



HAL
open science

Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions

Victor Allory, Aurélie Cambou, Patricia Moulin, Christophe Schwartz, Patrice Cannavo, Laure Vidal Beaudet, Bernard G. Barthès

► To cite this version:

Victor Allory, Aurélie Cambou, Patricia Moulin, Christophe Schwartz, Patrice Cannavo, et al.. Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) in situ or in laboratory conditions. *Science of the Total Environment*, 2019, 686, pp.764-773. 10.1016/j.scitotenv.2019.05.192 . hal-02136840v3

HAL Id: hal-02136840

<https://hal.science/hal-02136840v3>

Submitted on 14 Sep 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.

Quantification of soil organic carbon stock in urban soils using visible and near infrared reflectance spectroscopy (VNIRS) *in situ* or in laboratory conditions

Victor Allory^{a,b}, Aurélie Cambou^{a,b,*}, Patricia Moulin^{c,d}, Christophe Schwartz^a, Patrice Cannavo^b, Laure Vidal-Beaudet^b, Bernard G. Barthès^d

^a Université de Lorraine, INRA, Laboratoire Sols et Environnement, 54000 Nancy, France.

^b EPHOR, Agrocampus Ouest, IRSTV, 49045 Angers, France.

^c Imago, IRD, Dakar, Senegal.

^d Eco&Sols, Université de Montpellier, CIRAD, INRA, IRD, Montpellier SupAgro, 34060 Montpellier, France.

* Corresponding author: <aurelie.cambou@ird.fr>

Abstract

Urban soils, like other soils, can be sink or source for atmospheric carbon dioxide, and due to urban expansion, are receiving increasing attention. Studying their highly variable attributes requires high-density sampling, which can hardly be achieved using conventional approaches. The objective of this work was to determine the ability of visible and near infrared reflectance spectroscopy (VNIRS) to quantify soil organic carbon (SOC) concentration (gC kg^{-1}) and stock (gC dm^{-3} , or MgC ha^{-1} for a given depth layer) in parks and sealed soils of two French cities, Marseille and Nantes, using spectra collected on pit walls or in laboratory conditions (air dried, 2 mm sieved samples).

Better VNIRS predictions were achieved using laboratory than *in situ* spectra ($R^2 \approx 0.8-0.9$ vs. $0.7-0.8$ in validation), and for sample SOC concentration than stock (R^2_{val} up to 0.83 *in situ* and 0.95 in the laboratory vs. 0.78 and 0.89 , respectively). Stock was conventionally calculated according to four methods that variably account for coarse particles (> 2 mm); and it was better predicted when coarse particles were not taken into account. This was logical using laboratory spectra, collected on 2 mm sieved samples; but concerning *in situ* spectra, this suggested the operator tended to put the spectrometer beside the coarsest particles during spectrum acquisition. This point is worth considering for urban soils, often rich in coarse particles.

Stocks were then aggregated at the profile level: SOC stock prediction was more accurate at profile than sample level when using laboratory spectra ($R^2_{\text{val}} = 0.94$ vs. 0.89 , respectively), probably due to uncertainty compensation; but this was not the case when using *in situ*

spectra, possibly because samples collected for SOC analysis and corresponding VNIRS scans were not at the exact same location.

This work demonstrates VNIRS usefulness for quantifying SOC stock time- and cost-effectively, in urban soils especially.

Keywords

Diffuse reflectance spectroscopy; Soil organic carbon concentration; French cities; Sealed soils; Coarse particles

1. Introduction

Soils represent the largest terrestrial pool of organic carbon and they are in strong interaction with the atmosphere (Jacobson *et al.*, 2000; Scharlemann *et al.*, 2014). Soils can behave as a sink or source for atmospheric carbon dioxide (CO₂), depending on many factors such as land use and management (Dignac *et al.*, 2017). Thus soil organic carbon (SOC) maintenance is an important issue in the current context of climate change, beside the long-acknowledged and key role of SOC in soil physical, chemical and biological fertility (Reeves, 1997). Initiatives have been launched to support states and non-governmental actors in promoting better soil management, for a long-term maintenance of SOC. However, efforts have mainly been focusing on cultivated and forested soils (<http://4p1000.org/>¹; Paustian *et al.*, 2016). Currently, artificialized lands represent nearly 3% of terrestrial areas and this proportion will increase as a result of increasing urban population, which is expected to reach 66% of the World population by 2050 (Liu *et al.*, 2014; United Nations, 2014). Some attention has begun to be paid to SOC of artificialized soils since the 2000s, and several studies over the World have shown that SOC amount was generally higher in urban open soils than in cultivated soils and could be the same order of magnitude as under forest and grassland (Pouyat *et al.*, 2009; Edmondson *et al.*, 2012; Vasenev *et al.*, 2014; Cambou *et al.*, 2018). Urban soil attributes, including SOC, are highly variable spatially and temporally, and can be highly disturbed over short periods of time and short distances due to strong anthropic pressure (Bae & Ryu, 2015; Lorenz & Lal, 2015). Thus, quantifying soil attributes in urban areas requires high-density sampling and numerous analyses, which is time consuming and expensive. As a consequence, studies on SOC concentrations (gC kg⁻¹) and stocks (kgC m⁻² or MgC ha⁻¹ for a given soil depth) in urban areas have only been able to provide one-off states of play to date.

¹ Last accessed 21 Jan. 2019.

For a soil sample, SOC stock (gC dm^{-3}) is calculated as the product of SOC concentration and bulk density, or apparent volumetric mass (kg dm^{-3}), which is the ratio of dry soil mass to apparent volume, the latter being the total volume of soil *in situ*, including voids (pores). To date, the measurement of bulk density is tedious as it requires collecting undisturbed samples (e.g. using beveled cylinders of known volume, perfectly filled with undisturbed material); thus SOC stock determination is tedious. Moreover, determining SOC stock is particularly complex in urban soils, firstly because of their spatial and temporal variability, and also because sampling is difficult in sealed soils, which represent 70% of Europe's urban areas (European Commission, 2012). Properly addressing the question of SOC stock in urban areas would require time- and cost-effective methods.

In the last decades, much attention has been paid to visible and near infrared reflectance spectroscopy (VNIRS) for characterizing soil properties time- and cost-effectively (Stenberg *et al.*, 2010; Nocita *et al.*, 2015). VNIRS uses diffuse reflectance in the 350-2500 nm range for quantifying soil properties based on calibration models: these models use calibration samples, characterized both spectrally and conventionally, for expressing the properties considered as multivariate functions of VNIR spectrum; then the models can be applied to predict the properties considered on new samples from their VNIR spectra (Burns & Ciurczak, 2001). Numerous papers have reported the ability of VNIRS for quantifying SOC concentration (Brunet *et al.*, 2008; Stenberg *et al.*, 2010; Clairotte *et al.*, 2016; Viscarra Rossel *et al.*, 2016), even with spectrum acquisition *in situ* (Stevens *et al.*, 2008; Kusumo *et al.*, 2010; Nocita *et al.*, 2011; Gras *et al.*, 2014). To date, most work on SOC quantification by VNIRS has regarded SOC concentration, implying that SOC stock determination would still require the measurement of soil bulk density. However, a few pioneering studies have recently demonstrated that VNIRS could also be used to quantify SOC stock directly *in situ*, without having to determine bulk density specifically (Roudier *et al.*, 2015; Cambou *et al.*, 2016). To date, such work has been limited to agricultural fields.

The objective of the present work was to use VNIRS *in situ* for quantifying SOC concentration and stock for different urban uses (parks, fallows and sealed soils) in two French cities, Marseille and Nantes, which differ in climate, geology and history. Spectra were also acquired on air-dried, 2 mm sieved samples to compare predictions made from spectra acquired *in situ* vs. in laboratory conditions.

2. Materials and methods

2.1. Study sites

Sites were studied in two contrasted French cities: Marseille, in calcareous environment under Mediterranean climate; and Nantes, in acidic environment under oceanic climate (Table 1).

2.1.1. Marseille

Marseille is located in the south of France, along the Mediterranean Sea, with altitude ranging from 0 to 632 m a.s.l.. The city is settled in an Oligocene basin surrounded by reliefs or opened onto the sea. According to Köppen-Geiger classification, the climate is Csa (Mediterranean), with mean annual temperature and rainfall of 14.5°C and 518 mm, respectively. Two sites were studied in this city, the Borély Park and the Sainte-Marthe wasteland (Table 1).

The Borély Park was designed in 1860-1880 over an area of 54 ha, which has now been reduced to 17 ha. It is covered by lawn and managed tree groves and is located on recent river alluvium bedrock rich in carbonates and made of silts, sands, gravels and stones.

At the time of sampling (2017), the area of Sainte-Marthe had its largest part covered by grass and several trees (*ca.* 3.4 ha), and the other part covered by a grove (< 1 ha). The site has long been fallow, with no road or building until 2013, but was cleared between 2009 and 2011 except for some trees. The area is located on Lower Oligocene bedrock (Stampian), including clay and conglomerates rich in inorganic carbon.

2.1.2. Nantes

Nantes is located in the west of France, on the Loire River, about 50 km from the Atlantic Ocean, with altitude ranging from 0 to 55 m a.s.l.. Nantes is located at the confluence of several rivers, with outcroppings of the granitic bedrock possibly overlaid by loess deposits. Moreover, backfills are present in many parts of the city, with thickness up to 10 m. According to Köppen-Geiger classification, the climate is Cfb (temperate oceanic), with mean annual temperature and rainfall of 12.2°C and 725 mm, respectively. Two sites were also studied in Nantes, called respectively “cemetery” and “railway station”.

The cemetery site covers an area of 50 ha in the north of the city. The cemetery construction began in the 1950s and many trees were planted until 1969. The site included two distinct parts: one covered by well managed vegetation, mainly lawn, trees and hedges, the other by spontaneous grassland and urban woodland. The geology is characterized by loess deposits made of silts, gravels and clay.

The railway station site is located just north to the Nantes railway station and covers over 3000 m². It is characterised by sealed soils, mostly car parks and pavements, developed from sandy backfills that might have been brought and sealed since the 1960s.

2.2. Soil sampling and conventional determinations

At each site, three or four 2 m long, 1 m wide, 1 m deep pits were dug using an excavator. Each soil profile was divided into two to seven horizons according to macromorphology, and the horizons were described in terms of color, structure, texture and stoniness. In each horizon, two to four (disturbed) soil samples were collected with a knife, for laboratory analyses, and one undisturbed sample was collected for bulk density measurement, using a 0.25 L beveled cylinder pushed perpendicular to the pit wall, all roughly at the same depth (within a given pit); thus sampling was carried out on horizon basis. In total 15 pits were studied and 137 disturbed soil samples and 49 undisturbed soil samples were collected. Samples were individually conditioned in plastic bags then brought back to the laboratory (Pansu *et al.*, 2001).

The disturbed soil samples were air-dried then gently broken up and sieved to 2 mm to separate fine earth from coarse particles (Pansu *et al.*, 2001). SOC concentration was determined on 0.2 mm ground aliquots by two methods (Pansu & Gautheyrou, 2006): (1) by difference between total carbon analyzed by dry combustion (ISO, 1995a) using a CHN elemental analyzer (Flash EA 1112, CE Instruments, Rhodano, Italy; 25 mg aliquots were used) and soil inorganic carbon (SIC) analyzed by volumetric calcimetry (ISO, 1995b) using a Bernard calcimeter (Prolabo, ref. 05 215.00, Paris, France; 5 mL chlorhydric acid 6 M were used with 0.5 to 5 g soil aliquots depending on effervescence during a preliminary test); and (2) by dry combustion after decarbonatation by chlorhydric acid (HCl), which was achieved using a procedure close to that proposed by Nieuwenhuize *et al.* (1994): 10 µL HCl 4 M were slowly added to 25 mg of soil that had been weighed in a silver capsule (resistant to HCl), then the capsule was dried 4 h at 40°C; HCl addition then drying were repeated until gaseous emission ceased. The first method, by difference and presumably more accurate, was used for SIC-rich samples (SIC > 15 g kg⁻¹), and the second, direct after decarbonatation, for SIC-poor samples (< 15 g kg⁻¹; Pansu & Gautheyrou, 2006).

The undisturbed soil samples were also air-dried then gently broken up and sieved to 2 mm to separate fine earth from coarse particles (Pansu *et al.*, 2001), the dry mass of which was determined after 48 h oven-drying at 105°C (Gardner, 1986). The volume of coarse particles

was measured in a graduated cylinder partly filled with water, by difference between water levels before and after the addition of coarse particles.

Sample SOC stock was calculated according to four methods (Poeplau *et al.*, 2017):

$$\text{SSOC} = \text{SOC} \times (M_t / V_t) \quad (\text{Eq.1})$$

$$\text{SSOC} = \text{SOC} \times [(M_t - M_{>2}) / (V_t - V_{>2})] \quad (\text{Eq.2})$$

$$\text{SSOC} = \text{SOC} \times (M_t / V_t) \times [1 - (V_{>2} / V_t)] \quad (\text{Eq.3})$$

$$\text{SSOC} = \text{SOC} \times [(M_t - M_{>2}) / (V_t - V_{>2})] \times [1 - (V_{>2} / V_t)] \quad (\text{Eq.4})$$

where SSOC is the SOC stock (gC dm^{-3}), SOC the concentration (gC kg^{-1} soil < 2 mm), V_t and M_t the volume (0.25 L) and dry mass of total sample, and $V_{>2}$ and $M_{>2}$ the volume and dry mass of coarse particles (> 2 mm), respectively. The four methods differ in the way the coarse particles are taken into account: not at all in the first method, partially in the second and third methods, and completely in the fourth method, which has thus been considered more accurate (Poeplau *et al.*, 2017). It is worth noting that for each horizon, SOC concentrations determined on two to four disturbed soil samples were combined with masses and volumes measured on one undisturbed sample. The significance of differences ($p < 0.05$) between sample SOC stocks calculated according to the four methods was tested using the Wilcoxon signed-rank test, which is a paired difference test (R Core Team, 2017). In addition, the significance of differences in SOC concentration or stock between cities or between land uses was tested by ANOVA followed by a *post hoc* HSD Tukey test, after checking the normality of residuals distributions (R Core Team, 2017).

The SOC stock of a given soil horizon (in gC dm^{-2} , kgC m^{-2} or MgC ha^{-1}) was calculated as the product of sample SOC stock (in gC dm^{-3}) by horizon thickness (dm), averaged over the two to four samples of the horizon. Next, the SOC stock of the entire profile was calculated by adding up horizon SOC stocks.

2.3. Spectrum acquisition and analysis

The acquisition of visible and near infrared reflectance (VNIR) spectra was firstly carried out *in situ*, on the cleaned surface of pit walls, about 5 cm above and 5 cm below the location of each disturbed soil sample collected, in the same horizon, both spectra being then averaged. VNIR spectrum acquisition was secondly carried out in laboratory conditions, on air-dried, 2 mm sieved then oven-dried samples (48 h at 40°C), without replication (Barthès *et al.*, 2006). For both *in situ* and laboratory acquisitions, diffuse reflectance was measured from 350 to 2500 nm at 1 nm interval using a portable spectrophotometer LabSpec 2500 (ASD,

Boulder, CO, USA). In this device, light is delivered to the sample by a contact probe (about 80 mm² area), which then collects the reflected signal and transmits it to the spectrometer. After every spectral acquisition, the window of the contact probe was cleaned with lens paper and ethanol. The white reference standard, with nearly 100% reflectance, was a disk made of Spectralon (compressed polytetrafluoroethylene powder; Spectralon SRS-99 custom-made for ASD, 91 mm in diameter and 5.5 mm in thickness, Labsphere, North Sutton, NH, USA) and its reflectance was measured every 10 acquisitions. Each reflectance spectrum provided by the spectrometer resulted from the averaging of 32 co-added scans. Spectral data were recorded as (apparent) absorbance, which is the logarithm of the inverse of reflectance [$\log_{10}(1/\text{reflectance})$] (Stenberg *et al.*, 2010).

Spectrum analysis consisted in fitting the VNIR spectra to SOC concentration or stock as determined conventionally. This was done by partial least squares (PLS) regression, which is the most common procedure for such analysis to date (Bjørsvik & Martens, 2001; Stenberg *et al.*, 2010). It reduces the spectral data to a few orthogonal combinations of all absorbances, called latent variables (LV) or terms, which account for most spectral information and covary with the reference values. Spectral data analysis was conducted using The Unscrambler X 10.4 software (CAMO, Oslo, Norway).

Firstly, VNIR spectra were pretreated, which consists of mathematically transforming the signal in order to amplify its useful parts (*i.e.* relating to SOC) and reduce irrelevant information (*e.g.* resulting from light scattering). Pretreatment always involved smoothing with a Savitsky-Golay filter of order 2 and width 11 (Savitsky & Golay, 1964), alone or possibly combined with a second pretreatment: standard normal variate transformation (SNV), which reduces multiplicative effects; first- or second-order detrend (D1 and D2, respectively), which removes simple additive and multiplicative effects (Barnes *et al.*, 1989); or Savitsky-Golay first- or second-order derivation, with second order polynomial over 5, 11 or 25 points (denoted SG 1 2 5, 1 2 11 and 1 2 25, and SG 2 2 5, 2 2 11 and 2 2 25, respectively), which reduces baseline variation and enhances spectral features (Savitsky & Golay, 1964). Spectra were reduced to 400-2500 and 450-2500 nm after first- and second-order derivation, respectively, because their lower end was noisy.

A principal component analysis (PCA) was then carried out on all smoothed *in situ* spectra. The Hotelling's T² distance was computed on PCA scores, and five samples with distance $> 2 T^2_{\alpha=0.05}$ were removed as spectral outliers (Jensen & Ramirez, 2017). The remaining set, including 132 samples, was then divided into a calibration subset, used to build the prediction model, and a validation subset, used to test it. One pit per site was used for

validation, and the two or three other pits of the site for calibration. A PCA was performed on the smoothed *in situ* spectra of each site to select its validation pit, which should not have its samples too scattered and atypical. After deciding which pits would be used for calibration and validation, the calibration subset included 95 samples and the validation subset 37.

A cross-validation was then performed on the calibration subset, which was ranked according to sampling time then divided cyclically in six groups (*i.e.* the 1st, 7th, 13th, *etc.* samples in the first group, the 2nd, 8th, 14th, *etc.* samples in the second group, *etc.*, the 6th, 12th, 18th, *etc.* samples in the sixth group). The optimal number of PLS LV was the number of LV after which the root mean square error of cross-validation (RMSECV) no longer decreased meaningfully (Bjørsvik & Martens, 2001).

The performance of the prediction model was tested on the validation samples, which belonged to the pits that had not been used for calibration and could thus be considered independent. This performance was evaluated according to three figures of merit: the coefficient of determination R^2_{val} between observed and VNIRS-predicted values (SOC concentration or stock) over the validation subset; the root mean square error of prediction (RMSEP; see Eq.5) between observed and VNIRS-predicted values over the validation subset; and the RPD_{val} ratio, calculated by dividing the standard deviation (SD) of the validation subset by RMSEP. The latter was calculated as follows:

$$\text{RMSEP} = \sqrt{\frac{\sum_1^N (y_i - \hat{y}_i)^2}{N}} \quad (\text{Eq.5})$$

where y_i and \hat{y}_i are the observed and VNIRS-predicted values for sample i , respectively, and N the total number of samples in the validation subset.

Predictions were then averaged at the horizon level and summed up at the profile levels. RMSEP between observed and VNIRS-predicted horizon SOC stock was calculated according to Eq.5, with y_i and \hat{y}_i the observed and VNIRS-predicted values for horizon i , respectively, and N the total number of horizons. Similar calculations were then made for profiles.

3. Results

3.1. Reference data

Sample SOC stocks were calculated according to four methods (Eq.1 to Eq.4), which yielded closely correlated values (on 132 samples; five spectral outliers were removed, *cf.* Section 2.3). The closest correlations occurred between Eq.1 and Eq.2, between Eq.2 and Eq.3, and

between Eq.3 and Eq.4 ($R^2 \geq 0.98$, slope ≥ 0.90), and the weakest between Eq.1 and Eq.4 ($R^2 = 0.93$), which is consistent with the fact that coarse particles (> 2 mm) were increasingly and progressively taken into account from Eq.1 to Eq.4. Nevertheless, according to the Wilcoxon signed-rank test (R Core Team, 2017), there were significant differences ($p < 0.05$) between sample SOC stocks determined according to Eq.1, Eq.2, Eq.3 and Eq.4: their mean and SD were 20.1 ± 16.9 , 18.4 ± 15.4 , 17.7 ± 14.8 and 16.3 ± 13.7 gC dm⁻³, respectively. Thus, better accounting for coarse particles decreased SOC stock value significantly. Particles > 2 mm represented a noticeable proportion of the studied samples: 232 g kg⁻¹ in average (SD 161 g kg⁻¹); and this proportion tended to be higher in Nantes (284 g kg⁻¹ in average, SD 162 g kg⁻¹) than in Marseille (148 g kg⁻¹ in average, SD 118 g kg⁻¹), with no clear effect of land use (e.g. in Nantes, 291 g kg⁻¹ in sealed soils vs. 282 g kg⁻¹ in parks and fallows, in average).

The minimum, maximum, mean and SD of observed sample SOC concentration and stock are presented in Table 2, for Marseille (48 samples including 26 in parks and 22 in fallows) and Nantes (84 samples including 32 in parks, 31 in fallows and 21 in sealed soils), stock being calculated according to Eq.4 (which was considered the most relevant). According to ANOVA followed by a *post hoc* HSD Tukey test (R Core Team, 2017), sample SOC concentration and stock were significantly higher in Marseille, where soils are carbonated silty clay, than in Nantes, where soils are acidic and slightly sandier ($p < 0.05$). For one given city (Marseille or Nantes), sample SOC concentration or stock did not differ significantly between parks and fallows, possibly because sample depth was not taken into account, thus its effect masked that of land use (sampling based on pedological horizons did not allow comparisons between land uses at a given depth). In contrast, SOC concentration and stock in Nantes were significantly lower in sealed soils than under parks and under fallows.

3.2. Spectra

Some of the differences in sample SOC reported in the previous section, between cities or land uses, could also be found when considering spectra, or more precisely, PCA built from spectra. Two PCA were performed on spectra acquired *in situ* and in laboratory conditions, respectively (Figure 1). The PCA performed on laboratory spectra showed that Marseille samples had generally positive PC2 scores and Nantes samples negative PC1 scores, except sealed soil (in Nantes only), which had positive PC1 scores and negative PC2 scores in general. Indeed, the clearest feature was the tendency of sealed soil spectra to group, toward

high PC2 scores for *in situ* spectra and toward high PC1 scores and low (negative) PC2 scores for laboratory acquisitions.

3.3. Prediction of sample SOC concentration and stock (determined according to Eq.4)

On the whole, predictions were more accurate using laboratory spectra than *in situ* spectra, and for SOC concentration than SOC stock. Thus, the best predictions were achieved for SOC concentration, then for SOC stock in laboratory conditions, while results for SOC concentration and stock using *in situ* spectra differed less (best RPD_{val} reached 4.6, 3.1, 2.4 and 2.2, respectively; Table 3).

Among the spectrum pretreatments that were used in addition to smoothing, D1 and D2 yielded the best predictions of SOC concentration and stock when using *in situ* spectra; while this was achieved with SNV and SG 1 2 11 when using laboratory spectra. Smoothing alone yielded predictions that were among the worst in general.

Figure 2 compares conventional determinations and VNIRS predictions of SOC concentration and stock on the validation subset using *in situ* or laboratory spectra with pretreatments that yielded the best validation results. It shows poor predictions for several samples collected in sealed soils, especially for SOC stock and especially using *in situ* spectra, with residuals often larger than observed values, which were low (*cf.* Table 2).

3.4. Prediction of sample SOC stock according to its conventional determination procedure

The fitting of *in situ* spectra did not differ much according to the equation used for calculating SOC stock conventionally. Nevertheless, over a range of spectrum pretreatments that yielded good results in general, sample stock values tended to be more accurately predicted when calculated with Eq.2 than with the other equations, Eq.3 especially (Table 4). Moreover, with smoothing + D2, which was the most appropriate pretreatment when predicting SOC stock from *in situ* spectra, conventional stock values tended to be more accurately predicted when calculated with Eq.1 than with Eq.4, while Eq.2 and Eq.3 provided intermediate results.

The fitting of laboratory spectra was more affected by conventional SOC stock calculation: over a range of pretreatments that yielded good results in general, the prediction of sample SOC stock tended to be more accurate when calculated with Eq.1 than with Eq.4, while Eq.2 and Eq.3 provided intermediate results (Table 4). With smoothing + SG 1 2 11, which was the most appropriate pretreatment when predicting SOC stock from laboratory spectra, SOC stock tended to be more accurately predicted when calculated with Eq.2, and to a lesser extent Eq.3 than Eq.1, than with Eq.4.

3.5. Prediction of horizon and profile SOC stocks (determined according to Eq.4)

Considering SOC stocks at the horizon level, which included several samples, and at the profile level, which included several horizons, prediction accuracy was similar than at the sample level when *in situ* spectra were used: $RPD_{val} = 2.0$ and 2.3 vs. 2.2 , respectively (with smoothing + D2, which was the most appropriate pretreatment here; Table 5 for horizons and profiles, Table 3 for samples). When laboratory spectra were used, predictions were more accurate at the horizon level and even more at the profile level than at the sample level: $RPD_{val} = 4.0$ and 4.4 vs. 3.1 , respectively (with smoothing + SG 1 2 11, which was the most appropriate pretreatment here; Tables 5 and 3). Figure 3 compares observed and predicted profile SOC stocks using *in situ* or laboratory spectra with the most appropriate pretreatments. Considering the four validation profiles separately, observed and predicted stocks at Borély Park, Ste Marthe wasteland (Marseille), cemetery and railway station (Nantes) were respectively:

- 252 vs. 196, 310 vs. 356, 71 vs. 57, and 17 vs. 112 MgC ha⁻¹ using *in situ* spectra;
- 252 vs. 209, 310 vs. 313, 71 vs. 48, and 17 vs. 58 MgC ha⁻¹ using laboratory spectra.

Prediction was particularly poor for the validation profile of the railway station site, especially using *in situ* spectra. This was also the case, though to a lesser extent, for one out of the three calibration profiles from Nantes railway station (observed vs. predicted SOC stock using *in situ* spectra was 6 vs. 50 MgC ha⁻¹, respectively). Thus sample SOC stock in sealed soils was rather difficult to predict using *in situ* spectra, and this also tended to be the case using laboratory spectra, though to a much lesser extent (Figure 3).

4. Discussion

4.1. Overall considerations on predictions

Chang *et al.* (2001) considered that NIRS predictions of soil attributes with $RPD > 2$ were accurate, and in the present study, this could be achieved for both SOC concentration and stock, using both *in situ* and laboratory spectra. Among spectra pretreatments, which aim at reducing additive and/or multiplicative effects due to light scattering, detrending yielded the best predictions with *in situ* spectra, and SNV or first derivative the best predictions with laboratory spectra. Detrending and first derivative reduce the additive effects, which could thus be considered dominant in both *in situ* and laboratory spectra, while SNV additionally removes multiplicative effects (Swarbrick, 2016), which were thus also present in laboratory spectra.

Predictions were often poor for samples collected in sealed soils, either using *in situ* or laboratory spectra. This could not be attributed to their coarse particle content, which was not higher than in parks and fallows. Rather, this might be attributed to the under-representation of sealed soils in the calibration subset (13 out of 95 samples), in relation to their low proportion in the total sample and pit populations (*cf.* Table 1); while soils under parks and fallows were better represented (82 samples in the calibration subset). Considering that sealed soils cover large areas in urban areas, they should be sampled more extensively in future studies.

4.2. Comparison between VNIRS predictions of SOC *in situ* and in laboratory conditions

Better VNIRS predictions from laboratory than from *in situ* spectra have already been reported in the literature, for SOC concentration (Morgan *et al.*, 2009; Nocita *et al.*, 2011; Li *et al.*, 2015) and other soil properties (Mouazen *et al.*, 2006; Lagacherie *et al.*, 2008). And this seems rather intuitive, as explained by Stenberg *et al.* (2010). Indeed, laboratory spectra are most generally acquired on dry soil samples at room temperature, while spectrum acquisition *in situ* is made under variable conditions of soil moisture and temperature. Moreover, possible coarse particles (> 2 mm) are scanned *in situ* but removed by sieving before chemical analyses and spectrum acquisition in the laboratory, which homogenizes the laboratory samples and induces a difference with *in situ* samples. Variable moisture and temperature conditions and discrepancy between *in situ* and conventionally-analyzed samples complicate building a relationship between *in situ* spectra and conventionally-analyzed soil composition (calibration). Some authors however reported comparable predictions of SOC concentration from *in situ* and from laboratory spectra (Stevens *et al.*, 2008; Gras *et al.*, 2014). This could be explained by more replicates being generally carried out *in situ*, as a result of often visible heterogeneity, while 2 mm sieved samples scanned in the laboratory look much more homogeneous. This could also be attributed to higher sample cohesion *in situ*, which would improve the relationship between sample composition and absorbance (Gras *et al.*, 2014). Therefore, though more intuitive and apparently supported by more published studies, better VNIRS predictions of soil composition, SOC concentration especially, from laboratory than from *in situ* spectra is not a generalizable result. It may be assumed that rather homogeneous soil moisture and temperature conditions during the *in situ* spectroscopy campaigns, on the one hand, and low proportion of coarse particles, on the other hand, would help achieving as good and even better predictions from *in situ* than from laboratory spectra, as was the case for the study reported by Gras *et al.* (2014). In contrast, variable soil moisture and temperature

during *in situ* spectroscopy campaigns and high or variable proportion of coarse particles would result in more accurate predictions from laboratory than from *in situ* spectra, as was the case in the present study. Indeed, the soils tended to be drier in Marseille than in Nantes (mean and SD for moisture content were $283 \pm 118 \text{ g kg}^{-1}$ vs. $363 \pm 174 \text{ g kg}^{-1}$, respectively; data not shown), and to have smaller proportion of particles $> 2 \text{ mm}$ ($148 \pm 118 \text{ g kg}^{-1}$ vs. $284 \pm 162 \text{ g kg}^{-1}$, respectively). Moreover, *in situ* spectra were not acquired at the very location of samples collected for conventional SOC analysis (spectra were acquired in the same horizon but 5 cm above and 5 cm below each collected sample, then averaged; cf. Section 2.3), and could thus represent slightly different soil conditions; though the average of two spectra acquired 5 cm above and below each collected sample was *a priori* assumed to represent it properly. Actually, collecting *in situ* spectra and soil samples that match exactly was an issue because collected samples were several centimeters large while the soil scanned had an area $< 1 \text{ cm}^2$ (Clairotte *et al.*, 2016) and a thickness of no more than a few millimeters (Ollinger *et al.*, 2001). Conversely, aliquots used for laboratory VNIRS and conventional SOC analyses originated from the same homogenized sample and thus were more similar, which improved calibration.

In contrast, no comparison between SOC stock predictions using *in situ* and laboratory spectra has been published yet. Some studies reported VNIRS predictions of SOC stock using *in situ* spectra: Roudier *et al.* (2015) scanned undisturbed soil cores in one large field and achieved $\text{RPD}_{\text{val}} = 2.6$; while Cambou *et al.* (2016), in two fields 250 km apart, scanned disturbed soil cores and achieved $\text{RPD}_{\text{val}} = 1.8$. In laboratory conditions (air-dried, 2 mm sieved samples), Priori *et al.* (2016) achieved $\text{RPD}_{\text{val}} = 2.1$ for a sample set originating from nine fields in a small region.

4.3. Comparison between VNIRS predictions of SOC concentration and stock

Roudier *et al.* (2015) reported similar accuracy when predicting SOC concentration and SOC stock from VNIR spectra collected on intact soil cores ($\text{RPD}_{\text{val}} = 2.6$ vs. 2.6). But Cambou *et al.* (2016) reported better prediction of SOC concentration than SOC stock from VNIR spectra collected on disturbed cores ($\text{RPD}_{\text{val}} = 2.0$ vs. 1.8). *In situ* spectrum acquisition on undisturbed soil would thus result in similar VNIRS predictions of SOC concentration and stock. This was more or less the case in the present study, where spectra were acquired on pit walls (*i.e.* supposedly undisturbed soil). And indeed, prediction accuracy was comparable between SOC concentration and SOC stock ($\text{RPD}_{\text{val}} = 1.3\text{-}2.4$ vs. 1.8-2.2, respectively, depending on pretreatment), while it was noticeably better for SOC concentration than SOC

stock when using laboratory spectra ($RPD_{val} = 3.1-4.6$ vs. $2.1-3.1$, respectively; Table 3). Actually, the important shift in prediction accuracy was from *in situ* to laboratory spectra, and it was larger for SOC concentration than SOC stock (RPD_{val} increased from $1.3-2.4$ to $3.1-4.6$ for concentration, *i.e.* ≈ 2 , vs. $1.8-2.2$ to $2.1-3.1$ for stock, *i.e.* < 1 , respectively). This indicates that the benefit of stable moisture and temperature provided by laboratory conditions was clear when predicting SOC concentration; but 2 mm sieving destroys macrostructure thus reduces information on bulk density, which partly offset the benefit of stable conditions when predicting SOC stock.

4.4. Coarse particles and VNIRS prediction of sample SOC stock

Coarse particles (> 2 mm) are progressively and increasingly taken into account from Eq.1 to Eq.4 (*cf.* Section 2.2); thus better predictions of sample SOC stock calculated with Eq.1 than with Eq.4 indicated that coarse particles were not properly taken into account in the VNIRS prediction process. Such issue could be explained easily for predictions using laboratory spectra: indeed, they were acquired on 2 mm sieved samples (*i.e.* without coarse particles); and as a consequence, better SOC stock predictions were achieved when conventional stock calculation did not account for coarse particles than when it did. The fact that the equations might rank similarly when considering predictions from *in situ* spectra (with the most appropriate pretreatment) suggested that *in situ* spectrum acquisition might under-represent coarse particles. As a matter of fact, it is likely that during spectrum acquisition, the operator preferentially, almost "naturally", put the probe of the spectrometer beside the coarsest particles rather than on them. This underlines the need to acquire spectra that fully represent the soil considered, with its coarse fraction, which could be achieved through extensive replication, for instance according to a regular grid on pit walls. This point is worth considering for urban soils, which often contain more coarse particles than their natural and cultivated counterparts.

4.5. From sample to profile SOC stock

More accurate predictions at profile than horizon level and at horizon than sample level, as observed when using laboratory spectra, suggested that some uncertainties at the sample level could offset each other at the horizon then profile levels. Indeed, one undisturbed soil sample was collected in every horizon for measuring bulk density and two to four disturbed soil samples were collected at other locations of the horizon for conventional SOC analysis and laboratory VNIRS. As a consequence, each disturbed soil sample used for SOC concentration

analysis and spectrum acquisition in the laboratory did not represent exactly the same soil conditions as the corresponding undisturbed soil sample. Thus, it is likely that this undisturbed sample was better represented by average SOC concentration and average VNIR spectrum at the horizon level, hence better predictions at the horizon than sample level. Further but smaller improvement in prediction accuracy from horizon to profile level when using laboratory spectra might be due to offset between possible under-prediction in some horizons and over-prediction in other horizons, resulting from spatial variability at the horizon level. Such spatial variability would cause differences between the sample used for determining bulk density and those collected for spectral and SOC measurements in the laboratory.

In contrast, stock prediction was not improved from sample to horizon level when using *in situ* spectra, possibly because these spectra were not acquired at the exact location of samples collected for conventional SOC analysis (spectra were acquired 5 cm above and 5 cm below, then averaged; *cf.* Sections 2.3 and 4.2). Thus, the average spectrum at the horizon level possibly did not correspond exactly to the average of SOC concentrations determined on the samples of this horizon. Moreover, compensation between horizons at the profile level was limited for predictions based on *in situ* spectra, possibly because, again, soil samples used for SOC analysis and corresponding spectra were not collected at the exact same location.

5. Conclusions

Accurate VNIRS predictions could be achieved for SOC concentrations and stocks using *in situ* or laboratory spectra: $R^2_{\text{val}} \geq 0.78$, and even ≥ 0.94 when using laboratory spectra for predicting sample SOC concentration or horizon and profile SOC stock. Thus VNIRS, which has been used extensively for studying agricultural soils, can be used fruitfully for studying urban soils too.

Better predictions were achieved using laboratory than *in situ* spectra, especially for SOC concentration. Moreover, SOC concentration was better predicted than sample SOC stock, especially when using laboratory spectra. The benefit of stable laboratory conditions was thus clear for predicting SOC concentration; but for predicting SOC stock, this benefit was partly offset by 2 mm sieving, which reduces information on volumetric mass.

Predictions were less accurate for sealed soils than for soils under parks and fallows, especially using *in situ* spectra and for SOC stocks. This might be attributed to the low proportion of sealed soils in the calibration subset and in the total sample set. They cover large areas in urban areas, and should thus be sampled more extensively in future studies.

The accuracy of SOC stock prediction tended to decrease when coarse particles (> 2 mm) were better taken into account in conventional stock calculation (from Eq.1 to Eq.4). This was consistent when using laboratory spectra, acquired on 2 mm sieved samples. Observing the same tendency with *in situ* spectra suggested that coarse particles were under-represented in these spectra, probably because the operator tended to put the spectrometer probe beside, rather than on, the coarsest particles. This point is worth considering for urban soils, which often contain more coarse particles than their rural counterparts.

The accuracy of SOC stock prediction using laboratory spectra increased from sample to horizon then profile level, probably due to compensation of uncertainties at more integrative levels. In contrast, the accuracy of SOC stock prediction using *in situ* spectra varied little at more integrative levels, possibly because samples used for SOC analysis and corresponding spectra were not collected at the exact same location.

6. Acknowledgements

This work was supported by the SUPRA project funded by ADEME (Agence de l'environnement et de la maîtrise de l'énergie, which is a French government agency concerned with environmental protection and energy management; contract 1772C0021), and ADEME also funded the PhD grant of Aurélie Cambou, along with the Région Pays de la Loire. This work was also supported by IRD (Institut de recherche pour le développement, which is a French public research organization dedicated to southern countries).

The authors thank Béatrice Béchet and Liliane Jean-Soro, from the Laboratoire eau et environnement (Water and environment laboratory) of the GERS Department (Géotechnique, environnement, risques naturels et sciences de la Terre, which deals with Earth sciences, geotechnical engineering and natural hazards) at IFSTTAR (Institut français des sciences et technologies des transports, de l'aménagement et des réseaux, which is a French public research organization dedicated to transport and civil engineering) and IRSTV (Institut de recherche en sciences et techniques de la ville, which is a research federation studying urban environments and planning), 44340 Bouguenais, France, for fruitful exchanges and help to identify study sites and get permission for pit opening in Nantes.

Moreover, the authors acknowledge the technical assistance in soil physicochemical analyses provided by colleagues from both EPHor (Angers) and LSE (Nancy) research units: Claudie Mazzega, Yvette Barraud-Roussel, Dominique Renard, Lucas Charrois and Adeline Bouchard.

7. References

- Bae J., Ryu Y. (2015). Land use and land cover changes explain spatial and temporal variations of the soil organic carbon stocks in a constructed urban park. *Landscape Urban Planning*, 136, 57–67.
- Barnes R.J., Dhanoa M.S., Lister S.J. (1989). Standard normal variate transformation and detrending 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.
- Bjørsvik H.R., Martens H. (2001). Data analysis: calibration of NIR instruments by PLS regression. In: D.A. Burns & E.W. Ciurczak (Eds.), *Handbook of Near-infrared Analysis*, 2nd edition (pp. 185–207). Boca Raton, FL: Taylor and Francis.
- Brunet D., Bernoux M., Barthès B.G. (2008). Comparison between predictions of C and N contents in tropical soils using a Vis-NIR spectrometer including a fibre-optic probe versus a NIR spectrometer including a sample transport module. *Biosystems Engineering*, 100, 448–452.
- Burns D.A., Ciurczak E.W. (2001). *Handbook of Near-Infrared Analysis*, 2nd edition. Boca Raton, FL: Taylor and Francis.
- Cambou A., Cardinael R., Kouakoua E., Villeneuve M., Durand C., Barthès B.G. (2016). Prediction of soil organic carbon stock using visible and near infrared reflectance spectroscopy (VNIRS) in the field. *Geoderma*, 261, 151–159.
- Cambou A., Shaw R.K., Huot H., Vidal-Beaudet L., Hunault G., Cannavo P., Nold F., Schwartz C. (2018). Estimation of soil organic carbon stocks of two cities, New York City and Paris. *Science of the Total Environment*, 644, 452–464.
- Chang C.-W., Laird D.A., Mausbach M.J., Hurburgh C.R. (2001). Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 65, 480–490.
- Clairotte M., Grinand C., Kouakoua E., Thébault A., Saby N.P.A., Bernoux M., Barthès B.G. (2016). National calibration of soil organic carbon concentration using diffuse infrared reflectance spectroscopy. *Geoderma*, 276, 41–52.
- Dignac M.-F., Derrien D., Barré P., Barot S., Cécillon L., Chenu C., Chevallier T., Freschet G.T., Garnier P., Guenet B. (2017). Increasing soil carbon storage: mechanisms, effects of

- agricultural practices and proxies. A review. *Agronomy for Sustainable Development*, 37, 14.
- Edmondson J.L., Davies Z.G., McHugh N., Gaston K.J., Leake J.R. (2012). Organic carbon hidden in urban ecosystems. *Scientific Reports*, 2, 963.
- European Commission (2012). Guidelines on Best Practice to Limit, Mitigate or Compensate Soil Sealing. European Commission Staff Working Document. Luxembourg: Publications Office of the European Union.
- Gardner W.H. (1986). Water content. In: A. Klute (Ed.), *Methods of Soil Analysis: Part 1 – Physical and Mineralogical Methods*, 2nd edition (pp. 493–544). Madison, WI: American Society of Agronomy and Soil Science Society of America.
- Gras J.P., Barthès B.G., Mahaut B., Trupin S. (2014). Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoils. *Geoderma*, 214–215, 126–134.
- ISO (International Organization for Standardisation) (1995a). ISO 10694:1995 – Soil Quality – Determination of Organic and Total Carbon after Dry Combustion (Elementary Analysis). Geneva: ISO.
- ISO (International Organization for Standardisation) (1995b). ISO 10693:1995 – Determination of Carbonate Content – Volumetric Method. Geneva: ISO.
- Jacobson M., Charlson R.J., Rodhe H., Orians G.H. (2000). *Earth System Science: From Biogeochemical Cycles to Global Changes*. London: Academic Press.
- Jensen D.R., Ramirez D.E. (2017). Use of Hotelling's T^2 : Outlier diagnostics in mixtures. *International Journal of Statistics and Probability*, 6, 1–11.
- 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* (pp. 177–190). New York: Springer.
- 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.
- Li S., Shi Z., Chen S., Ji W., Zhou L., Yu W., Webster R. (2015). In situ measurements of organic carbon in soil profiles using vis-NIR spectroscopy on the Qinghai–Tibet Plateau. *Environmental Science & Technology*, 49, 4980–4987.
- Liu Z., He C., Zhou Y., Wu J. (2014). How much of the world's land has been urbanized, really? A hierarchical framework for avoiding confusion. *Landscape Ecology*, 29, 763–771.

- Lorenz K., Lal R. (2015). Managing soil carbon stocks to enhance the resilience of urban ecosystems. *Carbon Management*, 6, 35–50.
- 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.
- Nieuwenhuize J., Maas Y.E.M., Middelburg J.J. (1994). Rapid analysis of organic carbon and nitrogen in particulate materials. *Marine Chemistry*, 45, 217–224.
- 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.
- Nocita M., Stevens A., van Wesemael B., Aitkenhead M., Bachmann M., Barthès B., Ben Dor E., Brown D.J., Clairotte M., Csorba A., Dardenne P., Demattê J.A.M., Genot V., Guerrero C., Knadel M., Montanarella L., Noon C., Ramirez-Lopez L., Robertson J., Sakai H., Soriano-Disla J.M., Shepherd K.D., Stenberg B., Towett E.K., Vargas R., Wetterlind J. (2015). Soil spectroscopy: An alternative to wet chemistry for soil monitoring. *Advances in Agronomy*, 132, 139–159.
- Ollinger J.M., Griffiths P.R., Burger T. (2001). Theory of diffuse reflection in the NIR region. In: D.A. Burns & E.W. Ciurczak (Eds.), *Handbook of Near-infrared Analysis*, 2nd edition (pp. 19–51). Boca Raton, FL: Taylor and Francis
- Pansu M., Gautheyrou J., Loyer J.Y. (2001). *Soil Analysis – Sampling, Instrumentation, Quality Control*. Lisse, the Netherlands: Balkema Publishers.
- Pansu M., Gautheyrou J. (2006). *Handbook of Soil Analysis – Mineralogical, Organic and Inorganic Methods*. Berlin Heidelberg: Springer-Verlag.
- Paustian K., Lehmann J., Ogle S., Reay D., Robertson G.P., Smith P. (2016). Climate-smart soils. *Nature*, 532, 49–57.
- Poeplau C., Vos C., Don A. (2017). Soil organic carbon stocks are systematically overestimated by misuse of the parameters bulk density and rock fragment content. *SOIL*, 3, 61–66.
- Pouyat R.V., Yesilonis I.D., Golubiewski N.E. (2009). A comparison of soil organic carbon stocks between residential turf grass and native soil. *Urban Ecosystems*, 12, 45–62.

- Priori S., Fantappie M., Bianconi N., Ferrigno G., Pellegrini S., Costantini E.A.C. (2016). Field-scale mapping of soil carbon stock with limited sampling by coupling gamma-ray and Vis-NIR spectroscopy. *Soil Science Society of America Journal*, 80, 954–964.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reeves D.W. (1997). The role of soil organic matter in maintaining soil quality in continuous cropping systems. *Soil & Tillage Research*, 43, 131–167.
- Roudier P., Hedley C.B., Ross C.W. (2015). Prediction of volumetric soil organic carbon from field-moist intact soil cores. *European Journal of Soil Science*, 66, 651–660.
- Savitzky A., Golay M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedure. *Analytical Chemistry*, 36, 1627–1639.
- Scharlemann J.P., Tanner E.V., Hiederer R., Kapos V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Management*, 5, 81–91.
- 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.
- Swarbrick B. (2016). Near-infrared spectroscopy and its role in scientific and engineering applications. In: M. Kutz (Ed.), *Handbook of Measurement in Science and Engineering*, Vol. 3 (pp. 2583–2656). Hoboken, NJ: Wiley and Sons.
- United Nations (2014). *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)*. New York: United Nations.
- Vasenev V.I., Stoorvogel J.J., Vasenev I.I., Valentini R. (2014). How to map soil organic carbon stocks in highly urbanized regions? *Geoderma*, 226–227, 103–115.
- Viscarra Rossel R.A., Behrens T., Ben-Dor E., Brown D.J., Dematté 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., Changwen D., Chappell A., Fouad Y., Genot V., Gomez C., Grunwald S., Gubler A., Guerrero Maestre C., Hedley C.B., Knadel M., Morrás H.J.M., Nocita M., Ramirez-Lopez L., Roudier P., Rufasto Campos E.M., Sanborn P., Sellitto V.M., Sudduth K.A., Rawlins B.G., Walter C., Winowiecki L.A., Young Hong S. Ji W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, 198–230.

Table 1. Presentation of the studied sites and pits.

Site location, altitude, soil type and bedrock	Pit	Land use	Soil cover	Pit depth (cm)	No of soil horizons
Marseille, Borély Park	A	Park	Lawn	70	3
43.26073° N, 05.38384° E	B	Park	Lawn	80	3
Altitude 5 m a.s.l.	C	Park	Grove litter	85	3
Anthrosol on alluvial bedrock	D	Park	Grove litter	80	2
Marseille, Ste Marthe wasteland	A	Fallow	Grass	110	3
43.34325° N, 05.40054° E	B	Fallow	Grass	100	5
Altitude 116 m a.s.l.	C	Fallow	Brush	115	4
Anthrosol on marl and limestone					
Nantes, cemetery with park	A	Fallow	Grass near a tree	110	5
47.27217° N, 01.58297° W	B	Fallow	Grass	105	4
Altitude 38 m a.s.l.	C	Park	Lawn	125	4
Anthrosol on mica-schist	D	Park	Lawn	110	4
Nantes, railway station	A	Car park	Sealed soil	160	7
47.21725° N, 01.54372° W	B	Car park	Sealed soil	160	5
Altitude 9 m a.s.l.	C	Pavement	Sealed soil	115	6
Technosol on anthropic backfill	D	Pavement	Sealed soil	110	5

Table 2. Distributions of observed sample SOC concentration and stock according to location and land use (SOC stock was calculated according to Eq.4; five samples were removed as spectral outliers, *cf.* Section 2.3).

	Sample No	SOC concentration (gC kg ⁻¹)				SOC stock (gC dm ⁻³)			
		Min	Max	Mean	SD ^a	Min	Max	Mean	SD ^a
Marseille									
Parks	26	16.0	43.0	25.1	9.2	21.7	53.3	29.5	8.2
Fallow	22	16.9	60.8	27.4	11.4	21.9	54.6	33.3	8.7
Nantes									
Parks	32	3.4	24.1	11.0	6.9	2.2	27.1	11.0	7.1
Fallow	31	1.0	29.2	8.5	8.2	0.7	20.8	8.2	6.8
Sealed soils	21	0.4	10.8	1.9	2.9	0.3	10.7	2.1	2.9

^a Standard deviation

Table 3. Validation results of VNIRS predictions of sample SOC concentration and stock for *in situ* and laboratory conditions, achieved with pretreatment methods that yielded high RPD_{val} (37 validation samples were used; mean and SD were 12.4 and 12.4 gC kg⁻¹ for SOC concentration and 15.6 and 15.5 gC dm⁻³ for SOC stock, respectively).

Pretreatment	LV ^a	RMSEP ^b	Bias ^b	Slope ^a	R ² _{val} ^a	RPD _{val} ^a
Sample SOC concentration predicted from <i>in situ</i> spectra (gC kg⁻¹)						
Smoothing	5	9.8	3.4	0.46	0.40	1.3
Smoothing + D1	10	5.2	1.4	0.84	0.83	2.4
Smoothing + D2	9	5.3	1.6	0.81	0.82	2.3
Smoothing + SNV	10	5.2	1.7	0.82	0.83	2.4
Smoothing + SG 1 2 11	4	7.5	4.8	0.59	0.65	1.7
Sample SOC stock predicted from <i>in situ</i> spectra (gC dm⁻³)						
Smoothing	10	8.6	5.5	0.64	0.69	1.8
Smoothing + D1	10	7.3	4.0	0.75	0.77	2.1
Smoothing + D2	9	7.1	4.2	0.73	0.78	2.2
Smoothing + SNV	10	7.4	4.1	0.72	0.77	2.1
Smoothing + SG 1 2 11	6	7.3	3.9	0.76	0.77	2.1
Sample SOC concentration predicted from laboratory spectra (gC kg⁻¹)						
Smoothing	10	3.9	1.9	0.86	0.90	3.2
Smoothing + D1	8	4.1	2.0	0.84	0.90	3.1
Smoothing + D2	8	3.5	2.0	0.88	0.93	3.6
Smoothing + SNV	8	2.7	0.2	0.96	0.95	4.6
Smoothing + SG 1 2 11	9	3.2	2.9	0.85	0.94	3.9
Sample SOC stock predicted from laboratory spectra (gC dm⁻³)						
Smoothing	9	7.0	2.3	0.72	0.79	2.2
Smoothing + D1	8	7.5	3.5	0.67	0.76	2.1
Smoothing + D2	6	6.9	1.7	0.72	0.80	2.2
Smoothing + SNV	8	6.3	2.1	0.73	0.83	2.5
Smoothing + SG 1 2 11	9	5.1	1.9	0.79	0.89	3.1

^a Unitless; LV is the number of latent variables and RPD_{val} is the ratio of SD to RMSEP

^b Same unit as the variable considered (sample SOC concentration or stock)

Table 4. Accuracy of sample SOC stock prediction from *in situ* and laboratory spectra according to the conventional determination procedure (RPD_{val} is the ratio of SD to RMSEP; the number of latent variables LV used in the PLSR model is mentioned into brackets).

Pretreatment	RPD _{val} of sample SOC stock prediction using conventional data calculated according to			
	Eq.1	Eq.2	Eq.3	Eq.4
<i>In situ</i> spectra				
Smoothing	2.2 (13)	2.1 (13)	1.4 (7)	1.8 (10)
Smoothing + D1	2.3 (10)	2.2 (10)	2.2 (10)	2.1 (10)
Smoothing + D2	2.4 (9)	2.3 (9)	2.3 (9)	2.2 (9)
Smoothing + SNV	1.8 (7)	2.3 (11)	2.2 (10)	2.1 (10)
Smoothing + SG 1 2 11	1.7 (4)	1.8 (4)	1.8 (4)	2.1 (6)
Average ± SD ^a	2.1 ± 0.3	2.2 ± 0.2	2.0 ± 0.4	2.1 ± 0.2
Laboratory spectra				
Smoothing	2.7 (9)	2.4 (10)	2.4 (10)	2.2 (9)
Smoothing + D1	2.7 (8)	2.5 (8)	2.4 (8)	2.1 (8)
Smoothing + D2	3.2 (9)	2.5 (8)	2.7 (8)	2.2 (6)
Smoothing + SNV	3.3 (9)	2.9 (9)	3.0 (9)	2.5 (8)
Smoothing + SG 1 2 11	3.6 (9)	3.8 (9)	3.7 (9)	3.1 (9)
Average ± SD ^a	3.1 ± 0.4	2.8 ± 0.6	2.8 ± 0.5	2.4 ± 0.4

^a Standard deviation

Table 5. Accuracy of predictions of horizon and profile SOC stocks on the validation subset using spectra acquired *in situ* and in laboratory conditions (in MgC ha⁻¹ for the level considered; sample SOC stock was calculated with Eq.4; pretreatments were smoothing with D2 and with SG 1 2 11 for *in situ* and laboratory spectra, respectively).

Level	No	Mean ^a	SD ^a	RMSEP ^a	Bias ^a	Slope	R ² _{val}	RPD _{val} ^b
<i>In situ</i> spectra								
Horizon	14	46.4	46.5	22.9	12.5	0.84	0.76	2.0
Profile	4	162.4	140.5	60.3	47.0	0.82	0.78	2.3
Laboratory spectra								
Horizon	14	46.4	46.5	11.6	3.2	0.90	0.94	4.0
Profile	4	162.4	140.5	31.9	14.3	0.88	0.94	4.4

^a In MgC ha⁻¹

^b Ratio of standard deviation (SD) to RMSEP, unitless

Figure 1. PCA of smoothed *in situ* or laboratory absorbance spectra.

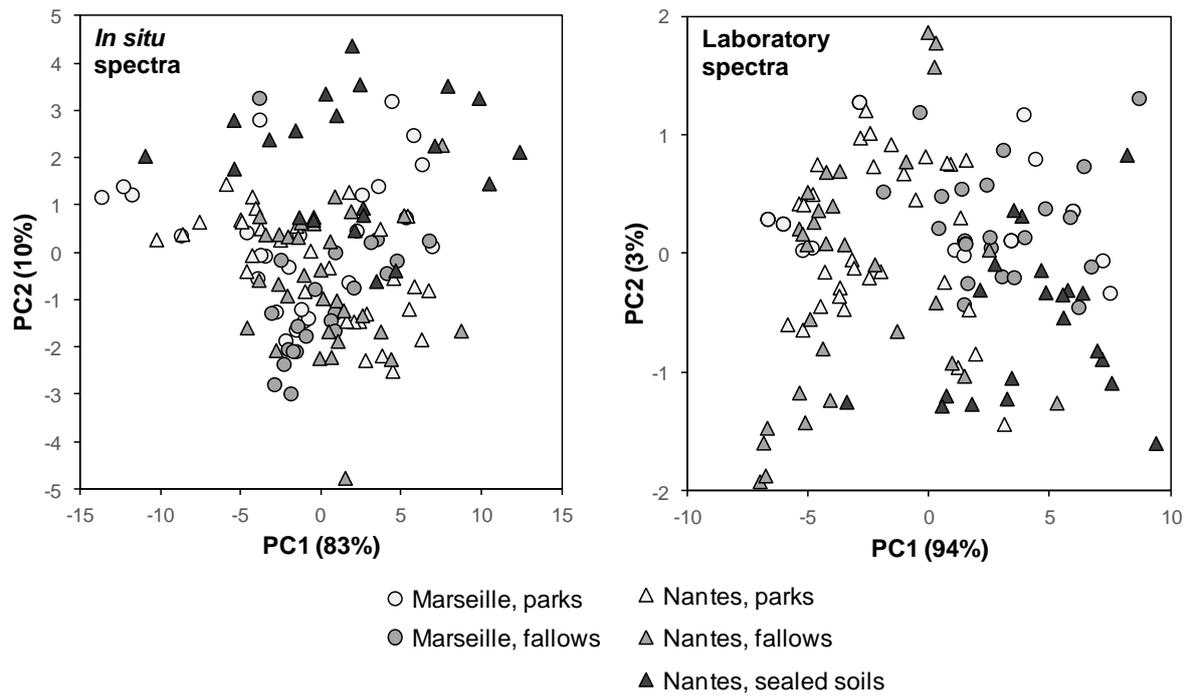


Figure 2. Comparison between observed and VNIRS-predicted sample SOC concentrations and stocks (calculated according to Eq.4) on the validation subset using the best spectrum pretreatment (*i.e.* smoothing with SNV for concentration; smoothing with D2 and with SG 1 2 11 for stock prediction *in situ* and in laboratory conditions, respectively).

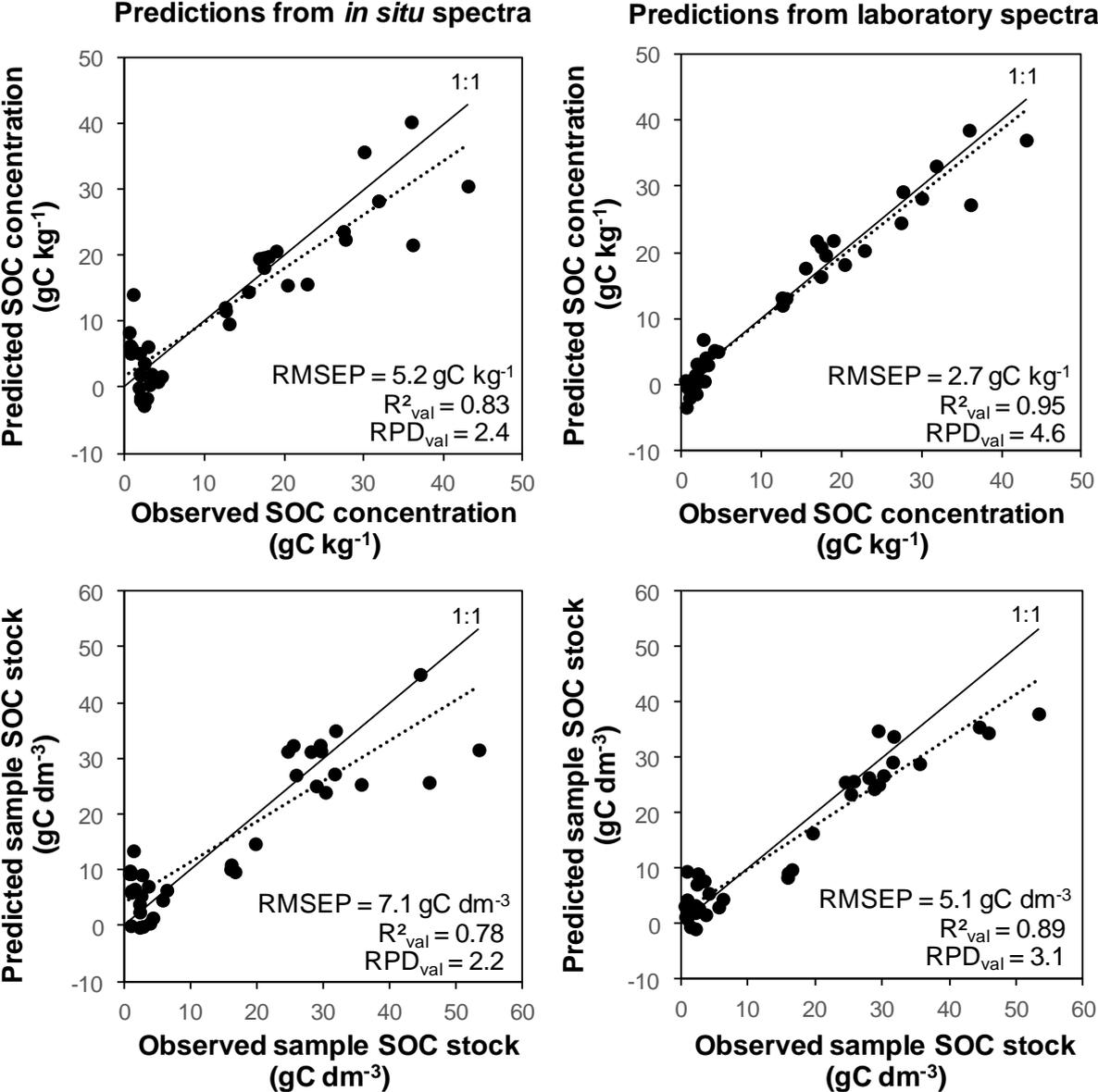


Figure 3. Comparison between profile SOC stocks observed and predicted using (a) *in situ* and (b) laboratory spectra (observed stock was calculated using Eq.4; *in situ* and laboratory spectra were pretreated with smoothing with D2 and with SG 1 2 11, respectively).

