European Journal of Soil Science

*European Journal of Soil Science*, April 2012, **63**, 141–151 doi: 10.1111/j.1365-2389.2012.01429.x

# Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS)

C. Grinand[a], B. G. Barthès[a], D. Brunet[a], E. Kouakoua[a], D. Arrouays[b], C. Jolivet[b], G. Caria[c] & M. Bernoux[a]

[a]*Institut de Recherche pour le Développement (IRD), UMR Eco&Sols (Montpellier SupAgro-CIRAD-INRA-IRD), 2 place Viala, bât. 12, 34060 Montpellier Cedex 2, France,* [b]*Institut National de la Recherche Agronomique (INRA), US 1106 InfoSol, 2163 avenue de la Pomme de Pin, CS 40001, Ardon, 45075 Orléans Cedex 2, France, and* [c]*INRA, US 010, Laboratoire d'analyses des sols, 273 rue de Cambrai, 62000 Arras, France*

## Summary

This work aimed to evaluate the potential of mid-infrared reflectance spectroscopy (MIRS) to predict soil organic and inorganic carbon contents with a 2086-sample set representative of French topsoils (0–30 cm). Ground air-dried samples collected regularly using a 16 × 16-km grid were analysed for total (dry combustion) and inorganic (calcimeter) carbon; organic carbon was calculated by difference. Calibrations of MIR spectra with partial least square regressions were developed with 10–80% of the set and five random selections of samples. Comparisons between samples with contrasting organic or inorganic carbon content and regression coefficients of calibration equations both showed that organic carbon was firstly associated with a wide spectral region around 2500–3500 $\text{cm}^{-1}$ (which was a reflection of its complex nature), and inorganic carbon with narrow spectral bands, especially around 2520 $\text{cm}^{-1}$. Optimal calibrations for both organic and inorganic carbon were achieved by using 20% of the total set: predictions were not improved much by including more of the set and were less stable, probably because of atypical samples. At the 20% rate, organic carbon predictions over the validation set (80% of the total) yielded mean $R^2$, standard error of prediction (SEP) and RPD (ratio of standard deviation to SEP) of 0.89, 6.7 g $\text{kg}^{-1}$ and 3.0, respectively; inorganic carbon predictions yielded 0.97, 2.8 g $\text{kg}^{-1}$ and 5.6, respectively. This seemed appropriate for large-scale soil inventories and mapping studies but not for accurate carbon monitoring, possibly because carbonate soils were included. More work is needed on organic carbon calibrations for large-scale soil libraries.

## Introduction

All countries are required to provide regular greenhouse gas inventories. Among them, countries with a quantified emission reduction objective (Annex B countries of the Kyoto Protocol to the United Nations Framework Convention on Climate Change) should produce estimates on a regular basis, including agriculture and forestry activities, using the 2006 Intergovernmental Panel on Climate Change guidelines (IPCC, 2006). While gross fluxes of methane ($CH_4$) and nitrous oxide ($N_2O$) are directly computed, carbon dioxide ($CO_2$) fluxes are derived from a mass-balance approach, from differences over time of variations in biomass carbon and soil organic carbon stocks. Annex B countries need, therefore, to implement systems to quantify soil organic carbon and changes in content. Soil carbon estimates over large areas, at national, continental or global scales (Batjes, 1996; Bernoux *et al*., 2002), have been proposed on the basis of classical digital maps linked with carbon estimates per map unit, using either soil classification (Batjes, 1996), vegetation or biome delineation (Woomer *et al*., 2004), or a combination of both (Bernoux *et al*., 2002). Other spatial estimates involve the interpolation of punctual data using either splines or geostatistical approaches (Bernoux *et al*., 2007). Saby *et al*. (2008) calculated that the number of sites needed to detect a 1% relative change in soil organic carbon stock (in MgC $\text{ha}^{-1}$) could reach 1 000 000 at the European scale. These calculations involve the organic fraction of soil carbon; inorganic soil carbon, which consists mainly of carbonates, is little involved with $CO_2$ exchanges. Soil organic carbon is often calculated by difference between total carbon and inorganic carbon, the latter being analysed separately. Soil inorganic carbon content provides important information on mineralogy and soil

Correspondence: B. G. Barthès. E-mail: bernard.barthes@ird.fr

Received 8 June 2011; revised version accepted 12 December 2011

chemical functioning involving pH and exchangeable cations. Thus both organic and inorganic soil carbon fractions are worth determining, but conventional procedures are costly and time consuming and not compatible with the implementation of large monitoring networks. Such networks have to be re-analysed regularly. There is thus an urgent need for time- and cost-effective methods that would provide accurate estimates of soil organic and inorganic carbon, in order to enable such networks.

Several methods have recently been proposed for cost-effective estimates of soil carbon. Cremers *et al*. (2001) tested laser-induced breakdown spectrometry (LIBS), which is based on atomic emission spectroscopy. Other spectroscopic methods for characterizing soil carbon include the analysis of light diffusely reflected by samples in two neighbouring electromagnetic domains, namely the near-infrared (0.8–2.5 μm) and mid-infrared (2.5–25 μm) ranges. Often this has involved near-infrared reflectance (NIR) spectroscopy (NIRS; Chang *et al*., 2001; Barthès *et al*., 2006; Brunet *et al*., 2007a), and more recently, visible and NIR (VisNIR, 400–2500 nm) spectroscopy (VisNIRS; Shepherd & Walsh, 2002; Brown *et al*., 2005, 2006; Brunet *et al*., 2008). A particularly interesting feature of VisNIRS, NIRS and MIRS is their capability to make rapid measurements (about one sample per minute or less), with no consumables or hazardous reagents required, and for limited costs. Recent reviews have highlighted the huge potential of diffuse infrared reflectance spectroscopy for agricultural applications regarding soil fertility assessment, and considered this to be an alternative to traditional soil analyses (Du & Zhou, 2009; Reeves, 2010).

The potential of infrared spectroscopy has been tested for decades in a wide range of settings, from field to global scales (Brown *et al*., 2005). It has gained recent interest because of the increasing need for soil data in environmental studies, for digital soil mapping (Grinand *et al*., 2008), or climate change mitigation (Feller & Bernoux, 2008). For two decades NIRS has mainly been used as a quantitative tool, to relate spectra to carbon concentrations statistically, whereas MIRS has long been used for qualitative studies and spectral interpretation (McCarty *et al*., 2002; Reeves, 2010). The MIR region mainly includes fundamental vibrations of chemical bonds, which allows visual discrimination of spectral features (peaks); in contrast, the NIR region includes overtones and combinations, which results in much overlapping and few identifiable features (Shepherd & Walsh, 2002). Recently, MIRS has also been used for quantitative characterization of soil properties (Janik & Skjemstad, 1995; Janik *et al*., 1998; Zimmermann *et al*., 2007), which can thus be performed using different spectral domains through field, laboratory, airborne or satellite measurements (Stevens *et al*., 2008). Quantitative spectroscopy most often requires calibration: this is usually a multivariate regression procedure that expresses a given property, determined by a conventional method, as a function of absorbance at all or selected wavelengths of the spectral region considered. The calibration equation can then be used to predict that property for new samples from their spectra only, provided that the calibration model has sufficient accuracy. Most determinations of organic and inorganic carbon by NIRS (Chang *et al*., 2001; Shepherd & Walsh, 2002; Barthès *et al*., 2006) and MIRS (Janik *et al*., 1998; McCarty *et al*., 2002) were accurate ($R^2 > 0.9$) and had a standard error close to the usual standard error for laboratory measurements (around 1 g $kg^{-1}$ for organic carbon). However, prediction accuracy was often less for heterogeneous sample sets, such as those collected from large regions, than for more homogeneous sets, originating from only a few hectares (McCarty *et al*., 2002; Brunet *et al*., 2007a).

The objective of the present paper was to evaluate the accuracy of MIRS predictions of soil organic and inorganic carbon contents at a country scale, using a national library including about 2000 samples representative of topsoils from France. The 0–30-cm soil depth is that most affected by agricultural practices such as tillage, and is thus considered in most models of soil organic matter dynamics from plot to national scales (Milne *et al*., 2007). In addition, this depth corresponds to the minimum requirement for carbon estimates at national scales according to the IPCC guidelines (IPCC, 2006).

## Materials and methods

### *Soil samples*

Soil samples were taken from the French Réseau de Mesures de la Qualité des Sols (RMQS). This is a national soil quality monitoring network, the first sampling campaign of which took place from 2002 to 2009. This network involves observations of soil properties on a 16 × 16-km regular grid across the French metropolitan territory (550 000 $km^2$), and was designed to represent climate, soil and land-use diversity over the territory. The inventory will eventually consist of observations from 2200 sites, but samples from 2086 sites were available at the time of the study. The sites were selected at the centre of each 16 × 16-km cell (Jolivet *et al*., 2006). At each site, 25 individual core samples were taken from the topsoil (0–30 cm) using an unaligned sampling design within a 400-$m^2$ square area. A systematic unaligned sample design combines features and advantages of both simple random and systematic sample designs, and a single sample plot is assigned to a randomly selected location within each cell of the systematic grid. This approach avoids the periodicities of systematic approaches, gives good coverage over an area, is efficient, and deals with most distributions (Caeiro *et al*., 2003). Core samples were bulked to obtain one composite sample for each site, which was then air-dried, gently crushed and sieved to pass a 2-mm mesh.

### *Conventional determination of soil organic and inorganic carbon contents*

Soil organic carbon content was calculated as the difference between total carbon and inorganic carbon contents. Total carbon content was determined by Dumas combustion analysis with an elemental analyser (Thermo Fisher Scientific CHN NA2000, Waltham, MA, USA), on about 25–30-mg aliquots of finely ground (<0.25 mm) air-dried soil samples, sealed into tin capsules, according to the standard French procedure NF ISO 10694 (AFNOR, 1995a).

Soil inorganic carbon content was calculated as 0.12× soil carbonate content. Carbonate content was determined on finely ground (<0.25 mm) air-dried soil samples using a Bernard calcimeter, according to the standard French procedure NF ISO 10693 (AFNOR, 1995b). The carbonate content was calculated after calibration with a pure calcium carbonate standard and was expressed as equivalent calcium carbonate content.

Distributions of both organic and inorganic carbon contents were characterized by their minimum, maximum, mean, median, standard deviation and skewness; the latter is defined as the third standardized moment:

$$\text{skewness} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \frac{(x_i - \bar{x})^3}{\mathrm{SD}^3}, \qquad (1)$$

where $n$ is the set size, $\bar{x}$ the mean and SD the standard deviation.

### Spectrum acquisition and pre-processing

Aliquots of about 0.5 g of <0.2-mm ground sample were placed in a 17-well plate and scanned from 4000 to 400 $\mathrm{cm}^{-1}$ (2500–25 000 nm) at 4 $\mathrm{cm}^{-1}$ resolution, using a Nicolet 6700 Diffusive Reflectance Fourier Transform Spectrophotometer (Thermo Fisher Scientific Instruments, Madison, WI, USA). Thirty-two co-added scans per sample were performed then averaged, and spectra were recorded as absorbance, which is the log-transform of the inverse of reflectance. Spectrum acquisition lasted about 1 minute per sample.

Data processing was performed using The Unscrambler 9.7 software (CAMO Technologies, Woodbridge, NJ, USA) for spectrum transformation and the R-Stat software (R Development Core Team, 2005) for partial least square (PLS) regression (PLSR) and validation. Several usual spectrum pre-processing methods such as first and second derivatives, standard normal variate (SNV) transform or multiplicative scatter correction (MSC) were tested but did not usefully improve predictions. The only beneficial spectrum transformations were the removal of 10 wavelengths at both spectrum ends and smoothing over a five-point segment, which helped reduce noise: this process was applied for subsequent analyses, either qualitative or quantitative. Limited benefit of MIR spectrum pre-processing was also reported by Minasny & McBratney (2008).

Principal component analysis was carried out on spectral data and showed that the two first components accounted for 84% of total variance.

### Calibration and validation

Outlier identification was carried out on the response variables (organic and inorganic carbon content) to remove extreme values that might over-influence calibration models. An arbitrary threshold defined by the mean plus eight times the standard deviation was used to identify these outliers. This led to the removal of two samples for organic carbon; no samples were removed for the inorganic carbon calibration.

Partial least square regression was used to fit sample spectra to conventional measurements of soil organic and inorganic carbon contents. It is a well-suited and widely used multivariate technique in quantitative spectroscopy. It can handle collinearity that affects infrared spectra by transforming absorbance values with respect to the response variable (organic or inorganic carbon content) in order to reduce the information from hundreds of variables to a few orthogonal factors. Ten-fold cross-validation was performed to calculate the standard error of cross-validation (SECV). This commonly used model setting was needed because of the large number of calibration samples, and enables grouping without altering the predictive ability of the cross-validated models. The minimum SECV was used to select the optimal number of PLS factors to retain for building the prediction model. Then calibration was performed on the whole calibration set, whereas the cross-validation results referred to a pooled combination of the results obtained from the samples left out during the cross-validation procedure. Finally, the prediction models were tested on the validation set, which included samples not used for model development. The number of samples used for calibration and validation is presented in the following section.

The ability of the MIRS-PLSR method to predict organic and inorganic carbon was evaluated by using parameters commonly used in quantitative infrared spectroscopy. For each dataset, coefficient of determination ($R^2$; Equation (2)) and standard error between MIRS predictions and conventional measurements were calculated. The former measures the proportion of the total variance accounted for by the model, and was presented for the validation set only. The latter is the standard deviation of the difference between measured and predicted values, expressed in g $\mathrm{kg}^{-1}$, and was presented for calibration (SEC), cross-validation (SECV) and validation (standard error of prediction, SEP; Equation (3)). Prediction accuracy on the validation set was also evaluated using the ratio of standard deviation to SEP (denoted RPD; Equation (4)) and the ratio of SEP to mean reference value over the validation set (relative SEP, denoted %SEP; Equation (5)). A good prediction result is assumed to have small standard error and large $R^2$ and RPD. As regards the predictions of soil properties using infrared spectroscopy, values proposed by Chang *et al*. (2001) have often been considered: models with RPD below 1.4 have no predictive ability, those with RPD between 1.4 and 2.0 are acceptable and improvable, and those with RPD above 2.0 have excellent predictive ability. The equations are:

$$R^2 = \frac{[\mathrm{cov}(y_i, x_i)]^2}{\mathrm{var}\,(y_i)\,\mathrm{var}\,(x_i)}, \qquad (2)$$

$$\text{SEC or SECV or SEP} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - x_i)^2}, \qquad (3)$$

$$\mathrm{RPD} = \mathrm{SD}/\mathrm{SEP}, \qquad (4)$$

and

$$\%SEP = SEP/mean, \quad (5)$$

where $x_i$ is the conventional measure and $y_i$ the MIRS prediction for the i$^{th}$ sample, $n$ the number of samples and SD the standard deviation.

### *Calibration set size and selection*

The effect of the calibration set size on prediction was tested by increasing the proportion of calibration samples from 10 to 80% of the total set, the remaining fraction being kept for validation. This test was referred to as sampling intensity analysis (Grinand *et al*., 2008). Regression was performed five times using different random iterations for the selection of calibration samples, in order to assess prediction robustness and reproducibility, which are two important criteria to build up a calibration model. Although it is known that the calibration set has to be representative of the total set, there are still issues regarding the way it is selected. Random sampling has been reported to produce varying results in heterogeneous sample populations (Islam *et al*., 2003; Brunet *et al*., 2007a) but has not been extensively studied yet (Brown *et al*., 2005). Here the minimum, maximum and mean model parameters (such as SECV, SEC, $R^2$, etc.) have been used to provide information on the extent of predictive ability and on significant differences between models.

## Results and discussion

### *Reference carbon data*

The soils used in this study were collected from a large geographical area, which covers a wide range of soil units and represents most of the soil types of France. According to the French soil classification, 33 soil reference groups were sampled, with a dominance of Brunisols (Cambisols according to IUSS Working Group WRB, 2006; 27% of the sample set), calcareous soils (Calcosols, 22%) and Luvisols (16%). As a result of the large variety of soil samples studied, variations in organic and inorganic carbon contents were huge, as measured using conventional procedures (Table 1). Organic carbon ranged from 0.6 to 170 g kg$^{-1}$ after removal of outliers (up to 250 g kg$^{-1}$), averaged 25.6 g kg$^{-1}$ and standard deviation was 78% of the mean. Inorganic carbon ranged from 0 to 104 g kg$^{-1}$, averaged 6.4 g kg$^{-1}$ and standard deviation was 250% of the mean. The calcium carbonate content of 1108 samples was below the 1 g kg$^{-1}$ detection limit, thus the median equalled 0 g kg$^{-1}$. Distributions of both organic and inorganic carbon were positively skewed (skewness was around 3 for both), with more small than large values; thus medians were smaller than the means, especially for inorganic carbon. Correlation between organic and inorganic carbon contents was not significant ($R = -0.01,\ P = 0.59$).

### *Semi-qualitative spectrum analysis*

Spectral measurements and laboratory results were first used in a qualitative way to identify peaks or absorption regions that characterized carbon fractions. Spectral regions associated with organic carbon variations were identified by testing the significance of difference in mean absorbance at each spectral data point between organic carbon-rich ($>$50 g kg$^{-1}$) and carbon-poor ($<$10 g kg$^{-1}$) samples from Cambisols (the soil class with the largest number of samples), using a Student $t$-test ($P < 0.001$). A similar procedure was carried out for inorganic carbon ($>$6.0 and $<$1.2 g kg$^{-1}$, which corresponds to $>$50 and $<$10 g calcium carbonate kg$^{-1}$) in the Calcosols.

The difference in mean absorbance between organic carbon-rich and carbon-poor topsoils from Cambisols was significant for a wide region between 3550 and 2080 cm$^{-1}$, and for bands around 1660 and 1150 cm$^{-1}$ (Figure 1a). Organic matter is a complex mixture of compounds, containing various proportions of chemical groups that are nearly all infrared active (Janik *et al*., 1998). Thus organic carbon cannot be identified with clearly separated peaks but as a whole spectral region with overlapping bands.

In contrast, numerous small spectral regions differed significantly in mean absorbance between inorganic carbon-rich and carbon-poor topsoils from Calcosols (Figure 1b). Among them, only two could be attributed to stretching or bending vibrations in carbonate molecules, around 2520 and 1800 cm$^{-1}$, as noted by Du & Zhou (2009). Other regions differing significantly in absorbance according to inorganic carbon level could not be attributed to carbonates but more probably to soil properties that correlated negatively with inorganic carbon. This was particularly the case for regions around 1880 and 2000 cm$^{-1}$, which relate to quartz (Du & Zhou, 2009).

In short, semi-qualitative analysis of spectra showed that absorption of both carbon fractions could be located in the MIR region. Nevertheless, though some carbon compounds have been assigned to specific spectral bands, broad spectral regions corresponding to C-O or C-C vibrations have to be taken into account to describe carbon contribution to MIR spectra fully. Thus a quantitative approach was needed to extract accurate information

**Table 1** Descriptive statistics of the reference dataset

| Soil property | Sample no | Min | Max | Mean | Median | SD | Skewness |
|---|---|---|---|---|---|---|---|
| Organic carbon / g kg$^{-1}$ | 2084 | 0.6 | 170.0 | 25.6 | 19.4 | 19.9 | 3.2 |
| Inorganic carbon / g kg$^{-1}$ | 2086 | 0.0 | 103.9 | 6.4 | 0.0 | 16.0 | 3.1 |

max = maximum; min = minimum; SD = standard deviation.

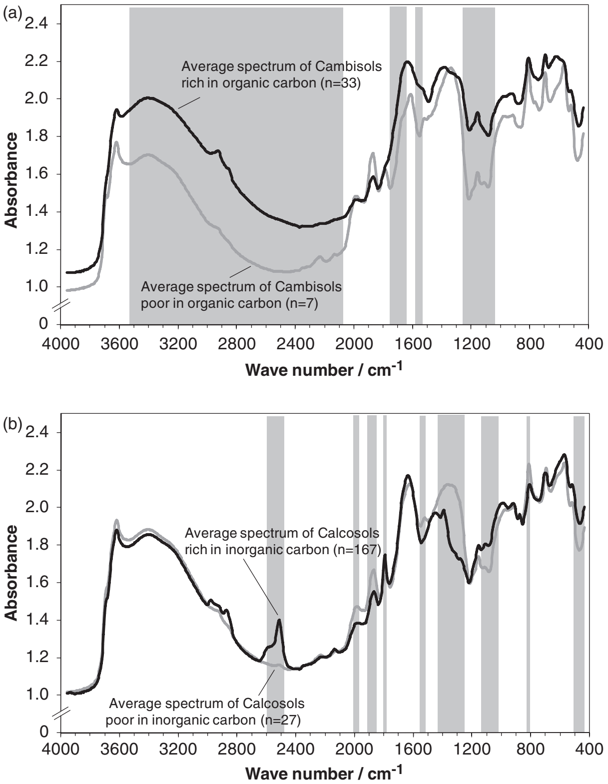


**Figure 1** Comparison of mean spectra (a) between Cambisols rich and poor in organic carbon ($>50$ or $<10$ g kg$^{-1}$), and (b) between Calcosols rich and poor in inorganic carbon ($>6$ or $<1.2$ g kg$^{-1}$; corresponding to $>50$ and $<10$ g calcium carbonate kg$^{-1}$, respectively); grey regions indicate significant difference at $P < 0.001$ (Student $t$-test).

regarding each carbon fraction. The identification of chemical bonds that affected MIRS prediction of soil organic and inorganic carbon contents was also carried out by studying the coefficients of regression of these contents on absorbance spectra.

### *Calibration intensity and model robustness*

The interest of diffuse reflectance spectroscopy for quantifying soil properties depends on the proportion of samples used for calibration; it also depends on the stability of predictions when different calibration sample selections are carried out. For organic carbon, mean validation RPD and $R^2$ increased with the proportion of samples used for calibration, markedly when the proportion increased from 10 to 20%, and then only slightly up to 80% (Figure 2a). Moreover, $R^2$ and RPD were less variable and thus more stable when calibration used 20% or 30% of the samples; at smaller or larger rates, the results were more affected by the selection of calibration samples. Reduced stability at large calibration rates could be explained by samples that were atypical for one reason or another. When included in the calibration set, atypical samples would reduce overall calibration quality and thus that of the prediction (validation), and this was more likely to occur when calibration intensity increased. When not included in the calibration set, atypical samples were not predicted well individually (validation), but they did not affect the overall performance of the calibration, and this was more likely to occur when calibration intensity decreased. At an intensity of 10%, it might be assumed that the calibration set was too small to be representative of the total set.

With inorganic carbon, mean and standard deviation of validation $R^2$ were much less affected by the size of the calibration set (Figure 2b). In contrast, the proportion of calibration samples affected validation RPD more clearly, with a noticeable increase from 10 to 20%, then again from 40 to 80%; it also affected the variability of validation RPD, with greater standard deviation when from 40 to 80% samples were used for calibration. Variations in validation RPD and in its variability with the size of the calibration set could be related to the very skewed distribution of conventional data for inorganic carbon, which resulted in a large standard deviation (it was three times greater than the mean; Table 1). However, RPD values were always large ($>5$) when compared with accuracy thresholds proposed for prediction models of soil properties using diffuse reflectance (Chang *et al*., 2001); thus the influence of the size of the calibration set on RPD and its variability was not important. Using 20% of the samples for calibrating both organic and inorganic carbon could thus be considered appropriate for the sample set studied. This indicated that the national soil sample population under study was correctly represented by 20% of the samples. Soil and soil-use diversity is large over all of France but not to the point where only the population of 2086 would be truly representative.

Very few studies have addressed such questions. Shepherd & Walsh (2002) reported a similar result for a very diverse set including about 1100 samples from seven countries and 10 soil orders in eastern and southern Africa. They observed that stable validation results for topsoil organic carbon could be achieved with 20% of the samples for calibration. They also noted that a global model might be more robust than local models in their ability to predict values for new samples even though geographically distinct samples were used in the calibration. Similarly, when discussing calibration sample size for a more homogeneous set including 283 samples from six sites with similar soils in north central Montana (USA), Brown *et al*. (2005) found an optimum when using 35% randomly selected samples for calibration. Increasing the calibration sample size did not result in a meaningful increase in prediction accuracy and thus was not worthwhile. The results of the present study also suggest that optimal calibration set size for heterogeneous soil sample populations could be estimated at 20 to 35% of the samples for organic carbon and inorganic carbon predictions. However, such calibration intensity cannot be considered appropriate for all soil sample sets: Brunet *et al*. (2007b) observed that optimal calibration sample size ranged between 50 and 75% for organic carbon for an extremely diverse set of 436 soil samples originating from about 300 locations in 40 countries and representing 13 soil orders. This strongly suggests that more work is needed to estimate optimal calibration sample size depending on the set diversity and the soil property considered.

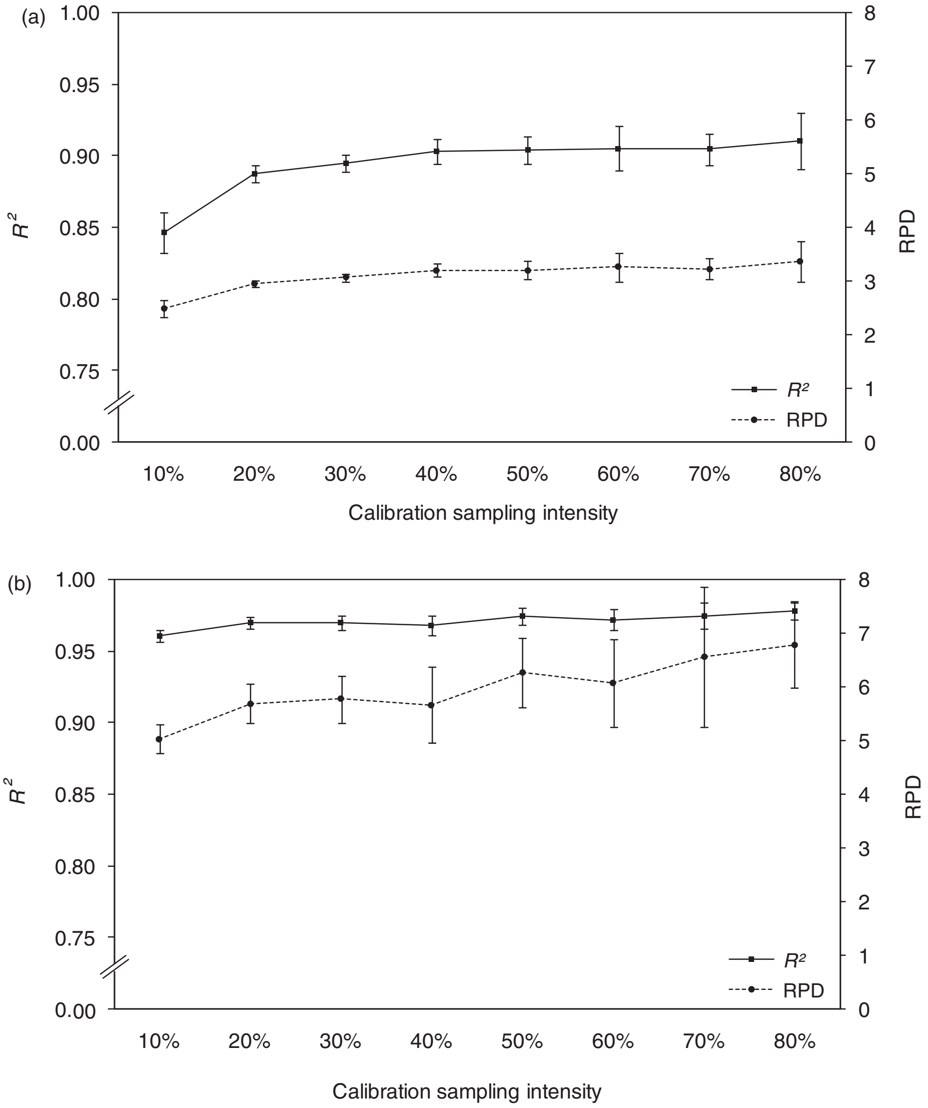


**Figure 2** Validation $R^2$ and RPD (ratio of standard deviation to standard error of prediction) for (a) organic carbon and (b) inorganic carbon depending on the proportion of calibration samples (mean and standard deviation over five replicates).

### *Quantifying soil inorganic carbon content at a national scale*

Calibration was built using 20% of the total number of samples (418 out of 2086), with five different random selections of calibration samples. For validation, $R^2$ was 0.97 and RPD averaged 5.6 and was greater than 5.3 whatever the replicate (Table 2). Thus MIRS prediction of inorganic carbon was excellent for the sample set considered, as also shown by the comparison between MIRS predictions and reference measurements presented in Figure 3(b) and in Figure 4(c,d). Depending on the selection of calibration samples, variation in $R^2$ was less than 0.01 and RPD ranged from 5.3 to 6.0, thus both were stable. Relative SEP nevertheless represented a large proportion of mean reference value for inorganic carbon (43 to 48%), mainly because of the very skewed distribution of inorganic carbon reference data. The mean reference value was small (6.4 g kg$^{-1}$) because many samples included no carbonates; relative prediction error could not be considered as an appropriate indicator of prediction accuracy for variables for which the average value is much smaller than the maximum. Not surprisingly, the map of MIRS predictions of inorganic carbon (Figure 4c), as did that of conventional measurements (Figure 4d), showed that the soils richest in inorganic carbon were located on calcareous substrates, especially in the chalk Champagne (northeast), in the calcareous Prealps (southeast), in areas of Jurassic limestone along the Aquitain Basin (southwest) and along the Pyrenees (south). Soils developed from calcareous materials were completely decalcified under wet climates (Jura and northern Prealps, centre-east).

Literature has also reported excellent predictions for inorganic carbon with diffuse reflectance spectroscopy. McCarty *et al.* (2002) achieved MIRS prediction of inorganic carbon with

**Table 2** Calibration and validation statistics for MIRS predictions of soil organic and inorganic carbon contents; min, max and mean refer to minimum, maximum and mean values measured for five random partitionings of calibration and validation sets; standard errors of calibration (SEC), of cross-validation (SECV) and of prediction (SEP) are in g kg$^{-1}$

| | | | Calibration | | | Validation | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Soil property | | Factors No | Sample No | SEC[a] | SECV[a] | Sample No | SEP[a] | $R^{2}$[b] | RPD[c] | %SEP[d] |
| Organic carbon | Min | 20 | 418 | 4.4 | 6.1 | 1666 | 6.3 | 0.88 | 2.7 | 0.25 |
| | Max | 24 | 418 | 6.3 | 8.0 | 1666 | 7.2 | 0.90 | 3.1 | 0.28 |
| | Mean | 23 | 418 | 5.1 | 6.9 | 1666 | 6.7 | 0.89 | 3.0 | 0.26 |
| Inorganic carbon | Min | 16 | 418 | 1.3 | 1.8 | 1668 | 2.6 | 0.97 | 5.3 | 0.43 |
| | Max | 24 | 418 | 3.1 | 4.1 | 1668 | 2.9 | 0.97 | 6.0 | 0.48 |
| | Mean | 20 | 418 | 2.1 | 2.9 | 1668 | 2.8 | 0.97 | 5.6 | 0.45 |

[a] SEC, SECV and SEP are standard errors of calibration, cross-validation and prediction, respectively.

[b] $R^2$ is the coefficient of determination over the validation set.

[c] RPD is the ratio of SEP to standard deviation of the reference value over the validation set.

[d] %SEP is the ratio of SEP to the mean reference value over the validation set.

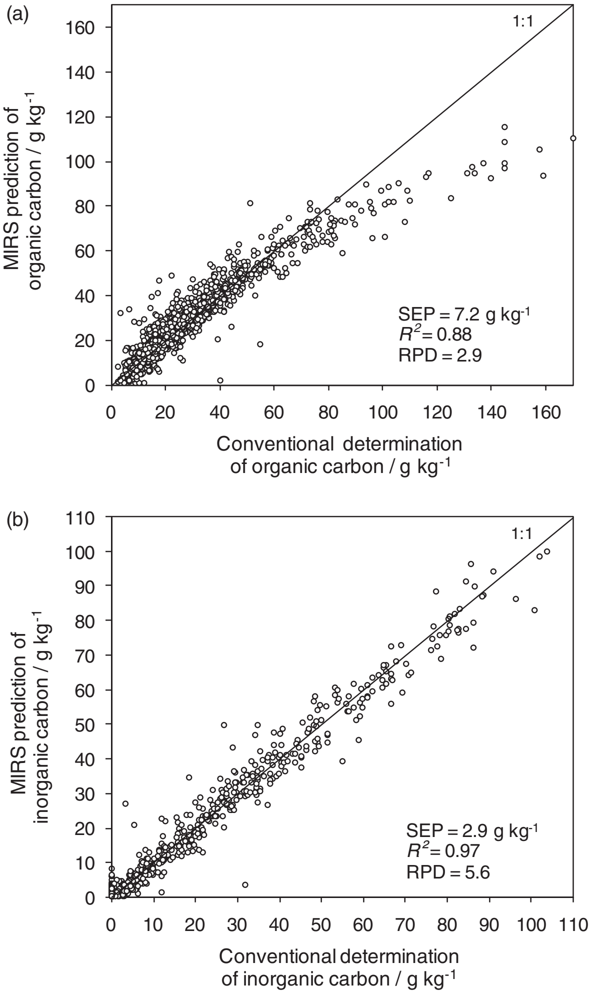
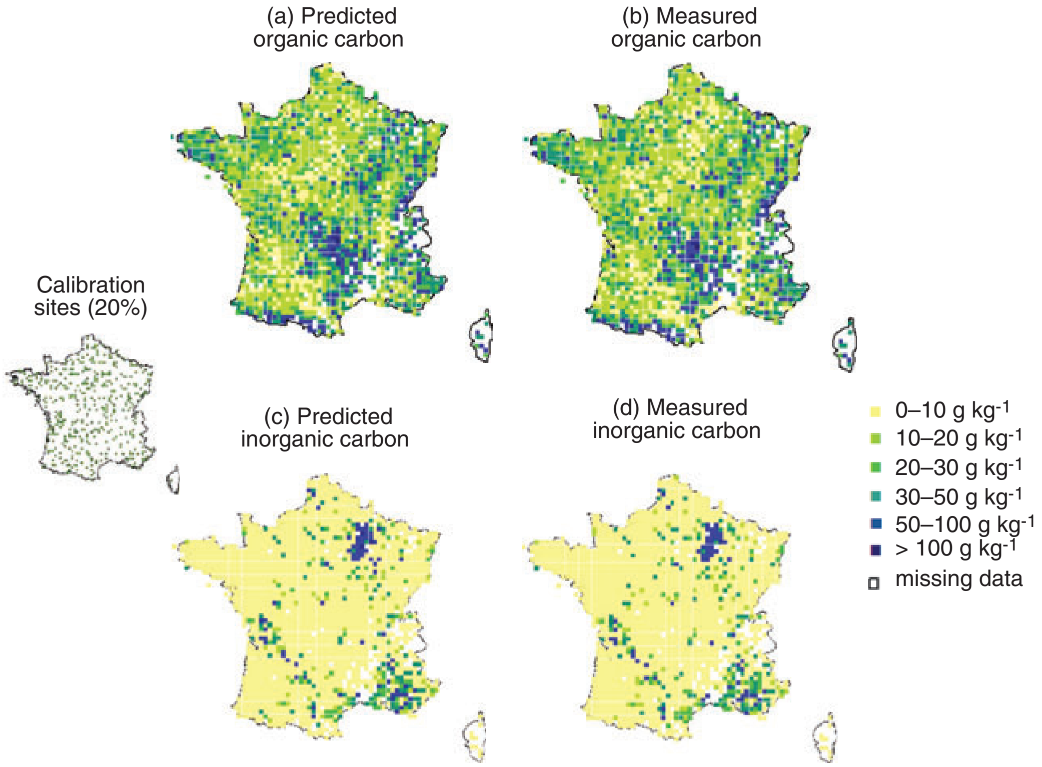


**Figure 3** Comparisons between reference measurements and MIRS predictions of (a) organic carbon and (b) inorganic carbon over the validation set (1666 samples for organic carbon and 1668 for inorganic carbon, after random selection of 418 calibration samples; SEP is the standard error of prediction and RPD the ratio of standard deviation to SEP).

$R^2 = 0.98$, RPD = 9.0 and %SEP = 19% with a more intensive calibration (75%) over a diverse set of 237 soil profile samples originating from nine states from central USA and having a fairly wide range in inorganic carbon (0–65 g kg$^{-1}$). Brown *et al.* (2005), studying a more homogeneous set with a narrow range in inorganic carbon (0–26 g kg$^{-1}$) and intensive calibration (70%), achieved VisNIRS prediction of inorganic carbon with $R^2 =$ 0.94–0.96, RPD = 4.0–4.9 and %SEP = 19–24% depending on the selection of the calibration set (five replicates). Brown *et al.* (2006) also reported cross-validation $R^2 = 0.83$ (RPD and relative error could not be calculated) for a very large, diverse set originating from four continents and having a wide range in inorganic carbon (0–129 g kg$^{-1}$). In contrast, somewhat disappointing VisNIRS cross-validation of inorganic carbon was reported by Summers *et al.* (2011) for a fairly diverse but small topsoil sample set from south Australia with a narrow range in inorganic carbon (0–51 g kg$^{1}$; 75 samples, $R^2 = 0.69$, RPD = 2.1, %SECV > 100%). In the latter case the use of reflectance instead of absorbance could be a possible reason for less accurate predictions than in other studies. Considering only those studies that used absorbance, relative prediction error tended to increase and $R^2$ and RPD to decrease when the range in inorganic carbon increased.

### *Quantifying soil organic carbon content at a national scale*

Calibration for this used 20% of the total set size, with five replicates. For validation, $R^2$ and RPD averaged 0.89 and 3.0 and were greater than 0.88 and 2.7, respectively, whatever the replicate (Table 2). The SEP ranged from 6.3 to 7.2 g kg$^{-1}$, representing 25–28% of the mean reference value. This indicated very good MIRS prediction of organic carbon for the sample set considered; however, predictions were less accurate than for inorganic carbon. Examples of comparisons between measured and MIRS-predicted organic carbon values are presented in Figures 3(a) and 4(a,b). The maps of MIRS predictions (Figure 4a) and conventional measurements (Figure 4b) of organic carbon were very similar, and strongly driven by climate and land use. The soils richest



(a) Predicted organic carbon

(b) Measured organic carbon

Calibration sites (20%)

(c) Predicted inorganic carbon

(d) Measured inorganic carbon

**Figure 4** Maps of MIRS predictions (a, c) and reference determinations (b, d) of topsoil organic and inorganic carbon content (0–30 cm) in samples from across France; 20% randomly selected sites, located on the small map, were used for calibration.

in organic matter were located in the mountains (Alps in the southeast, Pyrenees in the extreme southwest, Massif Central in the south-centre, Jura in the centre-east), in other cool regions extensively covered by forests and pastures such as Burgundy and Lorraine (centre-east), and in intensive livestock production areas such as Brittany and Normandy (northwest). In contrast, soils poor in organic matter were located in areas of large-scale crop cultivation such as the Parisian Basin (north-centre) and Aquitain Basin (southwest), and in vineyard areas such as along the Mediterranean coast (southeast).

Other authors have reported MIRS prediction of organic carbon with PLSR. McCarty *et al*. (2002) found similar results for organic carbon prediction over a fairly heterogeneous sample set from central USA ($R^2 = 0.94$, RPD = 4.1, %SEP = 26%) but used more samples for calibration (75%). Viscarra Rossel *et al*. (2006) using a leave-one-out cross-validation achieved worse results for organic carbon for a very homogeneous set of 118 topsoil samples collected in an Australian field, with much smaller $R^2$ (0.73) and RPD (1.7), though relative standard error of cross-validation was also small (11%). In contrast, Zimmermann *et al*. (2007), also using a large calibration intensity (85%), found better results for organic carbon ($R^2 = 0.94$, RPD = 4.1, %SEP = 10%) for a diverse set of 111 topsoil samples representative of Switzerland. Comparable results have been achieved by using VisNIRS with PLSR. For instance, for a large and very diverse topsoil sample set with a large calibration intensity (67%), Shepherd & Walsh (2002) reported validation $R^2$ of 0.91 and %SEP of 19%. Brown *et al*. (2005) achieved validation $R^2 = 0.81$ and RPD = 2.2 when averaging five replicates over a smaller and more homogeneous set and with a large calibration intensity (70%). For a very large and very diverse set (3794 samples from four continents) and using boosted regression trees, Brown *et al*. (2006) obtained cross-validation $R^2$ values of 0.82. Thus comparison with published results suggests that the results of the present study were very satisfactory: the accuracy of the organic carbon prediction model was comparable to others for similarly heterogeneous soil sample sets but using much fewer samples for calibration.

In the present study, large organic carbon values were noticeably under-predicted by MIRS. Studying a large and very diverse sample set with VisNIRS, Shepherd & Walsh (2002) attributed such poor predictions to possible errors in the laboratory analytical methods rather than to genuine lack of VisNIRS predictive power. Moreover, it is worth noting that in the present study organic carbon reference data were not measured but calculated by difference between total carbon and inorganic carbon, which caused error accumulation. As a consequence, the less than perfect MIRS predictions of organic carbon ($R^2 = 0.89$) might result from imperfect reference data. In contrast, inorganic carbon, for which reference data were obtained by direct measurement, was much more accurately predicted, especially at large values. Other studies involving MIRS or VisNIRS predictions of organic and inorganic carbon similarly reported more accurate results for

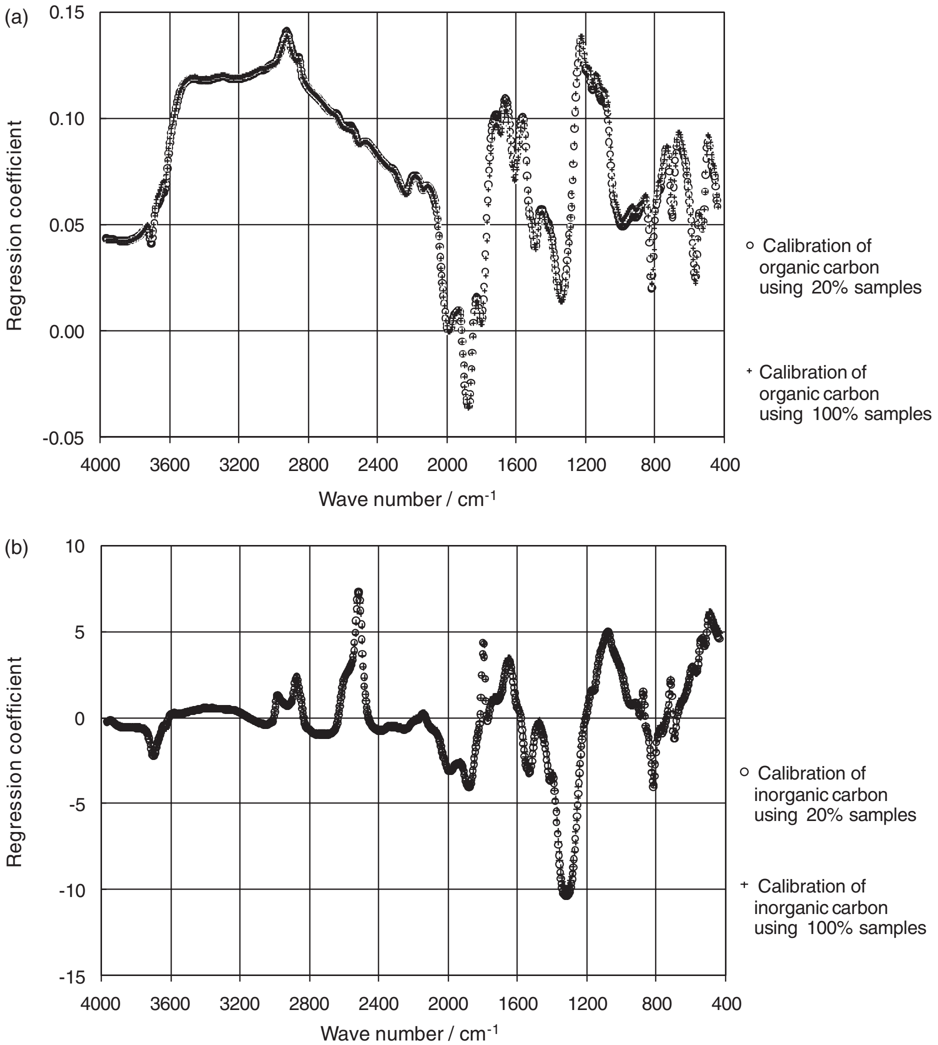


**Figure 5** Coefficients of regression of (a) organic and (b) inorganic carbon content on absorbance spectra, models being developed by using 20% or 100% of the samples from across France.

inorganic than for organic carbon. McCarty *et al*. (2002) measured organic carbon after soil acidification, which they acknowledged might remove some organic carbon. Brown *et al*. (2005) calculated organic carbon reference data by difference between total and inorganic carbon. Both procedures might result in some imprecision. In addition, McCarty *et al*. (2002) reported that carbonate absorption bands in the MIR domain could mask spectral features important for organic carbon calibration. As far as sample sets that include carbonated soils are considered, calibration models for organic carbon will not be completely satisfying because of either spectral perturbations by carbonates or imperfect reference data for organic carbon.

### *Spectral regions of greatest relevance*

The chemical compounds that affected MIRS predictions strongly were identified by studying the coefficients of PLS regression of organic and inorganic carbon contents on absorbance at every wave-number (Figure 5). Regressions built with 20% or 100% samples yielded similar coefficients, which confirmed that 20% samples represented the total set properly.

Some spectral regions contributed heavily to MIRS prediction of soil organic carbon content. Positively contributing regions were from 2700 to 3600 cm$^{-1}$ and especially around 2920 cm$^{-1}$, which has been assigned to aliphatic C-H stretch of carboxylic acids, humic acids and humin especially, and around 1230 cm$^{-1}$, which has been ascribed to C-O stretch and O-H bending in humic acids and aromatic amines (Silverstein & Webster, 1998; Tan, 2003). These regions were comparable to those where absorbance was significantly greater for Cambisols rich in organic carbon than for those with small contents (Figure 1a). The region around 1880 cm$^{-1}$, attributed to quartz (Du & Zhou, 2009), had a heavy negative contribution to organic carbon prediction, which is consistent with the fact that sandy soils are poor in organic matter in general (Barthès *et al*., 2008). This region is one of the very few where absorbance was less for organic-rich Cambisols than for those that were organic poor.

Spectral regions that had a major positive contribution to MIRS prediction of soil inorganic carbon content were (i) around 2520 and 1800 $\text{cm}^{-1}$, which have been ascribed to carbonate (Du & Zhou, 2009; D'Acqui *et al*., 2010); (ii) around 1660 $\text{cm}^{-1}$, assigned to asymmetric stretching of O-C-O in carbonate (Su & Suarez, 1997) but also to 2:1 clays such as montmorillonite (Du & Zhou, 2009); (iii) around 1080 $\text{cm}^{-1}$, assigned to symmetric stretching of O-C-O in carbonate (Su & Suarez, 1997) but also to 1:1 clays such as kaolinite (Du & Zhou, 2009); and (iv) around 490 $\text{cm}^{-1}$, ascribed to Fe-O bending vibrations (Russell, 1979), which might be related to iron oxide-rich residual clays that result from weathering of calcareous materials. The region around 1340 $\text{cm}^{-1}$, which has been attributed to quartz (D'Acqui *et al*., 2010), had a substantial negative contribution to carbonate prediction, which is consistent with the fact that carbonated soils include little quartz. On the whole, regions that had large positive (or negative, respectively) contributions to carbonate prediction were those where absorbance was significantly larger (or smaller, respectively) for Calcosols rich in, than for Calcosols poor in, inorganic carbon (Figure 1b).

## Conclusion

The present work demonstrated that MIRS yielded accurate predictions of topsoil organic and inorganic carbon contents at the scale of a medium-sized country (France), which has not been demonstrated so far. Analysis of variance between spectra of soils rich and poor in organic or inorganic carbon, as well as regression coefficients of calibration equations, showed that the MIRS signature of organic carbon involved a wide spectral band around 2500–3500 $\text{cm}^{-1}$, while that of inorganic carbon was related to some narrow bands, especially around 2520 $\text{cm}^{-1}$.

Varying the proportion of samples used for calibration showed that 20% could be considered an optimal level for MIRS prediction of both organic and inorganic carbon contents, for a sample set representing the topsoil diversity of France. This calibration rate resulted in accurate and stable predictions: ranges of SEP, validation $R^2$ and RPD were 6.3–7.2 g $\text{kg}^{-1}$, 0.88–0.90 and 2.7–3.1 for organic carbon, and 2.6–2.9 g $\text{kg}^{-1}$, 0.97 and 5.3–6.0 for inorganic carbon, respectively. The similarity between regression coefficients of calibration equations that used 20 and 100% of the total set confirmed that 20% was the optimal calibration intensity for the studied set. This was an important result, which has not been established clearly so far.

An important and interesting feature of MIRS (and other infrared spectroscopies) is that numerous soil properties, such as nitrogen and clay contents, pH and cation exchange capacity, can be inferred from the same spectrum, as long as corresponding reference data are available for calibration. Moreover, once determined, calibrations can be used to make predictions with samples from new sampling campaigns (with a possible need for calibration strengthening through spiking; Guerrero *et al*., 2010), which makes diachronic studies easier. However, though MIRS is appropriate for large-scale soil inventories and for many environmental and mapping studies, the accuracy of MIRS predictions presented here did not seem to be good enough for monitoring soil organic carbon precisely. This might be attributed to the noticeable proportion of carbonated samples included in the sample set, as carbonates complicate spectral and conventional characterizations of organic carbon and thus its calibration. More work is needed to improve organic carbon calibration in large-scale soil libraries including carbonated soils, for instance through sample set stratification or spectrally local calibrations.

## Acknowledgements

This study was supported by the French ADEME (Agence de l'Environnement et de la Maîtrise de l'Energie; contract 0675C0102), by IRD (Institut de Recherche pour le Développement) and by INRA (Institut National de la Recherche Agronomique). RMQS soil sampling and physico-chemical analyses were supported by a French Scientific Group of Interest on soils: the 'GIS Sol', involving the French Ministry for Ecology and Sustainable Development, the French Ministry of Agriculture, The French National Forest Inventory (IFN), and ADEME, IRD and INRA. We thank all the soil surveyors and technical assistants involved in sampling the sites. Two anonymous referees and the editor are also thanked for their helpful comments.

## References

AFNOR (Association Française de Normalisation) 1995a. *Qualité du Sol – Dosage du Carbone Organique et du Carbone Total après Combustion Séche (Analyse Elémentaire)*. NF ISO 10694, AFNOR, Paris.

AFNOR (Association Française de Normalisation) 1995b. *Qualité du Sol – Détermination de la Teneur en Carbonate – Méthode Volumétrique*. NF ISO 10693, AFNOR, Paris.

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., Kouakoua, E., Larré-Larrouy, M.C., Razafimbelo, T.M., De Luca, E.F., Azontonde, A. *et al*. 2008. Texture and sesquioxide effects on water-stable aggregates and organic matter in some tropical soils. *Geoderma*, **143**, 14–25.

Batjes, N.H. 1996. Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*, **47**, 151–163.

Bernoux, M., Carvalho, M.D.S., Volkoff, B. & Cerri, C.C. 2002. Brazil's soil carbon stocks. *Soil Science Society of America Journal*, **66**, 888–896.

Bernoux, M., Arrouays, D., Cerri, C.E.P. & Cerri, C.C. 2007. Regional organic carbon storage maps of western Brazilian Amazon based on prior soil maps and geostatistical interpolation. In: *Digital Soil Mapping – An Introductory Perspective* (eds P. Lagacherie, A. McBratney & M. Voltz), pp. 497–506. Elsevier, Amsterdam.

Brown, D.J., Bricklemyer, R.S. & Miller, P.R. 2005. Validation requirements for diffuse reflectance soil characterization models with a case study of VNIR soil C prediction in Montana. *Geoderma*, **129**, 251–267.

Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M. & Reinsch, T.G. 2006. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*, **132**, 273–290.

Brunet, D., Barthès, B.G., Chotte, J.L. & Feller, C. 2007a. Determination of carbon and nitrogen contents in Alfisols, Oxisols, and Ultisols form Africa and Brazil using NIRS analysis: effects of sample grinding and set heterogeneity. *Geoderma*, **139**, 106–117.

Brunet, D., Szwarc, M., Babre, D., Alary, K. & Barthès, B. 2007b. Intérêt de la spectrométrie proche infrarouge pour caractériser les propriétés physico-chimiques d'une gamme variée de sols. In: *Actes des 9$^e$ Journées Nationales de l'Etude des Sols, Angers (France), 3–5 April 2007* (ed. J.P. Rossignol), pp. 413–414. Association Française pour l'Etude des Sols & Institut National d'Horticulture, Angers.

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.

Caeiro, S., Painho, M., Goovaerts, P., Costa, H. & Sousa, S. 2003. Spatial sampling design for sediment quality assessment in estuaries. *Environmental Modelling & Software*, **18**, 853–859.

Chang, C.W., Laird, D.A., Mausbach, M.J. & Hurburg, C.R. 2001. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Science Society of America Journal*, **65**, 480–490.

Cremers, D.A., Ebinger, M.H., Breshears, D.D., Unkefer, P.J., Kammerdiener, S.A., Ferris, M.J. *et al*. 2001. Measuring total soil carbon with laser-induced breakdown spectroscopy (LIBS). *Journal of Environmental Quality*, **30**, 2202–2206.

D'Acqui, L.P., Pucci, A. & Janik, L.J. 2010. Soil properties prediction of western Mediterranean islands with similar climatic environments by means of mid-infrared diffuse reflectance spectroscopy. *European Journal of Soil Science*, **61**, 865–876.

Du, C. & Zhou, J. 2009. Evaluation of soil fertility using infrared spectroscopy: a review. *Environmental Chemistry Letters*, **7**, 97–113.

Feller, C. & Bernoux, M. 2008. Historical advances in the study of global terrestrial soil organic carbon sequestration. *Waste Management*, **28**, 734–740.

Grinand, C., Arrouays, D., Laroche, B. & Martin, M.P. 2008. Extrapolating regional soil landscapes from an existing soil map: sampling intensity, validation procedures, and integration of spatial context. *Geoderma*, **143**, 180–190.

Guerrero, C., Zornoza, R., Gómez, I. & Mataix-Beneyto, J. 2010. Spiking of NIR regional models using samples from target sites: effect of model size on prediction accuracy. *Geoderma*, **158**, 66–77.

IPCC (Intergovernmental Panel on Climate Change) 2006. *IPCC Guidelines for National Greenhouse Gas Inventories*, *Volume 4: Agriculture, Forestry and Other Land Use*. IPCC, Geneva.

Islam, K., Singh, B. & McBratney, A. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research*, **41**, 1101–1114.

IUSS (International Union of Soil Science) Working Group WRB (World Reference Base) 2006. *World Reference Base for Soil Resources 2006*. World Soil Resources Reports No 103, FAO, Rome.

Janik, L.J. & Skjemstad, J.O. 1995. Characterization and analysis of soils using mid-infrared partial least-squares. II. Correlations with some laboratory data. *Australian Journal of Soil Research*, **33**, 637–650.

Janik, L.J., Merry, R.H. & Skjemstad, J.O. 1998. Can mid infrared diffuse reflectance analysis replace soil extractions? *Australian Journal of Experimental Agriculture*, **38**, 681–696.

Jolivet, C., Arrouays, D., Boulonne, L., Ratié, C. & Saby, N. 2006. Le Réseau de mesures de la qualité des sols de France (RMQS). Etat d'avancement et premiers résultats. *Etude et Gestion des Sols*, **13**, 149–164.

McCarty, G.W., Reeves, J.B. III, Reeves, V.B., Follett, R.F. & Kimble, J. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. *Soil Science Society of America Journal*, **66**, 640–646.

Milne, E., Al Adamat, R., Batjes, N.H., Bernoux, M., Bhattacharyya, T., Cerri, C.C. *et al*. 2007. National and sub-national assessments of soil organic carbon stocks and changes: the GEFSOC modelling system. *Agriculture Ecosystems & Environment*, **122**, 3–12.

Minasny, B. & McBratney, A.B. 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemometrics & Intelligent Laboratory Systems*, **94**, 72–79.

R Development Core Team 2005. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna.

Reeves, J.B. III. 2010. Near-versus mid-infrared diffusive reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: where are we and what needs to be done? *Geoderma*, **158**, 3–14.

Russell, J.D. 1979. Infrared spectroscopy of ferrihydrite: evidence for the presence of structural hydoxyl groups. *Clays & Clay Minerals*, **14**, 109–114.

Saby, N.P.A., Bellamy, P.H., Morvan, X., Arrouays, D., Jones, R.J.A., Verheijen, F.G.A. *et al*. 2008. Will European soil monitoring networks be able to detect changes in topsoil organic carbon content? *Global Change Biology*, **14**, 2432–2442.

Shepherd, K.D. & Walsh, M.G. 2002. Development of reflectance spectral libraries for characterisation of soil properties. *Soil Science Society of America Journal*, **66**, 988–998.

Silverstein, R.M. & Webster, F.X. 1998. *Spectrometric Identification of Organic Compounds*, 6th edn. John Wiley & Sons, New York.

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.

Su, C. & Suarez, D.L. 1997. *In situ* infrared speciation of adsorbed carbonate on aluminum and iron oxides. *Clays & Clay Minerals*, **45**, 814–825.

Summers, D., Lewis, M., Ostendorf, B. & Chittleborough, D. 2011. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecological Indicators*, **11**, 123–131.

Tan, K.H. 2003. *Humic Matter in Soil and the Environment. Principles and Controversies.* Marcel Dekker, New York.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. & Skjemstad, J.O. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, **131**, 59–75.

Woomer, P.L., Touré, A. & Sall, M. 2004. Carbon stocks in Senegal‘s Sahel transition zone. *Journal of Arid Environments*, **59**, 499–510.

Zimmermann, M., Leifeld, J. & Fuhrer, J. 2007. Quantifying soil organic carbon fractions by infrared-spectroscopy. *Soil Biology & Biochemistry*, **39**, 224–231.