potassium using spectra from fresh and wet samples than from dried and sieved samples (cross-validation; RPD = 1.8 vs. 2.4, 1.4 vs. 1.7, and 1.2 vs. 1.3, respectively; scanning procedure was not specified). Similarly, Nocita et al. (2011) reported less accurate VNIRS predictions of organic carbon using spectra acquired on the topsoil surface (apparently, because the scanning procedure was not specified) than on 2-mm sieved air-dried samples (cross-validation; RPD = 2.4 vs. 2.8).

As regarded moisture conditions specifically, on a loamy sample set originating from different depth layers in one field, Chang et al. (2005) found less accurate NIRS predictions of inorganic carbon, total nitrogen, and organic carbon using spectra from field-moist samples than from air-dried, both 2-mm sieved (cross-validation; RPD = 5.6 vs. 6.1, 3.6 vs. 4.0, 3.2 vs. 3.3, respectively); they did not propose possible explanations. By contrast, studying a wide range of Swedish soils, Stenberg (2010) reported that rewetting of 2-mm sieved dry samples (20% or 30% v/v) improved VNIRS prediction of soil organic matter by 20-30%, and that improvement was mainly for the more sandy samples. Though he recognized that mechanisms behind this improvement remained largely unexplained, he suggested that rewetting could reduce the reflectance contrast between dark organic matter and highly reflecting sand particles (mainly featureless quartz and feldspar). These results involved standardized soil rewetting, which of course was not achieved in the present study; nevertheless spectrum acquisition in the field occurred in winter, in moist but rarely saturated conditions (which render field work very tedious), therefore soil moisture was not very variable. Thus negative effects of rather variable moisture conditions could possibly be offset by positive effects of higher moisture content in the field than in the laboratory. These considerations, and the paucity of explanations proposed in the cited papers, underline the need for fundamental works on interactions between light and soil, and possibly water, in order to help understand the effects of soil sample density and moisture on NIRS prediction accuracy.

Briefly, the present study and literature results suggested that VNIRS predictions of soil properties might have comparable accuracy using spectra acquired in the field or on 2-mm sieved air-dried samples. For total nitrogen, available phosphorus and exchangeable potassium, all studies achieved more (sometimes slightly) accurate predictions on sieved dried samples (Chang et al., 2005; Mouazen et al., 2006; this work); for calcium carbonate, most studies obtained better results on sieved dried samples (Chang et al., 2005; Lagacherie et al.,

2008; Morgan et al., 2009; but not the present study); for organic matter, half of the studies had better predictions in the field or on wet samples (Chang et al., 2005; Stevens et al., 2008; Morgan et al., 2009; Stenberg, 2010; Nocita et al., 2011; this work); for clay (which was not studied here), most studies obtained better results in the field (Waiser et al., 2007; Lagacherie et al., 2008; Viscarra Rossel et al., 2009). Differences between studies could be attributed to variable field scanning procedures, to field conditions (very variable, always dry, always wet), to the sample set (homogeneous or heterogeneous as regarded e.g. mineralogy, texture, depth), chemometrics (type of regression, possible outlier removing, mathematical preprocessing, cross or independent validation, etc.), to procedures carried out in the lab, to possible replications, and probably to other sources.

3.4. Effect of the preprocessing method

Spectrum preprocessing influenced prediction accuracy noticeably in general, but to an extent that depended on the variable considered and on spectrum acquisition procedure (Figure 2). Second derivatives yielded poor predictions and are not considered here. The effect of preprocessing on prediction accuracy was particularly important with CLODav_siev, and to a lesser extent, with CLODav_dry (i.e. on dry samples), and in some other cases (e.g. for organic matter with CLODraw). Thus choosing the preprocessing procedure mattered, especially in laboratory conditions. Considering COREraw and COREcut, which were in general the most appropriate scanning procedures in the field, best predictions were achieved (data not shown):

- for calcium carbonate, with SNVD 0 and 1 (i.e. SNVD with no or first derivatives), and SNV 1, MSC 1, WMSC 1 and IMSC 1 (first derivatives); SNVD 1 provided especially good results;
- for total nitrogen, with SNV 0 and 1, MSC 0 and 1 (no or first derivatives), and D 0 (no derivatives);
- for organic matter, with MSC 1 (first derivatives) and SNV 0 (no derivatives);
- for available phosphorus, with None 0 and 1 (no or first derivatives), D 0 (no derivatives), SNV 0 and 1, IMSC 0 and 1, WMSC 0 and 1, SNVD 0 and 1, and MSC 0 and 1 (no or first derivatives);
- for exchangeable potassium, with None 0 and D 0 (no derivatives); None 0 provided especially good results.

This was estimated according to the absence of significant difference with the lowest SECV achieved for the scanning procedure and soil property considered (Fisher test at $p < 0.05$).

In addition, when using first derivatives, large gaps (15 or 20) often yielded better results than a small one (5).

In short, MSC 1, and to a lesser extent SNV 1, were useful for predicting calcium carbonate, total nitrogen and organic matter (but the former was well predicted whatever the pretreatment); and None 0, and to a lesser extent D 0, for predicting exchangeable potassium (and available phosphorus, which was nevertheless poorly predicted). Comparing None and MSC with first or second derivatives, Ben-Dor et al. (2008) also selected MSC with first derivative to optimize field VNIRS prediction of organic matter. Morgan et al. (2009) also obtained better field VNIRS prediction of organic carbon with first derivatives than with second or without derivatives; however, they obtained better predictions of inorganic carbon without derivatives.

4. Conclusion

Good VNIRS cross-validation could be achieved in the field for calcium carbonate, total nitrogen, organic matter and exchangeable potassium, over a topsoil sample set representing the main areas of large-scale crop cultivation in France: depending on scanning procedure and mathematical pretreatment, RPD reached up to 9.1, 2.9, 2.8 and 3.0, respectively. Very good results for calcium carbonate suggested that conventional laboratory determination is not useful anymore for this property, except for maintaining calibration databases. By contrast, available phosphorous was poorly predicted ($RPD \leq 1.7$), but this soil property is hardly predictable using laboratory VNIRS too, except over very homogeneous sample sets (e.g. samples originating from one field).

The main objective of the work was to identify the most appropriate scanning procedure for field VNIRS prediction of soil properties. Indeed, prediction accuracy depended on the spectrum acquisition method, but to an extent that varied with the property: choosing the appropriate procedure was crucial for organic matter (the right procedure could increase RPD by 0.7); it was useful but less crucial for calcium carbonate (RPD could increase by 2.2; but all scanning methods yielded excellent predictions), total nitrogen (by 0.5), available phosphorus (by 0.3), and exchangeable potassium (by 0.4). Very generally, best predictions were achieved when spectra were acquired on cores, either raw or cut, while predictions from spectra acquired on the soil surface or on clods (resulting from core crumbling) were less accurate.

Spectra acquired in the field are affected by variations in moisture, temperature, stoniness, etc., which render calibration more difficult a priori. Counterintuitively, comparable

prediction accuracies were achieved for laboratory (i.e. on 2-mm sieved air-dried samples) and field VNIRS (on cores). This could be due, at last partly, to the larger number of replicates carried out in the field than in the lab (4 to 12 vs. 2 to 4). As replication is easier in the field than in the lab, this discrepancy represented a difference between field and lab approaches rather than an artifact. Another possible reason for similar prediction accuracies for field and lab VNIRS could be the higher cohesion of samples scanned in the field (in relation with higher density and moisture), which would be beneficial for calibrations and would thus offset the detrimental effects of variable field conditions. Fundamental works on light-soil interactions are needed to clarify the mechanisms involved.

In addition, mathematic preprocessing of spectra affected calibrations, but to a lesser extent using field- than laboratory-acquired spectra. For spectra acquired on soil cores (in the field), best calibrations of calcium carbonate, total nitrogen and organic matter were achieved with MSC, and to a lesser extent SNV, both with first derivatives; best calibrations of exchangeable potassium and available phosphorus were achieved without pretreatment, and with detrend without derivatives to a lesser extent. Using second derivatives always yielded the poorest calibrations.

5. Acknowledgements

The authors thank two anonymous reviewers for their helpful comments.

6. References

- AFNOR (Association Française de Normalisation), 1995a. Qualité du Sol – Détermination de la Teneur en Carbonate – Méthode Volumétrique. NF ISO 10693. AFNOR, Paris.
- AFNOR (Association Française de Normalisation), 1995b. Qualité du sol – Dosage du Phosphore – Dosage Spectrométrique du Phosphore Soluble dans une Solution d'Hydrogénocarbonate de Sodium. NF ISO 11263. AFNOR, Paris.
- AFNOR (Association Française de Normalisation), 1998a. Qualité du Sol – Détermination de la Teneur Totale en Azote par Combustion Sèche ("Analyse Élémentaire"). NF ISO 13878. AFNOR, Paris.
- AFNOR (Association Française de Normalisation), 1998b. Qualité du Sol – Dosage du Carbone Organique par Oxydation Sulfochromique. NF ISO 14235. AFNOR, Paris.
- AFNOR (Association Française de Normalisation), 2002. Qualité des Sols – Détermination des Cations Ca^{++} , Mg^{++} , K^+ , Na^+ Extractibles par l'Acétate d'Ammonium – Méthode par Agitation. NF X31-108. AFNOR, Paris.
- Barnes, R. J., Dhanoa, M.S., Lister, S.J., 1989: Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy* 43, 772-777.
- Barthès, B.G., Brunet, D., Ferrer, H., Chotte, J.L., Feller, C., 2006. Determination of total carbon and nitrogen content in a range of tropical soils using near infrared spectroscopy: influence of replication and sample grinding and drying. *Journal of Near Infrared Spectroscopy* 14, 341-348.

- Barthès, B.G., Brunet, D., Rabary, B., Ba, O., Villenave, C., 2011 Near infrared reflectance spectroscopy (NIRS) could be used for characterization of soil nematode community. *Soil Biology & Biochemistry* 43, 1649-1659.
- Baumgardner, M.F., Silva, L.F., Biehl, L.L., Stoner, E.R., 1985. Reflectance properties of soils. *Advances in Agronomy* 38, 1–44.
- Ben-Dor, E., Banin, A., 1995. Near-infrared analysis as a rapide method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal* 59, 364-372.
- Ben-Dor, E., Heller, D., Chudnovsky, A., 2008. A novel method of classifying soil profiles in the field using optical means. *Soil Science Society of America Journal* 72, 1113–1123.
- Bjørsvik, H.R., Martens, H., 2001. Data analysis: calibration of NIR instruments by PLS regression, in: Burns, D.A., Ciurczak, E.W. (Eds.), *Handbook of Near-infrared Analysis*, 2nd edition. Taylor and Francis, Boca Raton, FL, USA, pp. 185-207.
- Brunet, D., Barthès, B.G., Chotte, J.L., Feller, C., 2007, Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* 139, 106-117.
- Candolfi, A., De Maesschalck, R., Jouan-Rimbaud, D., Hailey, P. A. and Massart, D.L., 1999: The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra. *Journal of Pharmaceutical and Biomedical Analysis*, 21, 115-132.
- Cécillon, L., Barthès, B.G., Gomez, C., Ertlen, D., Genot, V., Hedde, M., Stevens, A., Brun, J.J., 2009. Assessment and monitoring of soil quality using near infrared reflectance spectroscopy (NIRS). *European Journal of Soil Science* 60, 770-784.
- Chang, C.W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R.J., 2001. Near-infrared reflectance spectroscopy - Principal components regression analyses of soil properties. *Soil Science Society of America Journal* 65, 480-490.
- Chang, G.W., Laird, D.A., Hurburgh, G.R., 2005. Influence of soil moisture on near-infrared reflectance spectroscopic measurement of soil properties. *Soil Science* 170, 244–255.
- Coûteaux, M.M., Berg, B., Rovira, P., 2003. Near infrared reflectance spectroscopy for determination of organic matter fractions including microbial biomass in coniferous forest soils. *Soil Biology & Biochemistry* 35, 1587–1600.
- Dhanoa, M.S., Lister, S.J., Sanderson, R., Barnes, R.J., 1994. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *Journal of Near Infrared Spectroscopy* 2, 43-47.
- Dunn, B.W., Beecher, H.G., Batten, G.D., Ciavarella, S., 2002. The potential of near infrared reflectance spectroscopy for soil analysis – a case study from the Riverine Plain of south-eastern Australia. *Australian Journal of Experimental Agriculture* 42, 607-614.
- European Commission, 2006. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions: Thematic Strategy for Soil Protection. COM(2006) 231 final, Brussels.
- Fystro, G., 2002. The prediction of C and N content and their potential mineralisation in heterogeneous soil samples using Vis–NIR spectroscopy and comparative methods. *Plant and Soil* 246, 139–149.
- Geladi, P., MacDougall, D., Martens, H., 1985: Linerization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy* 39, 491-500.
- Gubler, A., 2011. Quantitative Estimations of Soil Properties by Visible and Near Infrared Spectroscopy – Applications for Laboratory and Field Measurements. Ph.D. Dissertation, Bern University, Switzerland.
- Huber, S., Prokop, G., Arrouays, D., Banko, G., Bispo, A., Jones, R.J.A., Kibblewhite, M.G., Lexer, W., Möller, A., Rickson, R.J., Shishkov, T., Stephens, M., Van den Akker, J.J.H., Varallya, G., Verheijen, F.G.A. (Eds.), 2008. *Environmental Assessment of Soil for*

- Monitoring. Volume I: Indicators and Criteria. EUR 23490 EN/1. Office for the Official Publications of the European Communities, Luxemburg.
- IUSS-ISRIC-FAO (International Union of Soil Science, International Soil Reference and Information Centre, Food and Agriculture Organization of the United Nations), 2006. World Reference Base for Soil Resources 2006. FAO, Rome.
- Kusumo, B.H., Hedley, M.J., Tuohy, M.P., Hedley, C.P., Arnold, G.C., 2010. Predicting soil carbon and nitrogen concentrations and pasture root densities from proximally sensed soil spectral reflectance, in: Viscarra Rossel, R.A., McBratney, A.B., Minasny, B. (Eds.), *Proximal Soil Sensing*. Springer, New York, pp. 177-190.
- Lagacherie, P., Baret, F., Feret, J.B., Madeira Netto, J., Robbez-Masson, J.M., 2008. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sensing of Environment* 112, 825–835.
- Malley, D.F., Martin, P.D., Ben-Dor, E., 2004. Application in analysis of soils, in: Roberts, C.A. (Ed.), *Near-Infrared Spectroscopy in Agriculture*. Agronomy Monograph n°44. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, Wisconsin, USA, pp. 729-784.
- Morgan, C.L.S, Waiser, T.H., Brown, D.J., Hallmark, C.T, 2009. Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma* 151, 249-256.
- Mouazen, A.M., De Baerdemaeker, J., Ramon, H., 2006. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 14, 189-199.
- Nocita, M., Kooistra, L. Bachmann, M., Müller, A., Powell, M., Weel, S., 2011. Predictions of soil surface and topsoil organic carbon content through the use of laboratory and field spectroscopy in the Albany Thicket Biome of Eastern Cape Province of South Africa. *Geoderma* 167-168, 295-302.
- Osborne, B.G., Fearn, T., Hindle, P.T., 1993. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*, 2nd edition. Longman Scientific and Technical, Harlow, Essex, UK.
- Shenk, J.S., Westerhaus, M.O., 1991. Population structuring of near infrared spectra and modified partial least square regression. *Crop Science* 31, 1548–1555.
- Sørensen, L.K., Dalsgaard, S., 2005. Determination of clay and other soil properties by near infrared spectroscopy. *Soil Science Society of America Journal* 69, 159–167.
- Stenberg, B., 2010. Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma* 158, 15-22.
- Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. *Advances in Agronomy* 107, 163–215.
- Stevens, A., van Wesemael, B., Bartholomeus, H., Rosillon, D., Tychon, B., Ben-Dor, E., 2008. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma* 144, 395–404.
- Sudduth, K.A., Hummel, J.W., 1993. Soil organic-matter, CEC, and moisture sensing with a portable NIR spectrophotometer. *Transactions of the ASAE* 36, 1571–1582.
- Udelhoven, T., Emmerling, C., Jarmer, T., 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant and Soil* 251, 319–329.
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma* 150, 253–266.

Waiser, T.H., Morgan, C.L.S., Brown, D.J., Hallmark, C.T., 2007. In situ characterization of soil clay content with visible near-infrared diffuse reflectance spectroscopy. *Soil Science Society of America Journal* 71, 389–396.

Table 1. Distribution of measured properties over the studied topsoil (0-20 cm) sample population.

Soil property	Min	Max	Mean	SD ^a
Calcium carbonate (g 100 g ⁻¹)	0.0	84.9	16.1	27.3
Total nitrogen (g kg ⁻¹)	0.04	0.46	0.16	0.07
Organic matter (g 100 g ⁻¹)	1.07	9.23	3.22	1.42
Available phosphorous (mg P ₂ O ₅ kg ⁻¹)	17.5	177.5	65.5	31.0
Exchangeable potassium (mg K ₂ O kg ⁻¹)	31.0	684.0	236.5	156.0

^a SD is the standard deviation.

Table 2. Cross-validation statistics using preprocessing methods that yielded best predictions of the soils properties considered at 0-20 cm depth (SECV and bias are in the unit of the soil property; RPD, slope and R² are unitless; for preprocessing, see section 2.3.). The scanning procedures were (cf. section 2.2): on the soil surface (SURFACE); on raw/cut auger cores (COREraw/cut); on clods crumbled from cores separately (CLODraw) or together (CLODav), after a half-day (CLODav+0.5; field), the next morning (CLODav+1; field), after drying (CLODav_dry; lab), and after drying and sieving (CLODav_siev; lab).

Soil property	Sampling procedure	Preprocessing	SECV ^a	RPD ^b	Slope	Bias	R ²
Calcium carbonate (g 100 g ⁻¹)							
	SURFACE	SNV 1 20 5	3.60	7.4	1.00	-0.06	0.98
	COREraw	SNV 1 15 5	3.01	9.1	1.00	0.07	0.99
	COREcut	SNVD 1 20 5	3.20	8.5	0.99	0.01	0.99
	CLODraw	SNV 1 20 5	3.34	8.1	1.00	0.02	0.98
	CLODav	SNV 1 20 5	3.90	7.1	1.00	-0.09	0.98
	CLODav+0.5	SNV 1 20 5	3.74	7.3	0.99	-0.09	0.98
	CLODav+1	MSC 1 20 5	4.01	6.9	1.00	-0.03	0.98
	CLODav_dry	MSC 1 20 5	3.25	8.4	0.99	0.03	0.99
	CLODav_siev	WMSC 1 15 5	3.16	8.6	1.00	-0.06	0.99
Total nitrogen (g 100g ⁻¹)							
	SURFACE	SNV 0 0 5	0.024	2.6	0.99	0.000	0.85
	COREraw	SNV 1 15 5	0.023	2.7	0.98	-0.001	0.87
	COREcut	SNV 0 0 1	0.022	2.9	1.00	0.001	0.88
	CLODraw	MSC 0 0 1	0.024	2.7	0.98	0.004	0.86
	CLODav	MSC 1 20 5	0.024	2.6	0.97	-0.001	0.86
	CLODav+0.5	MSC 1 20 5	0.025	2.5	0.99	0.000	0.84
	CLODav+1	MSC 1 20 5	0.025	2.6	0.98	0.000	0.85
	CLODav_dry	WMSC 1 20 5	0.022	2.9	0.98	0.001	0.88
	CLODav_siev	WMSC 1 15 5	0.022	3.0	0.98	0.000	0.89
Organic matter (g 100 g ⁻¹)							
	SURFACE	SNV 0 0 1	0.48	2.4	0.98	-0.020	0.82
	COREraw	MSC 1 20 5	0.42	2.8	0.99	-0.002	0.86
	COREcut	MSC 1 20 5	0.42	2.7	0.97	-0.002	0.86
	CLODraw	MSC 1 20 5	0.45	2.4	0.96	-0.015	0.83
	CLODav	SNV 1 20 5	0.54	2.2	0.96	0.002	0.78
	CLODav+0.5	D 0 0 5	0.55	2.1	0.96	0.012	0.77
	CLODav+1	D 1 15 5	0.53	2.2	0.94	-0.002	0.80
	CLODav_dry	WMSC 1 15 5	0.48	2.3	0.96	0.010	0.81
	CLODav_siev	MSC 0 0 1	0.51	2.3	0.98	0.013	0.80
Available phosphorous (mg P ₂ O ₅ kg ⁻¹)							
	SURFACE	SNV 0 0 5	18.3	1.6	0.94	-0.2	0.59
	COREraw	None 0 0 5	17.5	1.7	0.95	-0.1	0.65
	COREcut	D 0 0 5	18.3	1.6	0.97	-0.2	0.59
	CLODraw	SNV 0 0 5	17.8	1.6	0.94	-0.1	0.62
	CLODav	MSC 0 0 5	17.9	1.6	0.95	0.1	0.61
	CLODav+0.5	SNVD 1 15 5	17.8	1.6	0.93	0.1	0.62
	CLODav+1	SNV 0 0 5	17.3	1.7	0.94	-0.1	0.63
	CLODav_dry	D 0 0 1	17.0	1.7	0.89	-0.0	0.64
	CLODav_siev	SNV 0 0 5	16.1	1.8	0.97	0.0	0.70
Exchangeable potassium (mg K ₂ O kg ⁻¹)							
	SURFACE	None 0 0 5	53.8	2.9	1.00	0.8	0.88
	COREraw	None 0 0 5	53.1	2.9	0.98	0.6	0.88
	COREcut	None 0 0 1	53.0	2.9	0.98	0.5	0.88
	CLODraw	None 0 0 1	51.2	3.0	0.98	0.9	0.89
	CLODav	D 1 20 5	55.1	2.8	1.00	1.8	0.88
	CLODav+0.5	IMSC 1 20 5	55.6	2.7	1.00	1.1	0.85
	CLODav+1	D 1 20 5	53.6	2.9	0.97	-0.4	0.88
	CLODav_dry	MSC 1 20 5	50.1	3.1	0.97	-0.6	0.90
	CLODav_siev	MSC 1 15 5	49.1	3.2	0.98	0.7	0.90

^a SECV is the standard error of cross-validation ^b RPD is the ratio of standard deviation to SECV.

Figure 1. Comparison between measured and cross-validation predicted values of soil contents in calcium carbonate, total nitrogen, organic matter, available phosphorous, and exchangeable potassium at 0-20 cm depth (mentioning the most appropriate scanning procedure in the field, either COREraw on raw cores, COREcut on cut cores, or CLODraw on clods crumbled from cores separately, cf. section 2.2).

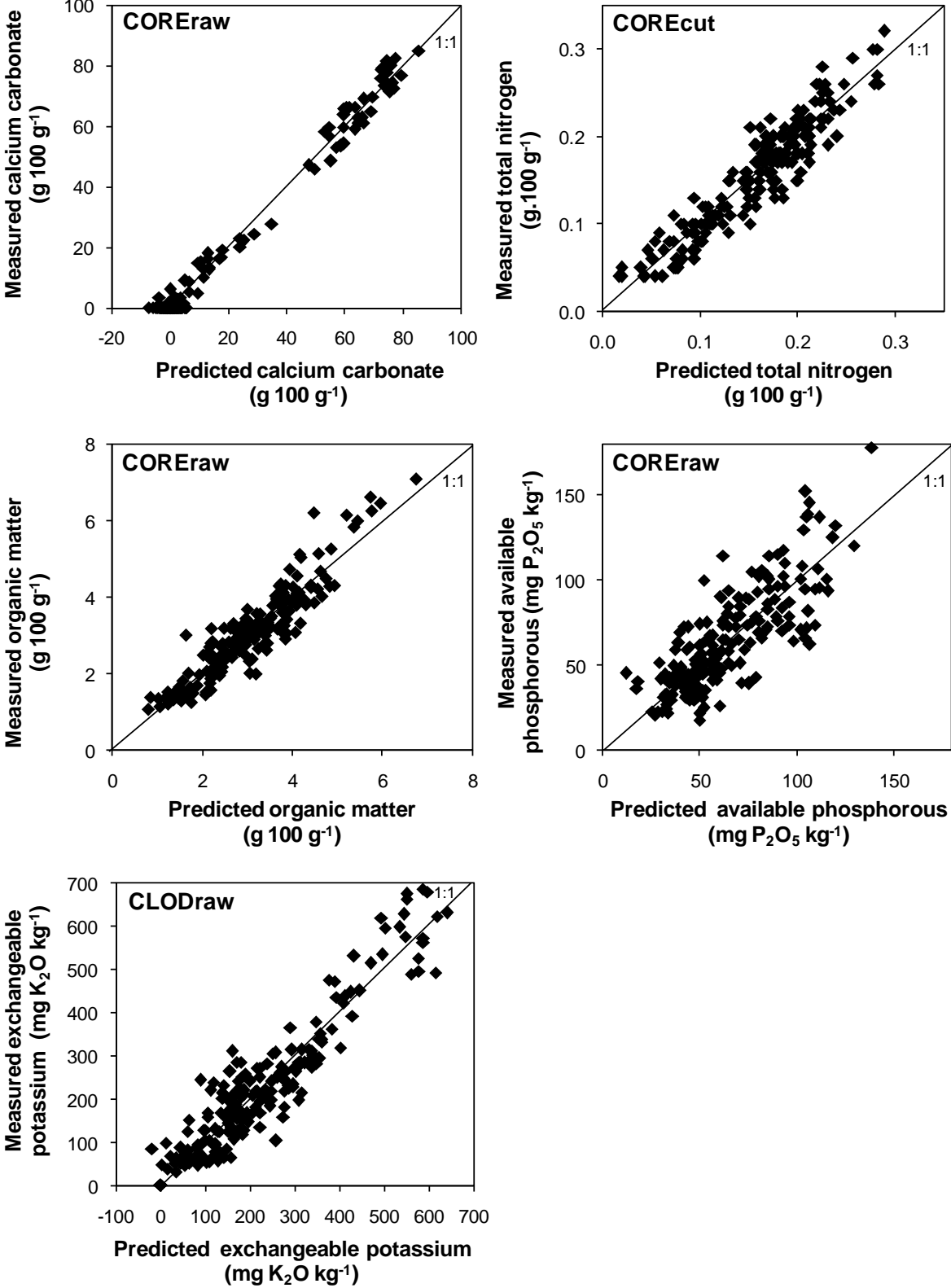


Figure 2. Minimum, maximum and mean cross-validation RPD for each soil property studied and every scanning procedure, depending on preprocessing procedures (second derivatives being excluded due to very poor predictions). The scanning procedures were (cf. section 2.2): on the soil surface (SURFACE); on raw/cut auger cores (COREraw/cut); on clods crumbled from cores separately (CLODraw) or together (CLODav), after a half-day (CLODav+0.5; field), the next morning (CLODav+1; field), after drying (CLODav_dry; lab), and after drying and sieving (CLODav_siev; lab).

