Near infrared reflectance spectroscopy (NIRS) could be used for characterization of soil nematode community

Bernard G. Barthès, Didier Brunet, Bodovololona Rabary, Oumar Ba, Cécile Villenave

Near infrared reflectance spectroscopy (NIRS) could be used for characterization of soil nematode community

Bernard G. Barthès [1], Didier Brunet [1], Bodovololona Rabary [2], Oumar Ba [3], Cécile Villenave [1]

[1] IRD, UMR Eco&Sols (INRA-IRD-Montpellier SupAgro), 2 place Viala, bâtiment 12, 34060 Montpellier Cedex 2, France.
[2] FOFIFA, URP SCRID (CIRAD, FOFIFA, University of Antananarivo), BP 230, Antsirabe 110, Madagascar.
[3] IRD, UMR Eco&Sols, LEMSAT, Centre IRD-ISRA, route des Hydrocarbures, BP 1386, Dakar, Senegal.
+ Corresponding author: bernard.barthes@ird.fr

Abstract

Studying soil nematofauna provides useful information on soil status and functioning but requires high taxonomic expertise. Near infrared reflectance (NIR) spectroscopy (NIRS) has been reported to allow fast and inexpensive determination of numerous soil attributes. Thus the present study aimed at assessing the potential of NIRS for determining the abundance and diversity of soil nematodes in a set of 103 clayey topsoil samples collected in 2005 and 2006 from agricultural soils in the highlands of Madagascar.

The morphological characterization of soil nematofauna involved extraction through elutriation then counting under binoculars and identification at family or genus level using microscopy, on ca. 150-g fresh soil samples. Taxa were assigned to five trophic groups, namely bacterial feeders, fungal feeders, obligate plant feeders, facultative plant feeders, and omnivores and predators (together). In addition, four ecological indexes were calculated: the Enrichment index, Structure index, Maturity index, and Plant parasitic index.

Oven-dried (40°C) < 2-mm sieved 5-g soil subsamples were scanned in the NIR range (1100-2500 nm), then spectra were fitted to nematofauna data using partial least square regression. Depending on the sample set considered (year 2005, year 2006, or both years), NIRS prediction of total nematode abundance was accurate (ratio of standard deviation to standard error of cross validation, i.e. RPD, ≥ 2) or acceptable (RPD ≥ 1.6). Predictions were accurate, acceptable, or quasi-acceptable (RPD ≥ 1.4) for several of the six most abundant taxa, and to a larger extent, for most trophic groups (except facultative plant feeders); but they could not be

made for taxa present in a small number of samples or at low abundance. By contrast, NIRS prediction of relative abundances (in proportion of total abundance) was poor in general, as was also the prediction of ecological indexes (except for the 2006 set). On the whole, these results were less accurate than NIRS predictions of soil attributes often reported in the literature. However, though not very accurate, NIRS predictions were worthwhile considering the labour-intensity of the morphological characterization. Most of all, NIRS analyses were carried out on subsamples that were probably too small (5 g) to allow representative sampling of nematofauna. Using larger samples for NIRS (e.g. 100 g) would likely result in more accurate predictions, and is therefore recommended. Scanning un-dried samples could also help improve prediction accuracy, as morphological characterization was carried out on samples not dried after sampling.

Examining wavelengths that contributed most to NIRS predictions, and chemical groups they have been assigned to, suggested that NIRS predictions regarding nematofauna depended on constituents of both nematodes and preys' food. Predictions were thus based on both nematofauna and soil organic properties reflected by nematofauna.

Research highlights

- NIRS characterization of soil nematofauna is a promising approach.
- Abundance prediction was acceptable for total community and most trophic groups.
- Predictions seemed to be based on nematode constituents and on soil properties reflected by nematofauna.
- Scanning large and un-dried samples would probably increase prediction accuracy.

Key-words

Soil nematodes; Trophic groups; Abundance; Near infrared reflectance spectroscopy (NIRS).

Introduction

Soil nematodes possess several characteristics that allow their utilisation as bio-indicators of soil functioning (Bongers and Ferris, 1999): they are abundant in most terrestrial ecosystem whatever the climatic area and the vegetation (Yeates, 2003); they have high taxonomic diversity and high functional diversity in relation to soil processes (Ekschmitt et al., 2001); moreover, several nematofaunal indexes linked to soil functioning have been developed and have proven to be powerful tools for characterizing the soil food-web in ecosystems and

agrosystems (Bongers, 1990; Ferris et al., 2001). Indeed, nematodes belong to different trophic groups that are present at several levels of the food-web: microbivorous (bacterial feeders and fungal feeders), plant feeders (obligate or facultative ones), omnivores, and predators of other nematodes. Thus nematode community structure integrates a lot of information on the soil micro-food-web: microbial compartment (fungal and bacterial parts), microfauna and mesofauna, which are responsible for the decomposition and mineralization of soil organic matter hence nutrient release (Freckman et al., 1997; Ekschmitt et al., 2001; Villenave et al., 2004; Sanchez-Moreno and Ferris, 2007).

Characterization of soil nematofauna involves extraction from soil samples by elutriation then counting and identification to genus or family level under microscope (ISO, 2007). This method, if not expensive, is time-consuming and requires expertise in taxonomy, which impedes the study of large numbers of samples. Molecular methods are being developed (Perry and Jones, 1998; Floyd et al., 2002); though they are currently used to detect some specific taxa or in population dynamics studies, they are not sufficiently efficient yet to allow quantification of all taxa present in soil (Donn et al., 2008). To our knowledge, spectrometric approaches have not been tested to date.

Near infrared reflectance (NIR) spectroscopy (NIRS) is a physical non-destructive, rapid, reproducible and low-cost approach that characterizes materials according to their reflectance in the wavelength range between 800 and 2500 nm (Roberts et al., 2004). The analysis of NIR spectra relies on calibration, which in general is a multivariate regression procedure that expresses a given property, determined using a conventional method, as a function of absorbance at all or selected wavelengths of the NIR region. The calibration equation can then be used to predict that property on new samples from their NIR spectra only, the acquisition of which is time- and cost-effective (< 1 min per sample, no consumables required). The application of NIRS to soil has been mentioned from the 1960s (Bowers & Hanks, 1965) and it has been used extensively to determine soil content in carbon and nitrogen (Al-Abbas et al., 1972; Chang et al., 2001; Barthès et al., 2006). It has also proven useful for characterizing soil fractions such as NMR species (i.e. alkyl, O-alkyl, carboxylic and aromatic C; Terhoeven-Urselmans et al., 2006), organic size fractions (Barthès et al., 2008) and microbial biomass (Palmborg and Nordgren, 1993; Chang et al., 2001), as well as microbial biomarkers based on phospholipid fatty acids (Zornoza et al., 2008) and microbial activities such as carbon and nitrogen mineralization (Palmborg and Nordgren, 1993; Chang et al., 2001; Terhoeven-Urselmans et al., 2006). However, few attempts have been made to apply quantitative NIRS for characterizing soil fauna.

The present study aimed at assessing the usefulness of soil NIR spectra for determining the abundance and functional diversity of soil nematodes in a clayey Ferralsol under soybean / rice rotation in the Madagascar highlands.

Materials and methods

Studied site and sample collection

The studied site was located at Bemasoandro, near Antsirabe, in the Madagascar highlands (19°47'S, 47°06'E, ca. 1600 m a.s.l.). The climate is altitude tropical; mean annual rainfall and temperature are 1300 mm and 16°C, respectively. The soil is developed on volcano-lacustrine alluvia (Raunet, 1981) and classified as andic Dystrustept (Soil Survey Staff, 2003). It is clayey, acidic (> 70% clay and pH ≈ 5 in the topsoil) and includes kaolinite, and to a lesser extent, gibbsite, quartz, hematite, and goethite. The topsoil contains ca. 8-9% organic matter at 0-10 cm depth.

The studied experiment was installed in 1996 by the French agricultural research centre CIRAD (Centre de coopération internationale en recherche agronomique pour le développement), the Malagasy NGO TAFA (i.e. Land and development) and the FOFIFA (i.e. National centre of applied research for rural development). The cropping system was a soybean / rice rotation (one crop per year) with either manual ploughing using a large spade (called angady) or no tillage, crop residues being removed in the former case but returned to the soil surface as mulch in the latter. Each plot was divided into three subplots with either no inputs, bovine manure application (5 Mg ha^{-1} yr^{-1}), or both mineral fertilizer (70N-30P-40K for rice and 30N-30P-40K for soybean) and bovine manure application (5 Mg ha^{-1} yr^{-1}). The six tillage × input treatments were replicated three times, resulting in 18 elementary plots, 13.5 m^2 each. Further information on the site and experiment has been provided by Razafimbelo (2005).

Topsoils were sampled during the rainy season, in January 2005 under soybean and in February 2006 under rice (Villenave et al., 2009a). Three composite soil samples were collected at 0-5 cm depth toward the upper part, the middle and the lower part of every elementary plot, yielding 54 composite samples yearly. Each composite sample resulted from the grouping and thorough mixing of five neighbouring samples, collected using 100-cm^3 cylinders. An aliquot of every composite sample was air-dried then gently crushed to pass a 2-mm sieve. Due to the loss of some samples, a total set of 103 samples was studied.

Morphological characterization of soil nematofauna

Soil nematofauna was analyzed following a standardized procedure ISO 23611-4 (ISO, 2007). For each of the 103 fresh samples, nematodes were extracted from approximately 150 g of wet soil by elutriation followed by an active pass through a filter for 48 h at room temperature, according to the Seinhorst method (Seinhorst, 1962); they were then counted using a binocular microscope. After fixing in a formalin-glycerol mixture and transferring to mass slides, the composition of soil nematofauna was determined at family or genus level through microscopic observation at 400× magnification. On average, 111 nematodes were identified per mass slide. The nematode taxa were then assigned to trophic groups modified from Yeates et al. (1993): bacterial feeders (BF), fungal feeders (FF), facultative plant feeders (FPF), obligate plant feeders (OPF), and omnivores and predators (OMPR). Nematodes were also allocated to colonisation-persistence (*c-p*) classes following Bongers (1990): the colonisation-persistence scale ranges from 1 (colonizers) to 5 (persisters); it varies within trophic groups, so that combination between trophic groups and *c-p* classes defines feeding guilds such as BF_1 (bacterial feeders with a *c–p* class of 1) and FF_4 (fungal feeders with a *c–p* class of 4). In addition, two nematode ecological indexes were calculated after Ferris et al. (2001):

Enrichment index: $EI = 100 \times \{e/(e+b)\}$

Structure index: $SI = 100 \times \{s/(b+s)\}$

where *e*, *b*, and *s* are the sum of weighted abundances of guilds BF_1 and FF_2 (enrichment component), BF_2 and FF_2 (basal component), and $BF_{3\text{-}5}$, $FF_{3\text{-}5}$ and $OMPR_{2\text{-}5}$ (structural component), respectively. Also, the Maturity index (MI) and the Plant parasitic index (PPI) were calculated as $\Sigma v_i p_i$, where v_i is the *c-p* value assigned to the family i and p_i its relative abundance in the sample, considering only free-living nematodes for MI and plant-feeding nematodes for PPI (Bongers, 1990).

Spectrum acquisition and pre-processing

Overnight oven-dried (40°C) < 2-mm sieved soil samples of ca. 5 g were scanned in the NIR region between 1100 and 2500 nm at 2 nm intervals using a Foss NIRSystems 5000 spectrophotometer (Silver Spring, MD, USA) in order to determine their reflectance. Each sample spectrum, automatically averaged from 32 spectra, was recorded as absorbance, which is the logarithm of the inverse of reflectance (log[1/R]). For every sample, two subsamples were scanned and averaged. More than 200 spectra could be acquired daily, without any

consumable. Data analysis was conducted using the Winisi III-v1.61e software (Foss NIRSystems/Tecator International, LLC, Silver Spring, MD, USA).

In order to reduce spectral data, spectra were condensed by keeping the first out of four adjacent spectral points, yielding 173 data points per spectrum (Barthès et al., 2006). Several usual spectrum pre-processing methods were tested: no derivation (denoted 01), first- or second-order derivation with 4-, 5-, 7-, 8- or 10-point gap and smoothing (denoted 14, 15, 17, 18, 110 and 24, 25, 27, 28 and 210, respectively) alone (denoted None) or in conjunction with standard normal variate transform (SNV), detrend (D), both SNV and detrend (SNVD), or standard multiplicative scatter correction (MSC). Indeed, it has often been observed that such procedures could increase the signal-to-noise ratio thus improve the prediction of sample properties using NIR spectra: derivation reduces baseline variation and enhance spectral features (Reeves et al., 2002), SNV transform reduces the particle size effect (Barnes et al., 1989), detrend removes the linear or curvilinear trend of each spectrum (Barnes et al., 1989), and MSC removes additive and/or multiplicative signal effects (Martens et al., 1983).

Processing methods

Statistical analyses were performed on three datasets, regarding samples collected in 2005, in 2006, and over both years (i.e. total set), respectively. A principal component analysis (PCA) was carried out on each spectral dataset to calculate the Mahalanobis distance H (Mark and Tunnell, 1985). Samples with $H > 3$ were considered spectral outliers and eliminated from further investigations (Shenk and Westerhaus, 1991a). Modified partial least square (mPLS) regression was used to fit spectral data to nematofaunal data (Shenk and Westerhaus, 1991b). The mPLS regression combines PCA and multiple regression in order to reduce a complex spectral matrix to a few orthogonal terms. Cross validation was performed to determine the optimum number of terms to be used in the model, in order to avoid overfitting. This was done by dividing the set into six groups, five being used for developing the model and one for prediction. The procedure was performed six times to use all samples for both model development and prediction, then the residuals of the six predictions were pooled to calculate the standard error of cross validation (SECV). The outliers for calibration (i.e. samples with $t > 2.5$) were removed and another cross validation was performed. This procedure was carried out twice. The number of factors giving the lowest final SECV determined the optimal number of terms to be used for calibration. Then the final model was applied to all remaining samples (i.e. outliers being excluded). The accuracy of the cross validation was assessed using SECV, the part of variance explained (i.e. 1 – residual variance, usually denoted Q^2), and the

ratio of standard deviation to SECV (usually denoted RPD). As regards site-specific soil studies, NIRS prediction models with RPD ≥ 2 have been considered accurate and those with 1.6 ≤ RPD < 2.0 acceptable (Dunn et al., 2002). We considered those with 1.4 ≤ RPD < 1.6 quasi-acceptable; in our opinion, the notion of "quasi-acceptable" makes sense when the reference method is tedious or produces rather imprecise results. Calibration accuracy was assessed according to standard error of calibration (SEC) and determination coefficient between measured and predicted values (R^2).

Results

Abundance and community structure of nematodes

Nematofauna abundance and composition, and the way they were affected by cropping systems in the experimental design, have been described in Villenave et al. (2009a) and are summarized here. In the 103 samples analyzed in 2005 and 2006, the abundance of nematodes ranged from 0 to 24 individuals g^{-1} dry soil and averaged 4-5 ind g^{-1} dry soil (Table 1).

On the whole, 49 taxa were distinguished. Bacterial feeders represented 34% of the nematode community and 22 taxa, four being abundant (> 15 ind 100 g^{-1} dry soil in average), *Acrobeloides* only being abundant and frequent (i.e. present in more than 70% of the samples). The obligate plant feeders (plant parasites) represented 33% of the community and nine taxa were isolated: five were abundant but *Pratylenchus* and *Xiphinema* only were abundant and frequent. Facultative plant feeders (root hair feeders) represented 8% of the community and belonged to a unique taxon, the family Tylenchidae, which was abundant and frequent. Fungal feeders represented 11% of the community and belonged to eight taxa, among which two were abundant, but *Aphelenchoides* only was abundant and frequent. Omnivores and predators represented 14% of the community and nine taxa were isolated, among which the family Dorylaimidae was abundant and frequent.

No tillage, crop residue mulching and the application of mineral fertilizers and manure caused an increase in the abundance of the soil nematofauna and modified its composition: facultative plant feeders and omnivores-predators were more abundant, and the community was more complex, under no tillage with residue mulching than under conventional tillage with residue removing; in addition most trophic groups were more abundant when manure and mineral fertilizers were applied. By contrast, there was little difference in nematofauna between soils under soybean (2005) and under rice (2006).

NIRS prediction of nematode abundances

Abundances were addressed in absolute values, in number of individuals per gram of dry soil. Calibration was hardly possible for taxa that were present in few samples thus prediction was only carried out for the six most frequent taxa, which were also abundant: *Acrobeloides*, *Aphelenchoides*, *Xiphinema*, *Pratylenchus*, Tylenchidae, and Dorylaimidae. Prediction was also carried out for trophic groups. Considering total population, trophic groups and frequent taxa in 2005, 2006 and over both years, 36 abundances were predicted by NIRS (Table 2). The spectrum pre-processing method that yielded the best predictions depended on the nematode group: for the 36 nematode groups considered, the best predictions were achieved with no derivation for 12 groups, with first-order derivation for 18 groups, and with second-order derivation for six groups; they were achieved with no other transformation for seven groups, with SNV for 15 groups, with SNVD for eight groups, with D for two groups, and with MSC for four groups; SNV with no derivation (SNV 01; six groups) and with first-order derivation (SNV 14, 15 and 18; five groups) were the pre-processing methods that yielded the best predictions most often. Very few samples were spectral outliers: none for the 2005 set, none for the 2006 set, and possibly three for the total set but never more than one for a given pre-treatment. Most outliers were calibration outliers, the proportion of which ranged from 0 to 23% depending on the set and variable, and averaged 12% (Table 2). The prediction of total abundance was accurate for 2005 (Q^2 = 0.75, SECV = 41% of the mean, RPD = 2.0) and acceptable for 2006 and the total set (Q^2 = 0.63-0.66, SECV = 53-54%, RPD = 1.6-1.7; Figure 1). The prediction of abundance was acceptable for several taxa: *Acrobeloides*, *Aphelenchoides* and *Pratylenchus* in 2005, and *Xiphinema* and Dorylaimidae in 2006 (Q^2 = 0.63-0.72, SECV = 57-92%, RPD = 1.6-1.9); it was quasi-acceptable for *Pratylenchus* over both years and Dorylaimidae for 2005 and over both years (Q^2 = 0.52-0.55, SECV = 93-109%, RPD = 1.4-1.5). Trophic groups were better predicted than taxa: the prediction of abundance was accurate for fungal feeders in 2005 and omnivores-predators in 2006 (Q^2 = 0.76-0.77, SECV = 45-57%, RPD = 2.0-2.1); it was acceptable for bacterial feeders in 2005, obligate plant feeders in 2006, and omnivores-predators over both years (Q^2 = 0.64-0.65, SECV = 65-70%, RPD = 1.7), and quasi-acceptable for bacterial feeders in 2006, obligate plant feeders in 2005 and over both years, and omnivores-predators in 2005 (Q^2 = 0.52-0.58, SECV = 63-67%, RPD = 1.4-1.5). On the whole, predictions were thus more accurate for "yearly sets" (i.e. 2005 or 2006) than over the total set, and considering yearly sets, the abundance of most trophic groups was predicted in an accurate, acceptable or quasi-acceptable manner (except fungal feeders in 2006 and facultative plant feeders in 2005 and in

2006). Comparison between conventional determination and NIRS prediction of trophic group abundance in 2005 is presented in Figure 2.

Nematode abundances were also addressed in relative values, in proportions of total abundance; but NIRS prediction of relative abundances was poor (except for bacterial feeders and for omnivores-predators in 2006) and was less accurate than absolute abundance prediction (data not shown).

NIRS prediction of nematode ecological indexes

Prediction was carried out for the four main ecological indexes of nematodes in 2005, 2006, and over both years (i.e. 12 cases). The best predictions were achieved with no derivation for six cases and with first-order derivation for the six other cases; they required no other transformation for one case, SNV for seven cases, D for two cases, and MSC for two cases; SNV with first-order derivation (SNV 14, 17 and 18; four cases) and with no derivation (SNV 01; three cases) were the pre-processing methods that yielded the best predictions most often. The proportion of calibration outliers ranged from 4 to 23% and averaged 12% (Table 3). Predictions using NIRS were acceptable for the Plant parasitic index in 2006 ($Q^2 = 0.62$, SECV = 11% of the mean, RPD = 1.6) and quasi-acceptable for the Structure index and the Maturity index in 2006 ($Q^2 = 0.54$-0.59, SECV = 14-27%, RPD = 1.5). Thus, most indexes were acceptably or quasi-acceptably predicted in 2006 (except the Enrichment index) but not in 2005 or over both years. It is worth noting that the variability of these indexes was low in general (the ratio of standard deviation SD to mean averaged 29%) when compared with the variability of abundances (SD / mean ≈ 110% in average). Consequently, the error of prediction (SECV) represented a relatively high proportion of SD (i.e. small RPD) but a low proportion of the mean (23% in average), which could result in contradictory appreciations of prediction accuracy regarding these indexes.

Most contributing wavelengths

Spectra being condensed (one data point every 8 nm), 173 absorbances were considered per NIR spectrum. The figure 3 presents the coefficients of linear regression of total and trophic group abundances on absorbance at every data point in the 2005 and 2006 samples.

Out of the 10 wavelengths that had the heaviest weight in NIRS prediction of total abundance in 2005, five were in the range 1748-1796 nm (including the first three in the range 1748-1764 nm) and two in the range 1300-1308 nm; by contrast, in 2006, seven were in the range

2044-2108 nm (including the first four in the range 2076-2100 nm) and two in the range 1132-1140 nm.

Most contributing wavelengths of bacterial feeder abundance in 2005 were in the ranges 1812-1860 and 1724-1740 nm (seven and three out of 10, respectively), and for 2006 samples, in the ranges 1756-1796 and 2116-2132 nm (six and three out of 10, respectively).

Main predictors of fungal feeder abundance were more dispersed: when ranked in decreasing weight order, in 2005 they were in the ranges 1388, 1996-2020, 2460-2476 and 1836-1860 nm (one, four, two and three out of the 10 heaviest contributors, respectively), and in 2006, in the ranges 2364-2380, 1708-1732 and 1676-1684 nm (two, three and two out the 10 heaviest contributors, respectively).

Similarly, main predictors of obligate plant feeder abundance in 2005 were dispersed: 2316-2348, 2460-2484 and 1764-1788 nm (two, two and two out of the 10 heaviest contributors, respectively). By contrast, for 2006 samples, most were in the range 1980-2092 nm (six out of the 10 heaviest ones, including the first two, 2084 and 2076 nm).

Spectral ranges that contributed most to facultative plant feeder abundance were comparable in 2005 and 2006: 1484-1540 nm included the five heaviest contributors for both years, and 1452-1524 nm even included the 10 heaviest contributors for 2006.

Main predictors of omnivores-predator abundance were in neighbouring ranges in 2005 and 2006: 1988-2052 nm for 2005 (five out of the 10 heaviest contributors, including the first two, 1996 and 1988 nm) and 2044-2092 nm for 2006 (five out of the 10 heaviest contributors, including the first two, 2084 and 2076 nm); in addition, the range 1796-1820 nm was important in 2005 (four out of the 10 heaviest contributors).

Spectral ranges that contributed most to NIRS predictions often differed for 2005 and 2006 samples, but they were sometimes similar in both years (e.g. 1484-1524 nm for facultative plant feeders and 2044-2052 nm for omnivores-predators). In addition, some ranges were much involved in the prediction of the abundance of several groups:

- 1132-1156 nm for total abundance in both years, and obligate plant feeders and omnivores-predators in 2006 (and to a lesser extent, facultative plant feeders in 2005);
- 1372-1388 nm for fungal feeders and facultative plant feeders in 2005 (and 2006), and omnivores-predators (and total abundance) in 2006;
- 1468-1492 nm for facultative plant feeders in both years, and obligate plant feeders and omnivores-predators in 2006 (and total abundance, obligate plant feeders and omnivores-predators in 2005, and bacterial feeders and fungal feeders in 2006);

- 1556-1572 nm for obligate plant feeders and omnivores-predators in 2006 (and total abundance and bacterial feeders in 2006, and omnivores-predators in 2005);
- 1748-1804 nm (1748-1764 nm especially) for total abundance, bacterial feeders, obligate plant feeders and omnivores-predators in both years (in 2005 especially);
- 1820-1836 nm for bacterial feeders, fungal feeders and omnivores-predators in 2005 (and obligate plant feeders in 2006);
- 1980-1996 nm for total abundance and obligate plant feeders in both years, fungal feeders in 2005, and omnivores-predators in 2005 (and 2006);
- 2044-2092 for omnivores-predators in both years, and total abundance and obligate plant feeders (and bacterial feeders and fungal feeders) in 2006;
- 2316-2340 nm for total abundance and obligate plant feeders in 2005, and fungal feeders in 2006 (and bacterial feeders in both years, facultative plant feeders in 2005, and obligate plant feeders and omnivores-predators in 2006);
- 2444-2476 nm for fungal feeders in both years, obligate plant feeders and facultative plant feeders in 2005 (and total abundance and omnivores-predators in 2005, and facultative plant feeders in 2006).

Discussion

Nematofauna abundance and composition in the studied samples and the way they were affected by cropping systems have been discussed by Villenave et al. (2009a) and are not re-discussed here. Nevertheless it is important to notice that nematofauna abundance and composition varied markedly according to the cropping system, as observed elsewhere in Madagascar (Villenave et al., 2009b) and in other countries (Lenz and Einsenbeis; 2000; Okada and Harada, 2007; Rahman et al., 2007). From an ecological viewpoint, this nematode dataset thus presented diversity and variations that made it an appropriate basis for a NIRS study.

NIRS prediction of nematode abundance

Among the pre-processing methods tested, SNV (whereby spectra are mean-centred and variance-scaled) and first-order derivation, which both correct the baseline and enhance weak signals, were most often the best pre-processing methods, as often observed for NIRS predictions of soil properties (Nduwamungu et al., 2009; Stenberg et al., 2010). As often reported too, there is no pre-processing method that works with all soil data sets, and selecting the most appropriate method for a given data set requires trial-and-error procedures

(Nduwamungu et al., 2009; Stenberg et al., 2010). The reasons why a given pre-processing method yields the best predictions of a given soil property, and how this depends on the property and data set considered, are unclear and have not really been addressed in the literature to date.

The small number of spectral outliers resulted from the homogeneity of the sample set, as all samples originated from a relatively small area and from the same depth layer. The proportion of calibration outliers was relatively high, but similar levels have been reported for low concentration analytes (e.g. Brunet et al., 2009; Barthès et al., 2010).

On the whole, NIRS determination of nematode abundances (total, trophic groups, dominant and frequent taxa) and ecological indexes could hardly be considered accurate as RPD rarely reached 2, which has been considered a threshold for accurate prediction of soil properties (Chang et al., 2001; Dunn et al., 2002). However, this overall moderate accuracy of NIRS prediction of nematode abundance has to be nuanced. Firstly, though they were not estimated (no laboratory replication), it is likely that variations in nematofauna and in its morphological characterization, for a given sample, were not negligible. As a consequence, apparent imprecision of NIRS predictions might partly reflect nematofauna heterogeneity, and possibly, imperfect repeatability of the morphological characterization. Moreover, several parameters considered were accurately predicted by NIRS: total abundance in 2005 and abundances of fungal feeders in 2005 and omnivores-predators in 2006. Other parameters were acceptably or quasi-acceptably predicted so that: (i) the abundance of most trophic groups was fairly well predicted in 2005 and 2006 (except facultative plant feeders); and (ii) most ecological indexes were fairly well predicted in 2006 (except the Enrichment index). This underlines that NIRS predictions were more accurate for yearly sets, which included samples collected under a given crop, than for the total set, which included samples collected under rice or soybean. This confirmed that the accuracy of NIRS predictions tends to increase with sample set homogeneity (Brunet et al., 2007). In addition, results were better with 2006 than with 2005 samples, possibly relating to the plant cropped (rice in 2006 vs. soybean in 2005).

In addition, the accuracy of NIRS predictions tended to decrease from total abundance to trophic group abundance then to ecological indexes and taxon abundances. In addition, no-frequent taxa (represented in less than 70 samples) were hardly predictable. The results led to address the following question: what did NIRS “see” about soil nematodes?

It is possible that nematodes include specific constituents (e.g. fatty acids) that make them quantifiable by NIRS, and the same hypothesis might be proposed for trophic groups and even

for taxa. Soil content in such constituents is very low but several works have demonstrated that NIRS could determine soil content in compounds that are present at low concentration (e.g. magnitude of ppm; Coûteaux et al., 2003; Brunet et al., 2009). If so, decreasing accuracy of NIRS prediction of nematode abundance from total community to trophic groups then taxa might be attributed to higher soil content in specific constituents in the former than in the latters. Indeed, the 5-g samples than were scanned included on average 21 nematodes, 1.4 to 6.9 individuals from each trophic group, and 0.9 to 2.3 individuals from each of the six dominant taxa. However, to date, no specific constituents have been associated with nematodes in general or with any trophic group or taxon. Alternatively, or additionally, NIRS predictions regarding nematofauna might be indirect, relating to soil properties and/or land uses that determined nematofauna abundance and composition. Indeed, nematofauna abundance and composition reflect soil conditions reliably and were found to provide good indicators of soil functioning, especially in the experiment considered (Villenave et al., 2009a): differences in the soil nematode community were linked to differences in soil properties due to differences in land management. The point is that NIR spectra also provide integrated information on soil conditions (Cécillon et al., 2009). It is worth noting that NIRS predictions of nematofauna abundance and composition did not result from NIRS prediction of soil content in organic carbon (or total nitrogen) and correlations between the latter and nematofauna parameters. Indeed, correlations between nematofauna parameters and soil organic carbon (or total nitrogen) were much smaller than corresponding Q^2 and/or close to 0 (data not shown); thus they were not sufficient for explaining indirect NIRS prediction of nematofauna parameters as a result of NIRS prediction of soil organic carbon (or total nitrogen) content.

The abundance of facultative plant feeders was less accurately predicted by NIRS than that of other trophic groups. This might be caused by the small size of the nematodes considered. The fact that the abundance of a group including small individuals was poorly predicted suggested that prediction might be based to some extent on nematode constituents rather than on soil conditions. However, this trophic group was the only one that was not found to vary significantly according to tillage, inputs or sampling year (Villenave et al., 2009a). This suggested that parameters varying with soil conditions might be better predicted by NIRS, thus that NIRS might consider soil conditions possibly reflected in nematofauna rather than nematofauna directly.

The subsamples scanned were ca. 5 g, thus much smaller than those used for morphological determination (> 100 g), and as mentioned above, included an average of 1 to 7 individuals of

each trophic group. This was probably too small to allow representative sampling of trophic groups and taxa, and might be an important reason for overall moderate accuracy of NIRS predictions here. Thus it is recommended to scan larger samples (e.g. 100 g), using larger sample containers. This was not possible for the present study because spectrum acquisition was carried out after nematofauna analysis, and little sample mass was available for scan.

In addition, soil samples had not been dried and were almost at field moisture when nematode extraction was carried out, while NIRS analysis was carried out after sample air-drying. Air-drying of course affected nematodes and killed many of them while others survived in anhydrobiotic state. Therefore, it is likely that scanning un-dried samples would improve NIRS prediction of nematofauna abundance and composition as determined by morphological characterization on moist samples, though NIRS predictions of soil attributes have been reported to be adversely affected by soil water (Chang et al., 2005).

Beside probable improvement in prediction accuracy through the scan of larger and un-dried samples, the validity of calibrations should be extended. Indeed, the calibrations presented here have been built using a set of clayey topsoil samples originating from a small area, and their validity is mainly limited to comparable soil samples. Extending the validity of such calibrations is a challenge that requires joint acquisition of nematofaunal and spectral data over a wide range of soil conditions (mineralogy, texture, etc.) and uses (native vegetation, pastures, etc.). Applications of NIRS to the characterization of soil microbial community composition, as proposed by Zornoza et al. (2008), has also been confronted to the challenge of extending calibration validity, which is a recurrent concern in NIRS studies, especially for soil applications (Brown et al., 2006).

As regards nematodes, studying more samples would additionally allow predictions regarding taxa that were present in too few samples here for allowing calibration.

Most contributing wavelengths

According to literature (mainly Shenk et al., 2001, and Workman and Weyer, 2008), many of the spectral regions that contributed heavily to NIRS prediction of nematofauna abundance and composition corresponded to the vibration of chemical bonds, often in particular molecules, as presented in Table 4. This allowed interpreting NIRS prediction models. In short, predictions involved spectral ranges attributed:

- in the case of bacterial feeders, to aliphatic hydrocarbons and cellulose, and secondarily to amines, cyclic amides and proteins;

- in the case of fungal feeders, to aliphatic and aromatic hydrocarbons, amines (including aromatic ones), amides (including proteins), and lipids;
- in the case of obligate plant feeders, to amides, especially in animal proteins (including digestive enzymes), to aliphatic and aromatic hydrocarbons, to lipids, and to a lesser extent, to amines;
- in the case of facultative plant feeders, to amines, amides, aliphatic and aromatic hydrocarbons, and lipids;
- in the case of omnivores-predators, to amides, animal proteins (including digestive enzymes), amines (especially aromatic ones), to cellulose and polysaccharides, and to a lesser extent, to aromatic hydrocarbons;
- and finally, as regarded the total nematode community, to amides, animal proteins (including digestive enzymes), aliphatic hydrocarbons, and to a lesser extent, to aromatic hydrocarbons and amines.

This suggested that NIRS prediction of bacterial and fungal feeder abundances depended, at last partly, on the substrates utilized by the preys of the trophic groups considered: mainly aliphatic hydrocarbons for bacteria, but both aliphatic and aromatic hydrocarbons for fungi (which can use lignin as substrate). By contrast, the weight of animal proteins in the predictions of obligate plant feeders, omnivores-predators and total abundances suggested NIRS predictions were influenced by nematode constituents. The importance of nitrogen-containing molecules such as amines, amides and proteins for predicting the abundance of all groups also suggested that NIRS predictions might be affected by nematode constituents, as animal matters include more nitrogen than plant and microbial matters. On the whole, NIRS prediction of nematofauna abundance and composition seemed to take account of constituents of both nematodes and preys' food, thus to be based on both nematofauna and soil organic properties reflected by nematofauna abundance and composition.

Differences in heavily contributing wavelengths of some groups' abundance between years might be due to differences in soil conditions, especially in recent organic matter, which could mask or alter some spectral features.

Conclusion

Worthwhile NIRS predictions of nematofauna abundance and composition were achieved for the studied sample set. The prediction of abundances was fairly good for the total community and acceptable for most trophic groups and some frequent and abundant taxa. In contrast, ecological indexes that describe nematode community structure were poorly predicted in

general. Some elements suggested NIRS predictions might be direct, based on nematode constituents; other elements suggested they might be indirect, based on soil attributes reflected by nematofauna.

These results are promising. Indeed, they show that morphological characterization of nematofauna, which requires high taxonomic expertise, could be used for building calibrations that would allow nematofauna characterization of many new soil samples just by scanning them. However, the calibrations presented here have been built using a sample set originating from a small area; their validity is limited and has to be extended over a wider range of soil conditions and uses. Improving the accuracy of NIRS predictions of nematofaunal data is also a challenge, which could be addressed by scanning larger and un-dried soil samples. Improvement in prediction accuracy and extension of calibration validity should optimize the performance of NIRS for soil nematofauna characterization. It will then be possible to assess how operational NIRS is as regards soil nematode community.

Acknowledgements

The sampling work was funded by the French National Program ACI-FNS Continental Ecosphere ECCO-2004 Muten. Nematofauna analyses were conducted in the project Nemageco-Icones funded by the French ADEME.

References

Al-Abbas, A.H., Swain, P.H., Baumgardner, M.F., 1972. Relating organic matter and clay content to the multispectral radiance of soils. Soil Science 114, 477-485.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Applied Spectroscopy 43, 772-777.

Barthès, B.G., Brunet, D., Ferrer, H., Chotte, J.L., Feller, C., 2006. Determination of total carbon and nitrogen content in a range of tropical soils using near infrared spectroscopy: influence of replication and sample grinding and drying. Journal of Near Infrared Spectroscopy 14, 341-348.

Barthès, B.G., Brunet, D., Hien, E., Enjalric, F., Conche, S., Freschet, G., d'Annunzio, R., Toucet-Louri, J., 2008. Determining the distributions of soil carbon and nitrogen in particle size fractions using near infrared reflectance spectrum of bulk soil samples. Soil Biology & Biochemistry 40, 1533-1537.

Barthès, B.G., Brunet, D., Brauman, A., Fromin, N., Lensi, R., Volant, A., Laclau, J.P., Blavet, D., Chapuis-Lardy, L., 2010. Determination of potential denitrification in a range of

tropical topsoils using near infrared reflectance spectroscopy (NIRS). Applied Soil Ecology 46, 81-89.

Bongers, T., 1990. The maturity index: an ecological measure of environmental disturbance based on nematode species composition. Oecologia 83, 14-19.

Bongers, T., Ferris, H., 1999. Nematode community structure as a bioindicator in environmental monitoring. Tree 14, 224-228.

Bowers, S.A., Hanks, R.J., 1965. Reflection of radiant energy from soils. Soil Science 100, 130-138.

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

Brunet, D., Woignier, T., Lesueur-Jannoyer, M., Achard, R., Rangon, L., Barthès, B.G., 2009. Determination of soil content in chlordecone (organochlorine pesticide) using near infrared reflectance spectroscopy (NIRS). Environmental Pollution 157, 3120-3125.

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

Cécillon, L., Barthès, B.G., Gomez, C., Ertlen, D., Genot, V., Hedde, M., Stevens, A., Brun, J.J., 2009. Assessment and monitoring of soil quality using near infrared reflectance spectroscopy (NIRS). European Journal of Soil Science 60, 770-784.

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

Chang, C.W., Laird, D.A., Hurburgh, C.R.J., 2005. Influence of soil moisture on near-infrared reflectance spectroscopic measurements of soil properties. Soil Science 170, 244-255.

Coûteaux, M.M., Berg, B., Rovira, P., 2003. Near infrared reflectance spectroscopy for determination of organic matter fractions including microbial biomass in coniferous forest soils. Soil Biology & Biochemistry 35, 1587-1600.

Cozzolino, D., Restaino, E., Fassio, A., 2010. Discrimination of yerba mate (*Ilex paraguayensis* St. Hil.) samples according to their geographical origin by means of near infrared spectroscopy and multivariate analysis. Sensing and Instrumentation for Food Quality and Safety 4, 67-72.

Donn, S., Griffiths, B.S., Neilson, R., Daniell, T.J., 2008. DNA extraction from soil nematodes for multi-sample community studies. Applied Soil Ecology 38, 20-26.

Dunn, B.W., Beecher, H.G., Batten, G.D., Ciavarella, S., 2002. The potential of near infrared reflectance spectroscopy for soil analysis – a case study from the Riverine Plain of south eastern Australia. Australian Journal of Experimental Agriculture 42, 607-614.

Ekschmitt, K., Bakonyi, G., Bongers, M., Bongers, T., Boström, S., Dogan, H., Harrison, A., Nagy, P., O'Donnell, A.G., Papatheodorou, E.M., Sohlenius, B., Stamou, G.P., Wolters, V., 2001. Nematode community structure as indicator of soil functioning in European grassland soils. European Journal of Soil Biology 37, 263-268.

Ferris, H., Bongers, T., de Goede, R., 2001. A framework for soil food web diagnostics: extension of the nematode faunal analysis concept. Applied Soil Ecology 18, 13-29.

Floyd, R., Abebe, E., Papert, A., Blaxter, M., 2002. Molecular barcodes for soil nematode identification. Molecular Ecology 11, 839-850.

Freckman, D.W., Blackburn, T.H., Brussaard, L., Hutchings, P., Palmer, M.A., Snelgrove, P.V.R., 1997. Linking biodiversity and ecosystem functioning of soils and sediments. Ambio 26, 556-562.

ISO (International Organization for Standardization), 2007. Soil Quality - Sampling of Soil Invertebrates – Part 4: Sampling, Extraction and Identification of Soil-Inhabiting Nematodes, ISO 23611-4. International Organization for Standardization, Geneva, Switzerland.

Lenz, R., Eisenbeis, G. , 2000. Short-term effects of different tillage in a sustainable farming system on nematode community structure. Biology and Fertility of Soils 31, 237-244.

Mark, H.L., Tunnell, D., 1985. Qualitative near-infrared reflectance analysis. Analytical Chemistry 58, 379-384.

Martens, H., Jensen, S.A., Geladi, P., 1983. Multivariate linearity transformation for near infrared reflectance spectrometry, in: Proceedings of the Nordic Symposium on Applied Statistics. Stokkand Forlag Publishers, Stavanger, Norway, pp. 205-234.

Nduwamungu, C., Ziadi, N., Parent, L.E., Tremblay, G.F., Thuriès, L., 2009. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review. Canadian Journal of Soil Science 89: 531-541.

Okada, H., Harada, H., 2007. Effects of tillage and fertilizer on nematode communities in a Japanese soybean field. Applied Soil Ecology 35, 582-598.

Palmborg, C., Nordgren, A., 1993. Modelling microbial activity and biomass in forest soil with substrate quality measured using near infrared reflectance spectroscopy. Soil Biology & Biochemistry 25, 1713-1718.

Perry, R.N., Jones, J.T., 1998. The use of molecular biology techniques in Plant Nematology: past, present and future. Russian Journal of Nematology 6, 47-56.

Rahman, L., Chan, K.Y., Heenan, D.P., 2007. Impact of tillage, stubble management and crop rotation on nematode populations in a long-term field experiment. Soil & Tillage Research 95, 110-119.

Raunet, M., 1981. Le Milieu Physique de la Région Volcanique Ankaratra - Vakinankaratra - Itasy (Madagascar). Aptitude à la Culture du Blé Pluvial. Ministère du Développement Rural et de la Réforme Agraire, Antananarivo, Madagascar.

Razafimbelo, T.M., 2005. Stockage et Protection du Carbone dans un Sol Ferrallitique sous Systèmes en Semis Direct avec Couverture Végétale des Hautes Terres Malgaches. Ph.D. Thesis, Ecole Nationale Supérieure Agronomique de Montpellier, Montpellier, France.

Reeves III, J.B., McCarty, G.W., Mimmo, T., 2002. The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. Environmental Pollution 116, S277-S284.

Roberts, C.A., Workman Jr., J., Reeves III, J.B., 2004. Near-Infrared Spectroscopy in Agriculture. ASA, CSSA, SSSA, Madison, WI.

Sanchez-Moreno, S., Ferris, H., 2007. Suppressive service of the soil food web: Effects of environmental management. Agriculture, Ecosystems & Environment 119, 75-87.

Seinhorst, J.W., 1962. Modifications of the elutriation method for extracting nematodes from soil. Nematologica 8, 117-128.

Shenk, J.S., Westerhaus, M.O., 1991a. Population definition, sample selection and calibration procedures for near infrared reflectance spectroscopy. Crop Science 31, 469-474.

Shenk, J.S., Westerhaus, M.O., 1991b. Population structuring of near infrared spectra and modified partial least square regression. Crop Science 31, 1548-1555.

Shenk, J.S., Workman Jr., J., Westerhaus, M.O., 2001. Application of NIR spectroscopy to agricultural products, in: Burns, D.A., Ciurczak, E.W. (Eds.), Handbook of Near-Infrared Analysis, second ed. CRC Press, New York, pp. 419-474.

Soil Survey Staff, 2003. Keys to Soil Taxonomy. Natural Conservation Service, US Department of Agriculture, Washington DC.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science, in: Sparks, D.L. (Ed.), Advances in Agronomy, Vol. 107. Academic Press, Burlington, MA, pp. 163-215.

Terhoeven-Urselmans, T., Michel, K., Helfrich, M., Flessa, H., Ludwig, B., 2006. Near-infrared spectroscopy can predict the composition of organic matter in soil and litter. Journal of Plant Nutrition and Soil Science 169, 168-174.

Villenave, C., Ekschmitt, K., Nazaret, S., Bongers, T., 2004. Interactions between nematodes and microbial communities in a tropical soil following manipulation of the soil food web. Soil Biology & Biochemistry 36, 2033-2043.

Villenave, C., Ba, A.O., Rabary, B., 2009a. Analyse du fonctionnement biologique du sol par l'étude de la nématofaune : Semis direct versus labour sur les hautes terres près d'Antsirabé (Madagascar). Etude et Gestion des Sols 16: 369-378.

Villenave, C., Rabary, B., Chotte, J.L., Blanchart, E., Djigal, D., 2009b. Impact of direct seeding mulch-based cropping systems on soil nematodes in a long-term experiment in Madagascar. Pesquisa Agropecuaria Brasileira 44, 949-953.

Workman Jr, J., Weyer, L., 2007. Practical Guide to Interpretive Near-Infrared Spectroscopy. CRC Press, Boca Raton, FL.

Yeates, G.W., 2003. Nematodes as soil indicators: functional and biodiversity aspects. Biology and Fertility of Soils 37, 199-210.

Yeates, G.W., Bongers, T., de Goede, R.G.M., Freckman, D.W., Georgieva, S.S., 1993. Feeding habits in soil nematode families and genera - An outline for soil ecologists. Journal of Nematology 25, 315-331.

Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V., Mataix-Beneyto, J., 2008. Near infrared spectroscopy for determination of various physical, chemical and biochemical properties in Mediterranean soils. Soil Biology & Biochemistry 40, 1923-1930.

Table 1. Soil nematode abundance (dominant and frequent taxa, trophic groups, and total community) and ecological indexes for the studied soil sample sets (year 2005, year 2006, and total 2005+2006) according to morphological characterization.

	2005				2006				Total set			
	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD
	(individuals 100 g^{-1} dry soil)											
Abundance of most abundant taxa												
Acrobeloides	0	184	33	37	0	234	22	42	0	234	28	40
Aphelenchoides	0	207	38	46	0	108	22	25	0	207	30	38
Xiphinema	0	131	22	31	0	148	24	34	0	148	23	32
Pratylenchus	0	355	72	102	0	217	34	51	0	355	54	83
Tylenchidae	0	178	38	40	0	280	38	61	0	280	38	51
Dorylaimidae	0	176	19	35	0	206	38	54	0	206	28	46
Abundance of trophic groups												
Bacterial feeders	0	732	183	195	0	1408	141	235	0	1408	163	215
Fungal feeders	0	273	56	61	0	452	48	65	0	452	52	63
Obligate plant feeders	0	1191	197	220	0	617	122	137	0	1191	161	187
Facultative plant feeders	0	178	38	40	0	280	38	61	0	280	38	51
Omnivores and predators	0	412	51	70	0	410	87	101	0	412	68	88
Total abundance	0	2102	525	453	0	2389	436	484	0	2389	482	468
Enrichment index (EI)	0	93	36	23	0	69	35	18	0	93	36	21
Structure index (SI)	0	96	73	17	0	97	69	26	0	97	71	22
Maturity index (MI)	1.3	4.1	2.8	0.5	2.0	3.9	3.0	0.6	1.3	4.1	2.9	0.6
Plant parasite index (PPI)	2.2	4.2	3.1	0.4	2.0	4.2	3.1	0.5	2.0	4.2	3.1	0.5

Min is the minimum, Max the maximum, and SD the standard deviation.

Table 2. Cross-validation and calibration statistics regarding NIRS prediction of soil nematode abundance (dominant and frequent taxa, trophic groups, and total community) for yearly and total sample sets using the pre-treatment that provided the best results (cf. the Spectrum acquisition and pre-processing subsection of the Materials and methods section).

	Set	Pre-treatment	*N*	Outliers (%)	Mean (individuals 100 g^{-1} dry soil)	SD	SEC	SECV	R^2_{cal}	Q^2	RPD
Most abundant and frequent taxa											
Acrobeloides	2005	None 01	48	9.4	26.5	25.4	14.9	15.4	0.65	0.63	1.6
	2006	SNV 25	42	16.0	14.2	19.2	9.7	15.7	0.75	0.35	1.2
	total	None 01	96	4.0	19.6	22.2	16.8	19.7	0.42	0.21	1.1
Aphelenchoides	2005	MSC 01	47	11.3	31.0	33.8	16.8	17.8	0.75	0.72	1.9
	2006	SNV 210	43	14.0	18.9	19.3	9.5	16.1	0.76	0.33	1.2
	total	None 14	84	16.0	24.8	26.7	25.9	26.9	0.06	0.00	1.0
Xiphinema	2005	None 14	45	15.1	19.1	26.9	18.2	20.5	0.54	0.42	1.3
	2006	D 01	44	12.0	21.3	30.0	14.6	17.0	0.76	0.68	1.8
	total	SNV 01	93	7.0	18.1	24.0	17.7	18.5	0.46	0.40	1.3
Pratylenchus	2005	SNV 15	43	18.9	61.1	91.8	40.1	56.4	0.81	0.64	1.6
	2006	SNVD 17	47	6.0	27.2	36.7	26.8	31.7	0.47	0.25	1.2
	total	SNVD 210	98	2.0	45.0	70.4	41.6	49.1	0.65	0.52	1.4
Tylenchidae	2005	MSC 14	47	11.3	33.7	29.8	19.0	24.0	0.59	0.36	1.2
	2006	SNV 01	44	12.0	20.3	21.7	20.1	21.1	0.14	0.07	1.0
	total	SNV 14	85	15.0	26.4	25.7	20.2	23.6	0.38	0.16	1.1
Dorylaimidae	2005	SNVD 18	41	22.6	9.1	12.6	7.0	8.5	0.69	0.55	1.5
	2006	SNVD 110	44	12.0	25.9	34.4	17.8	20.6	0.73	0.64	1.7
	total	None 14	92	8.0	19.3	28.8	15.7	19.3	0.70	0.55	1.5
Abundance of trophic groups											
Bacterial feeders	2005	None 01	49	7.5	154.0	168.8	91.3	100.6	0.71	0.64	1.7
	2006	SNVD 18	43	14.0	81.8	78.6	39.4	54.3	0.75	0.52	1.4
	total	SNV 01	89	11.0	107.6	112.1	72.9	82.9	0.58	0.45	1.4
Fungal feeders	2005	SNV 01	42	20.8	44.1	40.9	16.6	19.7	0.84	0.77	2.1
	2006	D 14	42	16.0	39.1	30.9	16.8	28.0	0.70	0.20	1.1
	total	None 01	86	14.0	40.8	35.4	28.8	32.0	0.34	0.20	1.1
Obligate plant feeders	2005	MSC 14	49	7.5	171.2	167.5	77.1	110.9	0.79	0.57	1.5
	2006	SNV 210	43	14.0	98.7	109.9	41.8	65.0	0.86	0.65	1.7
	total	SNV 18	98	2.0	143.7	146.4	85.8	95.9	0.66	0.57	1.5
Facultative plant feeders	2005	MSC 14	47	11.3	33.7	29.8	19.0	24.0	0.59	0.36	1.2
	2006	SNV 01	44	12.0	20.3	21.7	20.1	21.1	0.14	0.07	1.0
	total	SNV 14	85	15.0	26.4	25.7	20.2	23.6	0.38	0.16	1.1
Omnivores and predators	2005	SNVD 18	45	15.1	40.5	39.5	18.8	25.6	0.77	0.58	1.5
	2006	SNVD 210	50	0.0	86.5	100.8	35.1	49.7	0.88	0.76	2.0
	total	SNVD 18	97	3.0	58.7	69.7	33.2	41.3	0.77	0.65	1.7
Total abundance	2005	SNV 18	51	3.8	488	400	151	201	0.86	0.75	2.0
	2006	SNV 210	47	6.0	349	321	134	189	0.83	0.66	1.7
	total	SNV 01	94	6.0	413	358	198	218	0.70	0.63	1.6

N is the number of samples (outliers being deleted); SD is the standard deviation; SEC and SECV are standard errors of calibration and cross validation, respectively; R^2_{cal} and Q^2 are determination coefficients of calibration and cross validation, respectively; RPD is the ratio of SD to SECV.

Table 3. Cross-validation and calibration statistics regarding NIRS prediction of nemafaunal indexes for yearly and total sample sets using the pre-treatment that provided the best results (cf. the Spectrum acquisition and pre-processing subsection of the Materials and methods section).

Index	Set	Pre-treatment	*N*	Outliers (%)	Mean	SD	SEC	SECV	R^2_{cal}	Q^2	RPD
Enrichment index (EI)	2005	MSC 14	49	7.5	35.4	21.9	14.0	16.3	0.59	0.45	1.3
	2006	D 01	43	14.0	32.8	17.0	10.5	13.1	0.62	0.41	1.3
	total	SNV 14	94	6.0	35.3	19.8	16.7	18.4	0.29	0.15	1.1
Structure index (SI)	2005	SNV 01	41	22.6	74.6	12.2	9.7	11.8	0.36	0.06	1.0
	2006	SNV 01	46	8.0	67.4	26.4	16.9	17.9	0.59	0.54	1.5
	total	MSC 01	83	17.0	76.1	16.6	15.4	15.7	0.13	0.11	1.1
Maturity index (MI)	2005	SNV 17	51	3.8	2.79	0.54	0.33	0.42	0.62	0.37	1.3
	2006	SNV 01	43	14.0	2.92	0.61	0.34	0.40	0.69	0.59	1.5
	total	SNV 18	89	11.0	2.94	0.55	0.49	0.50	0.23	0.19	1.1
Plant parasitic index (PPI)	2005	SNV 14	47	11.3	3.02	0.40	0.25	0.30	0.60	0.44	1.3
	2006	D 14	44	12.0	3.08	0.54	0.26	0.34	0.76	0.62	1.6
	total	None 01	95	5.0	3.06	0.47	0.33	0.36	0.51	0.40	1.3

N is the number of samples (outliers being deleted); SD is the standard deviation; SEC and SECV are standard error of calibration and cross validation, respectively; R^2_{cal} and Q^2 are determination coefficient for calibration and cross validation, respectively; RPD is the ratio of SD to SECV.

Table 4. Wavelengths contributing heavily to NIRS prediction of soil nematode groups, and chemical functions assigned to them (according to Shenk et al., 2001, and Workman and Weyer, 2008, except when otherwise mentioned).

Wavelengths	Assigned groups	Predicted nematofaunal groups
1143	Aromatic hydrocarbons	Total abundance in both years; obligate plant feeders and omnivores-predators in 2006
1300	C-H bond (Cozzolino et al., 2010)	Total abundance in 2005
1370 1390	Aromatic hydrocarbons Aliphatic hydrocarbons	Fungal feeders in 2005; omnivores-predators in 2006; facultative plant feeders in both years
1460-1530	Amines, amides or proteins	Obligate plant feeders and omnivores-predators in 2006; facultative plant feeders in both years
1570	Amides	Obligate plant feeders and omnivores-predators in 2006
1680-1685	Aromatic hydrocarbons	Fungal feeders in 2006
1727 1735 1738 1740	Aliphatic hydrocarbons Amines Proteins Thiols	Bacterial feeders in 2005; fungal feeders in 2006
1762 1780	Aliphatic hydrocarbons Cellulose	Total abundance in 2005; bacterial feeders in 2006; obligate plant feeders and omnivores-predators in both years
1820	Cellulose	Bacterial feeders, fungal feeders and omnivores-predators in 2005
1978-1990	Aromatic amines, amides and proteins	Fungal feeders in 2005; total abundance, obligate plant feeders and omnivores-predators in both years
2030-2070	Proteins or amides	Total abundance and obligate plant feeders in 2006; omnivores-predators in both years
2075 2080 2083 2090-2100	Amides in animal ribonucleases (digestive enzymes) Urea and alcohols Animal proteins OH, COH or COOH (in polysaccharides especially)	Total abundance, obligate plant feeders and omnivores-predators in 2006
2120 2127	Cyclic amides Polyamides	Bacterial feeders in 2006
2320-2350	Complex sugars (e.g. cellulose)	Total abundance and facultative plant feeders in 2005; fungal feeders in 2006; obligate plant feeders in both years
2363 2380	Aliphatic hydrocarbons Lipids	Fungal feeders in 2006
2445, 2463, 2470 2458, 2470 2470 2477	Proteins Aliphatic hydrocarbons Lipids Aromatic hydrocarbons	Obligate plant feeders in 2005; fungal feeders and facultative plant feeders in both years

Figure 1. Comparison between morphological determination and NIRS prediction of total nematode abundance for the yearly sets and for the total set.

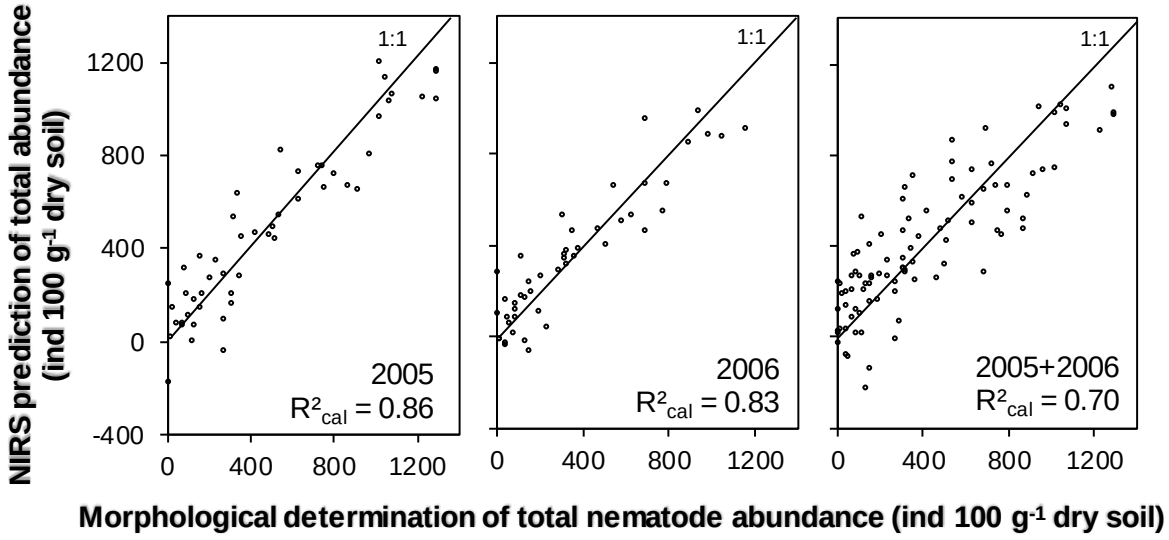

Figure 2. Comparison between morphological determination and NIRS prediction of trophic group abundances for the 2005 set.

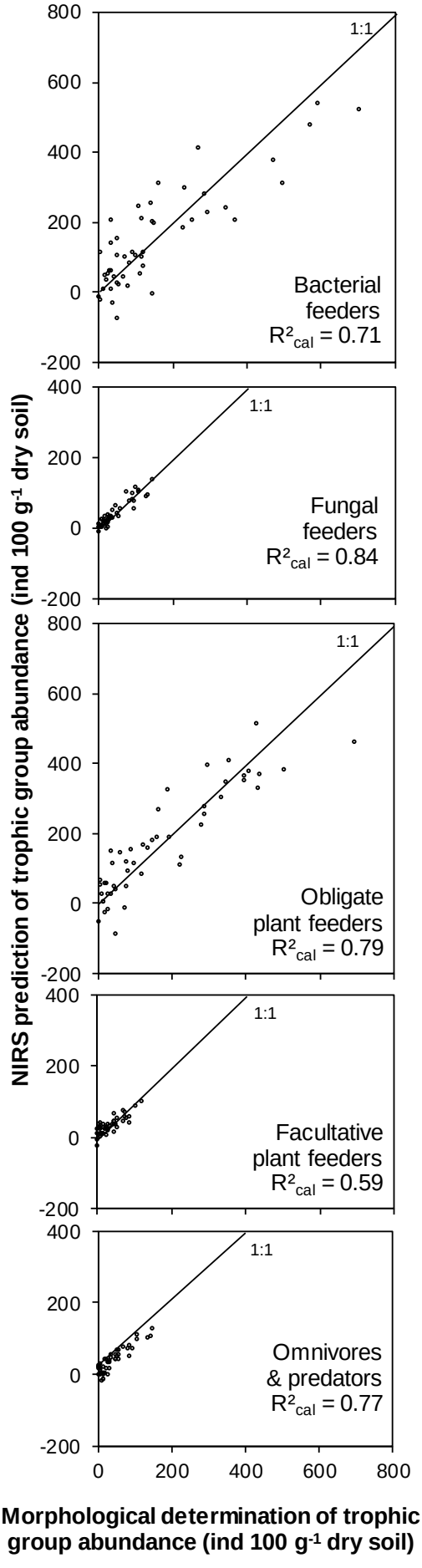

Figure 3. Coefficient of regression of total and trophic group abundances on absorbance at every data point (8-nm interval) of the NIR region for 2005 and 2006 sample sets.